Cornish-Fisher Distributions
Theory and Financial Applications

A thesis presented

by

Unai Ansejo Barra

to

Departamento de Fundamentos

del Análisis Económico II

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Quantitative Finance

Universidad del País Vasco

Bilbao, España

September 2006

Abstract

Cornish-Fisher Distributions: Theory and Financial Applications. Ph.D Thesis 2006. Unai Ansejo Barra. Departamento de Fundamentos del Análisis Económico II. Universidad del País Vasco.

Every financial theory is based on a model about the statistical nature of fluctuations of financial variables, being normality the standard statistical assumption in this area. In this work we propose several improvements to this normality hypothesis in order to incorporate in a better way extremal events, where fundamentally financial risks reside in.

Despite its popularity in empirical and theoretical finance, standard semi-parametric distributions based on Cornish-Fisher Expansions (Cornish and Fisher 1937), which allow the introduction of asymmetries and heavy tails, as Gram-Charlier (Charlier 1905) or Edgeworth (Edgeworth 1905), have long been recognized to be unsatisfactory due to their derivation of negative probabilities and lack of flexibility. These difficulties seem to be largely overcome by our new system, the Cornish-Fisher Density (CFD), which is also based on Cornish-Fisher Expansions. In the first Chapter we introduce the CFD functions in both its univariate and multivariate forms and analyze their contribution to static and dynamic models of asset price movements, which present stylized facts as conditional heteroskedasticity (GARCH, Bollerslev 1986) and dynamic conditional correlations (DCC, Engle 2002). We study the theoretical statistical properties of these distributions and analyze the in-sample goodness of fit of the different models to financial data as exchange rates and market indexes, comparing the results with other models standard in the literature. Our results show that these distributions are highly flexible, posses properties that are of interest for financial series like unimodality and, on the other hand, present good estimation results providing a good framework to analyze standard problems in mathematical finance.

Motivated by the results from Chapter 1, in Chapter 2 we study different areas of interest in mathematical finance: option valuation, measure of risk through the Value at Risk (VaR) and optimal portfolio selection under the assumption that financial variables follow a CFD.

In the optimal portfolio Section we test the Markowitz (Markowitz 1959) hypothesis that a mean-and-variance based analysis to construct optimal portfolios is enough to maximize investors' expected utility function. Many authors (e.g. Arditti 1967 and Samuelson 1970) have argued that the expected utility function may be more appropriately approximated by a function of higher moments, where investors are supposed to like positive skewness and dislike fat-tailedness as measured by kurtosis. On the other hand, early empirical evidence (e.g. Levy and Markowitz 1979 and Pulley 1981) suggests that a mean-variance optimization results in allocations that are similar to the ones obtained using a direct optimization of expected utility. Therefore, in order to provide more evidence on this issue, we have analyzed whether the inclusion of higher order moments improves the asset allocation in terms of utility using a Multivariate CFD model, finding that in many cases higher order moments cannot be discarded in the asset allocation process.

In the Value at Risk (Jorion 2000) Section we develop both an analytical and a simulation based framework to calculate the Value at Risk of portfolios of assets which follow a CFD, providing a graphic interface to perform easily these calculations. We compare the quality of these VaR calculations with other standard market models using a Backtesting, which clearly shows the out-performance of our model.

In the Options Section we generalize the Black and Scholes option model (Black and Scholes 1973) to include underlyings which follow a CFD density and obtain an analytical pricing model for vanilla options, which incorporates stylized facts as volatility smiles and has a simple interpretation. We also obtain analytical expression for the hedging parameters, which do not show anomalies present in other semi-parametric option pricing models (e.g. Corrado and Su 1997b and Jarrow and Rudd 1982). In addition, we compare in and out-sample estimations of prices options using Spanish options data, and find that our model out-performs the standard Black-Scholes model.

Resumen

Distribuciones de Cornish-Fisher : Teoría y Aplicaciones Financieras. Tesis 2006. Unai Ansejo Barra. Departamento de Fundamentos del Análisis Económico II. Universidad del País Vasco.

Toda teoría financiera esta basada en un modelo acerca de la naturaleza estadística de las fluctuaciones de las variables financieras, siendo la normalidad la hipótesis estadística estándar en este área. En este trabajo proponemos varias mejoras a esta hipótesis de normalidad para incorporar de una forma más adecuada los eventos extremos, que es donde residen fundamentalmente los riesgos financieros

A pesar de su popularidad en finanzas, las distribuciones semiparamétricas estándar basadas en Expansiones de Cornish-Fisher (Cornish and Fisher 1937) que permiten la existencia de asimetrías y colas pesadas, como Gram-Charlier (Charlier 1905) o Edgeworth (Edgeworth 1905), han sido permanentemente consideradas insatisfactorias debido a la presencia de probabilidades negativas y a su carencia de flexibilidad. Estas dificultades parecen haber sido superadas por nuestro nuevo sistema, la función de Densidad Cornish-Fisher (CFD), que también están basada en Expansiones Cornish-Fisher. En el primer capítulo introducimos las funciones CFD en sus formas univariante y multivariante y analizamos su contribución estática y dinámica a los movimientos de activos financieros, que presentan características

como la heterocedasticidad condicional (GARCH, Bollerslev 1986) y correlaciones condicionales dinámicas (DCC, Engle 2002). Estudiamos las propiedades estadísticas teóricas de estas distribuciones y analizamos la capacidad de ajuste de los diferentes modelos a datos financieros como tipos de cambio e índices de mercado bursátiles comparando los resultados con otros modelos estándar en la literatura. Estas distribuciones demuestran ser un buen marco para analizar problemas en matemáticas financieras dado que nuestros resultados muestran que son muy flexibles, presentan propiedades que son de interés para las series financieras como la unimodalidad y obtienen buenos resultados de estimación.

Motivados por los resultados del Capítulo 1, en el Capítulo 2 estudiamos diferentes áreas de interés en matemáticas financieras: la valoración de opciones, la medida del riesgo a través del Valor en Riesgo (VaR) y la selección óptima de carteras bajo la hipótesis de que las variables financieras siguen una distribución CFD.

En la Sección de selección de carteras testeamos la hipótesis de Markowitz (Markowitz 1959) de que una construcción de carteras óptimas basada en un análisis de media-varianza es suficiente para maximizar la utilidad esperada de los inversores. Varios autores (por ejemplo Arditti 1967 y Samuelson 1970) han argumentado que la función de utilidad esperada puede ser aproximada más apropiadamente por una función de los momentos de orden superior, donde se supone que los inversores valoran la asimetría positiva y evitan la presencia de eventos extremos, medida por el coeficiente de curtosis. Por otro lado, evidencia empírica (por ejemplo Levy and Markowitz 1979 y Pulley 1981) sugiere que los resultados de una optimización media-varianza proporcionan asignaciones similares a las obtenidas mediante la optimización directa de la función de utilidad. Por lo tanto, con el fin de proporcionar más evidencia sobre esta cuestión, hemos analizado si la inclusión de momentos de orden superior mejora la asignación de activos en términos de utilidad esperada utilizando una función multivariante CFD, encontrando que en varios casos los momentos de orden superior no pueden ser descartados en el proceso de asignación de activos.

En la Sección del Valor en Riesgo (Jorion 2000) desarrollamos dos marcos, uno analítico y otro basado en simulaciones, para calcular el Valor en Riesgo de carteras cuyos activos siguen una distribución CFD multivariante, proporcionando una aplicación informática gráfica para realizar fácilmente estos cálculos. Comparamos la calidad de las estimaciones del VaR con otros modelos estándar de mercado utilizando un Backtesting que muestra claramente la mejora de nuestro modelo.

En la Sección de opciones generalizamos el modelo de valoración de Black y Scholes (Black and Scholes 1973) para incluir subyacentes que posean una densidad CFD y obtenemos una fórmula analítica para la valoración de opciones eu-

ropeas vanilla, que incorpora características como la sonrisa de volatilidad y tiene una interpretación simple. Asimismo, obtenemos una expresión analítica para los parámetros de cobertura o Griegas que no muestran las anomalías presentes en otros modelos de valoración semiparamétricos (por ejemplo Corrado and Su 1997b y Jarrow and Rudd 1982). Adicionalmente, comparamos el modelo de valoración con estimaciones dentro y fuera de la muestra utilizando datos de opciones españoles, y encontramos que el modelo mejora al modelo estándar de Black-Scholes.

Acknowledgments

I would like to thank my thesis advisor, Prof. Aitor Bergara, for his limitless support, confidence, and for offering me the original idea from where all this work comes from. Specially, I appreciate his friendship.

I want also to thank the organization and members of the Quantitative Finance Ph.D program for their effort in creating an extraordinary Ph.D program in mathematical finance. Much of the ideas of this work were born in its lessons. Also, I am very greatful to all my colleagues of the QF program for their help during bad times.

I am also very thankful to the members of RiskLab Toronto, specially Prof. Luis Seco, Alejandro de los Santos, Marcos Escobar and Janko Hernandez, for their support and many interesting discussions.

I am also grateful for suggestions, comments, and contributions from: Fernando Tusell, Gonzalo Rubio, Angel León, Antonio Rubia, Miguel Angel Martinez and Francisco Javier Mencia.

I would like also to acknowledge the economic support of the Fundación Ramon Areces.

Finally, I thank my mother for giving me the opportunity to study, my friends and family, and Amaia for her unconditional support and for offering me peace of mind.

Contents

Introduction

In these thesis we develop a new distributional function family, the so called **Cornish-Fisher Distributions,** analyzing their theoretical properties studying their application in different areas of interest in mathematical finance: option valuation, measure of risk through the Value at Risk (VaR) and optimal portfolio selection. The search and development of new distributional families is an interesting mathematical task in itself and, in particular, the search of semi-parametric distributions which do not present negative probabilities as the traditional examples of Gram-Charlier (Charlier 1905) and Edgeworth (Edgeworth 1905). However, the principal motivation of this work is to analyze the application of this distributional family to financial theory, given that, as we will see next, the potential contribution of new statistical models in this area is specially significant.

The 1900 dissertation of Louis Bachelier, *The Theory of Speculation,* was the first attempt to model asset prices movements. He introduced the Brownian motion as the driver of the dynamics of asset prices, suggesting first that prices tended to follow a random walk and, second, that new prices are governed by a gaussian probability law. Therefore, the gaussian function was the first and became the most important distribution function, as most models conforming the traditional financial theory are based on the assumption of normally distributed returns. For example, the seminal work of Markowitz 1952, where the trade-off of risk and reward in the context of a portfolio of financial assets was made explicit, or others such as Sharpe 1966, Lintner 1965, and Ross 1976, who used equilibrium arguments to develop asset pricing models such as the capital asset pricing model (CAPM) and the arbitrage pricing theory (APT), relating the expected return of an asset to other risk factors. Even the classic Black and Scholes option pricing theory (Black and Scholes 1973) assumes that the return distribution of the underlying asset is normal.

One of the biggest drawbacks of gaussian distributions is that a gaussian variable bears the statistical characteristic that "large deviations" are extremely rare. For example, a gaussian variable departs from its most probable value by more than 2σ only 5% of the times, or more than 3σ in 0.2% of the times, whereas a fluctuation of 10σ has a probability of less than 2×10^{-23}, in other words, it almost never happens. However, market movements of $10\sigma,$ as in the market crash of October 1998, are not so uncommon in reality: one of the most spectacular examples of non-gaussianity in markets was the default of the hedge fund called Long Term Capital Management which was managed under gaussian risk models and did not survive the (for the managers) statistically unexpected 1998 crash, loosing $100 billions in two months (Lowenstein 2000). Therefore, more general statistical models allowing for these "large deviations" where needed.

The first model including heavy tails in finance dates from the sixties and is due to Benoit Mandelbrot (Mandelbrot 1963), where he assumed a Lévy stable distribution for the independent price changes, and since then there have been numerous studies[1] devoted to overcome limitations imposed by the popular normality assumption for stock returns. Almost every known distribution allowing for heavy tails has been proposed to model returns of financial assets given that financial data, just from a purely statistical point of view, provide a rich source of variables with diverse distributional characteristics. Approximations

[1] Instead of enumerating a complete list of references we recommend Rachev 2003 for an excellent and up-to-date revision on the publications and topics related to heavy tails in finance.

used in the literature to model the non-normality observed in financial time series follow very different strategies depending on the concrete purpose of the research. Either one can model the unconditional distribution of returns or the distribution conditioned by past returns[2]. On the other hand, one can also propose a continuous time model or a discrete time model and one can focus either on a univariate or on a multivariate model. Actually, it would be too long to list all possible strategies considered in the literature, but every model relays in the end on an assumption involving an hypothesis over an unconditional or a conditional distribution function. Actually, even continuous time models have an underlying distribution governing the behavior of the assets[3].

Moreover, when applying this model to a certain financial application this ultimate distributional assumption is generally the one about the researcher is concerned. For example, when one is constructing a statistical model to manage market risk via the Value at Risk (see Section 2.2) one finally comes up with a particular distribution where the multivariate gaussian case is the most common assumption. Or when building an optimal portfolio in some sense (see Section 2.1) one also realizes that independently on the particular model used (continuous or discrete time) one also ends with a random variable, usually the wealth portfolio, that follows a particular distribution. The same applies to the field of option valuation (see Section 2.3) where the assumption of a continuous time process for the underlying also derives a (risk neutral) distribution for the underlying at the expiration date

[2] Unconditional distributions are very useful when we are interested in the long-term and our investment horizons are such that far future returns are not determined by recent past ones, because such temporal dependence will be diluted with passage of time. However, if we have a short-term interest the conditional dynamic distributions would be more appropriate, where the conditional mean and volatility are very relevant variables.

[3] For example, the Brownian process is based on independent gaussian increments and the stable Lévy process is based on independent stable increments.

and this distribution is the necessary input for valuation purposes. Therefore, developing new distributional families is a good way of giving new answers to old questions in finance.

Chapter 1

Financial Modeling with Cornish-Fisher Distributions

In this Chapter we will present the *Univariate Cornish-Fisher Density Function*, which will be defined in terms of Cornish-Fisher Expansions, and, as we will see, can also be derived considering a variable transformation to normality. We also propose two *Multivariate Cornish-Fisher Density Functions* which incorporate different dependence structures. Next, we perform a descriptive analysis of the two databases (consisting of twelve series of exchange rates and five market indexes) that will be used along this work. After analyzing the non-normality of these time series, we will present three different estimation methods to quantify the parameters in the proposed model, where static and dynamic frameworks will be considered. Finally, our results will be compared with other models proposed in the literature.

1.1 Introduction

There have been numerous studies devoted to overcome the limitations imposed by the popular normality assumption for stock returns, which is rejected in the empirical financial literature. Nonetheless, any extension of the gaussian assumption should satisfy two crucial requirements: modeling flexibility and analytical tractability. Both needs are satisfied by the semi-parametric class of distributions, which are based on an expansion-like augmentation of the normal density through the inclusion of higher order moments such as skewness or kurtosis. The use of these expansions in this context is conceptually similar to Taylor expansions but applies to functions. In a conventional Taylor expansion some function is approximated at a given point by a simpler polynomial. Here, the density is approximated by an expansion around a, usually, normal density. A further difference is that expansions are usually made to obtain simplifications whereas here the approximation, by involving parameters which can be varied, allows us to generate more complicated functions. These semi-parametric densities have the theoretical appeal that, in principle, any density function could be arbitrarily approximated just by adding more terms and parameters in the expansion around the gaussian density and, therefore, offer an increasingly modeling flexibility. This characteristic of being an intrinsically approximative distribution is of special interest given that, as many authors remark, there could be no best true distributional model[4].

Most popular examples of semi-parametric densities are the Gram-Charlier (Charlier 1905, Jondeau and Rockinger 2001) and Edgeworth distributions (Edgeworth 1905) which are based on Cornish-Fisher Expansions, being a very useful tool in financial analysis, for example, in the field of option pricing by Jarrow and Rudd 1982, Corrado and Su 1997b, Corrado and Su 1996a, Capelle-Blanchard, Jurczenko, and Maillet 2001 and Jurczenko, Maillet, and Negrea 2002 or to describe deviations from normality of innovations in a GARCH framework (Gallant and G.Tauchen 1989). Nevertheless, both approximations,

---

[4] For example, Gourieraux (2000) explains this point: "*There does not exist a true model for each market that yields the best approximation of its price dynamics, captures most of the evidenced stylized facts, is valid at different frequencies, and provides a unifying framework to portfolio management, derivative pricing, forecasting, and risk control.*"

Gram-Charlier and Edgeworth, share the theoretical drawback of yielding negative density function values for certain parameter ranges. Besides, they cannot be used with moments of order higher than four, because estimation of higher order moments corresponding to observed data are often low accurate and their increasing flexibility characteristic is lost. This limitation implies a lack of flexibility for capturing high degrees of non-normality, so that only kurtosis up to six are achievable in practice.

As a solution of the negativeness associated to Gram-Charlier expansions, Jondeau and Rockinger 2001 propose the restriction of the parametrical space of Gram-Charlier distributions so that the density remains positive and, recently, León, Mencía, and Sentana 2005 analyze the use of the *semi-nonparametric* distribution, which is derived from a transformation of the Edgeworth distribution, which also ensures its positiveness. These two new approximations, despite of maintaining the analytical flexibility of the Edgeworth-Gram-Charlier moment expansion while solving the negativeness problem, inherit the lack of flexibility for capturing high degrees of kurtosis and skewness. León, Mencía, and Sentana 2005 demonstrate that in absence of skewness the maximal kurtosis that can be achieved using these expansions is eight, which turns out to be very restrictive.

To solve this problem Rockinger and Jondeau 2002 introduced the so called *entropy densities* that increase the modeling flexibility but at the cost of involving computationally costly optimizations in the density definition. Moreover, these approaches derive multimodal density functions for certain parameter values and this fact should be treated carefully in a financial framework as it implies that investors would find more than one return value as expectable.

In this Chapter we propose the use of a new semi-parametric distribution function, which turns out to overcome the difficulties of the traditional Gram-Charlier and Edgeworth distribution functions. While bearing the same tractability, characteristic of the semi-parametric class of density functions, it extends their modeling flexibility in terms of covered range of skewness-kurtosis possibilities and ensures unimodality and positiveness, becoming an excellent tool for financial modeling.

In particular, we study the parametric properties of the Cornish-Fisher Expansion (Cornish and Fisher 1937), initially proposed as an approximation method to estimate quantiles for distributions with known moments[5]. Truncating the Cornish-Fisher Expansion up to a fixed order we define a new class of densities that will be referred to as univariate *Cornish-Fisher density* (CFd). The first part of this Chapter will be devoted to the study of the theoretical properties of the univariate CFd.

In addition, in order to propose a multivariate model with CFD distributed marginals, we can use the fact that Cornish-Fisher Densities are based on transformations to normality[6]. Our proposed Cornish-Fisher density is equivalent to a transformation based on a Taylor series expansion of the QQ-Plot around normality, which is very appropriate for financial modeling and indirectly has already been applied in a very different science field, namely, in structural reliability analysis (Hong 1998). However, to our knowledge, no rigorous analysis and definition of the density implied by this transformation have been

---

5 For example, this expansion has been applied in finance to approximate the percentile of the profit and loss distribution in delta-gamma approximations for VaR calculation of portfolios containing options (Jorion 2000).

6 The concept of such transformations dates from the end of the XIX-th century and were put forward by Edgeworth 1898. Although Edgeworth considered only transformations which could be represented by polynomials, Johnson 1949 extended those transformations to other functions (e.g., the logarithmic one, proposing the log-normal density function).

presented yet. In this framework, imposing a multivariate gaussian behavior on the fictitious gaussian variables that arise from the transformation to normality we will propose the *Copula-Based Multivariate Cornish-Fisher density* (CB-MCFD), which exhibits non-gaussian marginals with a gaussian dependence structure. Additionally, we also define and analyze another multivariate distribution, namely, the *Variance-Covariance-Based Multivariate Cornish-Fisher density* (VCB-MCFD), that presents CFD-like marginals, easily incorporates standard first and second order dynamics models and, in opposition to the CB-MCFD model, allows for the occurrence of simultaneous extreme events in different marginal variables.

Many authors have highlighted the importance of incorporating first and second order dynamics in order to capture stylized facts of financial returns such as volatility clustering and volatility and correlation persistence, e.g. Engle 1982, Bollerslev 1986, Koedijk, Campbell, and Kofman 2002 and Engle 2002. In this work, we will choose the Dynamic Conditional Correlation (DCC) with univariate GARCH processes model to capture these features, because, as proven by Engle 2002, with this model just two parameters are required to capture correctly correlation dynamics. However, one difficulty of those models is that conditional residuals very often remain heavy tailed. Therefore, in order to capture both dynamic features and heavy tails, we will introduce two different multivariate dynamic models: the *Dynamic Copula-Based Multivariate Cornish-Fisher density* (DCB-MCFD) which incorporates a gaussian copula model with a DCC model and the *Dynamic Variance-Covariance-Based Multivariate Cornish-Fisher density* (DVCB-MCFD) which incorporates the dependence via a VCB-MCFD model and the dynamics through a DCC model. Although we have focused on the DCC model, any other multivariate dynamics, as the BKKK model proposed by Engle and Kroner 1995, could be easily incorporated as well.

This Chapter is organized as follows: in Sections 1.2 and 1.3 we will study statistical properties of the univariate and multivariate CFD distributions, including moments and standardized versions, and we will show that the CFD covers a more extended region in the skewness-kurtosis plane than other semi-parametric densities. In Section 1.4 we will perform a descriptive analysis of the two databases that will be used along this work. In Section 1.5 we will discuss three methods for the estimation of static models which are identically distributed as a CFD, and we will analyze the goodness-of-fit performance of this density using both exchange rates and financial indexes data. We also present univariate models with dynamic first and second moments and innovations following a CFD, and illustrate its application also analyzing the performance with respect to other standard models. Finally, in Section 1.6 we discuss estimation procedures and results for both Multivariate Cornish-Fisher Densities, the CB-MCFD and the VCB-MCFD, considering both static and dynamic frameworks.

1.2 Univariate Cornish-Fisher Density Function

1.2.1 Definition

In order to introduce the Cornish-Fisher Density Function, first we will present the Cornish-Fisher Expansion as defined by Cornish and Fisher 1937 and the Edgeworth and Gram-Charlier distributions, which are closely related to Cornish-Fisher Expansions.

A Cornish-Fisher Expansion approximates an unknown quantile of a distribution function F in terms of the quantiles of the gaussian distribution and the cumulants of the distribution F. This expansion becomes a very useful approximation technique, because there are available very efficient algorithms to calculate the quantiles of the gaussian distribution.

To be more explicit, let R be a quantile of a non-gaussian variable which we want to approximate and X the quantile of a gaussian variable. Formally, the Cornish-Fisher Expansion can be seen as a polynomial expansion of the quantile R in terms of the quantile X:

$$R = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + ... \tag{1.1}$$

where the parameters a_i depend on the cumulants of the distribution F. Actually, in order to obtain the Cornish-Fisher Expansion as can be found on any statistics book one has to re-group the infinite series of Equation 1.1 by a criteria motivated by the central limit theorem[7]. Consider a variable n that measures the approximation degree to the validity of the central limit theorem (n can be seen as a "sample size"), so that if variable n tends

[7] The central limit theorem states that the limit sum of independent variables with finite variance converges to a gaussian variable (Gnedenko and Kolmogorov 1954).

to infinity then R becomes a gaussian distributed variable. In terms of this variable n the expansion can be written as:

$$R = X + \sum_{k=1}^{\infty} n^{-k/2} \zeta_k(X) \tag{1.2}$$

where $\zeta_k(X)$ is the collection of all terms corresponding to the k-th power of $n^{-1/2}$, and can be written in terms of the cumulants κ_k of the distribution function R. Therefore, the first terms of the Cornish-Fisher Expansion are:

$$R = m + \sigma \left(X + \frac{1}{6}\frac{\kappa_3}{\sigma^3}(X^2 - 1) + \frac{1}{24}\frac{\kappa_4}{\sigma^4}(X^3 - 3X) - \frac{1}{36}\left(\frac{\kappa_3}{\sigma^3}\right)^2 (2X^3 - 5X) + ... \right) \tag{1.3}$$

where m and σ stand for the mean and standard deviation of the variable R.

In order to be more explicit, let us consider the following example: let R be distributed with the following *gamma distribution*, $f(R)$, with parameter p:

$$f(R) = \frac{1}{\Gamma(p)} e^{-R} R^{p-1}$$

where $\Gamma(p)$ is the Gamma function (Abramowitz and Stegun 1964). It is easy to derive that the general expression corresponding to the cumulants of this distribution, κ_k, is given by $\kappa_k = p(k-1)!$. Therefore, in this case Equation 1.3 leads to:

$$R = p + p^{1/2} \left(X + \frac{1}{3p^{1/2}}(X^2 - 1) + \frac{1}{4p}(X^3 - 3X) - \frac{1}{9p}(2X^3 - 5X) + ... \right) \tag{1.4}$$

According to this, as a practical example, if we wish to find the value of R whose distribution function is $F(R) = 0.99$ when $p = 15$, the gaussian standard percentile corresponding to such value is found to be 2.326 and replacing this value in Equation 1.4 we find $R = 25.45$, which is exact to two places of decimals.

Although Cornish-Fisher Expansions are applied to theoretically determined distributions (with known moments), they are directly related to the Edgeworth form of distribution[8]. Edgeworth (and Gram-Charlier) distributions are a family of distributions which provide an explicit relationship between the quantiles of the distribution and the moments or cumulants (see Appendix A.1 for more details). In practice, it is unusual to use moments higher than the fourth one in fitting an Edgeworth (or Gram-Charlier) expansion. This is mainly because the possibility of negative values (and multimodality) becomes more probable as higher terms are added, but also because, with observed data, estimation of higher moments is often of much lower accuracy. As mentioned in the Introduction, León, Mencía, and Sentana 2005 demonstrate that, in absence of skewness, the maximal kurtosis that can be achieved using these expansions up to fourth order is eight, which turns out to be very restrictive for financial modeling.

With the aim of obtaining a semi-parametric distribution, which does not show the limitations of the traditional ones, in this work, instead of using the expansion as an approximation method, *we propose a parametric use of the Cornish-Fisher Expansion to develop a new family of distributions*: **the univariate Cornish-Fisher Distributions**. As it is demonstrated below, besides being always unimodal, these new distributions are much more flexible than the Edgeworth, Gram-Charlier or *semi non-parametric* densities of similar order.

---

8 Edgeworth and Gram-Charlier distributions have been implemented in very different fields to model financial returns. As an example, in the field of option pricing we can cite the works of Jarrow and Rudd 1982, Corrado and Su 1997b, Corrado and Su 1996a, Capelle-Blanchard, Jurczenko, and Maillet 2001 and Jurczenko, Maillet, and Negrea 2002.

In order to derive these distributions we will proceed as follows: we truncate the Cornish-Fisher Expansion up to order m and consider the coefficients a_i in Equation 1.1 as the parameters of the distribution. Therefore, in order to fit this new family of distributions to financial data, we will seek the parameters a_i in Equation 1.1 that best fit observed data. With this parametrization, Equation 1.1 can be rewritten as:

$$R = \sum_{i=0}^{m} a_i X^i \equiv Q_m(X) \tag{1.5}$$

and, as will be demonstrated below, with this formulation and a polynomial degree as small as three we will be able to capture the main features of unconditional unidimensional financial distributions, namely, asymmetry and heavy tails. Basically, this new parametrization can be understood as a summation of the series made in a different and more efficient order. In order to clarify this point, we will consider again the example presented above. First, Equation 1.4 can be rewritten as:

$$R = \left(p - \frac{1}{3}\right) X^0 + \left(\sqrt{p} - \frac{7}{36}\frac{1}{\sqrt{p}}\right) X^1 + \left(\frac{1}{3}\right) X^2 + \left(\frac{1}{36}\frac{1}{\sqrt{p}}\right) X^3 + ... \tag{1.6}$$

Comparing Equations 1.5 and 1.6 we can find the expressions for the coefficients a_i corresponding to the Cornish-Fisher Expansion up to third-order. The main point here is that if we take more terms in the expansion of Equation 1.2 we would find that higher terms in X appear, like X^4 or X^5, but also (and more importantly) more factors containing higher order cumulants must be added in the coefficients $a_{0,}$ $a_{1,}$ a_2 and a_3. Therefore, taking more terms in the expansion above would be required to improve the accuracy of the first coefficients. As a conclusion, if we parameterize directly these first coefficients, which interestingly are supposed to be more important than the higher order ones, we will be gaining efficiency in terms of the number of parameters used to model the distribution. Therefore, instead of using the first terms of the Cornish-Fisher Expansion to model the distribution of returns as a function of the first moments m, σ, κ_3 and κ_4:

$$\begin{aligned} R \;=\; & \left(m - \frac{1}{6}\frac{\kappa_3}{\sigma^2} + ...\right) + \sigma\left(1 - \frac{3}{24}\frac{\kappa_4}{\sigma^4} + \frac{5}{36}\left(\frac{\kappa_3}{\sigma^3}\right)^2 + ...\right) X + \\ & \left(\frac{1}{6}\frac{\kappa_3}{\sigma^3} + ...\right) X^2 + \sigma\left(\frac{1}{24}\frac{\kappa_4}{\sigma^4} - \frac{2}{36}\left(\frac{\kappa_3}{\sigma^3}\right)^2 + ...\right) X^3 + ... \end{aligned} \tag{1.7}$$

we use a direct parametrization of the first coefficients:

$$R = a_0 + a_1X^1 + a_2X^2 + a_3X^3 + ...$$

For the following, a variable R that is distributed as a Cornish-Fisher Expansion of order m will be referred to as a variable with a m-th order *Cornish-Fisher Distribution* (CFD) or a m-th order *Cornish-Fisher density* (CFd).

1.2.2 Statistical properties

Next, we will discuss statistical properties of the Cornish-Fisher Distribution (CFD). Considering that X is the standard normal distribution with distribution function $\Phi(X)$:

$$\Phi(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} e^{-\frac{1}{2}t^2} dt,$$

the distribution function of a Cornish-Fisher variable, R, defined by Equation 1.5:

$$R = Q_m(X)$$

will be denoted by $CF_m(R)$ and can be expressed in the following way:

$$CF_m(R) = \Phi\left[Q_m^{-1}(R)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Q_m^{-1}(R)} e^{-\frac{1}{2}t^2} dt, \tag{1.8}$$

where Q_m is the m-th order polynomial and Q_m^{-1} is the inverse function of Q_m[9]. Furthermore, derivating the later expression with respect to R, one can easily find that the density function of a Cornish-Fisher variable, denoted by $cf_m(R)$, is given by:

$$cf_m(R) = \frac{d\left[Q_m^{-1}(R)\right]}{dR} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[Q_m^{-1}(R)\right]^2} \tag{1.9}$$

In this work we will mainly base our analysis on a third-order polynomial, since it is sufficiently appropriate to fit experimental data and it is the first non-trivial approximation that makes sense[10]. As it will be seen in the Univariate Inference Section 1.5, by means of a third-order polynomial we already achieve a very high adjusting performance, measured through the Kolmogorov-Smirnov statistic. On the other hand, the analyticity of the inverse of a third-order polynomial is also specially interesting, since becomes the basic ingredient for the density and distribution functions (Equations 1.9 and 1.8). The explicit form of the third-order CFD function is given in Appendix A[11]. However, if we wish to consider higher-order Cornish-Fisher Densities, although no explicit form is available for its density or distribution function, numeric procedures like Newton-Raphson (Abramowitz and Stegun 1964) could be used in order to obtain $Q^{-1}(R)$ and its derivative $d\left[Q^{-1}(R)\right]/dR$.

On the other hand, it is also interesting to note that parameters a_i of the third-order CFD have to be restricted in order to ensure that the density is properly defined. From the expression of the density of a CF variable (Equation 1.9) we can see that in order to

---

[9] Q^{-1} will be always defined for any non decreasing continuous function Q.

[10] As we are interested in modeling financial returns, we have not considered a quadratic polynomial. These returns are variables defined over an infinite support, which restricts the choose of polynomials to those of odd order, given that with an even-order polynomial we would map the real line corresponding to the support of the gaussian variable X onto the positive real segment instead of the entire real line.

[11] In the following, if we do not mention the order of the CFD, it will be assumed to be three. Moreover when writing the polynomial Q function, if we do not maintain the opposite it will be supposed to be a third-order polynomial Q_3.

be well defined it is sufficient and necessary to impose the existence and uniqueness of Q^{-1}. For a third-order polynomial this condition is equivalent to have a strictly increasing polynomial Q[12]. In the following Lemma we will find the conditions on the parameters a_i which guarantee the existence of a third-order CFD. The proof of this and the following Lemmas and Propositions will be presented in the Appendix B.

Lemma 1 *Let $cf_3(R)$ be a third-order Cornish-Fisher density function defined by Equation* 1.9*, with coefficients a_i, ($i = 0, 1, 2, 3$), then the sufficient and necessary conditions on the coefficients to guarantee the existence of the CFD are*

$$a_3 > 0 \quad , \quad a_1 > 0 \quad , \quad -\sqrt{3a_3a_1} < a_2 < \sqrt{3a_1a_3} \tag{1.10}$$

Therefore, a third-order CFD becomes completely defined by Equations 1.5, 1.9 and 1.10. It is interesting to note that third-order CFDs contain as special cases the gaussian distribution ($a_3 = a_2 = 0$), the χ^2 distribution ($a_1 = a_3 = 0$) and the non-central χ^2 ($a_3 = 0$)[13]. On the other hand, the following proposition will provide us an analytical expression to calculate the non-centered moments, μ'_r, of any m-th order CFD distribution:

Proposition 2 *Let $cf_m(R)$ be a m-th order Cornish-Fisher density function defined by Equation* 1.9 *for a random variable R, then non-centered r-th order moments, μ'_r, are given by:*

$$\mu'_r = E\left[R^r\right] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^r e^{\frac{1}{2}J^2}\Bigg|_{J=0}$$

where $Q\left(\frac{\partial}{\partial J}\right) = \sum_{i=1}^{m} a_i \frac{\partial^i}{\partial J^i}$ is a differential operator.

[12] Indeed, this is also true for any even polynomial.

[13] In the latter cases, although density functions will be given by Equation 1.9, it is important to note that variables X and R would not have the same support, i.e. X would be defined on the real segment $(-\infty, \infty)$ while R only on the positive real segment $(0,\infty)$.

Using this expression we can easily derive that the first four non-centered moments of a third-order CFD are:

$$\mu_1' = a_2 + a_0$$

$$\mu_2' = 15a_3^2 + a_0^2 + 2a_2a_0 + 6a_3a_1 + a_1^2 + 3a_2^2$$

$$\mu_3' = 9a_2a_1^2 + 15a_2^3 + a_0^3 + 45a_3^2a_0 + 3a_1^2a_0 + 315a_3^2a_2 + 9a_2^2a_0 + 18a_3a_1a_0 + 90a_3a_2a_1 + 3a_2a_0^2$$

$$\mu_4' = \begin{pmatrix} 105a_2^4 + 60a_0a_2^3 + 18a_0^2a_2^2 + 4a_0^3a_2 + 3a_1^4 + 10\,395a_3^4 + 6a_0^2a_1^2 + 1260a_1a_2^2a_3 + \\ 36a_0a_1^2a_2 + 90a_0^2a_3^2 + 3780a_1a_3^3 + 36a_0^2a_1a_3 + 60a_1^3a_3 + 90a_2^2a_1^2 + \\ 630a_1^2a_3^2 + 1260a_0a_2a_3^2 + 5670a_2^2a_3^2 + 360a_0a_1a_2a_3 + a_0^4 \end{pmatrix}$$

Accordingly to these expressions, the first four centered moments, μ_r, are given by:

$$\mu_1 = 0 \tag{1.11a}$$

$$\mu_2 = 6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2 \tag{1.11b}$$

$$\mu_3 = 72a_3a_2a_1 + 8a_2^3 + 270a_3^2a_2 + 6a_2a_1^2 \tag{1.11c}$$

$$\mu_4 = \begin{pmatrix} 10\,395a_3^4 + 60a_2^4 + 3a_1^4 + 60a_3a_1^3 + 3780a_3^3a_1 + \\ 936a_3a_2^2a_1 + 4500a_3^2a_2^2 + 630a_3^2a_1^2 + 60a_2^2a_1^2 \end{pmatrix} \tag{1.11d}$$

Henceforth, skewness and kurtosis coefficients are given by:

$$\xi(R) = \frac{72a_3a_2a_1 + 8a_2^3 + 270a_3^2a_2 + 6a_2a_1^2}{(6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2)^{3/2}}$$

$$\kappa(R) = \frac{\begin{pmatrix} 10\,395a_3^4 + 60a_2^4 + 3a_1^4 + 60a_3a_1^3 + 3780a_3^3a_1 + \\ 936a_3a_2^2a_1 + 4500a_3^2a_2^2 + 630a_3^2a_1^2 + 60a_2^2a_1^2 \end{pmatrix}}{(6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2)^2}$$

These Equations can be used to introduce the standardized third-order CFD, i.e. a CFD with zero mean and unit variance, and the corresponding conditions for its existence.

Lemma 3 *Let $cf_3(R)$ be a third-order Cornish-Fisher density function defined by Equation 1.9 for the random variable R, with coefficients a_i ($i = 0, 1, 2, 3$). Then, one can*

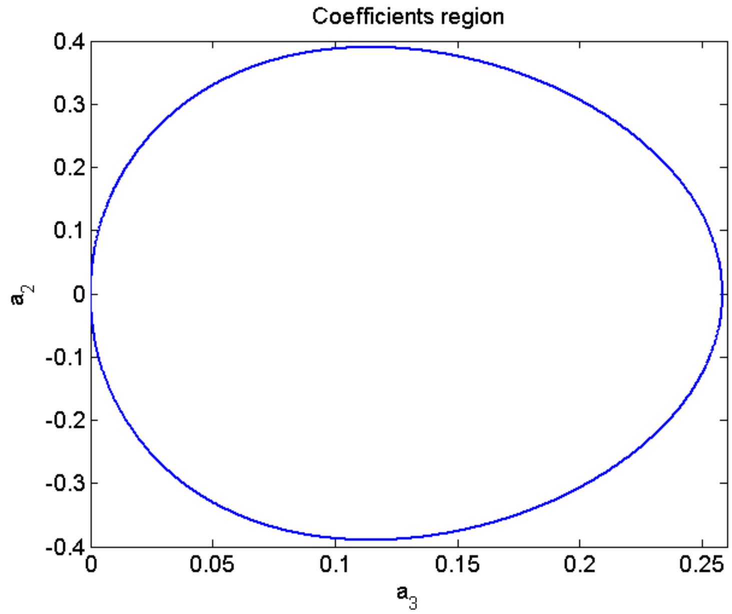


Fig. 1.1. The range of validity for the coefficients a_2 and a_3 of a standardized third-order CFD is represented by the region bounded by the blue line. The boundary corresponds to the Equations 1.13 and 1.14.

construct a standardized variable R with zero mean and unit variance imposing

$$a_0 = -a_2 \quad , \; a_1 = \sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3 \tag{1.12}$$

with the following conditions on a_2 and a_3 to guarantee the existence of the $cf_3(R)$:

$$0 \; < \; a_3 < \frac{1}{\sqrt{15}} \tag{1.13}$$

$$-\sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} \; < \; a_2 < \sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} \tag{1.14}$$

The validity region bounded by the limit cases of Equations 1.13 and 1.14 is plotted in Figure 1.1. Therefore, the standardized third-order CFD is completely defined by Equations 1.5, 1.9, 1.12, 1.13, 1.14. Although the permitted parameter range is bounded with just two shape parameters, a_2 and a_3, we are able to capture a kurtosis as high as

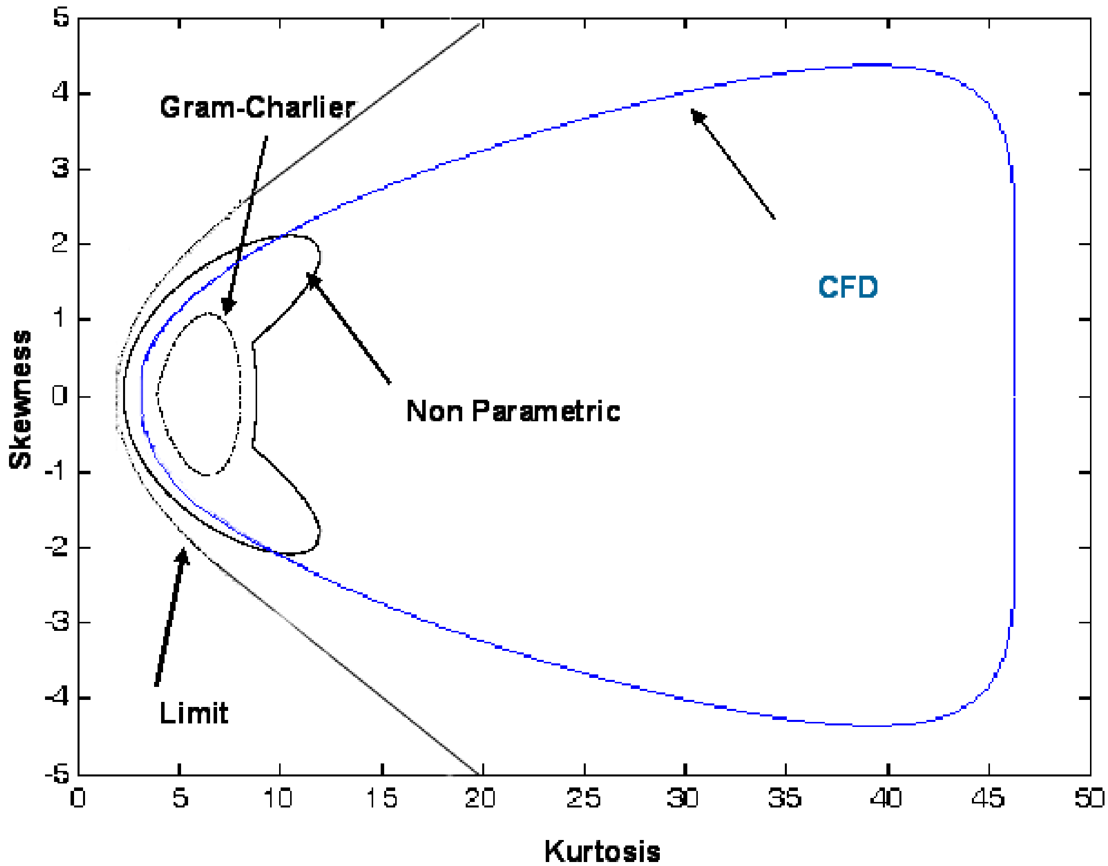


Fig. 1.2. Range of the skewness and kurtosis possibilities of the third-order Cornish-Fisher density. The blue line represents the skewness and kurtosis coefficients for the boundary plotted in Figure 1.1. We also present the limit for all distributions ($\kappa \geq \xi^2 - 2$, where κ and ξ are the kurtosis and skewness coefficients), and the regions covered by the Gram-Charlier distribution of Jondeau and Rockinger 2001 and the semi non-parametric distribution of León, Mencía, and Sentana 2005 of similar order. According to this Figure it is clear that the CFD is much more flexible than other semi-parametric approaches.

forty and a skewness up to a value of three as shown in Figure 1.2[14]. In this Figure we also plot the equivalent expanded region for two different semi-parametric distributions, the *semi non-parametric* one of León, Mencía, and Sentana 2005 and the Gram-Charlier distribution of Jondeau and Rockinger 2001, as well as the boundary limit for any distribution[15]. According to this Figure, it is interesting to observe that the extended region in the skewness-kurtosis plane that can be covered with the CFD model is much wider than the Gram-Charlier distribution or the *semi non-parametric* distribution and allows us to conclude that the CFD is much more flexible than other semi-parametric approaches.

The following transformation rule allows us to define a reparametrization of the third-order CFD in terms of the mean, μ, the volatility, σ, and the parameters a_2 and a_3.

Lemma 4 *Let R be a m-th order CFD distributed variable with parameters $\{a_i\}_{i=1}^m$ and consider the variable $Z = \mu + \sigma R$. Then, the new variable Z is also distributed as a CFD with parameters $\{a_i'\}_{i=1}^m$ given by*

$$
\begin{aligned}
a_i' &= \sigma a_i \qquad\qquad i = 1, ..., m \\
a_0' &= \sigma a_0 + \mu
\end{aligned}
$$

With this transformation rule we can re-define the function $Q(X)$ and, therefore, the Cornish-Fisher Density, using a new parametrization set, namely, m, σ, a_2 and a_3:

$$R = \sigma a_3 X^3 + \sigma a_2 X^2 + \sigma \left(\sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3 \right) X + m - \sigma a_2 \tag{1.15}$$

[14] The corresponding values of kurtosis and skewness coefficients for the boundary of the region in the a_2-a_3 plane given by Equations 1.13 and 1.14 are plotted in Figure 1.2.

[15] It is easy to see, that for every skewness coefficient $\xi = \mu_3/\mu_2^{3/2}$ and every kurtosis coefficient $\kappa = \mu_4/\mu_2^2 - 3$, the inequality $\kappa \geq \xi^2 - 2$ must hold.

This specification will be of special interest when modeling the dynamic behaviour of the conditional mean and volatility, as both parameters appear explicitly in the definition of the density function.

As mentioned in the Introduction, unimodality is a very desirable property for modeling financial returns that the other semi-parametric models do not share. In the following proposition we will demonstrate the unimodality of the CFD:

Proposition 5 *Third-order Cornish-Fisher Densities are unimodal.*

Figure 1.3 presents some possible shapes of the standardized third-order Cornish-Fisher Densities showing their flexibility to adjust different degrees of skewness and kurtosis. It is also interesting to point out that CFD includes the presence of heavy tails. Figure 1.4 shows the detail of the tails of the distribution in logarithmic scale. It can be observed that the CFD presents an almost linear behavior in the tails, as it corresponds to an exponential distribution. Besides that, it is also remarkable that the rate of decrease is lower than in the gaussian approximation (parabolic), so that a higher weight is assigned to the tails.

1.2.3 Relation of CFD with transformations and QQ-Plots

In order to gain some insight on the characteristics of the CFD one can view Equation 1.5:

$$R = Q_m(X)$$

as a percentile-percentile relation between a fictitious normal variable, X, and the non-normal variable, R, that we want to describe. We can consider that the value of the variable

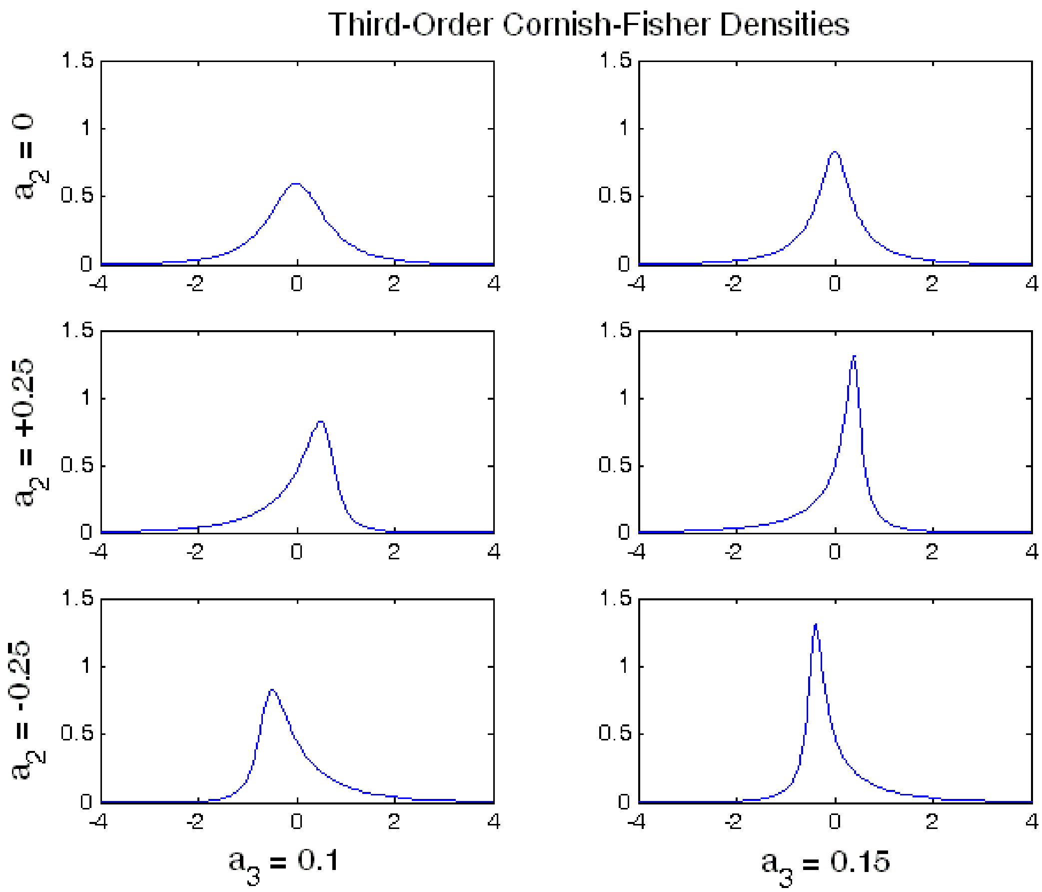


Fig. 1.3. Different shapes of the third-order CFD. All densities are standardized to zero mean and unit variance and, therefore, we only have two free parameters: a_2, which defines the skewness, and a_3, which incorporates the kurtosis.

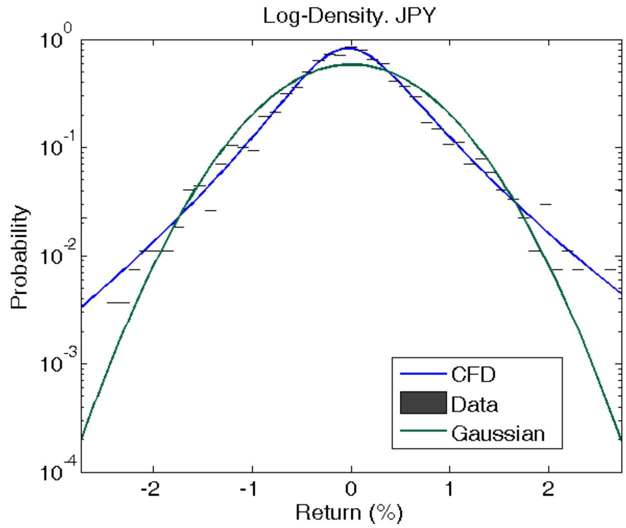


Fig. 1.4. Logarithmic scale graph of the fitted third-order Cornish-Fisher Density to the series of daily returns of the YEN/USD exchange rate for the period 04/01/1988-15/08/1997. The horizontal lines are the observations of the histogram, the green line is a fitted gaussian function density and the blue line is the fitted CFD.

X that we are fixing is the one that corresponds to a certain percentile α of the distribution[16]. In this way, Equation 1.5 relates percentiles of the non-normal distribution with percentiles of the normal distribution. Therefore, in order to estimate the parameters of the function $Q(X)$ in Equation 1.5 it will be reasonable to fit the function that relates the value of the percentile α of the empirical distribution in the ordered axis with the value of the same percentile of the standard normal distribution in the abscissas axis. As it is well-known, in statistical literature this representation is commonly denominated QQ-Plot and, therefore, the CFD function will be an appropriate model for financial series if they present a normal QQ-Plot polynomially shaped, which is generally the case, as we will see in the following example. In Figure 1.5 we present the QQ-Plot of the series of daily returns in the YEN/USD exchange rate for the period 04/01/1988-15/08/1997 against a standard

[16] For example, for a standard gaussian variable if $\alpha = 0.01$ then $X = -2.33$.

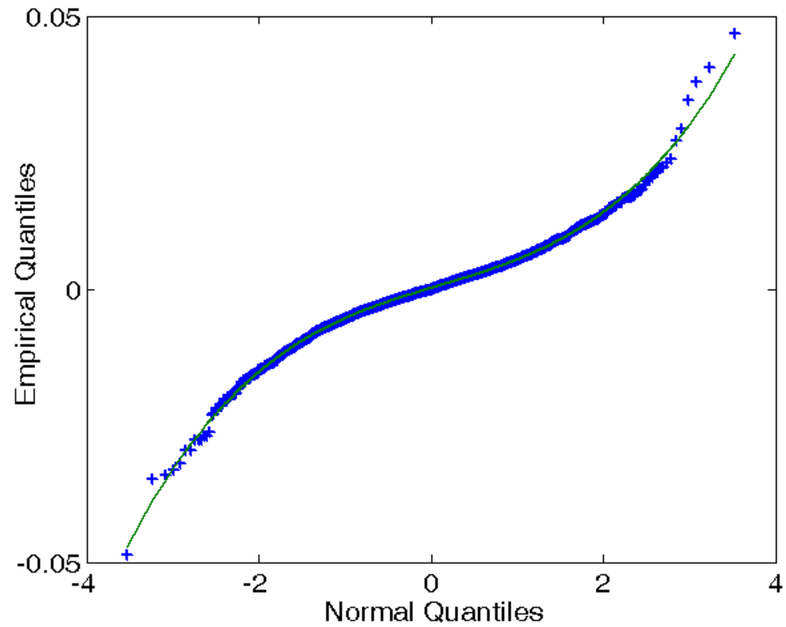


Fig. 1.5. QQ-Plot of the series of daily returns of the YEN/USD exchange rate for the period 04/01/1988-15/08/1997 against a standard normal. Crosses represent experimental data and the solid line represents the least squares fit by means of a third-order polynomial: $R = Q(X) = 0.06757X^3 + 0.01227X^2 + 0.3370X - 0.01077$.

normal one. Crosses represent experimental data and the full line corresponds to the fit of a third-order polynomial using the minimum least squares algorithm. As can be observed, it is remarkable the non-linear shape of the QQ-Plot that discards gaussianity and the high adjusting performance that can be reached with a third-order polynomial. In addition, in Figure 1.6 we present the histogram that corresponds to the USD/YEN exchange rate variable, whose QQ-Plot has been presented in Figure 1.5, along with the fitting of a third-order CFD and a gaussian. The series has been standarized so that it presents zero mean and unitary variance. In this Figure we can observe how the non-linear shape of the QQ-Plot derives in a leptokurtic distribution which is more peaked than the gaussian and

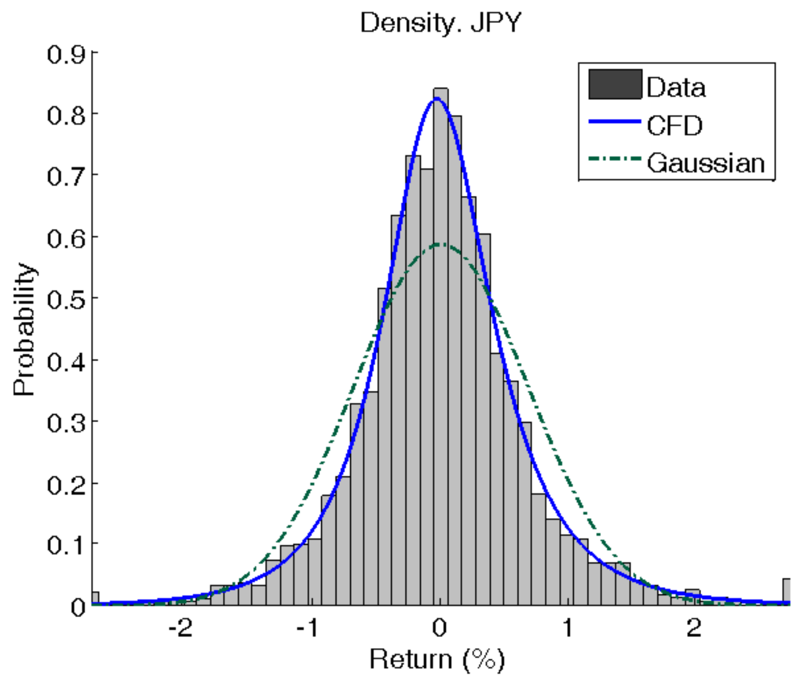


Fig. 1.6. Comparison between the fitting of a normal density function and the one corresponding to a third-order Cornish-Fisher Density. The fitting has been done by means of the QQ-Plot that appears in Figure 1.5. The data corresponds to the series of daily changes of the YEN/USD exchange rate for the period 04/01/1988-15/08/1997.

presents heavier tails. It is easy to observe that the CFD function is more adequate than the gaussian to model the USD/YEN exchange rate, as in real distributions intermediate events are less probable than in normal distributions. These intermediate events are distributed between near values very close to the mean and also around the tails. Therefore, the main effect of the kurtosis coefficient consists on assigning more probability to the tails.

Another point of view, that will be very interesting for proposing a multivariate distribution where the marginals are Cornish-Fisher distributed, is to interpret Equation 1.5 as a variable transformation in the style of Johnson 1949. The function Q contains the non-perturbative deviation with respect to the gaussian distribution, which is recovered when Q

is equal to the identity function. Defining a particular parametric form for the function Q, we will be supposing implicitly a certain parametric distribution function through Equation 1.5. Therefore, one can interpret the variable transformation as an alternative form of defining distribution functions, as pointed out by Johnson 1949 or Kendall, Stuart, and Ord 1994[17]. As it is shown in Figure 1.5, QQ-Plot of financial series, generally leptokurtic, show a clear deviation from the identity function, as corresponds to normally distributed variables. Actually, our starting point of using polynomials to represent the function Q came from the intuitive idea that considering the *simple shape of the QQ-Plot*, although containing strong deviation from normality, it would be possible to make a Taylor series expansion of the function Q around the identity function, where the terms with order higher than one contain the deviation from normality. Actually, as we will see in the Univariate Inference Section 1.4, third-order polynomials are enough to describe highly non-linear financial variables as interest rates or market indexes. This characteristic is of special interest, since the transformation based on a series expansion of powers of the QQ-Plot allows us to make a non-perturbative approximation of the distribution function that we want to model.

1.2.4 Simulation of Univariate Cornish-Fisher Variables

Equation 1.5 gives us a very simple way of simulating m-th order Cornish-Fisher variables R using the function Q. The algorithm is very simple:

1. First we simulate standard gaussian variables X,

[17] Something similar happens when one defines a density function $f(x)$ in terms of its characteristic function $g(k) = E(e^{ikx})$. For example, the α-stable distribution is defined through its characteristic function.

2. And second, we apply Equation 1.5, $R = Q_m(X)$, to each simulated observation. The so obtained variable R will be CF distributed.

This algorithm is standard to simulate non-normal market variables with distribution function F. First, one estimates the parameters describing F, and second, calculates the function $Q(X)$ using the expression:

$$Q\left(X\right) = F^{-1}\left[\Phi\left(X\right)\right]$$

and after simulating standard gaussian variables X, then applies $Q\left(X\right)$ to these observations to obtain the variables with distribution function F. It is easy to see that the function $Q\left(X\right)$, as it involves the inverse of the considered distribution function, F, will not be analytic for the most commonly used distribution functions and, therefore, in these cases the calculation of the transformation $Q\left(X\right)$ will add more time process to the already "expensive" Montecarlo method. For example, Hull and White 1998, used a density function of a mixture of two gaussians of parameters $(p, \mu_1, \mu_2, \sigma_1, \sigma_2)$, whose density function is given by:

$$f(x) = p\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} + (1-p)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2}$$

and presents the disadvantage of being a transcendent equation so that it is not possible to obtain the inverse function, $x = f^{-1}(y)$, analytically. However, this is not the case of the third-order CFD that we are considering in this work, given that the function $Q(X)$ is directly modeled as a third-order polynomial.

1.3 Multivariate Cornish-Fisher Density Functions

In many financial applications, like portfolio selection or risk management, we are interested in modeling the distribution function of returns of security portfolios. Let us consider a portfolio with n assets and let $R_1, ..., R_n$ be their daily returns, F_i , the marginal distribution function for the i-th variable, so that $F_i(\alpha) = \text{Prob}(R_i \leq \alpha)$, and ω_i the weight of asset i in the portfolio. Given that the Profit and Loss function of a portfolio is defined as $P = \sum_{i=1}^{n} \omega_i R_i$, when considering a portfolio the statistical variable of interest will be the one formed by a linear combination of random variables R_i. Therefore, we have to make some hypothesis on the structural dependencies of variables R_i that will allow us to consider a joint distribution function for them.

When marginal distributions are normal the most common approach to the problem is to suppose that variables come from a multivariate normal distribution, in such a way that the sum of normal variables is again a normal variable whose variance is given in terms of the variance and covariance matrix of the normal variables. However, in more general non-normal contexts, it is well-known that a zero linear correlation between two variables does not imply that they are independent.

In this work, we propose two different multivariate distributions for non-normal variables. First, we consider a joint distribution with a structural dependence given by a gaussian copula whose marginals are Cornish-Fisher distributed. Secondly, we will describe a multivariate generalization specially valid to incorporate first and second order dynamics and tail dependence whose marginals are being described by Cornish-Fisher distributions. Basically, these two multivariate densities differ in the way to capture the dependence between the variables. In the first case it is done by considering the normal rank correlation matrix and in the second one through the variance-covariance matrix. In the following, distributions considering both approximations will be referred as a *Copula-Based Multivariate Cornish-Fisher Distribution* (CB-MCFD) and as a *Variance-Covariance-Based Multivariate Cornish-Fisher Distribution* (VCB-MCFD). Main properties of both density functions will be analyzed in this Section.

1.3.1 Copula-Based Multivariate Cornish-Fisher Density

Definition

Before describing the CB-MCFD we will outline some basic concepts, properties and results related to the copula theory. Basically, copula functions describe the part of a multivariate density function that univariate density functions do not include. A function $C : [0,1]^n \to [0,1]$ is defined as an n-copula if it satisfies the following mathematical properties:

- $\forall u \in [0,1], C(1,...,1,u,1,...,1) = u$.
- $\forall u_i \in [0,1], C(u_1,...,u_n) = 0$ if at least any u_i is equal to zero.
- C is grounded (has a minimum) and is n-increasing, i.e., $dC(...,u_i,..)/u_i \geq 0$.

It is clear from this definition that a copula is nothing but a multivariate distribution with support in [0,1] and with uniform marginals. The fact that such copulas can be very useful for representing multivariate distributions with arbitrary marginals is explained next.

Intuitively, a copula function C is defined in such a way that if $P_i(R_i)$ are marginal distribution functions and $P(R_1, ..., R_n)$ is the joint distribution function of these variables, the copula function C connects both of them in the following way:

$$P\left(R_1, ..., R_n\right) = C\left(P_1\left(R_1\right), ..., P_n\left(R_n\right)\right) \tag{1.16}$$

Given the marginals $P_i(R_i)$ and the joint distributions, $P(R_1, ..., R_n)$, the copula function is unique if the distributions are continuous as determined by the theorem of Sklar (Sklar 1959, Malevergne and Sornette 2001). Therefore, copulas allow us to separate the modeling of joint distributions in two parts: a first one only related to the marginals and a second which only captures the structure of dependencies. Many different copulas have been proposed and tested in the financial literature. For a quite complete list of copula functions and their mathematical properties see Nelsen 1999, and for an up-to-date revision of the publications on copula theory in finance read Cherubini, Luciano, and Vecchiato 2004 and Frees and Valdez 1998.

For our purposes, we will only focus on one of the most well-known copulas, the *gaussian copula*, which is derived from the multivariate Gaussian distribution. Let Φ denote the standard normal distribution and $\Phi_{n,\kappa}$ the n-dimensional normal distribution with correlation matrix κ. Then, the gaussian n-copula with *normal rank correlation matrix*[18] κ,

[18] Some authors refer to the correlation matrix of the gaussian copula as to the normal rank correlation (Buckley, Comezaña, Djerroud, and Seco 2005) while others (Nelsen 1999) refer to it as a generalized correlation matrix. The notion of normal rank correlation is a specific case from a generalization of the concept of correlation. The idea of this generalization consists on calculating the correlation, not between the real variables S_i and S_j with distributions P_i and P_j, but between a transformation of these into other fictitious random variables with cumulative distribution functions F_i and F_j, respectively:

$$\Theta\left(F_i, F_j, S_i, S_j\right) = E\left[F_i^{-1}\left(P_i\left(S_i\right)\right) \cdot F_j^{-1}\left(P_j\left(S_j\right)\right)\right]$$

Choosing different functions F_i and F_j this generalization includes the most usual structural dependence measures. For example, for $F_i = P_i$ $F_j = P_j$ we recover the standard correlation or correlation coefficient denoted by $G_\kappa(u_1, ..., u_n)$, is defined as:

$$G_\kappa(u) = \Phi_{n,\kappa}\left[\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)\right] \tag{1.17}$$

One basic property of the gaussian copula is that it does not exhibit *tail dependence*. Loosely speaking, tail dependence describes the limiting proportion that one margin exceeds a certain threshold given that the other margin has already exceeded that threshold[19]. The question whether financial returns present non zero tail dependence is still not clear. For example, Malevergne and Sornette 2003 find that most pairs of exchange rates and major stocks are compatible with the gaussian copula hypothesis, while this hypothesis can be rejected for the dependence between pairs of commodities. On the other hand, Chen, Fan, and Patton 2004 in its study with equity and exchange rates return data find strong evidence against the gaussian copula, and little evidence against the more flexible Student's copula[20]. Considering the main role that the paradigm of the gaussianity has played and still does in finance, it is natural to begin with the simplest structure of dependencies between

---

of Pearson:

$$\Theta(P_i, P_j, S_i, S_j) = \text{corr}(S_i, S_j)$$

If we choose for the functions F_i and F_j the identity function we obtain the Spearman's rho:

$$\Theta(1, 1, S_i, S_j) = E(P_i(S_i), P_j(S_j))$$

where 1 denote the cumulative distribution function of the uniform distribution. The normal rank correlation in then defined as:

$$\kappa(S_i, S_j) = \Theta(\Phi, \Phi, S_i, S_j) = E\left[\Phi^{-1}\left[P_i\left(S_i\right)\right] \cdot \Phi^{-1}\left[P_j(S_j)\right]\right]$$

Notice, that the normal rank correlation for a gaussian variable coincides with the usual correlation. Each of these measures, together with others like Kendall's tau (Kendall, Stuart, and Ord 1994), provide different alternatives of estimating the intuitive idea of dependence.

19 See Nelsen 1999 for more details on the definition of tail dependence.

20 The Student-t n-Copula is defined analogously to the gaussian counterpart:

$$Student_\kappa(u) = T_{\nu,\kappa}\left[T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_n)\right]$$

where $T_{\nu,\kappa}$ and T are the multivariate and univariate student-t distribution functions with ν degrees of freedom and generalized correlation κ. This Copula presents a non-zero tail dependence coefficient, characteristic that has been put forward by some authors (Nelsen 1999, Demarta and McNeil 2004) as desirable.

different random variables: the gaussian copula. This election being specially attractive if we keep in mind the theorem that assures that this hypothesis maximizes entropy among a big group of possibilities in the sense of Shannon[21] (Buckley, Comezaña, Djerroud, and Seco 2005, Sornette, Andersen, and Simonetti 2000b, Sornette, Andersen, and Simonetti 2000a). Intuitively, we can also consider that one is choosing the hypothesis that makes less assumptions. By means of the normal rank correlation or the gaussian copula we will be picking up a more general structural dependence than just keeping the standard correlation.

Therefore, we would be interested in a density with Cornish-Fisher distributed marginals with a gaussian copula. We can obtain the **Copula-Based Multivariate Cornish-Fisher Distribution (CB-MCFD)** substituting in Equation 1.16 the gaussian copula given by Equation 1.17 and the Cornish-Fisher distribution functions defined in Equation 1.8:

$$\begin{aligned} \text{CB-MCFD}\left(R_1,...,R_n\right) &= G_\kappa\left[CF_1\left(R_1\right),...,CF_n\left(R_n\right)\right] \\ &= \Phi_{n,\kappa}\left[\Phi^{-1}(\Phi\left(Q_1^{-1}(R_1)\right)),...,\Phi^{-1}\left(\Phi\left(Q_n^{-1}\left(R_n\right)\right)\right)\right] \\ &= \Phi_{n,\kappa}\left[Q_1^{-1}\left(R_1\right),...,Q_n^{-1}\left(R_n\right)\right] \qquad (1.18) \end{aligned}$$

Derivating with respect to R_i we obtain the *Copula-Based Multivariate Cornish-Fisher Density* (cb-mcfd):

$$\text{cb-mcfd}(R) = \frac{1}{\sqrt{(2\pi)^n \det\left[\kappa\right]}} \prod_{i=1}^{n} \frac{\partial\left[Q_i^{-1}(R_i)\right]}{\partial R_i} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} Q_i^{-1}(R_i)\left(\kappa^{-1}\right)_{ij} Q_j^{-1}(R_j)\right) \qquad (1.19)$$

[21] Consider a random vector (S_i, S_j) such that maximizes the entropy among all random vectors with S_i and S_j as marginal. Then the random variables S_i and S_j are independent if and alone if the normal rank correlation is zero.

where $(\kappa^{-1})_{ij}$ is the i, j element of the inverse of the normal rank correlation matrix, κ, and R is the vector $(R_1, ..., R_n)$. One of the most interesting properties of this distribution is that captures both skewness and kurtosis through marginal Cornish-Fisher densities and the non-linear dependence via the gaussian copula, while maintaining an analytical form which allows us to characterize any property of the distribution. To our knowledge, except for the trivial case of multivariate gaussian variables[22], the gaussian copula has only been used with marginal distributions that do not derive analytical results.

Statistical Properties

Next, we will present some results and properties of the *Copula-Based Multivariate Cornish-Fisher Density* that will be useful in consequential applications. The following proposition gives an analytical formula for the non-centered moments of a portfolio including assets which follow a m-th order CB-MCFD distribution:

Proposition 6 *Let $(R_i)_{i=1}^n$ be the daily returns of n assets that follow a Copula-Based Multivariate Cornish-Fisher Density CB-MCFD$_m$ and κ be the* normal rank correlation *matrix of these assets. The Profit and Loss (P&L) distribution of a portfolio with weights $(\omega_i)_{i=1}^n$ corresponding to this assets, constrained to $\sum_{i=1}^n \omega_i = 1$, is given by:*

$$P = \sum_{i=1}^{n} \omega_i R_i \tag{1.20}$$

[22] For example, Hull and White 1998 use the gaussian copula with a gaussian mixture, and Malevergne and Sornette 2004 use it with a family of modified Weibull distributions, they where are able to calculate formulas for the moments and cumulants but not for the density function.

Then, the k-order non-centered moments of the aleatory variable P are given by:

$$E\left[P^k\right] = \left[\sum_i \omega_i Q_i(\frac{\partial}{\partial J_i})\right]^k \cdot e^{\frac{1}{2}J\kappa J^t}\Bigg|_{J=0} \tag{1.21}$$

where $J = (J_i)_{i=1}^n$ is an auxiliary n-dimensional vector.

The relation between the linear or Pearson correlation coefficient ρ and the *normal rank correlation* κ for a third-order CB-MCFD is given by the following Lemma:

Lemma 7 *Let ρ_{ij} be the linear correlation coefficient and κ_{ij} the* normal rank correlation *between the variables R_i and R_j which follow a CB-MCFD$_3$, and $a_{i,j}$ the i-th coefficient in Equation* 1.5 *corresponding to the asset j determining the transformation. The relation between both correlation coefficients is given by:*

$$\rho_{ij} = \frac{6a_{3,i}a_{3,j}\kappa_{ij}^3 + 2a_{2,i}a_{2,j}\kappa_{ij}^2 + \left(3a_{1,i}a_{3,j} + a_{1,i}a_{1,j} + 9a_{3,i}a_{3,j} + 3a_{3,i}a_{1,j}\right)\kappa_{ij}}{\sqrt{\left(6a_{3,i}a_{1,i} + 15a_{2,i}^2 + 2a_{2,i}^2 + a_{1,i}^2\right)\left(6a_{3,j}a_{1,j} + 15a_{3,j}^2 + 2a_{2,j}^2 + a_{1,j}^2\right)}} \tag{1.22}$$

According to this result, it is easy to deduce that the bigger the non gaussian character of the assets the bigger will be the difference between the Pearson coefficient and the normal rank correlation. If both assets are gaussian, i.e. $a_{3,i} = a_{2,i} = 0$, the CB-MCFD reduces to a multivariate gaussian distribution and the Pearson coefficient and the normal rank correlation will be the same. Therefore, this distribution is not adequate for modeling data displaying univariate gaussian distributions but not a multivariate gaussian dependence.

Considering the Equations defined above, we can introduce the independent and standardized CB-MCFD$_3$, (i.e. a CB-MCFD with zero mean and unitary variance-covariance matrix) and the corresponding conditions for its existence. This new distribution will be denoted as I-MCFD.

Lemma 8 *Let CB-MCFD$_3$ be a third-order Copula-Based Multivariate Cornish-Fisher density defined by Equation* 1.19 *for the random vector of variables R_i, with coefficients* a_i, $i = 0, 1, 2, 3$. *Then, one can define a new vector of variables R_i' with zero mean and unitary variance-covariance matrix imposing κ to be the identity matrix and:*

$$a_{i,0} = -a_{i,2} \quad , \ a_{i,1} = \sqrt{1 - 6a_{i,3}^2 - 3a_{i,2}^2} - 3a_{i,3} \tag{1.23}$$

Actually, $a_{i,2}$ and $a_{i,3}$ need to satisfy the following conditions in order to guarantee the existence of the distribution:

$$0 < a_{i,3} < \frac{1}{\sqrt{15}} \tag{1.24}$$

$$-\sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2 + 1} - 6a_{i,3}\right)} < a_{i,2} < \sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2 + 1} - 6a_{i,3}\right)} \tag{1.25}$$

In Figure 1.7 we present some possible shapes for the bivariate third-order CB-MCFD showing its flexibility to describe rather different distribution patterns. In all the Figures the parameters $a_{i,2}$ are fixed to 0.25, so that both univariate distributions are positively skewed. In the two upper Figures we present two independent CB-MCFD ($\kappa = 0$) with different degrees of kurtosis. It is interesting to note that contour plots of independent multivariate gaussian distributions are circles and, therefore, we can see how the positive skew and heavy tails of the CFD modify these circles. The other four Figures show contour plots of CB-MCFD with positive and negative normal rank correlation, which as we can see are much more general than the multivariate gaussian ellipses.

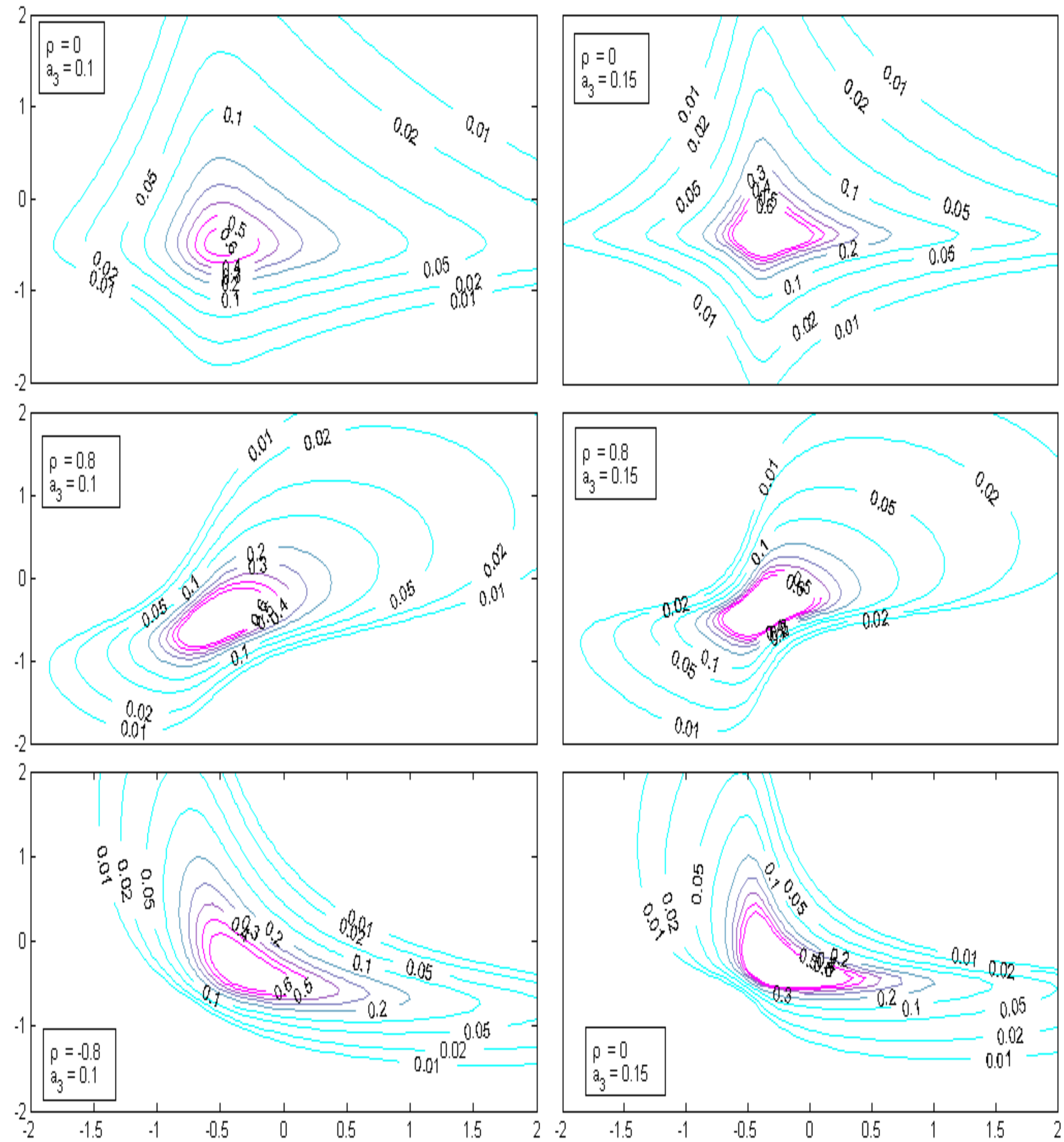


Fig. 1.7. Contour plots corresponding to different shapes of the bivariate third-order CB-M-CFD. All densities are standardized to zero mean and unit variance and, therefore, we only have three free parameters defining the distribution: a_2, which defines the skewness, a_3, which incorporates kurtosis, and the normal rank correlation ρ. In all the figures the parameter a_2 is set to 0.25, so that both marginal densities are positively skewed.

Simulation of CB-MCFD Variables

Equation 1.5 gives us a very simple way of simulating m-th CB-MCFD variables $(R_i)_{i=1}^n$ using the functions Q_i and the normal rank correlation, κ. The algorithm is as follows:

1. First, we simulate n-dimensional multivariate standard gaussian variables X_i, with correlation matrix equal to the normal rank correlation, κ.

2. And second, we apply Equation 1.5 to each variable R_i, $R_i = Q_i(X_i)$.

These algorithm is very fast, as very efficient multivariate gaussian random generators are available in each computation package.

1.3.2 Variance–Covariance-Based Multivariate Cornish-Fisher Density

Definition

In this Section we will present and describe basic properties of the *Variance-Covariance-Based Multivariate Cornish-Fisher Distribution,* which will be denoted as VCB-MCFD. Before going into the details of its definition, we will motivate briefly its introduction. Financial return series present dynamics in its second-order moments: correlations and variances are not constant over time and show stylized facts as persistence, i.e. periods with high(low) correlations or variances tend to predict high(low) correlations or variances[23].

[23] In Section 1.4 we analyze descriptive characteristics of our database and find significant conditional heteroskedasticity. Many authors have also pointed out this feature, see for example Engle 1982. On the other hand, Engle 2002 and Engle and Kroner 1995 show how correlation between returns is not constant over time.

Much of the research concerning the dynamics on the dependence between returns has focused on modeling the dynamics of the second-order moments (see for example, Engle 2002 where he analyzes different benchmark models). Therefore, it would be interesting to have a multivariate distribution with the variance-covariance matrix as input to allow us to incorporate easily these models, that, as we will next, the VCB-MCFD distribution does. Additionally, the VCB-MCFD presents marginal distributions which allow for incorporating asymmetry and heavy tails through the univariate Cornish-Fisher Density making and the occurrence of simultaneous extreme events (tail dependence) making this distribution specially interesting[24].

Now, let m and Σ be the mean vector and the variance-covariance matrix of variables R. Consider first the third-order Independent Multivariate Cornish-Fisher Density (i-mcfd) introduced above:

$$\text{i-mcfd}_3(z) = \frac{1}{\sqrt{(2\pi)^n}} \prod_{i=1}^{n} \frac{\partial \left[Q_i^{-1}(z_i)\right]}{\partial z_i} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left[Q_i^{-1}(z_i)\right]^2\right) \quad (1.26a)$$

$$Q_i(z_i) = a_{3,i}z_i^3 + a_{2,i}z_i^2 + \left(\sqrt{1-6a_{3,i}^2-3a_{2,i}^2} - 3a_{3,i}\right) z_i - a_{2,i} \quad (1.26b)$$

where variables z_i have by definition zero mean and unitary variance-covariance matrix.

We introduce the **Variance-Covariance-Based Multivariate Cornish-Fisher Distribution (VCB-MCFD)** as the distribution followed by variables R defined by:

$$R = \Sigma^{1/2} \cdot z + m \quad (1.27)$$

---

24 In Section 1.6.2 we will discuss and implement the Dynamic Conditional Correlation model (DCC), proposed by Engle 2002, to model second-order dynamics with the VCB-MCFD.

where $\Sigma^{1/2}$ denotes the Cholesky decomposition of Σ[25] and the vector $z = (z_1, ..., z_n)$ follows an i-mcfd. From this definition it is easy to see that the variables R have mean equal to m and variance-covariance matrix equal to Σ:

$$E[R] = E\left[\Sigma^{1/2} \cdot z + m\right] = \Sigma^{1/2} \cdot E[z] + m = m$$

$$\begin{aligned} E[(R-m)(R-m)'] &= E\left[\left(\Sigma^{1/2} \cdot z\right)\left(\Sigma^{1/2} \cdot z\right)'\right] = E\left[\Sigma^{1/2} \cdot z \cdot z' \left(\Sigma^{1/2}\right)'\right] \\ &= \left[\Sigma^{1/2} \cdot E[z \cdot z'] \left(\Sigma^{1/2}\right)'\right] = \left[\Sigma^{1/2} \cdot \left(\Sigma^{1/2}\right)'\right] = \Sigma \end{aligned}$$

We can calculate the density function of the vcb-mcfd considering that variables z_i follow an i-mcfd$_3$ and using the relationship between the variables R and z of Equation 1.27 making use of the density transformation theorem (see Johnson and Kotz 1972a). The density transformation theorem states that given a vector z with multivariate density function $f_z(z)$, then the multivariate density function $f_R(R)$ of a variable R defined by a transformation $R = g(z) = Az + m$, where A is a matrix is given by:

$$f_R(R) = \frac{1}{\det[A]} \cdot f_z(g^{-1}(R))$$

where $\det[A]$ denotes the determinant of matrix A. Therefore, applying this result to Equation 1.27 and considering that $\det\left[\Sigma^{1/2}\right] = \det[\Sigma]^{1/2}$ we obtain directly the expression for the vcb-mcfd$_3$:

$$\text{vcb-mcfd}_3(R) = \frac{1}{\sqrt{\det[\Sigma]}} \text{i-mcfd}_3(\Sigma^{-1/2}(R-m)) \tag{1.28}$$

[25] The Cholesky decomposition B of a matrix A is given by:

$$A = B * B'$$

The Cholesky decomposition exists only if A is definite positive.

Statistical Properties

The first interesting property of this multivariate density is that it is closed under convolution, i.e. the sum of two VCB-MCFD variables is also a VCB-MCFD[26].

Proposition 9 *Let $(R_i)_{i=1}^n$ be variables that follow a VCB-MCFD with parameters of the distribution given by $a_{3,i}$, $a_{2,i}$, μ and Σ. Then, any variable W defined as a sum of variables R_i is also a VCB-MCFD variable.*

This result is of special interest for financial applications, as portfolio management or portfolio risk assessment, where the sum of aleatory variables is usually involved. It is important to note here that the marginals of the VCB-MCFD are not exactly Cornish-Fisher distributed, only in the independent case (when Σ is diagonal) the marginal distributions are Cornish-Fisher. In order to check this consider two VCB-MCFD variables, R_1 and R_2, which are defined in terms of the I-MCFD variables, z_1 and z_2 with parameters $a_{3,i}$ and $a_{2,i}$. From Equation 1.27 we have:

$$\begin{aligned} R_1 &= w_{11} \cdot z_1(a_{1,3}, a_{1,2}) + w_{12} \cdot z_2(a_{2,3}, a_{2,2}) + m_1 \\ R_2 &= w_{21} \cdot z_1(a_{1,3}, a_{1,2}) + w_{22} \cdot z_2(a_{2,3}, a_{2,2}) + m_2 \end{aligned}$$

with $w_{ij} = \Sigma_{ij}^{1/2}$ and where we have written explicitly the dependence of the variables z_i. From this Equation it is obvious that the variable R_1 depends not only on the parameters $a_{1,j}$ but also on the parameters $a_{2,j}$ and, therefore, it cannot be distributed as univariate CFD. Only in the independent case ($w_{12} = w_{21} = 0$) we have $R_1 = R_1(a_{1,3}, a_{1,2})$. Although

[26] The better known example of a density closed under convolution is the case of the multivariate normal distribution but other distributions, as the multivariate α-stable distribution or the multivariate hyperbolic distributions, also exhibit this property.

no explicit expression is available for the density function of this marginal distribution, analytical formula for its multivariate moments are easy to calculate using tensor calculus:

Proposition 10 *Let $(R_i)_{i=1}^n$ be variables that follow a VCB-MCFD with parameters of the distribution given by $a_{3,i}$, $a_{2,i}$, μ and Σ. Then, the mean vector M_1 and the second, third and fourth centered multivariate moments, given by the variance-covariance matrix, M_2 the skewness tensor, M_3, and the kurtosis tensor, M_4, respectively, are given by:*

$$
\begin{aligned}
M_{1,i} &= \mu_i \\
M_{2,ij} &= \Sigma_{ij} \\
M_{3,ijk} &= \sum_{r=1}^{n} w_{ir} w_{jr} w_{kr} \mu_{3,r} \\
M_{4,ijkl} &= \sum_{r=1}^{n} w_{ir} w_{jr} w_{kr} w_{lr} \mu_{4,r} \\
&\quad + \sum_{r=1}^{n} \sum_{\substack{s=1 \\ s \neq r}}^{n} \begin{pmatrix} w_{ir} w_{jr} w_{ks} w_{ls} + w_{ir} w_{js} w_{kr} w_{ls} \\ + w_{ir} w_{js} w_{ks} w_{lr} \end{pmatrix}
\end{aligned}
$$

where $\mu_{3,r}$ and $\mu_{4,r}$ are the third and fourth order centered univariate moments for the variable r of the Cornish-Fisher Density, given by Equations 1.11*, and w_{ij} denotes the ij-element of the Cholesky decomposition of the covariance matrix, i.e, $\Sigma^{1/2} = (w_{ij})$, $i, j = 1, ..., n$.*

The Proposition above demonstrates that the higher centered moments depend on the variance-covariance matrix, Σ, and the individual higher moments, μ_i. For example, following the above bivariate example, the kurtosis coefficient of variable R_1 depends on the kurtosis coefficients of the variables z_1 and z_2. Therefore, through this specification, the structure of dependencies is not only captured through the covariance matrix Σ but also through the coefficients $a_{i,j}$ of each variable z_i. Therefore, this multivariate density allows for the occurrence of simultaneous extreme events (in opposition to the CB-MCFD) as can be seen from its definition: first we permit univariate extreme events (through z) and then we "mix" the fictitious independent CFD variables z to model the vector R (see Equation 1.27). In this way, the variables R may present simultaneous extreme events.

In Figure 1.8 we present some possible shapes for standardized third-order VCB-MCFD showing its flexibility to adjust rather different distribution patterns. It is interesting to note that with the same parameters as in Figures 1.7 we obtain very different contour plots compared with the CB-MCFD model: only for the independent case ($\Sigma = 0$) both densities are equivalent, as can be seen in the two upper Figures. When the variables are positively ($\Sigma_{ij} > 0$) or negatively ($\Sigma_{ij} < 0$) correlated, we observe that the VCB-MCFD is able to capture more extreme events than the CB-MCFD.

Simulation of VCB-MCFD Variables

Equation 1.27 along with the definition of an Independent MCFD gives us a very simple way of simulating m-th VCB-MCFD variables $(R_i)_{i=1}^{n}$ using the functions Q_i, the variance-covariance matrix, Σ, and the mean vector m. The algorithm is as follows:

1. First, we simulate n independent univariate standard gaussian variables X.

2. Second, we apply Equation 1.5, $z_i = Q_i(X_i)$, for each variable X_i to obtain variables the Independent MCFD variables, z_i.

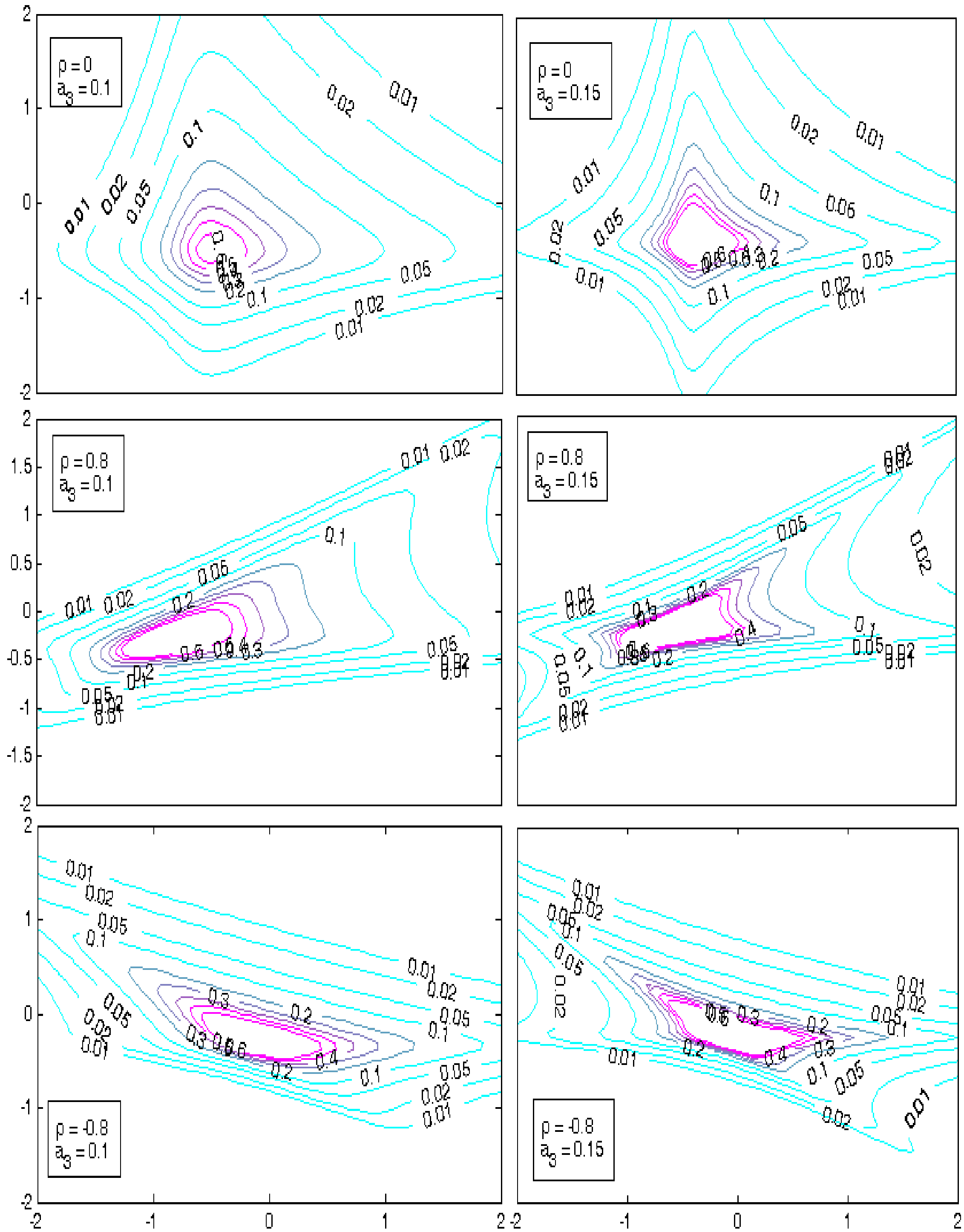


Fig. 1.8. Contour plots corresponding to different shapes of the bivariate third-order VCB-MCFD. All densities are standardized to zero mean and unit variance and, therefore, we only have three free parameters defining the distribution: a_2, which defines the skewness, a_3, which incorporates kurtosis, and the normal rank correlation ρ. In all the figures the parameter a_2 is set to 0.25, so that both marginal densities are positively skewed. The upper Figures show that this distribution is equivalent to the CB-MCFD when variables are independent ($\rho = 0$). The other Figures show that the VCB-MCFD captures more dependence on the tails than the CB-MCFD.

3. And finally, we calculate the Cholesky decomposition, $\Sigma^{1/2}$, of the variance-covariance matrix and apply Equation 1.27, $R = \Sigma^{1/2} \cdot z + m$, to the whole vector z to obtain the VCB-MCFD variables.

Again, this algorithm is very fast, as very efficient algorithms to calculate the Cholesky decomposition are available in each computation package.

1.4 Descriptive Data Analysis

Daily returns of the great majority of variables in financial markets, in particular those of exchange rates and market indexes, present a positive excess kurtosis (Bouchaud and Potters 2000). This behavior implies that extreme movements in market variables are much more frequent compared to predictions of a normal distribution. It is relatively common to find each year, that some data in daily series are over 10 standard deviations of the mean. A normal distribution would assign to this phenomenon a probability of order 10^{-23}, and in many cases the observed empirical probability can be of the order of 10^{-4}.

To illustrate the problem of non-normality in market variables we will analyze two different data bases[27]. The first one consists on daily exchange returns against the USD (United States dollar), for the twelve mayor currencies between January 4 of 1988 and the August 15 of 1997, making in total 2.245 observations. Foreign currencies are the Australian dollar (AUD), the Belgian franc (BEF), the Swiss franc (CHF), the German mark (DEM), the Danish crown (DKK), the Spanish peseta (ESP), the French franc (FRF), the

[27] Both data bases have been collected from a Bloomberg terminal and we are thankful to Consulnor for giving us access to it.

English pound (GBP), the Italian lira (ITL), the Japanese yen (JPY), the Dutch guilder (NGL) and the Swedish crown (SEK). The second data base consists on weekly returns (from Wednesday to Wednesday) for dollar denominated stock indexes for the main geographical areas: North America, Japan, Europe, Emerging Markets and Eastern Europe Emerging Markets, represented by the Standard and Poor's 500 Index (S&P) , the Nikkei-225 Stock Average (NKI), the Dow Jones EURO STOXX (STX), MSCI Emerging Markets Index (EM) and the MSCI Eastern Europe Emerging Market Index (EME)[28]. This data set consists of total logarithmic return indexes from January 4 of 1995 to the March 23 of 2005, making in total 519 observations.

As a preliminary investigation of the data, Table 1.1 presents a summary of the most important univariate statistics of the twelve exchange rates and the five index Series. For both data sets we begin estimating the first four moments and three tests of the null hypothesis of normality of univariate distributions. Standard errors of the moments are computed with the Generalized Moments Method (GMM) based procedure proposed by Bekaert and Harvey 1997[29]. Since the normality hypothesis is crucial to our analysis, we paid a particular attention to this test. Although a large number of test have been proposed in the

[28] Standard and Poor's 500 Index is an index consisting of 500 US stocks that have been selected because of the market size, liquidity and industry group representation, among other factors. The Nikkei-225 Stock Average is a price-weighted average of 225 top-rated Japanese companies listed in the First Section of the Tokyo Stock Exchange. Dow Jones EURO STOXX is a sub index of the Dow Jones STOXX 600 and it covers most of the European countries. The MSCI Emerging Markets Index is a float-adjusted market capitalization index. As of May 2005, it consisted of indices in 26 emerging economies: Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Jordan, Korea, Malaysia, Mexico, Morocco, Pakistan, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, Turkey and Venezuela. The Eastern Europe MSCI Index captures the capital markets behavior of Poland, Hungary, Russia and Czech Republic.

[29] The coefficients of skewness and kurtosis are jointly estimated along with the mean and variance in an exactly identified GMM system with four orthogonality conditions. The variance-covariance matrix of the parameters is heteroscedasticity consistent and corrects for serial correlation using a Bartlett kernel with an optimal band as in Andrews 1991.

literature, we focus on three well-known procedures: first the Jarque-Bera (JB) statistic proposed by Bera and Jarque 1982, which test whether skewness and excess kurtosis are jointly zero. This test is known to be suitable for large samples only, because skewness and kurtosis approach normality only very slowly. Second, a Wald test of the null hypothesis that the skewness and excess kurtosis coefficients are zero based on the GMM estimates, as proposed by Bekaert and Harvey 1997, which incorporates the approximated finite-sample distribution of skewness and kurtosis. Third, the Kolmogorov-Smirnov (KS) statistic which is based on the comparison between the theoretical and the empirical cumulative distribution functions.

Exchange rates data. Parameter estimates for the means are not significant for any of the daily series. The Jarque-Bera and Wald tests indicate that in none of the series the gaussianity hypothesis can be sustained. Daily variance ranges from 0.3136 for the Australian dollar to 0.5852 for the Swedish crown. Skewness for the exchange rates against the dollar tend to present a negative sign (9 negative and 3 positive) and are in general not significant at a 5% significance level. However, in the Australian dollar or the Japanese yen a significant skewness is observed. It is interesting to note that the AUD is the only series presenting positive skewness. On the other hand, the kurtosis coefficient reveals that the unconditional distribution of returns has heavier tails than the normal one for all exchange rates. This coefficient ranges from 14.46 for the Swedish crown to 5.05 for the Swiss franc. The Jarque-Bera, Wald and KS tests indicate that in none of the series the gaussianity hypothesis can be sustained with p-values close to zero. Therefore, it is necessary to consider a model that picks up both observed deviations with respect to normality.

Index data. Parameter estimates for the mean are only significant for the S&P and the STX. The weekly (annualized) return ranges from 9.36 for the S&P Index and a -5.148 for the Japan zone. Annualized standard deviations range from 21.237 for the S&P Index to 31.289 for the MSCI Emerging Markets Index. As we would expect, the most volatile indexes are those corresponding to Emerging Markets. The Jarque-Bera and Wald tests indicate that in none of the series the gaussianity hypothesis can be sustained. Although only Emerging Markets present significant negative skewness, with values -0.718 and -0.443, in general we find that the skewness coefficient is negative, indicating that crashes are more likely to occur than booms. The kurtosis coefficient, ranges from 3.9228 for the Nikkei to 5.997 for the E-STOXX. It is interesting to note that daily returns present heavier tails than weekly returns as should be expected considering the central limit theorem. Nevertheless, as in the exchange rates database presented above, the Jarque-Bera, Wald and KS tests indicate that the gaussianity hypothesis cannot be accepted with p-values close to zero.

For a better characterization of the data, a test for serial correlation of returns is also considered. We use the Q Ljung-Box statistic (Ljung and Box 1978) to test the null hypothesis of no serial correlation. Inspection of the Q statistics in Table 1.2 reveals that daily returns of exchange rates do not exhibit in general serial correlation (only two series, namely, ESP and ITL present significant correlation) in contrast to weekly returns of indexes, where in four of five series we find that the null hypothesis cannot be rejected. Next, we also test for heteroskedasticity by regressing squared returns on lagged squared returns using the Engle 1982 test. Table 1.2 provides evidence of second-order dynamics for all

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
Mean	-0.0010	-0.0050	-0.0065	-0.0056	-0.0050	-0.0142	-0.0054	-0.0062	-0.0172
	0.0108	0.0138	0.0154	0.0139	0.0136	0.0146	0.0134	0.0145	0.0139
σ^2	**0.3136**	**0.4844**	**0.5852**	**0.4719**	**0.4597**	**0.5258**	**0.4295**	**0.4626**	**0.4889**
	0.0178	0.0285	0.0254	0.0246	0.0253	0.0513	0.0233	0.0326	0.0477
ξ	**0.7808**	-0.1158	0.0323	-0.1301	-0.1012	-0.6084	-0.1305	-0.2532	**-1.0396**
	0.2000	0.1166	0.1238	0.1102	0.1138	0.4598	0.1271	0.1553	0.4703
κ	**7.3163**	**5.6157**	**5.0500**	**5.0835**	**5.1153**	**11.2490**	**5.3087**	**6.0214**	**11.8247**
	1.1286	0.3411	0.3577	0.2938	0.2777	4.6166	0.3101	0.4887	3.4787
Max	4.0870	3.0035	3.6436	2.8963	3.1232	3.8475	2.8195	4.1319	6.8213
Median	-0.0256	0.0000	-0.0160	0.0000	-0.0069	-0.0078	0.0000	0.0000	0.0000
Min	-2.0482	-3.6037	-4.0268	-3.3817	-3.1391	-7.9446	-3.3503	-3.2446	-2.9514
JB	2158	706	430	451	462	7122	552	960	8421
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald	15.52	60.20	37.62	52.15	58.02	4.96	56.09	42.34	7.40
p-val	0.000	0.000	0.000	0.000	0.000	0.083	0.000	0.000	0.024
KS	0.5091	0.3152	0.1418	0.5608	0.2800	0.2589	0.5447	0.4002	0.3297
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
Mean	0.0018	-0.0056	-0.0128	**0.180**	-0.099	**0.150**	0.027	0.110
	0.0137	0.0139	0.0140	0.0974	0.1282	0.1277	0.1494	0.2161
σ^2	**0.4689**	**0.4739**	**0.5026**	**5.6148**	**9.3441**	**8.6719**	**8.1140**	**18.832**
	0.0259	0.0258	0.0433	0.6436	0.7498	1.4039	0.8541	3.0010
ξ	0.2019	-0.1049	-0.8023	-0.085	-0.147	-0.414	**-0.718**	**-0.443**
	0.2444	0.1142	0.6548	0.1851	0.1462	0.2438	0.2291	0.2002
κ	**7.9843**	**5.2630**	**14.4600**	**4.509**	**3.928**	**5.997**	**5.098**	**5.604**
	1.0354	0.3048	6.8039	0.4064	0.3084	0.7616	0.7744	0.6794
Max	4.8744	3.0456	4.5953	10.182	11.307	14.461	8.589	21.236
Median	0.0000	0.0000	0.0000	0.332	-0.024	0.519	0.176	0.512
Min	-4.6809	-3.2734	-8.5299	-9.041	-10.281	-12.636	-13.978	-17.122
JB	2561	528	13720	49	21	209	140	164
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald	24.20	56.69	6.59	14.67	9.10	28.74	9.98	29.26
p-val	0.000	0.000	0.037	0.000	0.010	0.000	0.006	0.000
KS	0.4111	0.4057	0.4452	0.3430	0.1843	0.4473	0.4236	0.2916
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 1.1. Descriptive univariate statistics on returns. Mean, σ^2, ξ and κ denote the mean, variance, skewness, and kurtosis of returns, respectively. Estimates are not annualized and standard errors are computed with the GMM-based procedure proposed by Bekaert and Harvey 1997. Significant parameters at a 5% are in bold. Max, Median and Min, stand for the Maximum, Median and Minimum observation. JB, Wald and KS are the Jarque-Bera statistic (Bera and Jarque 1982), the Wald test of joint significance of skewness and kurtosis (Bekaert and Harvey 1997) and the Kolmogorov-Smirnov statistic under the null hypothesis of normality, respectively.

data sets expected for the Nikkei, confirming that there is a large amount of heteroskedasticity in both daily and weekly returns.

After analyzing preliminary univariate statistics, we will next turn to the multivariate analysis. In Tables 1.3 and 1.4 we show the correlation matrix of both data sets and the GMM-estimates of the finite sample standard deviation, below and above the diagonal respectively. Table 1.3 indicates that all exchange series are positively correlated except the Australian Dollar, which presents a non significant negative correlation coefficient. With respect to the indexes, all series are also positively correlated ranging from 0.2650 for the Japan and Eastern Europe areas to 0.7576 from the North American and European areas. Finally, we perform several multivariate normality tests. As compared to the univariate tests discussed above, these multivariate tests incorporate hypothesis on the co-skewness and co-kurtosis matrices. We consider five different tests for multivariate normality: the Shapiro-Wilk statistic (Royston 1982), the Mardia A and B statistics (Mardia 1985) which measure deviations in skewness and kurtosis matrices, respectively, the omnibus statistic (Doornick and Hansen 1994) based on the approximated finite-sample distribution of skewness and kurtosis, and a test proposed by Malevergne and Sornette 2003 which is based on the property that a sum of squared gaussian variables are χ^2 distributed. These tests are briefly described in Appendix C.1. As Table 1.5 shows, according to all the tests none of the data sets present a multivariate gaussian behavior.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
Q(1)	2.0672	1.2036	0.0069	0.0839	0.5618	**5.5558**	0.1054	0.1420	**5.3298**
	0.1505	0.2726	0.9338	0.7721	0.4535	0.0184	0.7454	0.7063	0.0210
Q(4)	9.3010	1.3448	0.6354	0.2815	0.7764	**12.5780**	1.9332	3.7215	6.5554
	0.0540	0.8537	0.9591	0.9910	0.9416	0.0135	0.7480	0.4450	0.1613
Q(10)	14.828	10.994	4.773	6.188	9.006	16.678	7.047	14.071	15.872
	0.1384	0.3580	0.9058	0.7992	0.5315	0.0818	0.7210	0.1698	0.1033
LM(1)	**24.15**	**42.52**	**26.61**	**31.52**	**15.62**	**6.34**	**30.99**	**10.25**	**18.76**
	0.0000	0.0000	0.0000	0.0000	0.0001	0.0118	0.0000	0.0014	0.0000
LM(5)	**56.96**	**75.37**	**32.42**	**54.75**	**53.65**	**167.32**	**63.76**	**104.40**	**208.11**
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LM(10)	**118.25**	**92.43**	**82.60**	**86.93**	**77.73**	**200.33**	**109.44**	**137.15**	**225.57**
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
Q(1)	0.3403	0.3960	1.0272	3.1854	0.9820	**7.4716**	2.9898	2.8533
	0.5597	0.5292	0.3108	0.0743	0.3217	0.0063	0.0838	0.0912
Q(4)	0.6746	0.6722	1.8337	5.0531	3.3629	**17.263**	**16.503**	**12.153**
	0.9544	0.9547	0.7663	0.2819	0.4990	0.0017	0.0024	0.0162
Q(10)	13.073	6.184	12.192	17.766	**25.907**	**25.537**	**22.049**	17.468
	0.2196	0.7995	0.2724	0.0590	0.0039	0.0044	0.0149	0.0646
LM(1)	**17.06**	**31.79**	**102.10**	**17.51**	0.7483	**52.03**	1.3503	**9.40**
	0.0000	0.0000	0.0000	0.0000	0.3870	0.0000	0.2452	0.0022
LM(5)	**30.60**	**58.26**	**119.75**	**41.65**	7.1681	**73.44**	**29.39**	**53.24**
	0.0000	0.0000	0.0000	0.0000	0.2084	0.0000	0.0000	0.0000
LM(10)	**43.61**	**81.57**	**124.75**	**43.59**	12.0315	**89.74**	**39.51**	**84.86**
	0.0000	0.0000	0.0000	0.0000	0.2830	0.0000	0.0000	0.0000

Table 1.2. Descriptive univariate statistics on returns (cont.). Q(1), Q(4), and Q(10) denote the Box-Ljung statistic for serial correlation, corrected for heteroscedasticity, computed with 1, 4 and 10 lags, respectively. Under the null hypothesis of no serial correlation, it is distributed as a $\chi^2(i)$. LM(1), LM(5) and LM(10) stand for the Engle 1982 test with 1, 4 and 10 squared lags. Under the null of no serial correlation of squared returns, the statistics are distributed as a $\chi^2(i)$. Significant not acceptance at a 5$\%$ are in bold.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		0.0310	0.0334	0.0328	0.0322	0.0221	0.0343	0.0354	0.0356	0.0367	0.0320	0.0267
BEF	-0.0529		0.0473	0.0553	0.0569	0.0584	0.0583	0.0490	0.0576	0.0402	0.0592	0.0482
CHF	-0.0570	0.8740		0.0497	0.0452	0.0409	0.0471	0.0454	0.0461	0.0411	0.0483	0.0380
DEM	-0.0457	0.9367	0.9204		0.0538	0.0479	0.0568	0.0505	0.0574	0.0414	0.0565	0.0455
DKK	-0.0450	0.9172	0.8599	0.9249		0.0559	0.0553	0.0489	0.0555	0.0396	0.0558	0.0489
ESP	0.0277	0.4315	0.3825	0.4164	0.4447		0.0591	0.0567	0.0664	0.0330	0.0588	0.0488
FRF	-0.0532	0.9413	0.8904	0.9554	0.9300	0.4388		0.0507	0.0606	0.0409	0.0580	0.0483
GBP	-0.1645	0.6505	0.6604	0.6920	0.6588	0.3116	0.6862		0.0663	0.0349	0.0494	0.0429
ITL	-0.1020	0.6710	0.6306	0.6869	0.6678	0.3886	0.7077	0.6228		0.0315	0.0577	0.0490
JPY	-0.0188	0.5885	0.6127	0.6201	0.5826	0.2456	0.6002	0.4438	0.3969		0.0405	0.0286
NGL	-0.0485	0.9532	0.9062	0.9757	0.9367	0.4429	0.9598	0.6807	0.6910	0.6113		0.0477
SEK	-0.0561	0.7140	0.6755	0.7254	0.7621	0.3693	0.7300	0.5571	0.5917	0.4537	0.7288	

Table 1.3. Correlations between the different exchange rates. Correlation estimates are under the diagonal and standard deviations are shown over the diagonal. Standard errors are computed with the GMM-based method (Bekaert and Harvey 1997).

	S&P	**NKI**	**STX**	**EM**	**EME**
S&P		0.0522	0.1217	0.0767	0.0696
STX	0.4037		0.0535	0.0545	0.0585
NKI	0.7576	0.4332		0.0866	0.0895
EM	0.5871	0.4543	0.6391		0.0993
EME	0.3733	0.2650	0.4230	0.6264	

Table 1.4. Correlations between indexes. Correlation estimates are below the diagonal and standard deviations are shown above the diagonal. Standard errors are computed with the GMM-based (Bekaert and Harvey 1997).

	Shapiro-Wilk	**Mardia A**	**Mardia B**	**Omnibus**	**KS-χ^2**
Exchange rate	0.7681 (0.000)	21221 (0.000)	649 (0.000)	39479 (0.000)	0.2527 (0.000)
Indexes	0.9591 (0.000)	159 (0.000)	24.04 (0.000)	335.0 (0.000)	0.1672 (0.000)

Table 1.5. Multivariate tests for normality. According to five different tests described in Appendix C.1, in this table we report the statistic value under the null hypothesis and its p-value, which show that none of the series present a multivariate gaussian behavior.

1.5 Univariate Inference

In this Section we will describe different methods that we have used to fit the univariate Cornish-Fisher function introduced in Section 1.2 and we will also analyze the goodness of fit of the CFD to the data sets presented in Section 1.4. First, we will describe three different estimation methods, namely, a quantile-quantile based estimation, the moments method and the Maximum Likelihood method. Next, we will analyze a static CFD model and analyze the in-sample performance of the CFD compared to other models. Finally, we will estimate a dynamic GARCH model with innovations distributed as CFD and also compare the in-sample results with other models.

1.5.1 Methods of Estimation

As we have already mentioned above, in this work we propose a third-order polynomial to model the function $Q_i(X)$ of Equation 1.5:

$$Q_i(X) = a_{3,i}X^3 + a_{2,i}X^2 + a_{1,i}X + a_{0,i} \tag{1.29}$$

We analyze different methodologies to fit the coefficients in the transformation function:

1. Perform a least square fitting of the experimental function Q_i arising from the QQ-Plot (QQ-estimates).

2. Fit the first four moments of the theoretical distribution to the experimental ones (MM-estimates).

3. Choose the Maximum Likelihood estimates that maximizes the logarithm of the likelihood function of the distribution CF_3 (ML-estimates).

It is interesting to point out that the first one is the most direct or computationally less expensive of the three methodologies because it only involves a least-squares algorithm that consists basically in matrix calculations. On the other hand, the fact that the density function, $cf_3(R)$, is defined through the function $Q_i(X)$ makes this methodology specially appropriate. The number of available points to make the regression will depend directly on the number of available historical data. In principle, the number of points in a QQ-Plot is chosen by the researcher but the most appropriate one consists on considering the interval for the percentiles defined by each one of the points of the QQ-Plot in such a way that it is the inverse of the number of sample data in the series. For example, if we had 100 data, the first point would correspond to the percentile of 1%, and if we had 1000 data it would correspond to the percentile of 0.1%.

In the second method we carry out the fitting based on the moments method, which basically consists on comparing the sample moments to the theoretical ones by means with relations derived in Equations 1.11, that relate the first four moments of the distribution with the coefficients in the third-order polynomial. However, this estimate involves a non-linear process of optimization, because Equations 1.11 show the moments as a function of the coefficients and the inverse would be required in order to estimate the coefficients as a function of the moments associated to the distribution. For this purpose, we have applied the macro FMINSEARCH of the MATLAB optimization toolbox and we use the

QQ-estimates as initial estimates in the iterative process for the search algorithm. It is interesting to point out that in all the cases considered the obtained convergence has always been fast. In Appendix C.1 we analyze in detail the developed algorithm for the optimization

Finally, the third method based on a maximum-likelihood fitting would not be in principle well adapted in our case, since, as can be seen in Equation A.1, coefficients of the third-order polynomial enter in the density function in a highly non-linear way, which greatly complicates the numeric process of optimization. However, we have seen that using adequate starting parameters for the non-linear optimization algorithm, namely the QQ-estimates, we have been able to achieve global maxima for the log-likelihood functions easily and quickly.

On the other hand, it has to be remembered that the function $Q_i(X)$ has to be invertible and, therefore, restrictions shown in Equation 1.10 must hold. Given that historical data present a positive excess kurtosis, as is generally verified in financial series, these conditions will be naturally satisfied when using the QQ-Plot based method. For the other two methods it is necessary to impose explicitly the restrictions in the optimization procedures. Nevertheless, using the estimates of the quantile-quantile based method and given the high flexibility of the third-order CFD (see Section 1.2) we always get inner solutions to the optimization problems. In contrast, Jondeau and Rockinger 2001 who use Gram-Charlier Densities, which are not so flexible, obtain many frontier solutions.

In Appendix C.2 these algorithms are tested and we also investigate the properties of the maximum-likelihood, moments and QQ-estimates when Cornish-Fisher Densities are used in an attempt to directly obtain higher moments that differ from the ones of the normal distribution using Montecarlo experiments. We consider the fit of CFDs to data generated with a Cornish-Fisher distribution and to data generated with a mixture of normals. Furthermore, in the latter type of simulation we distinguish the situation where parameters for the simulated data are in or out of the restricted domain given by the Figure 1.1. In our simulations we simulate N=100 series of length T=2000 of data standardized CF with zero mean and unit variance. According to these experiments we can conclude that in general the three algorithms are well behaved and that the estimations of the QQ and ML are sensibly better than the ones corresponding to the MM. Besides that, we find that ML-estimates are the most efficient when the true distribution is CFD. In addition, with the second experiments we find that the estimates errors and dispersions are very similar in the whole region of permitted values.

1.5.2 Static Framework

In this Section a first approach to the empirical analysis will be described. We estimate the parameters under the hypothesis that both data bases described in Section 1.4 are independent CFD distributed and the three estimation methods defined above will be considered. In each case we will test the kindness of the fit using four different statistics. The first one is the classical Kolmogorov-Smirnov statistic, that tests for the similarity between the empirical distribution function and the CFD distribution. For the other three tests we take advantage of the property that if R is a CFD variable, the variable X defined as $X = Q^{-1}(R)$ (see Equation 1.5) should be normal if the CFD specification is sufficiently correct. Therefore, we can apply on these fictitious X variables any of the usual normality tests to analyze if returns are CFD distributed or not. In particular, we use three different test: (1) the Jarque-Bera test that checks if the third and fourth moments of the sample are equal to the normal ones , (2) the Liliefors test, which is a KS-like test specially meant for testing normality and, therefore, looks at the central part of the cumulative distribution, and (3) the Anderson-Darling test which gives special attention on the tails of the distribution. In Appendix C.3 these tests are presented briefly.

Table 1.6 shows the QQ-estimates and the four different statistics while Table 1.7 and Table 1.8 report the results for the MM and ML methods. In each Table we report the estimates for the coefficients a_i with their standard errors, which have been calculated using a bootstrap method with 1000 simulations for the QQ and MM-estimates, and using the Hessian matrix evaluated at the Maximum Likelihood estimates for the ML method. We also present the Kolmogorov-Smirnov, Jarque Bera, Liliefors and Anderson-Darling statistics with their p-values, and non rejection of the null hypothesis that unconditional returns are samples of a CFD at a 5% significance level are in bold.

First, in general it should be noticed the high degree of non-rejection of the CFD hypothesis, which indicates that in almost all cases we have considered just four parameters are required to capture the non-normal (unconditional) behavior of financial series. Comparing the results obtained with the different methods, first we observe that the p-values for the KS test of the MM-estimates are systematically smaller than the p-values of the QQ-estimates and that both are smaller than the ML-estimates. As a consequence, using the KS statistic at a 5% significance level, only five, one and none of the 17 series considering the MM, QQ and ML-estimates, respectively, the null hypothesis that the data are Cornish-

	$\hat{a}_3$		$\hat{a}_2$		$\hat{a}_1$		$\hat{a}_0$		**KS-Test**		**JB-Test**		**Liliefors**		**AD-Test**	
	est.	std.	est.	std.	est.	std.	est.	std.	est.	p-v	est.	p-v	est.	p-v	est.	p-v
AUD	0.0404	0.0053	0.0403	0.0094	0.4265	0.0132	-0.0412	0.0111	**0.025**	0.07	**1.360**	0.50	0.024	<0.01	1.491	0.00
BEF	0.0494	0.0044	-0.0095	0.0099	0.5374	0.0147	0.0045	0.0132	**0.015**	0.64	**0.168**	0.91	**0.015**	>0.2	**0.474**	0.24
CHF	0.0444	0.0054	0.0044	0.0100	0.6247	0.0166	-0.0108	0.0145	**0.016**	0.54	**0.014**	0.99	**0.016**	0.10	**0.394**	0.37
DEM	0.0420	0.0040	-0.0107	0.0088	0.5535	0.0135	0.0051	0.0132	**0.013**	0.81	**0.181**	0.91	**0.013**	>0.2	**0.437**	0.29
DKK	0.0424	0.0039	-0.0084	0.0087	0.5429	0.0144	0.0034	0.0126	**0.016**	0.52	**0.227**	0.89	**0.016**	0.11	**0.525**	0.18
ESP	0.0669	0.0109	-0.0216	0.0151	0.5038	0.0251	0.0074	0.0150	**0.020**	0.29	**0.728**	0.69	0.019	0.04	0.751	0.05
FRF	0.0433	0.0041	-0.0102	0.0084	0.5170	0.0134	0.0047	0.0121	**0.014**	0.75	**0.213**	0.89	**0.013**	>0.2	**0.394**	0.37
GBP	0.0543	0.0048	-0.0184	0.0097	0.5044	0.0149	0.0122	0.0124	**0.020**	0.25	**0.619**	0.73	0.019	0.02	1.013	0.01
ITL	0.062	0.0096	-0.0420	0.0146	0.4886	0.0215	0.0248	0.0147	0.044	0.00	**3.410**	0.18	0.041	<0.01	2.713	0.00
JPY	0.0641	0.0077	0.0175	0.0130	0.4749	0.0181	-0.0157	0.0140	**0.015**	0.65	**0.244**	0.88	**0.015**	>0.2	**0.280**	0.64
NGL	0.0445	0.0043	-0.0086	0.0090	0.5466	0.0146	0.0030	0.0130	**0.012**	0.85	**0.177**	0.91	**0.012**	>0.2	**0.349**	0.47
SEK	0.0649	0.0127	-0.0226	0.0174	0.4917	0.0287	0.0098	0.0166	**0.022**	0.16	**1.327**	0.51	0.021	<0.01	0.710	0.06
S&P	0.1218	0.0302	-0.0406	0.0611	1.9906	0.0969	0.2205	0.0991	**0.035**	0.55	**0.145**	0.93	**0.034**	0.14	**0.496**	0.21
NKI	0.1103	0.0288	0.0547	0.0651	2.7170	0.1335	-0.1541	0.1286	**0.025**	0.90	**0.027**	0.98	**0.025**	>0.2	**0.350**	0.47
STX	0.2288	0.0484	-0.1930	0.0979	2.2029	0.1384	0.3428	0.1183	**0.034**	0.39	**0.407**	0.81	**0.037**	0.07	**0.758**	0.05
EM	0.1493	0.0358	-0.2617	0.0770	2.3624	0.1290	0.2881	0.1215	**0.026**	0.87	**0.092**	0.95	**0.026**	>0.2	**0.266**	0.69
EME	0.3391	0.0685	-0.2828	0.1257	3.2338	0.2313	0.3928	0.1724	**0.041**	0.35	**0.373**	0.83	0.039	0.05	0.710	0.06

Table 1.6. Quantile-Quantile (QQ) estimates of the exchange rate and indexes databases. We present the estimates of the parameters of the CFD: a_3, a_2, a_1 y a_0, and the standard errors of these estimates. Standard errors are calculated using a bootstrap re-sampling procedure with 1000 simulations. KS-Test denotes the Kolmogorov-Smirnov statistic for the CFD null hypothesis and JB-Test, Liliefors and AD-Test stand for the Jarque-Bera, Liliefors and Anderson Darling statistic that test different aspects of the normality of the fictitious normal variables defined by Equation 1.5. The critical value for the KS statistic (JB statistic) at a 5% level of significance is 0.0273 (5.9915) for the exchange rates database and 0.0593 (5.9915) for the index database. Cases for which the null hypothesis cannot be rejected at a 5% level of significance appear in bold in the column of the statistics.

	$\hat{a}_3$		$\hat{a}_2$		$\hat{a}_1$		$\hat{a}_0$		**KS-Test**		**JB-Test**		**Liliefors**		**AD-Test**	
	est.	std.	est.	std.	est.	std.	est.	std.	est.	p-v	est.	p-v	est.	p-v	est.	p-v
AUD	0.0427	0.0053	0.0482	0.0094	0.4177	0.0132	-0.0492	0.0111	0.034	0.00	**5.462**	0.06	0.030	<0.01	2.253	0.00
BEF	0.0431	0.0044	-0.0095	0.0099	0.5584	0.0147	0.0045	0.0132	**0.020**	0.24	**3.362**	0.18	0.019	0.03	0.878	0.02
CHF	0.0405	0.0054	0.0031	0.0100	0.6371	0.0166	-0.0095	0.0145	**0.018**	0.40	**0.851**	0.65	**0.017**	0.05	**0.534**	0.17
DEM	0.0365	0.0040	-0.0110	0.0088	0.5713	0.0135	0.0054	0.0132	**0.017**	0.40	**2.821**	0.24	**0.016**	0.08	**0.838**	0.03
DKK	0.0365	0.0039	-0.0084	0.0087	0.5624	0.0144	0.0034	0.0126	**0.021**	0.21	**3.480**	0.17	0.020	0.02	**0.968**	0.01
ESP	0.0854	0.0109	-0.0400	0.0151	0.4358	0.0251	0.0258	0.0150	0.042	0.00	20.70	0.00	0.034	<0.01	4.382	0.00
FRF	0.0374	0.0041	-0.0103	0.0084	0.5367	0.0134	0.0049	0.0121	**0.018**	0.35	**3.660**	0.16	**0.017**	0.06	**0.821**	0.03
GBP	0.0457	0.0048	-0.0198	0.0097	0.5333	0.0149	0.0136	0.0124	0.027	0.04	7.441	0.02	0.024	<0.01	2.184	0.00
ITL	0.0815	0.0096	-0.0665	0.0146	0.4188	0.0215	0.0493	0.0147	0.069	0.00	33.45	0.00	0.060	<0.01	9.374	0.00
JPY	0.0626	0.0077	0.0141	0.0130	0.4792	0.0181	-0.0123	0.0140	**0.013**	0.78	**0.178**	0.91	**0.012**	>0.2	**0.218**	0.84
NGL	0.0388	0.0043	-0.0088	0.0090	0.5653	0.0146	0.0032	0.0130	**0.017**	0.43	**2.985**	0.22	**0.015**	0.12	0.737	0.05
SEK	0.0981	0.0127	-0.0473	0.0174	0.3692	0.0287	0.0345	0.0166	0.072	0.00	50.29	0.00	0.056	<0.01	10.399	0.00
S&P	0.1008	0.0302	-0.0264	0.0611	2.0541	0.0969	0.2064	0.0991	**0.038**	0.43	**0.379**	0.82	0.037	0.07	**0.556**	0.15
NKI	0.0883	0.0288	0.0632	0.0651	2.7832	0.1335	-0.1628	0.1286	**0.023**	0.92	**0.235**	0.88	**0.023**	>0.2	**0.388**	0.38
STX	0.1914	0.0484	-0.1419	0.0979	2.3264	0.1384	0.2921	0.1183	**0.044**	0.24	**1.520**	0.46	0.042	0.03	0.943	0.01
EM	0.1188	0.0358	-0.2703	0.0770	2.4512	0.1290	0.2974	0.1215	**0.030**	0.72	**0.446**	0.80	**0.029**	>0.2	**0.328**	0.51
EME	0.2533	0.0685	-0.2316	0.1257	3.5225	0.2313	0.3423	0.1724	**0.054**	0.08	**4.043**	0.13	**0.051**	<0.01	1.485	0.00

Table 1.7. Moments method (MM) estimates of the exchange rates and indexes databases. The legend is the same as Table 1.6.

	$\hat{a}_3$		$\hat{a}_2$		$\hat{a}_1$		$\hat{a}_0$		**KS-Test**		**JB-Test**		**Liliefors**		**AD-Test**	
	est.	std.	est.	std.	est.	std.	est.	std.	est.	p-v	est.	p-v	est.	p-v	est.	p-v
AUD	0.0454	0.0047	0.0248	0.0070	0.4136	0.0114	-0.0282	0.0095	**0.014**	0.64	**1.781**	0.41	**0.014**	0.19	**0.488**	0.22
BEF	0.0544	0.0035	-0.0073	0.0056	0.5241	0.0083	0.0028	0.0109	**0.013**	0.75	**0.293**	0.86	**0.013**	>0.2	**0.360**	0.44
CHF	0.0474	0.0042	0.0036	0.0063	0.6166	0.0107	-0.0103	0.0131	**0.015**	0.57	**0.175**	0.91	**0.015**	0.13	**0.346**	0.48
DEM	0.0475	0.0032	-0.0093	0.0047	0.5387	0.0074	0.0040	0.0112	**0.009**	0.97	**0.416**	0.81	**0.009**	>0.2	**0.273**	0.66
DKK	0.0482	0.0034	-0.0065	0.0053	0.5275	0.0086	0.0037	0.0110	**0.012**	0.81	**0.422**	0.80	**0.011**	>0.2	**0.353**	0.46
ESP	0.0618	0.0058	-0.0081	0.0090	0.5168	0.0136	-0.0038	0.0119	**0.012**	0.81	**0.729**	0.69	**0.012**	>0.2	**0.345**	0.48
FRF	0.0479	0.0032	-0.0074	0.0048	0.5044	0.0074	0.0026	0.0105	**0.012**	0.80	**0.282**	0.86	**0.012**	>0.2	**0.273**	0.66
GBP	0.0653	0.0057	-0.0104	0.0081	0.4760	0.0122	0.0061	0.0107	**0.011**	0.92	**1.621**	0.44	**0.011**	>0.2	**0.353**	0.46
ITL	0.0573	0.0049	-0.0090	0.0078	0.5020	0.0119	-0.0029	0.0112	**0.021**	0.20	**4.069**	0.13	0.021	<0.01	**0.520**	0.18
JPY	0.0597	0.0051	0.0146	0.0083	0.4860	0.0122	-0.0137	0.0110	**0.012**	0.80	**0.186**	0.91	**0.012**	>0.2	**0.188**	0.90
NGL	0.0496	0.0031	-0.0071	0.0047	0.5327	0.0072	0.0018	0.0110	**0.008**	0.99	**0.332**	0.84	**0.008**	>0.2	**0.216**	0.84
SEK	0.0548	0.0043	-0.0131	0.0076	0.5184	0.0114	0.0015	0.0113	**0.011**	0.89	**1.135**	0.56	**0.011**	>0.2	**0.239**	0.77
S&P	0.1217	0.0324	-0.0926	0.0560	1.9908	0.0941	0.2654	0.0955	**0.026**	0.86	**0.316**	0.85	**0.025**	>0.2	**0.349**	0.47
NKI	0.1127	0.0378	0.0239	0.0706	2.7063	0.1268	-0.1271	0.1319	**0.022**	0.96	**0.112**	0.94	**0.021**	>0.2	**0.304**	0.56
STX	0.2200	0.0418	-0.3220	0.0837	2.2358	0.1057	0.4475	0.1080	**0.028**	0.77	**1.153**	0.56	**0.028**	>0.2	**0.365**	0.43
EM	0.1376	0.0380	-0.2239	0.0626	2.3839	0.1114	0.2560	0.1146	**0.020**	0.97	**0.114**	0.94	**0.020**	>0.2	**0.207**	0.86
EME	0.4829	0.0912	-0.3516	0.1250	2.8760	0.1804	0.4610	0.1448	**0.021**	0.97	**1.143**	0.56	**0.021**	>0.2	**0.231**	0.80

Table 1.8. Maximum Likelihood (ML) estimates of the exchange rates and indexes databases. The legend is the same as Table 1.6. In this case standard errors of estimates are calculated using the Hessian matrix.

Fisher distributed is rejected. For example, observing the characteristics of the exchange rate ITL/USD, which is the only not well captured by the QQ method, it is interesting to see (Table 1.1) that although the SEK/USD exchange rate presents the highest kurtosis ITL is the one with the highest skewness. On the other hand, considering the Jarque-Bera statistic similar results can be observed. According to this test we find a rejection of the CFD assumption in none of the series for the QQ and ML-estimates and in five for the MM, the latter being the same as the KS statistic. The p-values of the QQ and ML suggest that according to the Jarque-Bera statistic the QQ estimation method derives the best results, actually, in 14 out of 17 cases the p-value of the QQ is higher than in the ML method. The Liliefors test, which is a version of the Kolmogorov test, becomes the most restrictive of the four tests in order to characterize the fitting. According to this test, we find a rejection in one, six and nine cases for the ML, QQ and MM- estimates, respectively. Again we can conclude that the best estimation method in terms of p-values is the ML, where only the ITL series is rejected. Finally, the Anderson-Darling test gives also very similar results as only none, five and nine cases are rejected for the ML, QQ and MM- estimates, respectively.

According to the results presented above, it is easy to conclude that the Maximum Likelihood method (ML) is the most flexible in finding good fits and that the assumption of the Cornish-Fisher density becomes a great starting point to simulate financial data series from a statistical point of view. Therefore, from now on, just the ML will be considered for estimation purposes.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
									Likelihoods								
Gaussian	-2067	-2602	-2835	-2570	-2538	-2703	-2454	-2546	-2614	-2562	-2575	-2648	-1183	-1315	-1296	-1279	-1497
CFD	**-1909**	**-2473**	**-2751**	**-2471**	**-2434**	**-2504**	**-2340**	**-2374**	**-2409**	**-2364**	**-2466**	**-2452**	**-1170**	**-1309**	**-1260**	**-1256**	**-1454**
Johnson	-1915	-2476	-2753	-2475	-2438	-2505	-2343	-2383	-2418	**-2364**	-2469	-2454	-1171	**-1309**	-1263	**-1256**	-1459
CFE	-1996	-2483	-2757	-2475	-2438	-3645	-2346	-2384	-3578	-2560	-2472	-3503	-1171	**-1309**	-1268	-1258	**-1454**
									Akaike Criteria								
Gaussian	4142	5213	5679	5148	5084	5415	4917	5100	5236	5133	5159	5304	2375	2639	2600	2566	3003
CFD	**3827**	**4955**	**5511**	**4951**	**4877**	**5017**	**4688**	**4756**	**4827**	**4736**	**4941**	**4913**	**2349**	**2625**	**2529**	**2521**	**2916**
Johnson	3839	4961	5514	4957	4883	5019	4694	4775	4844	**4736**	4947	4916	2350	**2626**	2533	2520	2926
CFE	4001	4974	5523	4957	4883	7299	4700	4776	7164	5129	4952	7015	2351	2626	2545	2525	2917
									Bayesian Criteria								
Gaussian	4165	5237	5702	5172	5107	5439	4940	5123	5259	5156	5182	5327	2392	2656	2617	2583	3020
CFD	**3850**	**4978**	**5535**	**4974**	**4900**	**5040**	**4711**	**4779**	**4851**	**4760**	**4965**	**4937**	**2366**	**2642**	**2546**	**2538**	**2933**
Johnson	3862	4984	5537	4980	4907	5042	4717	4798	4867	**4760**	4971	4939	2367	2643	2550	2537	**2943**
CFE	4024	4997	5546	4981	4906	7323	4723	4799	7188	5153	4975	7038	2368	2643	2562	2542	2934

Table 1.9. Comparison between the CF, gaussian, Johnson and CFE distributions for the twelve exchange rates series. We present three different model selection criteria: the Likelihood, the Akaike and Bayesian criteria. According to the results of these tests, all criteria suggest that the CFD fits better in-sample returns than the gaussian, Johnson and the CFE distributions. This out-performance is specially remarkable in the exchange rates database.

Next, in order to have a more intuitive and graphical representation of the advantage of considering CFD to simulate financial series, CFD in-sample performance will be compared, both graphically and analytically, with other distributional models. Table 1.9 presents the values of the Maximum Likelihood function for the exchange rates and indexes databases considering four different distributions: gaussian, Johnson distribution (Appendix A), the third-order Cornish-Fisher Expansion (CFE, see Equation 1.7) and the CFD. We compare our model with these distributions for the following reasons: the gaussian distribution is a special case of the CFD (when $a_3 = a_2 = 0$) and is a standard market model, so that in our comparison it might be used as the first order approximation. The Johnson distribution, as the third-order CFD, is also a four parametrical distribution and is also very flexible allowing for heavy tails and asymmetry. Therefore, it is used to compare the CFD with a distribution with the same number of parameters (degrees of freedom). Finally, given that the CFD is related to the Cornish-Fisher Expansion, we use the expansion as a distribution although it is not usual[30] and compare the results with the CFD. In this table, we also report the Akaike and Bayesian model selection criteria, which penalize for an increase in complexity through the inclusion of more parameters. First, we find an overall improvement of the CFD compared to the Gaussian or the Johnson distribution. Given that the CFD nests the gaussian one imposing $a_i = 0$ for $i > 1$, we can perform a likelihood ratio test. For every exchange rate we find that we can reject the null hypothesis that the restricted model (gaussian) is the correct model relative to the unrestricted CFD model

[30] See the comments on the relationship between the CFE and the CFD in Section 1.2. Basically, to construct a distribution from the CFE we truncate the expansion up to order three and calculate the coefficients of the polynomial with the empirical moments of the data-series (see Equation 1.7). The resulting distribution is a CFD but with different coefficients a_i. To our knowledge, this is the first time that the CFE is used directly as a distribution.

with p-values almost equal to zero. It is interesting to note that both the Akaike and the Bayesian criteria suggest that the CFD fits better in-sample returns than the Johnson distribution and the CFE. Indeed, in some cases (i.e., the ESP, ITL or SEK) which correspond to series with higher kurtosis coefficients, we find that the CFE is the worst model, even worse than the gaussian. Therefore, we can conclude that considering directly the coefficients a_i as parameters instead of the moments or cumulants implies a significant improvement in terms of goodness of fit. It is also interesting to note that although CFD behaves better than Johnson distributions both have the same number of parameters allows to get and similar results everywhere but in the tails, where the CFD becomes more flexible and tractable to describe higher deviations from normality. Finally, we also find that the bigger the deviation from normality the bigger the out-performance of the CFD with respect to the other distributions. This conclusion can be drawn from the fact that the strongest differences are found in the exchange rates database where the deviation with respect to normality are the biggest (see Section 1.4).

On the other hand, Tables of Figures 1.10-1.13 show graphical representations corresponding to the fit for all twelve exchange rates series of the QQ-Plot for the CFD, the Johnson distribution, the CFE and a cubic spline fitted with a penalization parameter of 0.8. The spline function, as an example of a non-parametrical estimation method, is also included for comparison purposes. Besides the QQ-Plot, we also present histograms and the corresponding fitted density functions for both the gaussian and the CFD case. Because the Johnson and polynomial models have the same number of degrees of freedom, their respective graphical representations for the QQ-Plot are almost identical in all the cases, although the Akaike and Bayesian criteria select the CFD as the best model. The plot using splines is sometimes apparently different (as in the AUD or ITL series), but they do not posses statistical significance given the above test results.

1.5.3 Dynamic Framework

In this Section we estimate a dynamic GARCH model with innovations distributed as CFD and compare the fitting with different dynamic benchmark models. In spite of the good results obtained in the previous Section 1.5.2 with the static CFD model, we consider the inclusion of first and second order dynamics for the following two reasons. Descriptive results on the dynamic behavior of return series in Section 1.4 have shown statistical evidence on the existence of autocorrelation and volatility clustering and, therefore, we notice that our static model remains misspecified. And second, in the Multivariate Inference Section 1.6 we will find that static multivariate models are not enough flexible to model multivariate returns and introducing first and second order dynamics will be required to improve the fitting.

As a result of this, we make mean and variance of the series time varying, for example using an ARMA-GARCH type model as they include the possibility of time-changing volatility and mean. First, Engle 1982 proposed his ARCH (AutoRegressive Conditional Heteroskedasticity) model and Bollerslev 1986 extended it to GARCH (Generalized ARCH). Time varying volatility has led to a significant amount of literature, summarized in Bollerslev, Chou, and Kroner 1992. However, one difficulty of those models is that conditional residuals very often remain heavy tailed. In the model that we will present

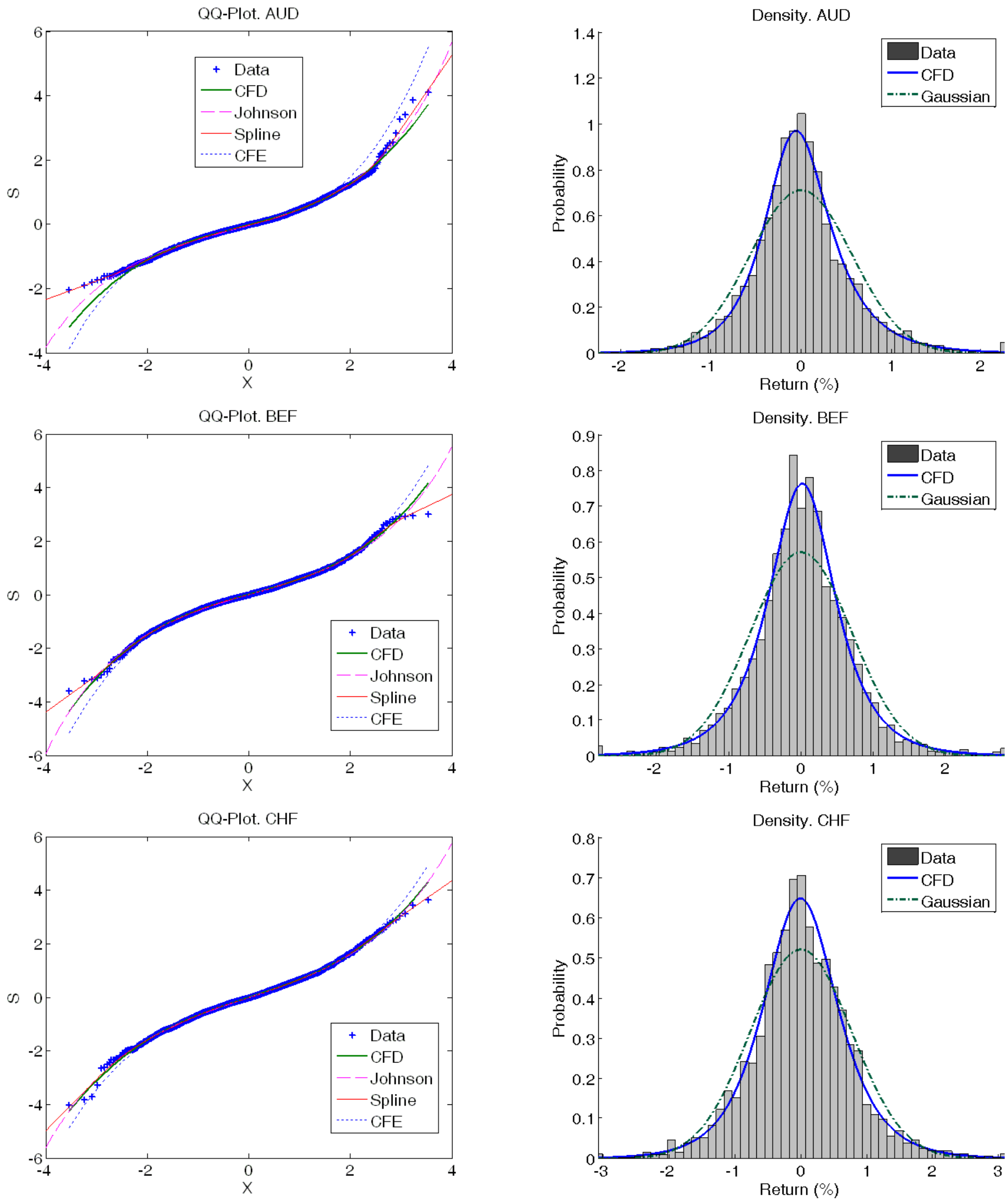


Table 1.10. Graphical representations of the experimental QQ-Plots and density functions for the AUD, BEF and CHF, with the corresponding fittings by using different distributions: the Cornish-Fisher (CFD), the Johnson, the spline and the Cornish-Fisher Expansion (CFE).

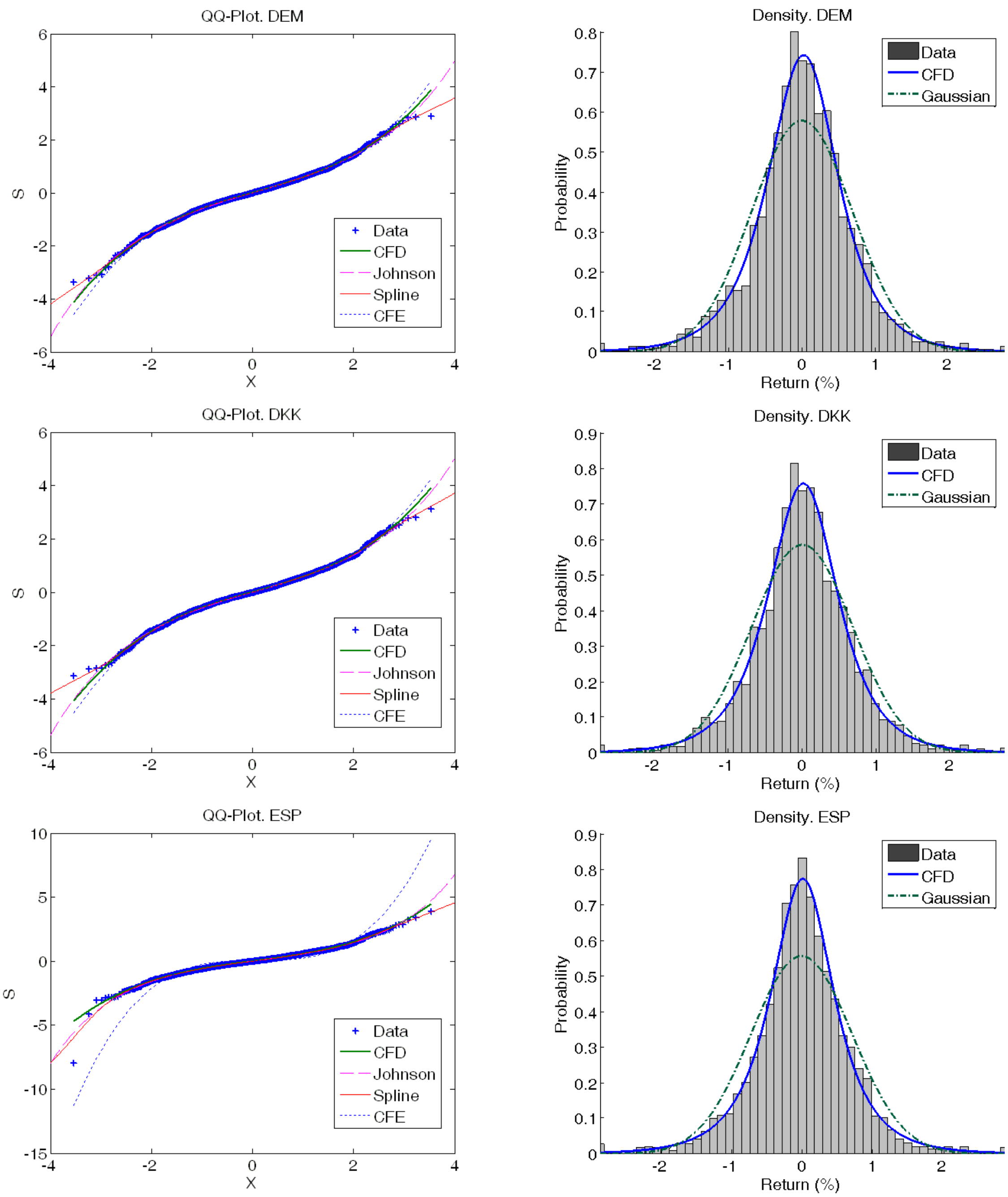


Table 1.11. Graphical representations of the experimental QQ-Plots and density functions for the DEM, DKK and ESP, with the corresponding fittings by using different distributions: the Cornish-Fisher (CFD), the Johnson, the spline and the Cornish-Fisher Expansion (CFE).

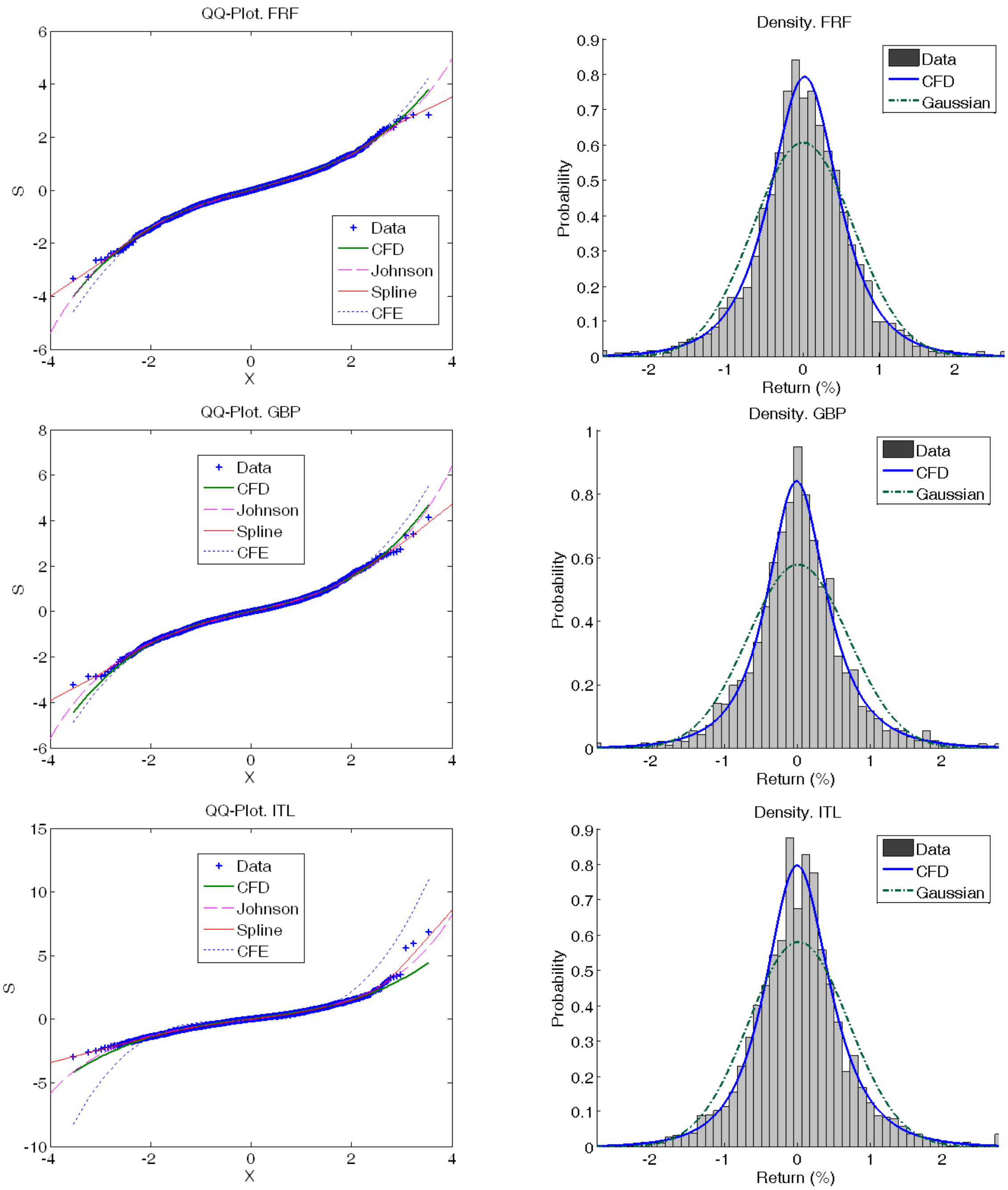


Table 1.12. Graphical representations of the experimental QQ-Plots and density functions for the FRF, GBP and ITL, with the corresponding fittings by using different distributions: the Cornish-Fisher (CFD), the Johnson, the spline and the Cornish-Fisher Expansion (CFE).

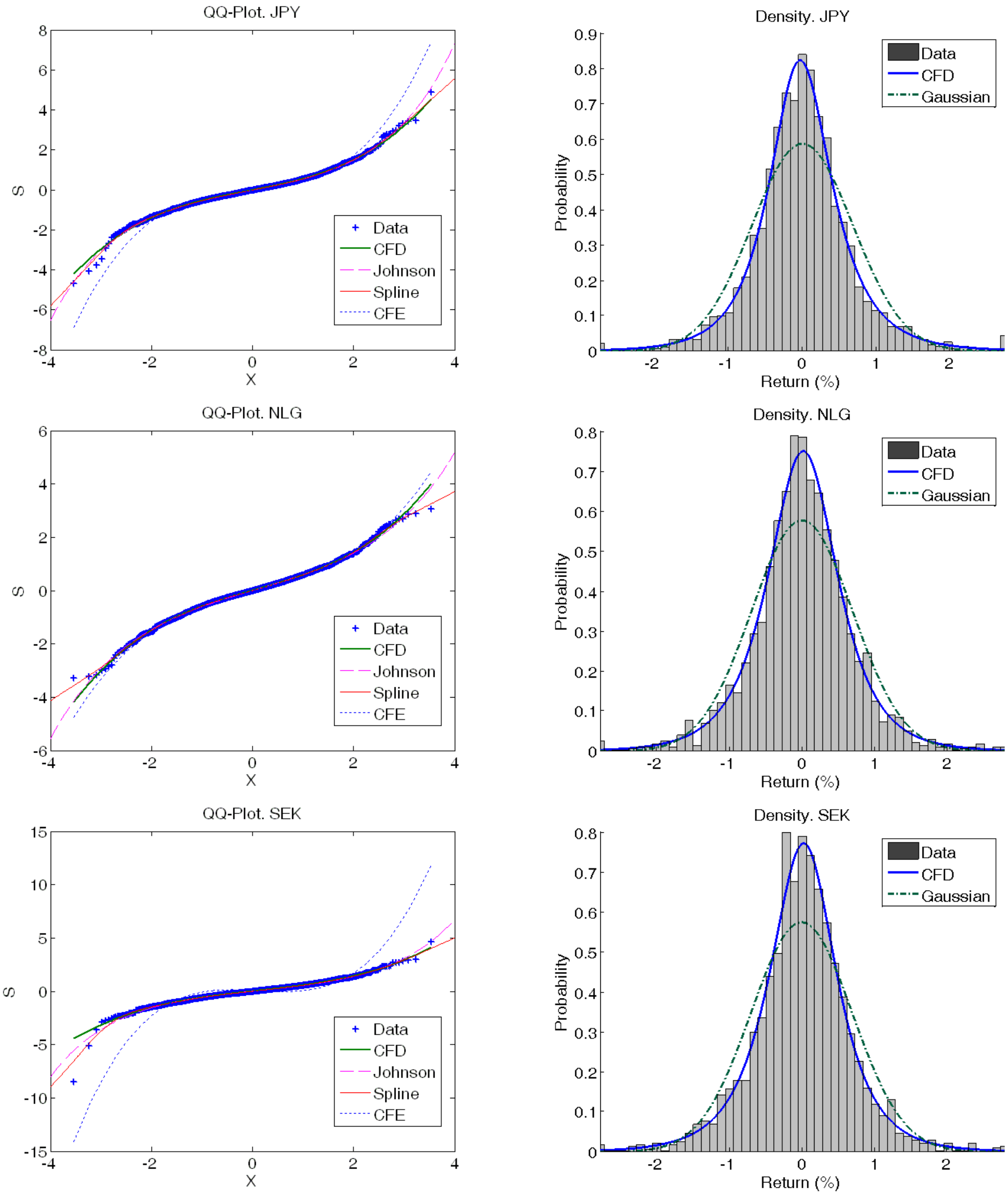


Table 1.13. Graphical representations of the experimental QQ-Plots and density functions for the JPY, NGL and SEK, with the corresponding fittings by using different distributions: the Cornish-Fisher (CFD), the Johnson, the spline and the Cornish-Fisher Expansion (CFE).

in this Section we will keep the usual GARCH-type parametrization of volatility, but we will allow a more general distribution than the gaussian to characterize the innovations. As has been described before, we propose that innovations follow a distribution which can be approximated by using a Cornish-Fisher Density.

Formally, we assume that the model for the prices returns, R_t, may be described by the following equations:

$$R_t = m_t + y_t \tag{1.30}$$

$$m_t = c + \text{AR} \cdot R_{t-1} \tag{1.31}$$

$$y_t = \sigma_t z_t \tag{1.32}$$

$$\sigma_t^2 = w + p y_{t-1}^2 + q \sigma_{t-1}^2 \tag{1.33}$$

$$z_t \sim CF(a_3, a_2) \tag{1.34}$$

The term m_t in Equation 1.30 corresponds to the conditional mean, which in our model will be described by a first order autoregressive model AR(1) presented in Equation 1.31[31]. Autoregressive models suppose that past returns tend to predict future ones and much literature (e.g. Lo and MacKinlay 1988, Lo and MacKinlay 1990) has highlighted that weekly or monthly returns of equities and indexes present an autoregressive pattern. On the other hand, the term y_t in Equation 1.30 represents the unexpected part of returns and is modeled by Equation 1.32. The variable σ_t is the conditional volatility and it will be described by a GARCH(1,1) representation given by Equation 1.33, although more

[31] Actually, we model separately the exchange rates and the indexes databases. Weekly series in the indexes database present autocorrelation, as verified in Table 1.2, while for the exchange rates database the conditional mean of the series do not present any dynamics. Therefore, we will only consider an AR(1) model for the index series while for the exchange rate we will consider a constant mean $\mu_t = c$. (i.e. $AR = 0$).

complicated processes could be trivially accommodated. In the GARCH(1,1) setting the parameter p captures the sensibility of present volatility to past returns (high or low returns tend to predict volatility) and the parameter q captures the volatility clustering phenomenon, i.e. high past volatility tends to predict high present volatility. However, in standard GARCH models it is assumed that the innovation, z_t, follows a normal standardized distribution, but in our model we will assume that innovations are distributed as a standard third-order Cornish-Fisher Density defined in Equation 1.12 with parameters a_3 and a_2. Therefore, in this new model mean and variance are allowed to be time-varying while the higher order moments associated to the innovations are kept constant, but not zero as in the gaussian case.

Given the static results presented in Section 1.5.2 above, we will only consider the ML estimation method. In Appendix C.1 we present the likelihood function that we have maximized to obtain the set of parameters of our model (c, AR, w, p, q, a_3 and a_2) and in Appendix C.2 we test our algorithm in a set of Montecarlo experiments. In these experiments we simulate 100 samples of length 2000 for 6 sets of parameters with different skewness and kurtosis characteristics. According to the results of these experiments we find that our estimation algorithms are accurate. Tables 1.14 and 1.15 present the ML-estimates of GARCH models with a normal density and a Cornish-Fisher density for the innovations, respectively.

Starting with Table 1.14, parameter p indicates that in general after a large return also volatility of the next period tends to be high. Parameter q indicates that a high volatility is also followed by high volatilities and, therefore, volatility remains quite persistent.

Although not reported, we have estimated the skewness and kurtosis coefficients for the innovations, and we have seen that in all cases kurtosis coefficients are significantly different from three and, therefore, incompatible with a gaussian distribution. However, we have also tested the adequacy of the gaussian hypothesis by using the same four tests presented in the static Section 1.5.2. In Table 1.14 we have presented p-values of the tests and have marked in bold the cases where normality assumption cannot be rejected. As can be seen, the four results imply an overall rejection of this hypothesis in both data series. As commented before, these results are already well established in the literature and suggest that GARCH models should be modeled with distributions for the innovations allowing for heavy tails. It is also important to notice that in all cases the likelihood function has increased with respect to its static counterpart shown in Table 1.9.

Table 1.15 shows the results for the estimations of a GARCH(1,1) model with innovations distributed as a Cornish-Fisher Density. The parameters for the variance dynamics, w, p and q, are similar to the ones presented for the gaussian case in Table 1.14. In addition, we also report the estimations of the parameters a_3 and a_2 of the standardized Cornish-Fisher distribution. As shown in the previous Section, the four tests also demonstrate that we cannot reject the null hypothesis that standardized innovations are compatible with the CFD. While for all series the a_3 coefficient is statistically different from zero, indicating that the true QQ Plot of innovations is non linear, in some cases we find that in many cases the a_2 coefficient is not significantly different from zero, as an indication that the conditional distribution of these series is highly symmetric. The highest asymmetries are found in the indexes database, indicating that even in filtered returns extreme crashes are more likely to happen than extreme booms, which will be an important feature when describing the portfolio selection to maximize returns in the next Chapter. In all the series the value of the Log-Likelihood function has highly improved with respect to the gaussian approximation and, moreover, according to this criteria the gaussian approximation is rejected[32]. Besides that, considering the CFD, a general decrease of standard errors associated to the parameters is obtained, and as a consequence, our estimation becomes more efficient. Although not reported, we find in general that the skewness and kurtosis coefficients implied from the estimation are closer to the sample values of the standardized innovations than in the gaussian case.

In order to analyze the goodness of in-sample fit corresponding to the dynamic CFD model presented in this work (GARCH-CFD), it is compared with two well known and accepted models to describe the non-gaussian innovations: the T-Student (GARCH-T) and the Skewed-T Student (GARCH-ST) distributions. The first one just accounts for symmetric heavy tails while the second one incorporates asymmetry between positive and negative returns (see Appendix A.1 for more details). In Table 1.16 the results of these estimations are presented along with the results of the static models used in the later Section for comparison purposes and the gaussian GARCH model (GARCH-G), where we report the Maximum Likelihood values, the Akaike Criteria and the Bayesian Criteria[33].

According to the likelihood test, we find that the GARCH-CFD model outperforms in 13 out of 17 cases the other models and the Akaike criteria also suggests that in 13 cases the GARCH-CFD is more accurate. In contrast, given that the Bayesian criteria penalizes more

32 In all cases the p-value is smaller than 10^{-4}.

33 Estimation details are available from authors upon request.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
c	-0.0041	0.0024	-0.0019	-0.0008	0.0022	0.0146	0.0024	-0.0149	-0.0156
std.	0.0103	0.0131	0.0146	0.0128	0.0128	0.0124	0.0121	0.0112	0.0117
w	**0.0050**	**0.0077**	**0.0130**	**0.0063**	**0.0049**	**0.0138**	**0.0066**	**0.0028**	**0.0072**
std.	0.0009	0.0016	0.0030	0.0015	0.0012	0.0026	0.0015	0.0006	0.0015
p	**0.0450**	**0.0424**	**0.0381**	**0.0418**	**0.0364**	**0.0719**	**0.0465**	**0.0377**	**0.0746**
std.	0.0051	0.0046	0.0056	0.0055	0.0046	0.0051	0.0061	0.0035	0.0063
q	**0.9390**	**0.9416**	**0.9391**	**0.9449**	**0.9530**	**0.9023**	**0.9384**	**0.9565**	**0.9122**
std.	0.0065	0.0066	0.0093	0.0075	0.0060	0.0090	0.0081	0.0039	0.0077
KS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
JB	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Liliefors	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Log-Lik	-1943	-2507	-2772	-2477	-2445	-2521	-2351	-2378	-2387

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
c	-0.0041	-0.0002	0.0148	0.1238	0.0481	0.1578	0.1287	0.1069
std.	0.0129	0.0130	0.0117	0.0937	0.1348	0.1016	0.1261	0.1634
AR				-0.0848	-0.0375	-0.0632	**0.1069**	0.0574
std.				0.0510	0.0438	0.0515	0.0520	0.0497
w	**0.0124**	**0.0065**	**0.0041**	**0.1135**	**0.2911**	**0.2371**	**0.2460**	**0.7152**
std.	0.0020	0.0015	0.0008	0.0728	0.3073	0.0994	0.0875	0.2543
p	**0.0472**	**0.0403**	**0.0497**	**0.1074**	0.0277	**0.2142**	**0.0736**	**0.0905**
std.	0.0044	0.0051	0.0046	0.0288	0.0159	0.0404	0.0200	0.0218
q	**0.9264**	**0.9460**	**0.9445**	**0.8782**	**0.9402**	**0.7723**	**0.8970**	**0.8707**
std.	0.0070	0.0070	0.0046	0.0308	0.0420	0.0414	0.0244	0.0300
KS	0.00	0.00	0.00	**0.19**	**0.44**	0.04	**0.08**	0.00
JB	<0.01	<0.01	<0.01	<0.01	**0.06**	<0.01	<0.01	<0.01
Liliefors	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AD	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Log-Lik	-2480	-2485	-2474	-1149	-1314	-1217	-1258	-1458

Table 1.14. GARCH estimates under normality assumption for both databases. c, w, p and q are the coefficients in Equations 1.30 to 1.33 and AR is the autoregressive coefficient (only for the indexes). Under each estimate the standard deviation is reported. KS, JB, Liliefors and AD represent, respectively, the p-values of the Kolmogorov-Smirnov, Jarque-Bera, Liliefors and Anderson-Darling tests for normality of standardized innovations. Finally, Log-Lik is given by the sum of all log-likelihoods.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
c	-0.0063	-0.0043	-0.0072	-0.0074	-0.0037	-0.0032	-0.0038	-0.0123	0.0081
std.	0.0102	0.0130	0.0147	0.0129	0.0127	0.0128	0.0121	0.0116	0.0120
w	**0.0050**	**0.0066**	**0.0117**	**0.0058**	**0.0046**	**0.0093**	**0.0069**	**0.0017**	**0.0060**
std.	0.0018	0.0027	0.0048	0.0024	0.0021	0.0032	0.0026	0.0010	0.0021
p	**0.0545**	**0.0469**	**0.0424**	**0.0443**	**0.0417**	**0.0547**	**0.0506**	**0.0422**	**0.0618**
std.	0.0107	0.0092	0.0097	0.0089	0.0085	0.0100	0.0101	0.0077	0.0102
q	**0.9313**	**0.9411**	**0.9383**	**0.9447**	**0.9497**	**0.9274**	**0.9347**	**0.9562**	**0.9273**
std.	0.0128	0.0117	0.0149	0.0113	0.0104	0.0134	0.0132	0.0075	0.0116
$\mathbf{a}_3$	**0.0668**	**0.0669**	**0.0550**	**0.0573**	**0.0598**	**0.0657**	**0.0585**	**0.0741**	**0.0602**
std.	0.0070	0.0068	0.0070	0.0071	0.0070	0.0063	0.0069	0.0070	0.0070
$\mathbf{a}_2$	**0.0435**	-0.0063	0.0080	-0.0107	-0.0059	-0.0030	-0.0085	0.0005	0.0084
std.	0.0123	0.0124	0.0119	0.0121	0.0122	0.0123	0.0120	0.0125	0.0122
KS	**0.77**	**0.98**	**0.93**	**0.95**	**0.86**	**0.72**	**0.56**	**0.30**	**0.67**
JB	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**>0.20**	0.04
Liliefors	**0.65**	**0.90**	**0.82**	**0.77**	**0.82**	**0.81**	**0.91**	**0.64**	**0.72**
AD	**0.66**	**0.72**	**0.55**	**0.89**	**0.88**	**0.76**	**0.78**	**0.82**	**0.75**
Log-Lik	-1834	-2409	-2710	-2411	-2371	-2412	-2276	-2264	-2309

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
c	-0.0047	-0.0012	0.0166	0.0954	0.0303	0.0769	0.0412	0.1270
std.	0.0128	0.0129	0.0121	0.0895	0.1359	0.1042	0.1189	0.1639
AR				**-0.0862**	-0.0347	-0.0684	0.0561	0.0236
std.				0.0458	0.0422	0.0461	0.0422	0.0421
w	**0.0143**	**0.0067**	**0.0063**	0.1187	0.2408	**0.2413**	0.0940	0.5766
std.	0.0047	0.0027	0.0022	0.0893	0.3808	0.1256	0.0700	0.3663
p	**0.0486**	**0.0446**	**0.0593**	**0.1110**	0.0261	**0.1814**	**0.0483**	**0.1053**
std.	0.0110	0.0092	0.0102	0.0369	0.0200	0.0452	0.0187	0.0410
q	**0.9203**	**0.9425**	**0.9290**	**0.8724**	**0.9475**	**0.7980**	**0.9395**	**0.8724**
std.	0.0177	0.0120	0.0115	0.0413	0.0525	0.0469	0.0221	0.0444
$\mathbf{a}_3$	**0.0758**	**0.0603**	**0.0596**	**0.0296**	**0.0354**	**0.0298**	**0.0417**	**0.0802**
std.	0.0064	0.0069	0.0056	0.0119	0.0147	0.0121	0.0116	0.0164
$\mathbf{a}_2$	0.0219	-0.0066	-0.0041	**-0.0666**	0.0004	**-0.0990**	**-0.0778**	**-0.0682**
std.	0.0125	0.0122	0.0120	0.0226	0.0239	0.0247	0.0241	0.0277
KS	**0.67**	**0.99**	**0.59**	**0.344**	**0.73**	**0.30**	**0.86**	**0.95**
JB	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**>0.20**	**0.15**	**>0.20**	**>0.20**
Liliefors	**0.78**	**0.84**	**0.27**	**0.97**	**0.97**	**0.75**	**0.88**	**0.76**
AD	**0.78**	**0.97**	**0.89**	**0.81**	**0.45**	**0.14**	**0.96**	**0.86**
Log-Lik	-2321	-2409	-2351	-1.1386	-1.3072	-1.2053	-1.2377	-1.4308

Table 1.15. GARCH estimates under CFD assumption for both databases. The legend is the same as in Table 1.14, with a_3 and a_2 being the coefficients of the standardized Cornish-Fisher Density.

over-parametrization, we find that in 11 cases the GARCH-T model is the best and only in three cases the GARCH-CFD appears to be better. Nevertheless, even according to the Bayesian Criteria the GARCH-CFD is always better than the GARCH-ST approximation, indicating that comparing models with the same number of parameters the CFD becomes more appropriate.

It is also interesting to note that the better fit of the GARCH-CFD is specially remarkable in the exchange rates database, which, as we have seen before (see Section 1.4), presents a higher non-gaussian behavior. This fact indicates that our distribution could be specially adequate in highly non-linear time series. Finally, as can be seen from the difference between values of static and dynamic models, we also want to emphasize the importance of allowing for first an second order dynamics. We have also performed likelihood ratio tests for the pairs GARCH-G, Gaussian and GARCH-CFD, finding in all cases a strong rejection of the dynamic model.

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
Likelihoods																	
GARCH-G	-1943	-2507	-2772	-2477	-2445	-2521	-2351	-2378	-2387	-2480	-2485	-2473	-1148	-1313	-1216	-1258	-1458
GARCH-CFD	**-1834**	**-2409**	**-2710**	**-2411**	**-2371**	**-2412**	**-2276**	**-2264**	**-2309**	**-2321**	**-2409**	**-2351**	-1138	**-1307**	-1205	**-1237**	**-1430**
GARCH-T	-1843	-2412	-2713	-2414	-2374	-2413	-2279	-2268	-2311	-2324	-2412	-2349	-1142	**-1307**	-1213	-1241	-1436
GARCH-ST	-1838	-2412	-2712	-2414	-2374	-2413	-2278	-2269	-2312	-2322	-2412	**-2351**	**-1136**	-1308	**-1203**	-1239	-1434
Gaussian	-2067	-2602	-2835	-2570	-2538	-2703	-2454	-2546	-2614	-2562	-2575	-2648	-1183	-1315	-1296	-1279	-1497
CFD	-1909	-2473	-2751	-2471	-2434	-2504	-2340	-2374	-2409	-2364	-2466	-2452	-1170	-1309	-1260	-1256	-1454
Johnson	-1915	-2476	-2753	-2474	-2437	-2505	-2343	-2383	-2418	-2364	-2469	-2454	-1171	-1309	-1263	-1256	-1459
Akaike Criteria																	
GARCH-G	3894	5022	5553	4962	4898	5051	4710	4765	4782	4968	4978	4955	2306	2637	2443	2526	2926
GARCH-CFD	**3680**	**4831**	**5433**	**4834**	**4755**	**4836**	**4564**	**4540**	**4630**	**4655**	**4831**	**4714**	2291	2628	2424	**2489**	**2875**
GARCH-T	3697	4835	5436	4839	4759	4836	4568	4546	4633	4658	4835	4709	2296	**2626**	2439	2495	2883
GARCH-ST	3689	4836	5437	4840	4761	4839	4569	4550	4636	4657	4836	4714	**2287**	2630	**2420**	2493	2883
Gaussian	4142	5213	5679	5148	5084	5415	4917	5100	5236	5133	5159	5304	2375	2639	2600	2566	3003
CFD	3827	4955	5511	4951	4877	5017	4688	4756	4827	4736	4941	4913	2349	2625	2529	2521	2916
Johnson	3839	4961	5514	4957	4883	5019	4694	4775	4844	4736	4947	4916	2350	2626	2533	2520	2926
Bayesian Criteria																	
GARCH-G	3918	5045	5577	4985	4922	5074	4733	4788	4806	4991	5001	4978	2328	2658	2464	2548	2947
GARCH-CFD	**3715**	4866	5467	4869	4790	4871	4599	**4575**	4664	4690	4865	4749	2321	2658	2454	**2519**	**2905**
GARCH-T	3726	**4864**	**5465**	**4868**	**4788**	**4865**	**4597**	**4575**	**4662**	**4687**	**4864**	**4738**	2322	**2652**	2465	2520	2909
GARCH-ST	3724	4871	5472	4875	4796	4873	4604	4585	4670	4691	4871	4748	**2317**	2660	**2450**	2523	2913
Gaussian	4165	5237	5702	5172	5107	5439	4940	5123	5259	5156	5182	5327	2392	2656	2617	2583	3020
CFD	3850	4978	5535	4974	4900	5040	4711	4779	4851	4760	4965	4937	2366	2642	2546	2538	2933
Johnson	3862	4984	5537	4980	4907	5042	4717	4798	4867	4760	4971	4939	2367	2643	2550	2537	2943

Table 1.16. Comparison between the GARCH model with CFD innovations (GARCH-CFD), gaussian innovations (GARCH-G), student-t innovations (GARCH-T), skewed-student-t innovations (GARCH-ST), the static gaussian, CFD and Johnson distributions. We present three different model selection criteria: the Likelihood, the Akaike and Bayesian criteria. According to the likelihood test, we find that the GARCH-CFD model outperforms in 13 out of 17 cases the other models and the Akaike criteria also suggest that in 13 cases the GARCH-CFD is more accurate. In contrast, given that the Bayesian criteria penalizes more over-parametrization, we find that in 11 cases the GARCH-T model is the best. The better fit of the GARCH-CFD is specially remarkable in the exchange rates database.

1.6 Multivariate Inference

In this Section we will analyze and characterize the estimation of Multivariate Cornish-Fisher Densities for the two databases in both of its forms: the CB-MCFD and the VCB-MCFD considering separately static and dynamic frameworks. In each case we describe first the estimation procedure and afterwards analyze the estimation results.

1.6.1 Static Framework

Static Copula-Based MCFD

Estimation. As was described in Section 1.3.1 the static CB-MCFD model assumes a structural dependence given by a gaussian copula and, generally, copula functions are estimated in two steps: first, one has to model and fit the univariate distributions and, second, one has to transform the real data onto the "copula scale data" , $X_{i,t}$, using the fitted univariate distributions and fit the parameters of the copula function. In the CB-MCFD model:

$$\text{cb-mcfd}_m(R_t) = \frac{1}{\sqrt{(2\pi)^n \det[\kappa]}} \prod_{i=1}^{n} \frac{\partial \left[Q_i^{-1}(R_{i,t})\right]}{\partial R_{i,t}} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} Q_i^{-1}(R_{i,t}) \left(\kappa^{-1}\right)_{ij} Q_j^{-1}(R_{j,t})\right)$$

if our real data are denoted as $R_{i,t}$, corresponding to the value of the variable R_i at time t, the corresponding "copula scale data" with our notation is given by $X_{i,t} = Q_i^{-1}(R_{i,t})$, which, under the CB-MCFD hypothesis, should be distributed as a multivariate gaussian with zero mean, unit variance and correlation matrix κ. We can prove this just calculating the distribution of the variables $X_{i,t}$ using the density transformation theorem presented in

Section 1.3.2 (see also Johnson and Kotz 1972a):

$$\begin{aligned} f(X_t) &= \text{cb-mcfd}_m(Q_i\left(X_t\right)) = \\ &= \frac{1}{\sqrt{(2\pi)^n \det\left[\kappa\right]}} \prod_{i=1}^{n} \frac{\partial\left[Q_i^{-1}(Q_i\left(X_{i,t}\right))\right]}{\partial X_{i,t}} * \\ & \quad \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} Q_i^{-1}(Q_i\left(X_{i,t}\right))\left(\kappa^{-1}\right)_{ij} Q_j^{-1}(Q_j\left(X_{j,t}\right))\right) \\ &= \frac{1}{\sqrt{(2\pi)^n \det\left[\kappa\right]}} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} X_{i,t}\left(\kappa^{-1}\right)_{ij} X_{j,t}\right) = \Phi_\kappa(X_t) \end{aligned}$$

As a consequence, in order to fit the CB-MCFD we have to calculate the correlation matrix of these fictitious variables $X_{i,t}$, which will be called as normal rank correlation. We will use the common estimator for correlations $\hat{\kappa}$:

$$\hat{\kappa}_{ij} = \frac{1}{T}\sum_{t=1}^{T} X_{i,t}X_{j,t} = \sum_{t=1}^{T} Q_i^{-1}(R_{i,t})Q_j^{-1}(R_{j,t})$$

where $X_{i,t}$ is the value of the fictitious gaussian variable X_i at time t and T is the sample size. In Appendix C.1 we show that this estimator is, as a matter of fact, the Maximum Likelihood estimator of the normal rank correlation, given that the univariate CFD specification is the correct one for each asset.

Results. The estimation results for the exchange rates and indexes databases corresponding to the CB-MCFD model are presented in Table 1.17 and 1.18, respectively. In both cases the estimates are presented under the diagonal and the standard errors above it, which are calculated using the numerical Hessian matrix. We can observe that estimates of the normal rank correlation are very similar to the estimates corresponding to the linear correlation presented in Table 1.3. Actually, as we could have deduced from the results of

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		0.0203	0.0203	0.0204	0.0203	0.0202	0.0203	0.0198	0.0201	0.0202	0.0204	0.0202
BEF	-0.0439		0.0030	0.0014	0.0020	0.0142	0.0014	0.0150	0.0141	0.0099	0.0009	0.0060
CHF	-0.0478	0.8800		0.0018	0.0034	0.0150	0.0027	0.0145	0.0143	0.0096	0.0022	0.0071
DEM	-0.0372	0.9418	0.9192		0.0018	0.0143	0.0010	0.0146	0.0140	0.0091	0.0005	0.0058
DKK	-0.0356	0.9217	0.8645	0.9261		0.0139	0.0017	0.0149	0.0141	0.0101	0.0014	0.0053
ESP	0.0330	0.4203	0.3781	0.4123	0.4377		0.0140	0.0180	0.0167	0.0176	0.0139	0.0154
FRF	-0.0460	0.9428	0.8919	0.9545	0.9301	0.4293		0.0146	0.0137	0.0096	0.0009	0.0057
GBP	-0.1545	0.6465	0.6648	0.6860	0.6542	0.2836	0.6805		0.0170	0.0161	0.0147	0.0149
ITL	-0.1077	0.6896	0.6558	0.6983	0.6862	0.3782	0.7183	0.6180		0.0163	0.0140	0.0137
JPY	-0.0028	0.5946	0.6146	0.6236	0.5886	0.2383	0.6061	0.4527	0.4209		0.0093	0.0127
NGL	-0.0417	0.9564	0.9073	0.9760	0.9371	0.4315	0.9585	0.6743	0.7058	0.6138		0.0056
SEK	-0.0620	0.7542	0.7155	0.7592	0.7852	0.3795	0.7661	0.5861	0.6377	0.4872	0.7646	

Table 1.17. Normal rank correlation between the different exchange rates. The Maximum Likelihood estimates are presented under the diagonal and the standard errors, which are calculated using the numerical Hessian matrix, above it.

	S&P	**NKI**	**STX**	**EM**	**EME**
S&P		0.0361	0.0184	0.0276	0.0373
STX	0.4037		0.0345	0.0342	0.0408
NKI	0.7488	0.4495		0.0246	0.0246
EM	0.5945	0.4569	0.6475		0.0246
EME	0.3705	0.2568	0.4197	0.5980	

Table 1.18. Normal rank correlations between indexes. Maximum likelihood-estimates are under the diagonal and standard errors above it.

Lemma 7, biggest differences correspond to series with highest non-gaussian behavior, as ITL and SEK.

In order to characterize the goodness of fit of the estimated CB-MCFD, we will apply the same tests already presented in Section 1.4: the KS-χ^2 test, the Omnibus test and the Mardia A and B tests. In the tests presented below, first, we will focus both on pairs of assets, i.e., on the dependence structures between just two random variables and on the whole sample. Actually, testing the gaussian copula hypothesis for two random variables gives useful information for a larger number of dependent variables constituting a large portfolio. Indeed, let us assume that each pair (R_i, R_j), (R_i, R_k) and (R_j, R_k) has a gaussian copula, then, the triplet (R_i, R_j, R_k) also has a gaussian copula and this result might be generalized to an arbitrary number of random variables. In Tables 1.20 and 1.19 we report the p-values from the bivariate tests for the exchange rates and the indexes databases and bold numbers show the cases where the null hypothesis that historical data are distributed as a CB-MCFD cannot be rejected at a 5% significance level. Additionally, in the lower part of the Tables we also present the results of the tests for the whole database.

Considering that CFD fit properly the marginal distributions of all the series, as we have seen in Section 1.5, these results show that the static gaussian copula is not flexible enough to characterize the multivariate distribution describing the exchange rates. In particular, we observe that only 10, 33, 34 and 9 of the 66 bivariate performed tests are positive at a 5% significance level. Moreover, we find that the tests for the whole sample clearly reject the CB-MCFD hypothesis. Additionally, although bivariate tests of the indexes database generally support the CB-MCFD hypothesis (10, 10, 7 and 7 out of 10 give positive results), almost all multivariate tests (only the Omnibus test is positive with a p-value of 0.06) reject the CB-MCFD. Given that the returns of the exchange rates are daily and the ones corresponding to the indexes are weekly, these results are an indication that weekly returns are more likely to present a static gaussian copula dependence structure than daily returns.

Static Variance-Covariance Based MCFD

Estimation. As shown in Section 1.3.2 in the VCB-MCFD model we are considering a dependence structure which allows the incorporation of simultaneous extreme events.

Omnibus \ KS-χ^2

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		0.03	**0.10**	0.03	0.03	**0.17**	**0.10**	**0.08**	**0.48**	0.01	**0.06**	**0.09**
BEF	**0.72**		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CHF	**0.73**	0.00		0.00	0.00	0.00	0.00	0.02	0.00	0.01	0.00	0.00
DEM	**0.69**	0.00	0.04		0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
DKK	**0.69**	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.01	0.00	0.00
ESP	**0.63**	**0.51**	**0.53**	**0.48**	**0.34**		0.00	0.01	0.00	0.04	0.00	0.00
FRF	**0.71**	0.00	0.00	0.00	0.00	**0.54**		0.02	0.00	0.01	0.00	0.00
GBP	**0.44**	**0.07**	**0.25**	**0.21**	**0.27**	**0.80**	**0.13**		0.00	**0.05**	0.01	**0.05**
ITL	**0.19**	0.00	0.00	0.00	0.00	**0.24**	0.00	**0.09**		**0.11**	0.00	0.00
JPY	**0.74**	**0.27**	0.12	0.06	**0.24**	**0.81**	0.10	**0.41**	**0.05**		0.00	0.00
NGL	**0.71**	0.00	0.00	0.00	0.00	**0.57**	0.00	**0.13**	0.00	**0.07**		0.00
SEK	**0.50**	0.00	0.00	0.00	0.00	**0.11**	0.00	0.00	0.00	0.00	0.00	

Mardia B\ Mardia A

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		**0.78**	**0.82**	**0.56**	**0.31**	**0.50**	**0.41**	**0.69**	**0.16**	**0.83**	**0.49**	**0.66**
BEF	**0.18**		**0.70**	0.00	0.00	**0.08**	0.00	0.00	0.00	0.01	0.00	0.00
CHF	**0.10**	0.00		**0.10**	0.00	**0.43**	**0.40**	**0.19**	0.00	**0.83**	0.01	0.00
DEM	**0.20**	0.00	0.00		0.00	0.08	0.00	0.00	0.02	0.02	**0.12**	0.00
DKK	**0.15**	0.00	0.00	0.00		**0.22**	**0.13**	**0.62**	0.00	0.02	0.00	**0.68**
ESP	**0.15**	0.00	0.00	0.00	0.00		**0.06**	**0.71**	**0.02**	**0.15**	**0.20**	**0.72**
FRF	**0.17**	0.00	0.00	0.00	0.00	0.00		0.00	0.01	0.02	0.00	0.00
GBP	**0.31**	0.00	0.00	0.00	0.00	0.00	0.00		**0.37**	**0.12**	0.00	**0.22**
ITL	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.02	0.00	**0.12**
JPY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.01	**0.23**
NGL	**0.22**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00
SEK	**0.11**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
3629	266	13339	0.25
0.00	0.00	0.00	0.00

Table 1.19. Tests for the static Copula-Based Multivariate Cornish-Fisher Density model corresponding to the exchange rate database. In the first panel, we present the p-values for the bivariate tests KS-χ^2 and Omnibus (over and below the diagonal, respectively). In the second panel, we report the p-values for the bivariate Mardia A and B tests (over and below the diagonal, respectively). In the third panel we present the joint multivariate tests for the static CB-MCFD model. Marked in bold are the cases where the null hypothesis of the CB-MCFD cannot be rejected at a 5% significance level.

Omnibus \ KS-χ^2

	S&P	NKI	STX	EM	EME
S&P		**0.34**	**0.73**	**0.11**	**0.89**
NKI	**0.68**		**0.89**	**0.43**	**0.49**
STX	**0.08**	**0.10**		**0.06**	**0.86**
EM	**0.21**	**0.94**	**0.06**		**0.36**
EME	**0.73**	**0.90**	**0.47**	**0.52**	

Mardia A \ Mardia B

	S&P	NKI	STX	EM	EME
S&P		**0.28**	0.04	0.00	**0.42**
NKI	**0.15**		0.00	**0.76**	**0.53**
STX	**0.07**	0.01		**0.08**	**0.42**
EM	**0.06**	0.04	0.00		**0.77**
EME	**0.51**	**0.46**	**0.13**	**0.69**	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
0.00	0.00	**0.06**	0.01

Table 1.20. Tests for the static Copula-Based Multivariate Cornish-Fisher Density model corresponding to the Indexes database, with the same structure as Table 1.19

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
m	-0.0010	-0.0050	-0.0065	-0.0056	-0.0050	-0.0142	-0.0054	-0.0062	-0.0172
std.	0.0108	0.0138	0.0154	0.0139	0.0136	0.0146	0.0134	0.0145	0.0139
σ^2	**0.3135**	**0.4842**	**0.5850**	**0.4717**	**0.4595**	**0.5256**	**0.4293**	**0.4625**	**0.4887**
std.	0.0162	0.0214	0.0251	0.0195	0.0196	0.0343	0.0187	0.0214	0.0324
$\mathbf{a}_3$	**0.0799**	**0.0751**	**0.0895**	**0.1101**	**0.1503**	**0.0950**	**0.1231**	**0.0775**	**0.0998**
std.	0.0062	0.0058	0.0055	0.0045	0.0036	0.0048	0.0043	0.0055	0.0042
$\mathbf{a}_2$	**0.0454**	-0.0110	**0.0295**	**-0.0215**	-0.0012	-0.0066	0.0005	**0.0309**	**0.0219**
std.	0.0095	0.0104	0.0114	0.0119	0.0097	0.0108	0.0109	0.0108	0.0119

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
m	-0.0041	-0.0002	0.0148	0.1800	-0.0996	0.1502	0.0271	0.1107
std.	0.0129	0.0130	0.0117	0.0974	0.1282	0.1277	0.1494	0.2161
σ^2	0.4687	0.4737	0.5024	5.6039	9.3262	8.6552	8.0984	18.7958
std.	0.0254	0.0201	0.0377	0.4776	0.7136	0.8670	0.7493	1.8055
$\mathbf{a}_3$	**0.0890**	**0.1596**	**0.1556**	**0.0505**	**0.0321**	**0.0646**	**0.0660**	**0.0645**
std.	0.0054	0.0031	0.0037	0.0117	0.0094	0.0120	0.0108	0.0118
$\mathbf{a}_2$	**0.0394**	0.0023	**-0.0153**	-0.0379	0.0071	-0.0098	**-0.0781**	-0.0358
std.	0.0102	0.0097	0.0085	0.0254	0.0267	0.0235	0.0218	0.0238

Table 1.21. Static VCB-MCFD estimates under both databases. m, σ denote the mean and variance while a_3 and a_2 denote the coefficients of the standardized Cornish-Fisher Density. Under each estimate its standard deviation calculated using the Hessian matrix is reported. Significant parameters at a 95% confidence appear in bold.

Parameters of the VCB-MCFD distribution (m_i, Σ_{ij}, $a_{3,i}$ and $a_{2,i}$) will be estimated using the Maximum Likelihood method. The details of the estimation as the log-likelihood function or the initial parameters are described in Appendix C.1.

Results. Next we present two sets of tables: in Table 1.21 we present the estimates for this distribution and the standard errors for both data sets, and in Tables 1.22 and 1.23 we show the test results for this model.

Analyzing the results of the bivariate tests (the KS-χ^2 test, the Omnibus test and the Mardia A and B tests) in Table 1.22 we observe 35, 25, 55 and 24 positive results (at a 5% confidence level) out of 66 pairs for the exchange rates database and 10, 10, 9 and 6 out of

Omnibus \ KS-χ^2

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		0.00	**0.24**	**0.67**	0.03	**0.07**	**0.28**	**0.07**	**0.16**	0.00	0.03	**0.54**
BEF	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CHF	0.00	0.00		**0.47**	0.00	**0.17**	**0.35**	**0.64**	**0.52**	**0.29**	0.00	**0.06**
DEM	**0.06**	0.00	0.00		0.00	0.03	0.00	0.02	0.02	0.01	0.00	0.03
DKK	0.00	0.00	0.00	0.00		**0.43**	0.03	**0.36**	**0.33**	**0.32**	0.00	**0.08**
ESP	**0.79**	0.00	0.01	**0.08**	0.00		**0.28**	0.01	0.03	**0.11**	**0.15**	**0.22**
FRF	**0.23**	0.00	0.00	0.02	0.00	**0.39**		**0.18**	**0.21**	**0.88**	0.00	**0.27**
GBP	**0.76**	0.00	0.01	**0.08**	0.00	**0.90**	**0.31**		0.00	0.01	**0.08**	**0.82**
ITL	**0.10**	0.00	0.00	0.01	0.00	**0.14**	0.03	**0.13**		**0.13**	**0.08**	**0.84**
JPY	**0.55**	0.00	0.00	**0.04**	0.00	**0.70**	**0.20**	**0.70**	**0.09**		**0.04**	**0.17**
NGL	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00		**0.08**
SEK	**0.49**	0.00	0.00	**0.05**	0.00	**0.65**	**0.33**	**0.65**	**0.07**	**0.43**	0.00	

Mardia B\ Mardia A

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		**0.20**	**0.62**	**0.09**	**0.65**	**0.57**	**0.75**	**0.21**	**0.53**	**0.65**	**0.33**	**0.39**
BEF	0.00		**0.09**	0.02	**0.19**	**0.49**	**0.29**	**0.12**	**0.05**	0.01	**0.42**	**0.35**
CHF	**0.12**	0.00		**0.06**	**0.63**	**0.08**	**0.84**	**0.82**	**0.06**	**0.21**	**0.65**	0.02
DEM	**0.20**	0.00	0.00		0.04	0.04	**0.06**	0.03	0.04	**0.09**	**0.05**	**0.09**
DKK	**0.09**	0.00	**0.65**	0.00		**0.99**	**0.55**	**0.99**	**0.13**	**0.43**	**0.56**	**0.72**
ESP	**0.28**	0.00	**0.03**	0.00	**0.13**		**0.96**	**0.85**	0.00	**0.79**	**0.69**	**0.59**
FRF	**0.49**	0.00	**0.84**	0.00	0.00	**0.07**		**0.33**	**0.16**	**0.55**	**0.86**	0.01
GBP	**0.10**	0.00	**0.26**	0.00	**0.14**	0.00	**0.01**		**0.72**	**0.61**	**0.70**	**0.78**
ITL	0.00	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00		0.03	**0.35**	**0.25**
JPY	0.00	0.00	**0.21**	0.00	**0.78**	0.01	0.04	0.00	0.00		**0.25**	**0.88**
NGL	0.03	0.00	**0.88**	0.00	0.00	0.03	0.00	**0.10**	0.00	**0.35**		**0.41**
SEK	**0.34**	0.00	**0.99**	0.00	0.00	**0.46**	0.00	0.02	0.00	**0.09**	0.01	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
800	84	813	0.16
0.00	0.00	0.00	0.00

Table 1.22. Tests for the static Variance-Covariance Based Multivariate Cornish-Fisher Density corresponding to the exchange rates database, with the same structure as Table 1.19

Omnibus \ KS-χ^2

	S&P	NKI	STX	EM	EME
S&P		**0.45**	**0.33**	**0.99**	**0.50**
NKI	**0.23**		**0.06**	**0.61**	**0.05**
STX	**0.08**	**0.30**		**0.86**	**0.08**
EM	**0.19**	**0.56**	**0.26**		**0.13**
EME	**0.16**	**0.47**	**0.20**	**0.41**	

Mardia A \ Mardia B

	S&P	NKI	STX	EM	EME
S&P		**0.51**	0.01	**0.90**	**0.11**
NKI	**0.31**		**0.06**	**0.83**	0.02
STX	**0.24**	**0.08**		**0.13**	0.00
EM	0.01	**0.51**	**0.11**		0.00
EME	**0.20**	**0.39**	**0.43**	**0.04**	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
0.00	0.00	**0.20**	0.01

Table 1.23. Tests for the static Variance-Covariance Based Multivariate Cornish-Fisher Density corresponding to the Indexes database, with the same structure as Table 1.19.

10 for the indexes database. Therefore, these results indicate that the static VCB-MCFD clearly outperforms the previous static CB-MCFD. This is the case because this density allows for the occurrence of simultaneous extreme events as mentioned in Section 1.3.2. In spite of this, the overall results of the fit are still somehow poor: the four tests give negative results with p-values smaller than 0.00 for the exchange rates database, while in the indexes database we find that only one test (the Omnibus) is positive at a 5% confidence level with a p-value of 0.20 and the KS-χ^2 is also positive at 1% confidence level.

Actually, the rejection of the static gaussian structure we have proposed with the CB-MCFD model (Section 1.6.1) and the rejection of the static VCB-MCFD can be explained considering that returns are more highly correlated during volatile markets and during market downturns (Longin and Solnik 2001 and Ang, Chen, and Xing 2002), given that we do not capture this feature with static models. As we show next, we believe that any appropriate approach to model correctly multivariate dependence in returns relays on assuming a dynamic dependence structure instead of considering more complex static copulas or dependence structures[34]. As we have seen in Section 1.4, univariate return series present dynamics in at least its first two moments and in Section 1.5.3 we have found that including the dynamics associated with the first two moments highly improves the quality of the fitting. Therefore, a similar behavior would be expected when including the dynamics to describe the multivariate structural dependence[35].

[34] In contrast, the empirical rejection of the gaussian copula has lead some researchers to propose alternative models of dependence like more general copulas as the t-student or Archimedean ones (Demarta and McNeil 2004, Malevergne and Sornette 2003, Henessy and H.E.Lapan 2002 or Frey, McNeil, and Nyfeler 2001).

[35] Not much research has been done in this direction but Dias and Embrechts 2005 is a good reference.

In order to understand our expectations on including dynamics it would be interesting to go deeply into the origin of the failures associated to the static copula. With this aim we present in Table 1.24 Figures corresponding to a bivariate scatter plot of three different pairs of exchange rates[36] and the QQ-Plot of the empirical z^2 [37] versus the theoretical χ^2. The three Figures show three different cases according to the validity of the gaussian copula assumption, in order to describe the bivariate structural dependence. For example, for the ITL/JPY pair it cannot be rejected the null hypothesis that dependence between these variables can be described by the gaussian copula at a 5% significance level, for the BEF/JPY it cannot be rejected at a 1% and for the BEF/CHF the gaussian copula is rejected. We have observed that the main data responsible of rejecting the null hypothesis are those which are outside the gaussian ellipse and lie in the X-positive Y-negative axis (these points are marked with a red circle in the Figure). In principle, one could impose a more general dependence structure allowing for extreme results to capture this points as well, as we have tried using the VCB-MCFD, but in our opinion it is more likely that in these points correlation is lower than for the rest of the sample.

Therefore, in Section 1.6.2 we will present a dynamic copula model with CB-MCFD innovations and a dynamic model with VCB-MCFD innovations.

---

[36] Actually, we plot the fictitious gaussian variables or "copula scale data" that corresponds to each asset in the CB-MCFD model.

[37] The statistic z^2 is defined as

$$z^2 = \sum_{i,j=1}^{n} Q_i^{-1}(R_i) \left(\kappa^{-1}\right)_{i,j} Q_j^{-1}(R_j)$$

and under the null hypothesis that returns follow a static CB-MCFD it is distributed as a χ^2 with n degrees of freedom. See Appendix C.3 for more details on this test.

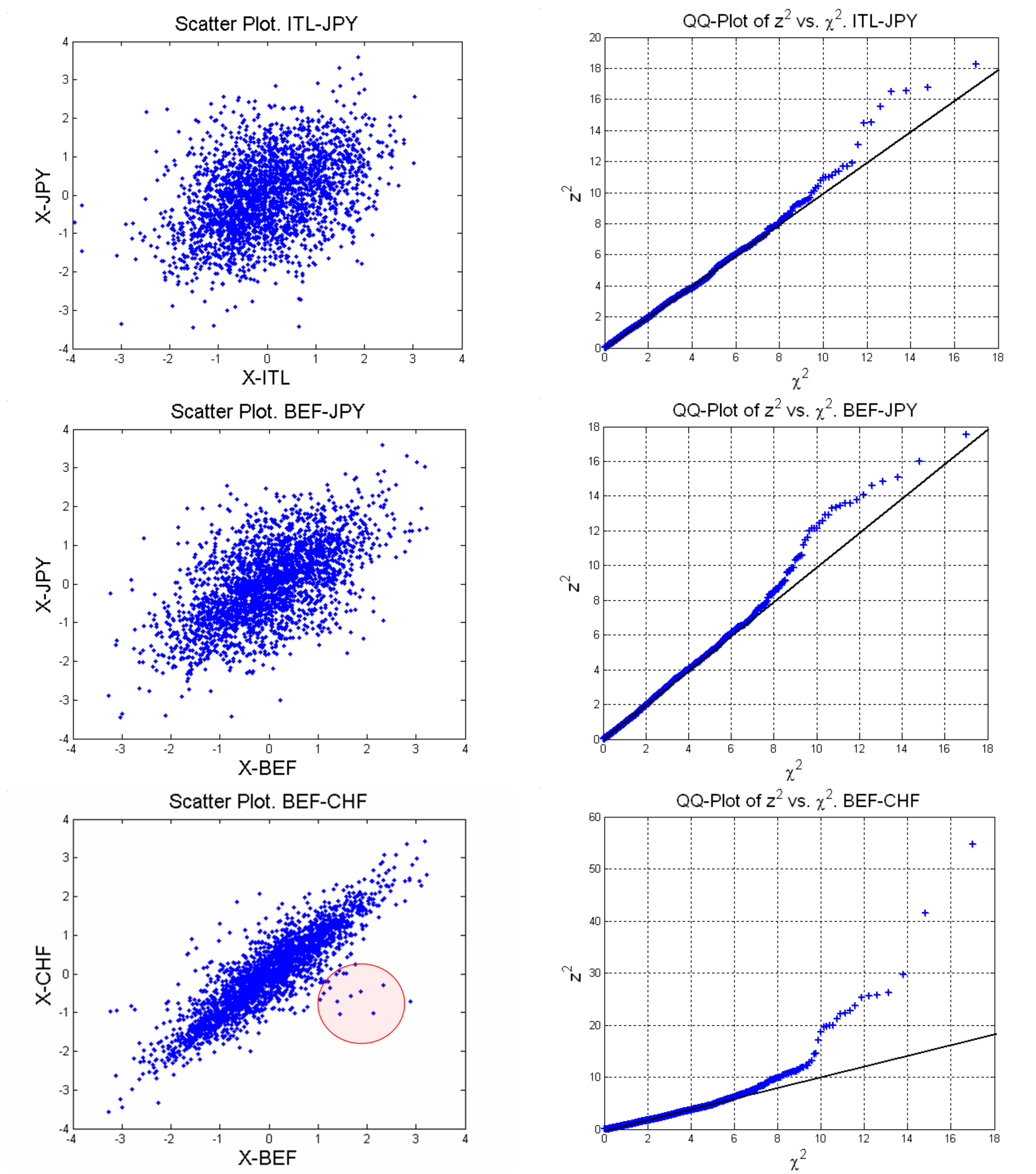


Table 1.24. Bivariate scatter plot of three different pairs of exchange rates and the QQ-Plot of the empirical z^2 versus the theoretical χ^2. If the data satisfy the null hypothesis that dependence between these variables can be described by the gaussian copula, the QQ-Plot should be a straight line. In the BEF/CHF case the main points responsible of rejecting the null hypothesis are marked with a red circle.

1.6.2 Dynamic Framework

As opposed to the univariate case, we find that static MCFD densities are not able to capture well the multivariate characteristics of financial returns. Therefore, we will consider two dynamic multivariate models: a dynamic copula model based on the CB-MCFD (dynamic CB-MCFD), and a Dynamic Conditional Correlation model (DCC) with Multivariate Cornish-Fisher innovations (dynamic VCB-MCFD).

Dynamic CB-MCFD model

Definition. As mentioned in Section 1.6.1, multivariate financial return series present dynamics in their dependence structures and there are a number of models which capture this feature: for example, the Constant Conditional Correlation model of Bollerslev 1990, the Dynamic Conditional Correlation model of Engle 2002 and the BKKK model proposed by Engle and Kroner 1995, among others. In this work, we choose the Dynamic Conditional Correlation model because, as proven by Engle 2002, with just two parameters captures correctly correlation dynamics. Due to this little number of extra parameters we avoid the dimensionality curse (large increase of the number of parameters as the number of dimensions increases) that occurs when one is confronted with modeling of a large number of assets in multivariate GARCH-like models.

Before presenting the dynamic CB-MCFD model, we will analyze first the usual DCC model as described by Engle 2002 and Engle and Sheppard 2001. Let $R_{i.t}$ be the rate of return of asset i from $t-1$ to t, $i = 1, ..., n$, and let $m_{i,t}$ be the conditional expected rate of return of asset i at time $t-1$. $\Sigma_t = \sigma_{ij,t}$ is the expected variance-covariance matrix conditional to the information available at $t-1$. The DCC model is designed to model both volatility persistence and time-varying correlations avoiding the dimensionality curse. In order to capture the dynamics of the expected mean $m_{i,t}$ an AR(1)[38] model in considered:

$$R_t = m_t + \Sigma_t^{1/2} z_t \tag{1.35}$$

$$m_t = c + \text{AR} \cdot R_{t-1} \tag{1.36}$$

where R_t is defined as the n-th dimensional vector with components $R_{i,t}$, m_t is an n-dimensional conditional mean vector and $\Sigma_t^{1/2}$ is a symbol for the Cholesky decomposition of Σ_t. The dynamics of the second moments associated to variances and correlations are treated by the following Equation:

$$\Sigma_t = D_t \Theta_t D_t \tag{1.37}$$

where D_t is the matrix with the conditional variances in its diagonal following a unidimensional GARCH(1,1) process defined by:

$$D_{ii,t} = \sqrt{\sigma_{ii}^2} \text{ and } D_{ij,t} = 0 \tag{1.38}$$

$$\sigma_{ii,t}^2 = w_i + p_i \sigma_{ii,t-1}^2 + q_i \varepsilon_{i,t-1}^2 \tag{1.39}$$

with $\varepsilon_t = \Sigma_t^{1/2} z_t$ being the vector of unexpected and correlated returns, and Θ_t being the conditional correlation matrix which follows a process defined by:

$$\Theta_t = [\text{diag}(\Omega_t)]^{-1} \times \Omega_t \times [\text{diag}(\Omega_t)]^{-1} \tag{1.40}$$

$$\Omega_t = (1 - \psi_1 - \psi_2)\,\bar{\Omega} + \psi_1 \left(u_{t-1} u_{t-1}'\right) + \psi_2 \Omega_{t-1} \tag{1.41}$$

$$\bar{\Omega}_{ij} = \rho_{ij} \tag{1.42}$$

[38] See the univariate empirical Section 1.5.3 for details on the AR(1) model.

where diag(Ω_t) means the diagonal of Ω_t, ρ_{ij} is the unconditional correlation of asset i and j and $u_t = D_t^{-1}\varepsilon_t$ denotes the vector of standardized and unexpected returns. With this specification, it easy to see that the conditional correlation matrix, Ω_t, follows a GARCH-like process where the parameters ψ_1 and ψ_2 are the equivalents to p_i and q_i in Equation 1.39. In order to keep the variances finite and the correlation matrix positive definite the following conditions on the parameters must be imposed: $p_i + q_i < 1$, $0 \leq \psi_1, \psi_2 \leq 1$ and $\psi_1 + \psi_2 = 1$.

In this Section we propose the dynamic CB-MCFD model where in order to capture and describe the dynamical evolution of dependencies, instead of the variance-covariance matrix, the normal rank correlation[39] is the one following a DCC model defined above. With this model we are able to capture the unidimensional non-gaussian behavior and also the dynamics of the mean, variances and dependencies. Explicitly, we are imposing a dynamical DCC model given by Equations 1.39 on the correlations of the fictitious gaussian "copula scale data", $X_{i,t}$, as opposed to the static CB-MCFD model of Section 1.6.1, where we were assuming a constant correlation between the variables $X_{i,t}$. Therefore, the density function at time t is formally identical to the CB-MCFD model (see Equation 1.19):

$$\text{cb-mcfd}(R_t) = \frac{1}{\sqrt{(2\pi)^n \det\left[\kappa_t\right]}} \prod_{i=1}^{n} \frac{\partial\left[Q_{t,i}^{-1}(R_{t,i})\right]}{\partial R_{t,i}} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} Q_{t,i}^{-1}(R_i) Q_{t,j}^{-1}(R_{t,j}) \left(\kappa_t^{-1}\right)_{ij}\right). \tag{1.43}$$

Given that the functions $Q_{t,i}$ are defined in terms of the variances (considering the second parametrization of the univariate CFD), implicitly included in $Q_{t,i}$,we have an univariate

39 See Section 1.3 for more details on the normal rank correlation and the gaussian copula.

GARCH processes for the variances[40] (Equations 1.38) and, on the other hand, we also consider that the normal rank correlation matrix, κ, follows a DCC process (Equations 1.39). Therefore, the full model can be formally denoted as:

$$
\begin{aligned}
R_t &\sim \text{cb-mcfd}(R_t) = \frac{1}{\sqrt{(2\pi)^n \det[\kappa_t]}} \prod_{i=1}^{n} \frac{\partial \left[Q_{t,i}^{-1}(R_{t,i}) \right]}{\partial R_{t,i}} \\
&\quad \exp\left(-\frac{1}{2} \sum_{i,j=1}^{n} Q_{t,i}^{-1}(R_i) Q_{t,j}^{-1}(R_{t,j}) \left(\kappa_t^{-1}\right)_{ij} \right) && \text{(1.44a)} \\
Q_{t,i} &= \sigma_{ii,t} \left(\begin{array}{c} a_{i,3} X_{t,i}^3 + a_{i,2} X_{t,i}^2 + \\ \left(\sqrt{1 - 6a_{i,3}^2 - 3a_{i,2}^2} - 3a_{i,3} \right) X_{t,i} - a_{i,2} \end{array} \right) + m_{i,t} && \text{(1.44b)} \\
\sigma_{ii,t}^2 &= w_i + p_i \sigma_{ii,t-1}^2 + q_i \varepsilon_{i,t-1}^2 && \text{(1.44c)} \\
m_{i,t} &= c_i + \text{AR}_i \cdot R_{i,t-1} && \text{(1.44d)} \\
\kappa_t &= [\text{diag}(\Omega_t)]^{-1} \times \Omega_t \times [\text{diag}(\Omega_t)]^{-1} && \text{(1.44e)} \\
\Omega_t &= (1 - \psi_1 - \psi_2) \bar{\Omega} + \psi_1 \left(u_{t-1} u_{t-1}' \right) + \psi_2 \Omega_{t-1} && \text{(1.44f)} \\
\bar{\Omega}_{ij} &= \rho_{ij} && \text{(1.44g)}
\end{aligned}
$$

Estimation. The parameters of this model are c_i and AR_i for the conditional mean (Equation 1.44d), w_i, p_i and q_i for the univariate GARCH processes (Equation 1.44c), $a_{i,2}$ and $a_{i,3}$ for the skewness and kurtosis parameters of the univariate CFD (Equation 1.44b), ψ_1 and ψ_2 for the DCC model for the normal rank correlation, κ, (Equation 1.44f) and the unconditional normal rank correlations, ρ_{ij} (Equation 1.44g). We estimate these parameters using a two steps Maximum Likelihood algorithm: we estimate first a set of n GARCH processes with CFD distributed innovations, following the method described in Section 1.5.3, to obtain the parameters c_i, AR_i, w_i, p_i, q_i, $a_{i,2}$ and $a_{i,3}$. If the model

[40] Time dependence in the functions Q_i is denoted by the subscript t: $Q_{i,t}$. Given that the variances follow a GARCH process we have a different transformation $R_{i,t} = Q_{i,t}(X_{i,t})$ at each time t.

Exchange rates			Indexes		
ψ_1	**0.0226**	(0.0019)	ψ_1	**0.0115**	(0.0027)
ψ_2	**0.9731**	(0.0026)	ψ_2	**0.9726**	(0.0072)

Table 1.25. Dynamic CB-MCFD model estimates for both databases. ψ_1 and ψ_2 are the coefficients of the DCC model in Equation 1.39.

is well specified, we can transform the variables R_t into the fictitious variables X_t and then, estimate the parameters of a standard DCC model (ψ_1 and ψ_2), with these fictitious variables using a Maximum Likelihood method as presented in Engle 2002. This method of estimation is fully described in Appendix C.1.4.

Results. Given that the estimation results for the parameters (c_i, AR$_i$, w_i, p_i, q_i, $a_{i,2}$ and $a_{i,3}$) are the same as in the univariate dynamic Section presented in Table 1.15, in Table **??** we just present the estimation for the parameters ψ_1 and ψ_2 of the DCC model. According to this Table, we find that for the exchange rates and indexes database the coefficients ψ_1 are equal to 0.0226 and 0.0115, with an standard error of 0.0019 and 0.0027, and that ψ_2 are equal to 0.9731 and 0.9726 with an standard error of 0.0026 and 0.0725, respectively. According to these results we can conclude that, first, given that $\psi_1 > 0$ the normal rank correlation is persistent, that is, if correlation yesterday was high, it is very likely that today's will be also high. Second, given that $\psi_2 > 0$ we have evidence for normal rank correlation clustering, that is, if returns were yesterday high correlation is also likely to be higher. These results hold for both daily and weekly databases, and are in concordance with the results obtained by Engle 2002 for the usual correlation matrix. The parameters $a_{i,3}$, being all positive and significant, show that the conditional distributions of returns (conditioned on the mean and variance) are leptokurtic.

As in the previous Section, we have considered the KS-χ^2 test, the Omnibus test and the Mardia A and B tests to describe the quality of fit for this model, and the results are presented in Tables **??** and **??**. We observe that at a 5% confidence level, 33, 14, 25 and 12 out of 66 pairs have positive results for the exchange rates database, and that 10, 8, 8 and 9 out of 10 for the indexes database are also positive. These pair results on the bivariate tests are similar to the static CB-MCFD[41]. However, it has to be pointed out that the overall results have significantly improved in the dynamic case with respect to the static one. Although the p-values for the exchange rates database are still lower than 0.00 they are much bigger than in the static case, and considering the indexes database we observe that 3 out of 4 tests are positive at a 1% confidence level and one at a 5% confidence level. Therefore, these results clearly show that dynamics plays a very important role to describe the dependence.

Dynamic VCB-MCFD model

Definition. In this Section we describe a dynamic version of VCB-MCFD presented in Section 1.3.2, which incorporates dynamics and also allows for extreme and asymmetric comovements. This model will be denoted as the *dynamic Variance-Covariance Based Multivariate Cornish-Fisher Density Model* (DVCB-MCFD). As was described in Section 1.3.2, the static VCB-MCFD is basically defined by assets, R, describing the following

[41] It may seem not obvious to obtain less positive pairs test results for the dynamic model than for the static, as the latter is a particular case of the first. This may be the case because we are only considering two parameters for the whole dependence structure ψ_1 and ψ_2 and therefore for some pairs a constant correlation may be a better model. Nevertheless, we will never find a worse overall result in the dynamic case than in the static.

Omnibus \ KS-χ^2

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		0.03	**0.12**	0.03	**0.07**	**0.07**	**0.14**	**0.23**	**0.23**	**0.09**	**0.10**	**0.17**
BEF	**0.91**		0.00	0.00	0.00	0.00	0.00	**0.08**	0.00	**0.06**	0.00	0.00
CHF	**0.94**	0.00		**0.07**	0.01	**0.06**	0.01	**0.07**	0.00	**0.23**	**0.10**	0.00
DEM	**0.94**	0.00	0.00		0.00	0.00	0.00	**0.05**	0.00	**0.07**	0.00	0.00
DKK	**0.91**	0.00	0.00	0.00		0.01	0.00	**0.15**	0.00	**0.27**	0.00	0.00
ESP	**0.82**	0.00	0.00	0.00	0.00		0.02	**0.37**	**0.26**	**0.18**	0.00	**0.13**
FRF	**0.94**	0.00	0.00	0.00	0.00	0.00		**0.09**	0.00	**0.16**	0.00	0.00
GBP	**0.74**	0.00	0.00	0.00	0.00	**0.66**	0.00		**0.05**	**0.07**	**0.15**	**0.13**
ITL	**0.90**	0.00	0.00	0.00	0.00	0.00	0.00	0.00		**0.08**	0.00	0.00
JPY	**0.79**	0.00	0.00	0.00	0.00	**0.07**	0.00	0.01	0.01		0.02	**0.12**
NGL	**0.94**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.84**		0.00
SEK	**0.32**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Mardia B\ Mardia A

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL	JPY	NGL	SEK
AUD		**0.46**	**0.55**	**0.19**	**0.11**	**0.57**	**0.09**	**0.10**	**0.35**	**0.45**	**0.19**	**0.80**
BEF	**0.14**		0.00	0.00	**0.07**	0.00	0.00	0.00	0.00	0.01	0.00	0.00
CHF	**0.21**	0.00		0.00	0.00	0.00	0.00	0.00	0.00	**0.66**	0.00	0.00
DEM	**0.28**	0.00	0.00		0.00	0.00	0.00	0.00	0.00	**0.11**	0.00	0.00
DKK	**0.11**	0.00	0.00	0.00		**0.38**	**0.73**	**0.19**	0.00	**0.09**	0.00	0.00
ESP	**0.07**	0.00	0.00	0.00	0.00		0.00	**0.66**	**0.65**	0.01	0.00	0.00
FRF	**0.18**	0.00	0.00	0.00	0.00	0.00		0.00	0.02	0.04	0.00	0.00
GBP	**0.49**	0.00	0.00	0.00	0.00	**0.85**	0.00		**0.50**	**0.19**	0.00	0.00
ITL	**0.09**	0.00	0.00	0.00	0.00	0.00	0.00	0.00		**0.05**	0.00	**0.06**
JPY	0.01	0.03	0.02	0.04	**0.10**	0.01	0.01	0.00	0.00		0.04	**0.14**
NGL	**0.21**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.10**		0.00
SEK	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
4630	198	9821	0.20
0.00	0.00	0.00	0.00

Table 1.26. Tests for the dynamic Copula-Based Multivariate Cornish-Fisher Density corresponding to the exchange rates database, with the same structure as Table 1.19.

Omnibus $\setminus$ KS-χ^2

	S&P	NKI	STX	EM	EME
S&P		**0.72**	**0.67**	**0.32**	**0.68**
NKI	**0.50**		**0.30**	**0.38**	**0.53**
STX	**0.50**	0.02		**0.69**	**0.94**
EM	**0.71**	0.02	**0.61**		**0.80**
EME	**0.88**	**0.93**	**0.93**	**0.17**	

Mardia A $\setminus$ Mardia B

	S&P	NKI	STX	EM	EME
S&P		**0.12**	**0.45**	**0.09**	**0.22**
NKI	**0.64**		0.00	0.00	**0.43**
STX	**0.07**	0.01		**0.41**	**0.47**
EM	**0.13**	**0.32**	**0.51**		**0.19**
EME	**0.21**	**0.21**	**0.55**	**0.30**	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
0.02	0.00	0.01	**0.10**

Table 1.27. Tests for the dynamic Copula-Based Multivariate Cornish-Fisher Density corresponding to the Indexes database, with the same structure as Table 1.19.

relation:

$$R = \Sigma^{1/2} \cdot z + m$$

where $\Sigma^{1/2}$ denotes the Cholesky decomposition of the variance-covariance matrix Σ and z follows an Independent Multivariate Cornish-Fisher Distribution:

$$\text{i-mcfd}_3(z) = \frac{1}{\sqrt{(2\pi)^n}} \prod_{i=1}^{n} \frac{\partial \left[Q_i^{-1}(z_i)\right]}{\partial z_i} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(Q_i^{-1}(z_i)\right)^2\right)$$

In this Section we propose to include a dynamic character by assuming a DCC model described in the previous Section for the variance-covariance matrix, Σ. In this way, the full model can be formally written as:

$$\begin{aligned}
R_t &= \Sigma_t^{1/2} \cdot z_t + m_t \\
m_{i,t} &= c_i + \text{AR}_i \cdot R_{i,t-1} \\
\Sigma_t &= D_t \Theta_t D_t \\
D_{ii,t} &= \sqrt{\sigma_{ii}^2} \text{ and } D_{ij,t} = 0 \\
\sigma_{ii,t}^2 &= w_i + p_i \sigma_{ii,t-1}^2 + q_i \varepsilon_{i,t-1}^2 \\
z_t &\sim \text{i-mcfd}_3(z_t) = \frac{1}{\sqrt{(2\pi)^n}} \prod_{i=1}^{n} \frac{\partial \left[Q_{i,t}^{-1}(z_{i,t})\right]}{\partial z_{i,t}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(Q_{i,t}^{-1}(z_{i,t})\right)^2\right) \\
Q_{i,t}(z_{i,t}) &= a_{i,3} z_{i,t}^3 + a_{i,2} z_{i,t}^2 + \left(\sqrt{1 - 6a_{i,3}^2 - 3a_{i,2}^2} - 3a_{i,3}\right) z_{i,t} - a_{i,2} \\
\Theta_t &= \left[\text{diag}(\Omega_t)\right]^{-1} \times \Omega_t \times \left[\text{diag}(\Omega_t)\right]^{-1} \\
\Omega_t &= (1 - \psi_1 - \psi_2)\bar{\Omega} + \psi_1 \left(u_{t-1} u_{t-1}'\right) + \psi_2 \Omega_{t-1} \\
\bar{\Omega}_{ij} &= \rho_{ij}
\end{aligned}$$

The parameters of this model are the same as in the DCB-MCFD model, namely, c_i and AR_i for the conditional mean, w_i, p_i and q_i for the univariate GARCH processes, $a_{i,2}$ and $a_{i,3}$

for the skewness and kurtosis parameters, ψ_1 and ψ_2 for the DCC model of the correlation matrix, Θ_t, and the unconditional correlations, ρ_{ij}.

Estimation. We estimate these parameters using a two steps Maximum Likelihood algorithm. First, we estimate a set of n standard GARCH processes with gaussian distributed innovations to obtain the parameters c_i, AR_i, w_i, p_i and q_i. Afterwards, we estimate the parameters of the standard DCC model (ψ_1 and ψ_2) together with the $a_{2,i}$ and $a_{3,i}$ parameters using a Maximum Likelihood method. This method of estimation is fully described in Appendix C.1.4.

Results. In Table 1.28 we present the estimation results for this model with their standard errors. According to these estimations, we basically observe the same behavior as in Engle 2002: persistence and clustering of correlation and volatility. Besides, we obtain significant parameters $a_{i,3}$ indicating the existence of multivariate heavy tails in the data.

In Tables 1.29 and 1.30 we present the results of the fit tests for this model. In the bivariate tests (the KS-χ^2 test, the Omnibus test and the Mardia A and B tests) we observe 50, 26, 36 and 20 positive results (at a 5% confidence level) out of 66 pairs for the exchange rates database and 10, 10, 9 and 10 out of 10 for the indexes database. It is interesting to remark that these results are the best compared to the other three multivariate methods.

Comparison of Multivariate Models

In Table 1.31 we present a summary of the main results of the estimation tests for the four multivariate models presented in this work (static CB-MCFD, dynamic CB-MCFD, static VCB-MCFD and dynamic VCB-MCFD) for comparison purposes. In addition, we

	AUD	BEF	CHF	DEM	DKK	ESP	FRF	GBP	ITL
c	-0.0010	-0.0050	-0.0065	-0.0056	-0.0050	-0.0142	-0.0054	-0.0062	-0.0172
std.	0.0108	0.0138	0.0154	0.0139	0.0136	0.0146	0.0134	0.0145	0.0139
w	**0.0058**	**0.0082**	**0.0134**	**0.0067**	**0.0054**	**0.0142**	**0.0070**	**0.0028**	**0.0074**
std.	0.0015	0.0021	0.0027	0.0019	0.0016	0.0029	0.0019	0.0011	0.0020
p	**0.0495**	**0.0435**	**0.0385**	**0.0431**	**0.0380**	**0.0728**	**0.0480**	**0.0378**	**0.0760**
std.	0.0066	0.0057	0.0055	0.0051	0.0045	0.0132	0.0057	0.0037	0.0093
q	**0.9319**	**0.9395**	**0.9378**	**0.9426**	**0.9504**	**0.9007**	**0.9358**	**0.9562**	**0.9103**
std.	0.0033	0.0022	0.0018	0.0017	0.0013	0.0078	0.0022	0.0014	0.0058
a_3	**0.0654**	**0.0647**	**0.0562**	**0.0875**	**0.1042**	**0.0798**	**0.0728**	**0.0591**	**0.0617**
std.	0.0061	0.0056	0.0059	0.0054	0.0051	0.0055	0.0061	0.0054	0.0053
a_2	**0.0465**	-0.0041	**0.0313**	-0.0095	0.0020	-0.0094	-0.0110	**0.0346**	0.0107
std.	0.0097	0.0106	0.0110	0.0118	0.0111	0.0111	0.0110	0.0097	0.0105

	JPY	NGL	SEK	S&P	NKI	STX	EM	EME
c	-0.0041	-0.0002	0.0148	0.1943	-0.1038	0.1682	0.0245	0.1020
std.	0.0129	0.0130	0.0117	0.1048	0.1351	0.1360	0.1337	0.1977
AR				**-0.0782**	-0.0434	**-0.1196**	**0.0759**	**0.0740**
std.				0.0351	0.0416	0.0290	0.0397	0.0353
w	**0.0125**	**0.0070**	**0.0043**	**0.0723**	**2.6415**	**0.2434**	**0.2658**	**0.8483**
std.	0.0026	0.0019	0.0020	0.0410	0.7463	0.0992	0.1357	0.3087
p	**0.0471**	**0.0416**	**0.0514**	**0.0814**	0.0409	**0.2054**	**0.0781**	**0.1064**
std.	0.0081	0.0052	0.0072	0.0148	0.0402	0.0420	0.0204	0.0275
q	**0.9261**	**0.9436**	**0.9425**	**0.9092**	**0.6764**	**0.7784**	**0.8920**	**0.8511**
std.	0.0039	0.0018	0.0027	0.0021	0.0364	0.0214	0.0037	0.0090
a_3	**0.0631**	**0.1126**	**0.1080**	**0.0274**	**0.0327**	**0.0281**	**0.0531**	**0.0655**
std.	0.0055	0.0049	0.0052	0.0100	0.0091	0.0130	0.0124	0.0108
a_2	**0.0262**	-0.0072	-0.0094	**-0.0664**	-0.0013	-0.0253	**-0.0435**	-0.0299
std.	0.0099	0.0115	0.0104	0.0232	0.0277	0.0209	0.0224	0.0258

	Exchange rates			Indexes	
ψ_1	**0.0210**	0.0012	ψ_1	**0.0138**	0.0106
ψ_2	**0.9755**	0.0016	ψ_2	**0.9431**	0.0591

Table 1.28. Dynamic VCB-MCFD model estimates for both databases. Legend is the same as in Table 1.14 with a_3 and a_2 the coefficients of the standardized Cornish-Fisher Density.

Omnibus \ KS-χ^2

	AUD	**BEF**	**CHF**	**DEM**	**DKK**	**ESP**	**FRF**	**GBP**	**ITL**	**JPY**	**NGL**	**SEK**
AUD		0.00	**0.71**	**0.97**	**0.15**	**0.48**	**0.99**	**0.02**	**0.20**	0.00	**0.37**	**0.60**
BEF	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CHF	**0.28**	0.00		**0.76**	**0.35**	**0.50**	**0.70**	**0.91**	**0.48**	**0.09**	**0.53**	**0.83**
DEM	**0.12**	0.00	0.04		**0.17**	**0.52**	**0.19**	**0.58**	**0.33**	**0.20**	0.00	**0.21**
DKK	**0.16**	0.00	**0.07**	0.04		**0.60**	**0.21**	**0.47**	**0.43**	**0.61**	0.04	**0.41**
ESP	**0.55**	0.00	**0.23**	**0.10**	**0.14**		**0.39**	**0.39**	**0.34**	**0.70**	**0.17**	**0.36**
FRF	**0.86**	0.00	**0.44**	**0.16**	**0.26**	**0.78**		**0.33**	**0.24**	**0.16**	**0.85**	**0.55**
GBP	**0.58**	0.00	**0.26**	**0.12**	**0.15**	**0.51**	**0.82**		0.01	0.01	**0.70**	**0.37**
ITL	**0.33**	0.00	**0.13**	**0.05**	**0.07**	**0.29**	**0.51**	**0.31**		**0.09**	**0.92**	**0.74**
JPY	0.01	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.00		**0.70**	**0.27**
NGL	0.02	0.00	0.00	0.00	0.00	0.01	0.04	0.02	0.01	0.00		**0.24**
SEK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Mardia B\ Mardia A

	AUD	**BEF**	**CHF**	**DEM**	**DKK**	**ESP**	**FRF**	**GBP**	**ITL**	**JPY**	**NGL**	**SEK**
AUD		**0.05**	**0.49**	0.02	**0.37**	**0.49**	**0.50**	**0.05**	**0.75**	**0.22**	**0.31**	0.02
BEF	0.00		0.01	0.00	**0.13**	**0.05**	**0.18**	**0.13**	0.02	0.00	0.03	0.00
CHF	**0.43**	0.00		0.03	**0.24**	**0.05**	**0.75**	**0.05**	0.00	0.00	0.01	0.00
DEM	**0.95**	0.00	0.00		0.02	**0.11**	0.01	0.03	0.01	0.01	0.03	0.00
DKK	**0.25**	0.00	**0.05**	0.00		**0.78**	**0.98**	**0.95**	**0.81**	**0.13**	**0.12**	0.00
ESP	**0.46**	0.00	0.03	**0.08**	**0.25**		**0.53**	**0.38**	**0.25**	**0.26**	**0.35**	0.01
FRF	**0.71**	0.00	**0.08**	0.00	0.00	**0.76**		**0.31**	**0.43**	**0.18**	**0.10**	0.00
GBP	**0.15**	0.00	**0.07**	0.00	**0.05**	**0.08**	0.00		**0.44**	**0.11**	**0.16**	0.03
ITL	0.02	0.00	0.02	0.00	0.04	0.01	0.01	0.00		0.03	**0.09**	0.00
JPY	0.00	0.00	0.00	0.00	**0.07**	0.03	0.00	0.00	0.00		0.01	0.01
NGL	**0.19**	0.00	0.03	0.00	0.00	**0.43**	0.00	**0.15**	**0.24**	**0.17**		0.00
SEK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Complete Tests

Mardia A	**Mardia B**	**Omnibus**	**KS-χ^2**
802	51	286	0.11
0.00	0.00	0.00	0.00

Table 1.29. Tests for the dynamic Variance-Covariance Based Multivariate Cornish-Fisher Density corresponding to the exchange rates database, with the same structure as Table 1.19.

Omnibus \ KS-χ^2

	S&P	NKI	STX	EM	EME
S&P		**0.52**	**0.98**	**0.70**	**0.71**
STX	**0.22**		**0.15**	**0.56**	**0.10**
NKI	**0.21**	**0.54**		**0.94**	**0.64**
EM	**0.31**	**0.70**	**0.69**		**0.38**
EME	**0.27**	**0.64**	**0.63**	**0.80**	

Mardia A \ Mardia B

	S&P	NKI	STX	EM	EME
S&P		**0.22**	**0.54**	**0.18**	**0.28**
STX	**0.38**		**0.62**	**0.32**	**0.14**
NKI	**0.13**	**0.11**		**0.70**	**0.97**
EM	**0.06**	**0.52**	**0.36**		0.02
EME	**0.20**	**0.27**	**0.62**	**0.24**	

Complete Tests

Mardia A	Mardia B	Omnibus	KS-χ^2
0.11	0.00	**0.53**	**0.09**

Table 1.30. Tests for the dynamic Variance-Covariance Based Multivariate Cornish-Fisher Density corresponding to the Indexes database, with the same structure as Table 1.19.

have estimated two benchmark models, a dynamic multivariate normal (dynamic MN) as in Engle 2002 and a dynamic multivariate Skew-T model (dynamic MSkT) as in Jondeau and Rockinger 2003. These two models present the same dynamic characteristics as the dynamic VCB-MCFD, namely, an AR(1)-GARCH(1,1) with a DCC(1,1) model:

$$\begin{aligned}
R_t &= \Sigma_t^{1/2} \cdot z_t + m_t \\
m_{i,t} &= c_i + \text{AR}_i \cdot R_{i,t-1} \\
\Sigma_t &= D_t \Theta_t D_t \\
D_{ii,t} &= \sqrt{\sigma_{ii}^2} \text{ and } D_{ij,t} = 0 \\
\sigma_{ii,t}^2 &= w_i + p_i \sigma_{ii,t-1}^2 + q_i \varepsilon_{i,t-1}^2 \\
\Theta_t &= [\text{diag}(\Omega_t)]^{-1} \times \Omega_t \times [\text{diag}(\Omega_t)]^{-1} \\
\Omega_t &= (1 - \psi_1 - \psi_2)\,\bar{\Omega} + \psi_1 \left(u_{t-1} u_{t-1}'\right) + \psi_2 \Omega_{t-1} \\
\bar{\Omega}_{ij} &= \rho_{ij}
\end{aligned}$$

but posses different distributional hypothesis for the innovations, z_t, namely, multivariate normality for the dynamic MN model and a multivariate Skew-T for the dynamic MSkT. As we already know, the multivariate normal distribution does not incorporate asymmetries or heavy tails while the multivariate Skew-T does (see Appendix A.1 for more details on the multivariate Skew-T distribution). In particular in Table 1.31, we present the number of positive pair tests for each test (KS-χ^2, Omnibus, Mardia A and Mardia B) as well as the test results of the whole database and the Log-Likelihood values (Log-Lk.) and the Akaike (AIC) and Bayesian (BIC) model selection criteria, which penalize for an increase in complexity through the inclusion of more parameters.

Model	Number of Positive Pair Tests				Results of Overall Fit Tests						
	KS-χ^2	Omni	M. A	M. B	KS-χ^2	Omni	M. A	M. B	Log-Lk.	AIC	BIC
	Exchange rates database										
	Statistic Values (all p-values are 0.00)										
st. CB-MCFD	10	33	34	9	0.25	13339	3629	266	-9681	19615	20347
st. VCB-MCFD	35	25	55	24	0.16	813	800	84	-7355	14963	15695
dy. CB-MCFD	33	14	25	12	0.20	9821	4630	198	-5782	11906	12650
dy. VCB-MCFD	50	26	36	20	0.11	286	802	51	-4376	9008	9752
dy. Gaussian	0	0	0	0	0.27	17604	27998	503	-7092	14996	14996
dy. Skew Student-T	41	27	48	25	0.11	164	700	45	-4343	9008	9752
	Indexes database										
	p-values										
st. CB-MCFD	10	10	7	7	0.01	0.06	0.00	0.00	-5892	11854	12003
st. VCB-MCFD	10	10	9	6	0.01	0.20	0.00	0.00	-5912	11894	12043
dy. CB-MCFD	10	8	8	9	0.10	0.01	0.02	0.00	-5825	11724	11881
dy. VCB-MCFD	10	10	9	10	0.09	0.53	0.11	0.00	-5799	11672	11830
dy. Gaussian	1	1	1	1	0.00	0.00	0.00	0.00	-5845	11878	11878
dy. Skew Student-T	10	10	9	10	0.06	0.58	0.09	0.00	-5802	11672	11830

Table 1.31. Summary of the test results for the four models: static and dynamic CB-MCFD and VCB-MCFD. In addition we present a dynamic multivariate gaussian and a dynamic multivariate Skew Student T model. The first four columns present the number of positive pair tests in the sample. For the exchange rates (indexes) database we have a total of 66 (10) possible pairs. Columns 5-8 stand for the test results of the whole database. For the exchange rates database we present the statistic value given that all associated p-values are 0.00, while in the indexes database we present the p-value. KS-χ^2, Omni, M. A and M. B stand for the Kolmogorov-Smirnov Test, the Omnibus Test, and the Mardia A and B Tests. Finally, Columns 9-11 present the Log-Likelihood (Log-Lk.), the Akaike Criteria (AIC) and the Bayesian Criteria (BIC). Lower (higher) AIC and BIC (Log-lk.) values imply a better fit.

Analyzing the results we can obtain the following conclusions:

1. In general, considering the selection criteria statistics (Log-Lk., AIC and BIC), we observe that the best models for both databases are the dynamic MSkT and the dynamic VCB-MCFD: the estimation tests (KS-χ^2, Omnibus, Mardia A and Mardia B) are a somewhat better for the dynamic MSkT, but model selection criteria do not distinguish between both models.

2. Dynamic models present much better results than static models. Therefore, we can conclude that both databases present second order (variances and correlations) dynamics.

3. The dynamic Multivariate Normal model is much worse than the dynamic MSkT and VCB-MCFD ones. Therefore, both databases present univariate heavy tails and asymmetries.

4. The dynamic Copula-Based model is much worse than the dynamic MSkT and the dynamic VCB-MCFD. Therefore, both databases present multivariate heavy tails and asymmetries.

5. In the indexes database we find a strong statistical support for the dynamic VCB-MCFD model. Actually, in 3 out of 4 estimation tests we find that we cannot reject the null hypothesis that returns are samples of a dynamic VCB-MCFD with p-values of 0.09, 0.53, 0.11 and 0.00 for the KS-χ^2 test, the Omnibus test and the Mardia A and B tests, respectively.

6. Considering the exchange-rates database we find that still more research has to be done to capture adequately the styled facts of these series given that we obtain p-values for all models equal to zero. We find a rejection of the dynamic VCB-MCFD model for this database, but p-values of this model are the highest (with the dynamic MSkT) of the six models as can be seen from the statistic values.

1.7 Conclusions

In this Chapter we have analyzed the theory of Cornish-Fisher distributions and we have applied them to financial data to model stylized facts as asymmetry, heavy tails, clustering and persistence of volatility and correlation dynamics.

These distributions are based on Cornish-Fisher Expansions which are applied to theoretically determined distributions (with known moments) to obtain the quantiles but have not been used to create a family of distributions. We have demonstrated that univariate CFD are very easy to estimate and we have provided three methods for the estimation of static models and developed a Maximum Likelihood estimator for a GARCH model with CFD distributed innovations. The interpretation of the parameters of the distribution have a natural and straightforward interpretation in terms of the coefficients of the third-order polynomial of a QQ-plot function. This distribution is more flexible than Gram-Charlier, Edgeworth or semi-nonparametric distributions of similar order, and capture well the main characteristics of univariate financial series performing better than standard market models as the skewed-t distribution.

We have also developed two families of multivariate distributions, namely Copula-Based and Variance-Covariance Based MCFD, which incorporate different dependence structures. The Copula-Based distribution incorporates a gaussian copula dependence and the Variance-Covariance permits multivariate heavy tails. We have observed that static multivariate models are not enough to model financial series and, therefore, we have introduced dynamics using a Dynamic Conditional Correlation model. According to our results, we have been able to model accurately the indexes database, but still more research has to be done to capture adequately the exchange rates database. A possible path to solve this problem would be to include dynamics not only in the first two moments, but also an autoregressive behavior could be imposed to the parameters a_2 and a_3 to model the dynamics on higher moments considering works like Rubio, Serna, and León 2006 and Jondeau and Rockinger 2005. In addition, other second order dynamics (as BKKK models) could be considered but this path seems not to be very promising given the results of Engle 2002. More interesting could be the analysis of other dependence structures based on the Extreme Value Theory.

Finally, it is interesting to note that, although not reported in this thesis, the dynamic Variance-Covariance Based Multivariate CFD has motivated more investigation and the model has been proven to be a good predictor of multivariate volatility (Bergara, Ansejo, and Rubia 2006).

Chapter 2
Financial Applications of Cornish-Fisher Distributions

In this Chapter we will present several applications of both the *Cornish-Fisher density function* and the *Multivariate Cornish-Fisher density function* described in the first Chapter for three different areas of interest in financial modeling: i) optimal portfolio selection, ii) measure of risk via the Value at Risk (VaR), and iii) option valuation.

2.1 Optimal Portfolio Selection

In this Section we analyze the Markowitz (Markowitz 1959) hypothesis, that a mean-and-variance based analysis to construct optimal portfolios is enough to maximize investors' expected utility function. Many authors (e.g., Arditti 1967 and Samuelson 1970) have argued that the expected utility function may be more appropriately approximated by a function of higher moments, where investors are supposed to like positive skewness and dislike fat-tailedness, as measured by the kurtosis, but, on the other hand, early empirical evidence (e.g., Levy and Markowitz 1979 and Pulley 1981) suggests that a mean-variance optimization results in allocations that are similar to the ones obtained using a direct optimization of expected utility. Therefore, in order to provide more evidence on this issue, we use a VCB Multivariate CFD model (Section 1.3) to analyze whether the inclusion of higher order moments improves asset allocation in terms of utility, finding that in many cases higher order moments are fundamental in the asset allocation procedure.

2.1.1 Introduction

The theory of portfolio optimization establishes a framework for asset selection through a competition between expected return and risk. Investors choose a bearable risk level depending on their aversion to it and look for the portfolio offering the highest expected return for that given level[1]. Markowitz's theory (Markowitz 1952) assumes that expected returns of assets in a portfolio follow a multivariate normal distribution, as many other theories that constitute the fundamentals of traditional financial mathematics. According to Markowitz's theory, an optimal investment behavior promotes diversification among assets in order to reduce the risk of a portfolio. This hypothesis became a good first approximation that as a result derived analytically tractable theories. It is well known that the mean-variance theory is only valid under the hypothesis of gaussian distributed returns or under the assumption that investors' utility function is quadratic. Nonetheless, as we have seen in Chapter 1, there is a strong empirical support against this normality hypothesis and, as pointed out by Tsiang 1972, the quadratic utility function may lead to absurd results. To this respect, in his Nobel Lecture Markowitz (Markowitz 1991) stated the following caveat: "*Equipped with database, computer algorithms and methods of estimation, the modern portfolio theorist is able to trace out mean-variance frontiers for large universes of securities. But, is this the right thing to do for the investor? In particular, are mean and variance proper and sufficient criteria for portfolio choice?*". Empirical and theoretical *attacks* to Markowitz's portfolio theory addressing this question have given impetus to both the investigation on moments of higher order than the mean and variance and how their

[1] Equivalently, investors can also select an expected return target and look for the portfolio that minimizes the risk.

inclusion in the definition of risk would affect investment decisions of rational investors. In this work we will mainly focus on the first part, namely, whether higher order moments are sufficient criteria for portfolio choice[2].

While some authors have argued that the expected utility function may be more appropriately approximated by a function of higher moments where investors are supposed to like positive skewness and dislike fat-tailedness as measured by kurtosis (Arditti 1967, and Samuelson 1970), early empirical evidence suggests that mean-variance criterion under certain assumptions, as small investment periods or high wealth to risk ratios, results in allocations that are very similar to the ones obtained using a direct optimization of the expected utility (Levy and Markowitz 1979, Pulley 1981, Tsiang 1972 and Kroll, Levy, and Markowitz 1984). An explanation of the good performance of the mean-variance criterion reported in these papers may be that, although returns are non-normal, they are driven by an elliptical distribution for which the mean-variance approximation of the expected utility remains exact for all utility functions (Owen and Rabinovitch 1983). However, under large departure from normality, in particular when the distribution is severely asymmetric (Flôres and de Athayde 2002, Athayde and Flôres 2004 and Chunhachinda, Danpadani,

[2] The second aspect, namely, which risk measure to use once it has been realized that variance is not a sufficiently good measure, has also attracted much interest among financial literature. For example, the Semivariance model (Markowitz 1959), incorporating only the negative returns, the Gini risk measure or Gini's mean difference (Yitzhaki 1982 and Shalit and Yitzhaki 1984), which is based on the expected value of the absolute difference between every pair of realizations, and the Shortfall probability (Roy 1952, Browne 1999), where the risk is defined over an expected return or reward. The Hodges ratio (Hodges 1998), the Omega ratio (Shadwick and Keating 2002) or the inferior partial moments (Bawa and Lindenberg 1977 Fishburn 1977) are also defined in terms of a weighted average of deviations. Likewise, given the success that VaR has raised, Alexander and Baptista 2002, describe a portfolio selection model based on the mean-VaR competition and, recently, a number of new risk measures with the desirable properties of subadditivity and coherence (Artzner, Eber, and Heath 1999), like the Expected Shortfall (Artzner, Eber, and Heath 1999) or the Spectral Risk Measures (Acerbi 2002, Acerbi 2004) have also been proposed. Nevertheless, if it would be proven that mean and variance would be a sufficient criteria for portfolio choice all these measures would reduce to the Markowitz's framework.

Hamid, and Prakash 1997), a three or four moment optimization strategy provides a better approximation of the expected utility. Therefore, the question stated by Markowitz still seems to be open and most of the recent literature on this subject tries to answer the question whether mean-variance portfolios are a good approximation to reality or if, in contrast, predictions of this theory are too rude[3].

On the other hand, recently a number of papers have highlighted the importance of modeling the dynamic behavior of higher order moments in order to capture the real investment opportunities at each date. For example, Ang and Bekaert 2002 and Guidolin and Timmerman 2005 use the Markov-switching approach, where the mean and variance of returns are allowed to vary over time as regimes change, Jondeau and Rockinger 2005 model the dynamics as DCC Model first proposed by Engle and Sheppard 2001 with innovations given by a Skewed T-Student and Harvey and Siddique 2000 shows how the conditional skewness is priced by investors.

In this Section we use the new model presented in Chapter 1, the dynamic VCB-MCFD, for the characterization of non-normality to provide more evidence on the issue whether the expected utility function may be more appropriately approximated by a function of higher moments than just by the first two moments. While keeping to capture

[3] In past research, this topic was discussed for example by Arditti 1967, Levy 1969, Jean 1971, Rubinstein 1973, Jean 1973, Arditti and Levy 1975, Francis 1975 Ingersoll 1975 and Kraus and Litzenberger 1976, but portfolio optimization taking into account more than the first two moments has been receiving renewed interest in recent years. Both on the theoretical side, including its links with the multimoment CAPM extensions, or on what relates to econometric tests or updates based on higher conditional moments. Works like Adcock 2002, Hung, Shackleton, and Xu 2004, Athayde and Flôres 2004, Flôres and de Athayde 2002, Jurczenko and B.Maillet 2001, Prakash, Chang, and Pactwa 2003, Dittmar 2002, Barone-Adesi, Gagliardini, and Urga 2000, Harvey and Siddique 2000, Guidolin and Timmerman 2005 and Jondeau and Rockinger 2005, far from exhausting the full list of contributions, pay good witness to the growing awareness of the importance of higher moments in both lines of research.

first-order moment dynamics and stylized facts like volatility clustering with a GARCH model via a Dynamic Conditional Correlation (DCC) as in Jondeau and Rockinger 2005, we investigate the consequences of assuming a Variance-Covariance Based Multivariate Cornish-Fisher Density (MCFD) for the returns. Main objectives of the inclusion of the VCB-MCFD model described in Section 1.3 are to address two characteristic limitations of the hypothesis that returns follow a multivariate normal distribution with time independent parameters throughout time: first, including asymmetry and fat tails and, second, incorporating non-linear correlations between the different assets of the investment set.

With this setting we try to give more evidence on the importance of higher moments in portfolio decision theory and, after presenting the model, we evaluate the inclusion of more terms in the Taylor expansion of the utility function through the cost of opportunity (Simaan 1993) in both an unconditional and a conditional framework for a portfolio of several international equity indexes. We also analyze the gain of considering a conditional setting instead of an unconditional one and the impact of restricting the investment possibilities prohibiting short-selling and borrowing. Basically, our results allow us to conclude that inclusion of higher order moments might be very relevant in asset allocation and considering time-varying moments highly improves the allocation in terms of cost of opportunity.

The structure of this Section is as follows. First, we introduce the VCB-MCFD in the portfolio selection problem and next, we analyze the optimal allocation problem with higher moments. Afterwards, in Sections 2.1.5 and 2.1.6 we present the optimization results both in a constant and a time-varying investment universe and, finally, we end this Section with our conclusions.

2.1.2 The Model

While the multivariate normal distribution has several attractive properties for modeling a portfolio there is considerable evidence that portfolio returns are non-normal. In the first Chapter we have presented a conditional set-up that incorporates most statistical features of stock market returns, like the well-known properties of volatility clustering (Bollerslev 1986), time-varying correlations (Engle and Sheppard 2001) and both the asymmetry and fat-tailedness that is often found in stock or index data. In particular, we will use the model named dynamic VCB-MCFD introduced in Chapter 1, as it is the one that best incorporates these mentioned features. With this setting we model the dynamics of the first two moments of the return distribution following the DCC model (Dynamic Conditional Correlation Multivariate GARCH) of Engle and Sheppard 2001 that has also been applied by Jondeau and Rockinger 2005 in the field of portfolio optimization, and we also capture the non-gaussianity behavior of the standardized returns proposing the use of the Multivariate Cornish-Fisher Density. As has been proven in Chapter 1, a remarkable feature of this distribution is that it is defined starting from a Cornish-Fisher Expansion and, therefore, in principle it is possible to fit as much as needed any univariate density adding more terms in the expansion. Besides, it is a straightforward extension of the multivariate normal distribution and the associated parameters have a quite natural interpretation.

For the ease of readiness we will present here the Equations defining this model. Let R_t be the vector of returns of the assets at time t, $R_t = \{R_{i,t}\}_{i=1}^n$, then the dynamic VCB-MCFD is defined as:

$$R_t = \Sigma_t^{1/2} \cdot z_t + m_t \tag{2.1a}$$

$$m_{i,t} = c_i + \text{AR}_i \cdot R_{i,t-1} \tag{2.1b}$$

$$\Sigma_t = D_t \Theta_t D_t \tag{2.1c}$$

$$D_{ii,t} = \sqrt{\sigma_{ii}^2} \text{ and } D_{ij,t} = 0 \tag{2.1d}$$

$$\sigma_{ii,t}^2 = w_i + p_i \sigma_{ii,t-1}^2 + q_i \varepsilon_{i,t-1}^2, \qquad \varepsilon_t = \Sigma_t^{1/2} \cdot z_t \tag{2.1e}$$

$$z_t \sim \text{i-mcfd}_3(z_t) = \frac{1}{\sqrt{(2\pi)^n}} \prod_{i=1}^n \frac{\partial \left[Q_{i,t}^{-1}(z_{i,t})\right]}{\partial z_{i,t}} e^{\left(-\frac{1}{2}\sum_{i=1}^n \left(Q_{i,t}^{-1}(z_{i,t})\right)^2\right)} \tag{2.1f}$$

$$Q_{i,t}(z_{i,t}) = a_{i,3} z_{i,t}^3 + a_{i,2} z_{i,t}^2 + \left(\sqrt{1 - 6a_{i,3}^2 - 3a_{i,2}^2} - 3a_{i,3}\right) z_{i,t} - a_{i,2} \tag{2.1g}$$

$$\Theta_t = [\text{diag}(\Omega_t)]^{-1} \times \Omega_t \times [\text{diag}(\Omega_t)]^{-1} \tag{2.1h}$$

$$\Omega_t = (1 - \psi_1 - \psi_2)\bar{\Omega} + \psi_1 \left(u_{t-1} u_{t-1}'\right) + \psi_2 \Omega_{t-1}, \quad u_t = D_t^{-1} \varepsilon_t \tag{2.1i}$$

$$\bar{\Omega}_{ij} = \rho_{ij} \tag{2.1j}$$

The term m_t in Equation 2.1a represents the conditional vector mean, which in our model will be described by a first order autoregressive model AR(1) (Equation 2.1b). On the other hand, the first term in Equation 2.1a, $\Sigma_t^{1/2} \cdot z_t$, represents the unexpected part of returns and consists on a time varying variance-covariance matrix, Σ_t, and a vector of innovations, z_t. The variance-covariance matrix is modeled by Equation 2.1c, which divides the contribution of covariance into variance (D_t) and correlation (Θ_t). The conditional variances of asset i, $\sigma_{ii,t}^2$, will be described by a GARCH(1,1) representation given by Equation 2.1e, although more complicated processes could be trivially accommodated. In the GARCH(1,1) setting, the parameter p_i captures the sensibility of present volatility to past returns (high or low returns tend to predict volatility) and the parameter q_i captures the volatility clustering phenomenon, i.e. high past volatility tends to predict high present volatility. In addition, we model the conditional correlation matrix, Θ_t,with a DCC model, as proposed by Engle and Sheppard 2001. This model is designed to model time-varying correlations avoiding the dimensionality curse. With this specification (Equations 2.1h-2.1j), it is easy to see that the conditional correlation matrix, Θ_t, follows a GARCH-like process where the parameters ψ_1 and ψ_2 are the equivalents to p_i and q_i in the GARCH Equation 2.1e.

Although in standard multivariate GARCH models it is assumed that the innovation, z_t, follows a multivariate normal standardized distribution, in our model we will consider that innovations are distributed as an independent multivariate third-order Cornish-Fisher Density defined in Equations 2.1f and 2.1g with parameters $a_{i,3}$ and $a_{i,2}$. As mentioned in Section 1.3, the variables $a_{i,2}$ capture the asymmetry and the variables $a_{i,3}$ capture the heavy tailedness. Therefore, in this new model first and second order moments are allowed to be time-varying while higher order moments associated to the innovations are kept constant, but not zero as in the gaussian case.

Summarizing, the parameters of this model are: c_i and AR_i for the conditional mean, w_i, p_i and q_i for the univariate GARCH processes, $a_{i,2}$ and $a_{i,3}$ for the asymmetry and fat-tailedness parameters, ψ_1 and ψ_2 for the DCC model of the correlation matrix Θ_t and the unconditional correlations, ρ_{ij}. Therefore, as the number of assets, n, increases the number of parameters in the model just increases only with $\frac{15}{2}n + \frac{1}{2}n^2 + 2$ [4], avoiding in this way the dimensionality curse. As an example, with 5 assets we would have 52 parameters.

We estimate these parameters using a two steps Maximum Likelihood algorithm: we estimate first a set of standard GARCH processes with gaussian distributed innovations to obtain the parameters c_i, AR_i, w_i, p_i and q_i, and afterwards, we estimate the parameters of a standard DCC model, ψ_1 and ψ_2, together with the $a_{i,2}$ and $a_{i,3}$ parameters using a Maximum Likelihood method. This method of estimation is fully described in Appendix C.1.4.

In the unconditional optimization Section 2.1.4 we will consider a static version of the dynamic VCB-MCFD model presented above. We can obtain this static model by setting in Equations 2.1a-2.1j the following restrictions: $\text{AR}_i = p_i = q_i = \psi_1 = \psi_2 = 0$.

2.1.3 Optimal Portfolio Selection with Higher Moments

When investors' beliefs about future returns strongly depart from normality or when investors' utility function differs from the quadratic one, the standard mean-variance criterion may be inappropriate, or even a bad approximation. One way of quantifying how bad this approximation is, is to compare the portfolio chosen with a mean-variance criteria with the optimal portfolio chosen in an expected utility framework corresponding to the fundamental theories of action under risk and uncertainty of Von Neumann and Morgenstern and L. J. Savage (Tsiang 1972, Neumann and Morgenstern 1953). Therefore, in

[4] We have $2n$ parameters for the coefficients c_i and AR_i in the conditional mean, $3n$ for the GARCH parameters, k_i, β_i and γ_i, $2n$ for the parameters $a_{i,2}$ and $a_{i,3}$, 2 for the DCC parameters, ψ_1 and ψ_2, and $n(n+1)/2$ parameters for the unconditional correlations, $\rho_{ij.}$

this work we will suppose that investors optimal selection is the one obtained through a maximization of the expected utility function. Markowitz's theory takes into account just the first two moments of the expected utility function, but, as put forward by many authors (e.g. Jondeau and Rockinger 2005, Ang and Bekaert 2002), incorporating the effect of higher moments than the mean and variance on the expected utility of investors (which corresponds to Markowitz's framework) would improve the moments approximation in the allocation of wealth and fill the gap between moments based and utility based selections.

Theoretical research suggests (under rather general conditions) that investors prefer high values of odd moments, and low values of even moments (e.g. Scott and Horvath 1980 or Pratt and Zeckhauser 1987). Hence, investors would prefer positive skewness, because they prefer positive extreme values and dislike negative extreme value and, in addition, they would avoid kurtosis, because it is a measure of dispersion and, therefore, of uncertainty. Since we are primarily interested in the effect of higher moments on the asset allocation, a convenient approach consists in approximating the utility function using a Taylor series expansion around the current value of the portfolio return, and afterwards to maximize this expansion to obtain the highest expected utility function. As we will see below, direct optimization of the utility function is in most cases computationally unfeasible as the number of assets in the investment universe increases, but with the Taylor expansion based approximation we will show that this problem can be solved efficiently in practice.

In the following, we will consider the general Von Neumann-Morgenstern utility function $U(W_{t+1})$ defined at the end of period wealth W_{t+1}[5]. Let W_t be the initial wealth of

[5] We do not consider here a multi-period investment problem. Ang and Bekaert 2002 have shown that even if portfolio weights may be slightly affected by the investment horizon, the opportunity cost of a myopic

the investor at time t, and let the investor universe consist on n risky assets with stochastic returns $R_{i,t+1}$ and one riskless asset at which he can borrow or lend at a rate r_f. The (random) end of period wealth is given by $W_{t+1} = W_t\left(1 + \left(1 - \omega_t e\right) r_f + \omega_t R_{t+1}\right)$where e is an n-dimensional vector of ones and ω are the different weights on the risky assets. This notation shows that if $e\omega_t = 1$ portfolio weights sum up to one, so that the investor can only invest in risky assets. On the other hand, if ω_i are forced to be positive, it is not possible to short sell assets.

In order to obtain the optimal allocation among the different assets, i.e. the optimal weights ω^*, the investor maximizes the expected utility function over the next period conditional on the information at time t:

$$\max_{\omega_t} E_t\left[U(W_{t+1})\right] = E_t\left[U(W_t\left(1 + \left(1 - \omega_t e\right) r_f + \omega_t R_{t+1}\right))\right] \tag{2.2}$$

The First Order Conditions (FOCs) necessary to solve this optimization problem are calculated identifying the first derivatives of this Equation with respect to $\omega_{i,t}$ to zero:

$$E_t\left[U'(W_{t+1})\left(R_{i,t+1} - r_f\right)\right] = 0, \qquad i = 1, ..., n \tag{2.3}$$

where $U'(W_{t+1})$ is the first derivative of the utility function evaluated in the future wealth $W_{t+1} = W_t\left(1 + \left(1 - \omega_t e\right) r_f + \omega_t R_{t+1}\right)$. In general, for non-normal returns the FOCs in Equation 2.3 do not have a closed-form solution and, in addition, the complexity of the problem increases with the number of assets, as pointed out by Jondeau and Rockinger 2002. It is not difficult to see that Equations 2.3 are very complicated as long as the returns R_t are not multivariate gaussian or the utility function is not quadratic. If we elaborate

strategy is negligible. This result suggests that hedging against unfavorable changes in the investment set does not result in any significant gain.

Equation 2.3 we find:

$$\int \cdots \int_{R_{1,t+1},\dots,R_{n,t+1}} U'(W_t\left(1+\left(1-\omega_t e\right)r_f+\omega_t R_{t+1}\right))\left(R_{i,t+1}-r_f\right) f(R_{1,t+1},...,R_{n,t+1})dR_{1,t+1}...dR_{n,t+1}$$

where $f(R_{1,t+1},...,R_{n,t+1})$ is the multivariate distribution of returns. It is easy to see that solving a set of n Equations like this involving multivariate integrations becomes a daunting task even for a limited number of assets as small as three. In the Markowitz case this problem becomes trivial as the utility function U is quadratic and, therefore, its derivative U' is the identity function. Therefore, in this case this Equation becomes:

$$\int \cdots \int_{R_{1,t+1},\dots,R_{n,t+1}} \left(1+\left(1-\omega_t e\right)r_f+\omega_t R_{t+1}\right)\left(R_{i,t+1}-r_f\right) f(R_{1,t+1},...,R_{n,t+1})dR_{1,t+1}...dR_{n,t+1}$$

and these integrals can be solved exactly.

A solution to the general non-gaussian problem has been provided by Tauchen and Hussey 1991 using Quadrature-Based Methods, but the number of nodes to be used in this approximation happens to increase exponentially with the number of assets getting rapidly unmanageable. Therefore, an approximative method for this calculation becomes necessary and very appropriate.

As we will see, the expansion in Taylor Series of the expected utility function (Equation 2.2), while it simplifies the optimization procedure, provides a natural link between the utility based problem and the moments based one. Actually, it is used as a justification of the use of higher moments in investment evaluation or multi-moment asset pricing models. Expanding the utility function, $U(W_{t+1})$, in Taylor series around the expected wealth $\bar{W}_{t+1} = W_t\left(1+\left(1-\omega_t e\right)r_f+\omega_t m_{t+1}\right)$[6], where $m_{t+1}$is the expected rate of return of the

[6] We have chosen the expected wealth, $\bar{W}_{t+1}$, as the point to expand the Taylor aleatorily. We could also have chosen, for example, the actual wealth W_t but, although the formulas would be different, main results

risky assets $m_{t+1} = E_t\left[R_{t+1}\right]$, we obtain:

$$\begin{aligned} U(W_{t+1}) &= U(\bar{W}_{t+1}) + U'(\bar{W}_{t+1})\left(W_{t+1} - \bar{W}_{t+1}\right) + \frac{1}{2!}U''(\bar{W}_{t+1})\left(W_{t+1} - \bar{W}_{t+1}\right)^2 + \\ & \frac{1}{3!}U'''(\bar{W}_{t+1})\left(W_{t+1} - \bar{W}_{t+1}\right)^3 + \frac{1}{4!}U^{(iv)}(\bar{W}_{t+1})\left(W_{t+1} - \bar{W}_{t+1}\right)^4 + O(W_{t+1}^5) \end{aligned}$$

where $O(W_{t+1}^5)$ is the remainder of the series of order W_{t+1}^5 or higher. Considering that the following equality holds,

$$E[U'(\bar{W}_{t+1})(W_{t+1} - \bar{W}_{t+1})] = U'(\bar{W}_{t+1})E\left[(W_{t+1} - \bar{W}_{t+1})\right] = 0,$$

the expected utility function of Equation 2.2 becomes:

$$\begin{aligned} E_t\left[U(W_{t+1})\right] &= U(\bar{W}_{t+1}) + \frac{1}{2!}U''(\bar{W}_{t+1}) \cdot \mu_2(W_{t+1}) + \\ & \frac{1}{3!}U'''(\bar{W}_{t+1}) \cdot \mu_3(W_{t+1}) + \frac{1}{4!}U^{(iv)}(\bar{W}_{t+1}) \cdot \mu_4(W_{t+1}) + O(W_{t+1}^5) \end{aligned} \tag{2.4}$$

with $\mu_2(W_{t+1})$ being the variance, $\mu_3(W_{t+1})$ being the third central moment, and $\mu_4(W_{t+1})$ the fourth central moment of the end-of-period wealth distribution.

At this stage we will consider various specifications of the utility function in order to investigate the resulting preferences. First, we will consider an investor with a constant relative risk aversion (CRRA) utility function:

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma}$$

where γ measures the investor's constant risk aversion. Using this expression we obtain the following expression for the approximated expected utility function (Equation 2.4), up

---

would not change significantly.

to the fourth order:

$$E_t\left[U(W_{t+1})\right] \approx \frac{\bar{W}_{t+1}^{1-\gamma}}{1-\gamma} - \frac{\gamma}{2}\bar{W}_{t+1}^{-\gamma-1}\mu_2(W_{t+1}) + \frac{\gamma\left(\gamma+1\right)}{6}\bar{W}_{t+1}^{-\gamma-2}\mu_3(W_{t+1}) - \frac{\gamma(\gamma+1)(\gamma+2)}{24}\bar{W}_{t+1}^{-\gamma-3}\mu_4(W_{t+1}). \quad (2.5)$$

Analogously, considering an investor with a constant absolute risk aversion (CARA) utility function:

$$U(W) = -\exp\left(-\eta W\right),$$

with η being the investor's absolute risk aversion, we obtain the following expression for the approximated expected utility function:

$$E_t\left[U(W_{t+1})\right] \approx -\exp\left(-\eta\bar{W}_{t+1}\right)\left(1 + \frac{\eta^2}{2}\mu_2(W_{t+1}) - \frac{\eta^3}{6}\mu_3(W_{t+1}) + \frac{\eta^4}{24}\mu_4(W_{t+1})\right) \quad (2.6)$$

Therefore, both approximations can be considered to obtain the optimal weights, ω_i^*, just maximizing directly Equation 2.5 or 2.6 for the CRRA or the CARA utility function, respectively. It is interesting to note that within this approximation, in order to take optimal investment decisions, agents only care about the first four moments instead of the whole distribution of returns. As a consequence, this approximation will be useful as long as we are able to obtain analytical expressions for the moments of various orders of future wealth. But, as we will see next, these moments can be worked out analytically in the case of the dynamic VCB-MCFD model.

In the following we will show that centered moments of various orders of future wealth, for instance $\mu_2(W_{t+1})$, $\mu_3(W_{t+1})$ and $\mu_4(W_{t+1})$, can be expressed in a very convenient way for the VCB-MCFD model. It is easy to relate the moments of future wealth,

W_{t+1}, with the moments of the profit and loss function as defined in Equation (1.11)[7]:

$$\begin{aligned}\mu_2(W_{t+1}) &= E\left[(\omega_t(R_{t+1}-m_{t+1}))^2\right] = \mu_2(\omega_t R_{t+1}) = \mu_2^P \\ \mu_3(W_{t+1}) &= E\left[(\omega_t(R_{t+1}-m_{t+1}))^3\right] = \mu_3(\omega_t R_{t+1}) = \mu_3^P \\ \mu_4(W_{t+1}) &= E\left[(\omega_t(R_{t+1}-m_{t+1}))^4\right] = \mu_4(\omega_t R_{t+1}) = \mu_4^P\end{aligned}$$

where μ_i^P is the i-th centered moment of the portfolio. To obtain the moments of a portfolio (μ_2^P, μ_3^P and μ_4^P) whose assets follow a VCB-MCFD we will follow a three steps procedure. First, since the vector of unexpected returns is defined as $\varepsilon_{t+1} = R_{t+1} - m_{t+1} = \Sigma_{t+1}^{1/2} z_{t+1}$ (see Equation 2.1a), its first multivariate moments, $\Sigma_{t+1,}$ $M_{3,t+1}$ and $M_{4,t+1}$, are defined as :

$$\begin{aligned}\Sigma_{t+1} &= E_t[\varepsilon_{t+1} \times \varepsilon'_{t+1}] \\ M_{3,t+1} &= E_t\left[\varepsilon_{t+1} \times \varepsilon'_{t+1} \otimes \varepsilon'_{t+1}\right] = \{s_{ijk,t+1}\} \\ M_{4,t+1} &= E_t\left[\varepsilon_{t+1} \times \varepsilon'_{t+1} \otimes \varepsilon'_{t+1} \otimes \varepsilon'_{t+1}\right] = \{\kappa_{ijkl,t+1}\},\end{aligned}$$

can be computed using matrix calculus, instead of numerical integration. Notice that instead of using tensors of third and fourth order to calculate the multivariate third and fourth order moments we define a (n, n^2) matrix as the co-skewness matrix and a (n, n^3) matrix for the co-kurtosis matrix . With these matrices, Σ_{t+1}, $M_{3,t+1}$ and $M_{4,t+1}$, we can calculate

---

[7] As an example, consider the centered third-order moment, $\mu_3(W_{t+1})$:

$$\begin{aligned}\mu_3(W_{t+1}) &= E\left[(W_{t+1} - \bar{W}_{t+1})^3\right] \\ &= E\left[(W_t\,(1 + (1-\omega_t e)\,r_f + \omega_t R_{t+1}) - W_t\,(1 + (1-\omega_t e)\,r_f + \omega_t m_{t+1}))^3\right] \\ &= E\left[(\omega_t(R_{t+1} - m_{t+1}))^3\right] \\ &= \mu_3\,(\omega_t R_{t+1})\end{aligned}$$

the portfolio moments (μ_2^P, μ_3^P and μ_4^P) using matrix calculus with the following Equations:

$$\begin{aligned}
\mu_2^P &= \omega' \times \Sigma_{t+1} \times \omega \\
\mu_3^P &= \omega' \times M_{3,t+1} \times (\omega \otimes \omega) \\
\mu_4^P &= \omega' \times M_{4,t+1} \times ((\omega \otimes \omega) \otimes \omega)
\end{aligned}$$

In the second step we calculate the multivariate moments, Σ_{t+1}, $M_{3,t+1}$ and $M_{4,t+1}$, for a VCB-MCFD model. We obviously have $E_t[\varepsilon_{t+1}] = 0$ and $\Sigma_{t+1} = E_t[\varepsilon_{t+1}\varepsilon_{t+1}]$ and we denote $\Sigma_{t+1}^{1/2} = (w_{ij,t+1})$, $i, j = 1, ..., n$ the Cholesky decomposition of the covariance matrix of returns. Then, in order to calculate $M_{3,t+1}$ and $M_{4,t+1}$ (i.e. $\kappa_{ijkl,t+1}$ and $s_{ijk,t+1}$), we use the results of Proposition 10 in Section 1.3:

$$\begin{aligned}
s_{ijk} &= \sum_r w_{ir}w_{jr}w_{kr}E\left[z_r^3\right] = \sum_r w_{ir}w_{jr}w_{kr}\mu_{3,r} \\
\kappa_{ijkl} &= \sum_{r,s,t,u} w_{ir}w_{js}w_{ks}w_{lu}E\left[z_r z_s z_t z_u\right] \\
&= \sum_r w_{ir}w_{jr}w_{kr}w_{lr}\mu_{4,r} + \\
&\quad \sum_r \sum_{\substack{s \\ s \neq r}} \left(\begin{array}{c} w_{ir}w_{jr}w_{ks}w_{ls} + w_{ir}w_{js}w_{kr}w_{ls} \\ +w_{ir}w_{js}w_{ks}w_{lr} \end{array} \right)
\end{aligned}$$

Finally, in the third step we calculate the four moments of a univariate CFD distribution (denoted by $\mu_{i,r}$, the i-th moment of the r -th asset) as given by Equations 1.11 in Chapter 1.

Therefore, once we have calculated the Cholesky decomposition of the variance-covariance matrix, Σ_{t+1}, and the univariate moments of the distribution, we can easily obtain the matrices $M_{3,t+1}$ and $M_{4,t+1}$, and afterwards the portfolio moments, μ_2^P, μ_3^P and

μ_4^P. It is interesting to note that considering this procedure it becomes straightforward to calculate the expected utility for a given vector of weights ω.

Summarizing, to solve the optimization problem given by Equation 2.2, where the utility function is approximated by a Taylor series expansion, we will maximize the following expression for the CRRA utility function:

$$\begin{aligned}\max_{\omega_t} \quad & \left(\frac{\bar{W}_{t+1}^{1-\gamma}}{1-\gamma} - \frac{\gamma}{2}\bar{W}_{t+1}^{-\gamma-1}(\omega_t' \times \Sigma_{t+1} \times \omega_t) - \right. \\ & -\frac{\gamma(\gamma+1)}{6}\bar{W}_{t+1}^{-\gamma-2}\left(\omega_t' \times M_{3,t+1} \times (\omega_t \otimes \omega_t)\right) \\ & \left. + \frac{\gamma(\gamma+1)(\gamma+2)}{24}\bar{W}_{t+1}^{-\gamma-3}\left(\omega_t' \times M_{4,t+1} \times ((\omega_t \otimes \omega_t) \otimes \omega_t)\right) \right)\end{aligned}$$

and the following expression for the CARA utility function:

$$\max_{\omega_t} \quad \left(-e^{\left(-\eta \bar{W}_{t+1}\right)} \left(\begin{array}{c} 1 + \frac{\eta^2}{2}(\omega_t' \times \Sigma_{t+1} \times \omega_t) - \\ -\frac{\eta^3}{6}\mu\left(\omega_t' \times M_{3,t+1} \times (\omega_t \otimes \omega_t)\right) \\ +\frac{\eta^4}{24}\left(\omega_t' \times M_{4,t+1} \times ((\omega_t \otimes \omega_t) \otimes \omega_t)\right) \end{array} \right) \right)$$

with $\bar{W}_{t+1} = W_t\left(1 + (1-\omega_t e)\, r_f + \omega_t m_{t+1}\right)$.

2.1.4 Unconditional Investment Under Non-normality

In order to provide evidence on how the CRRA and the CARA utility functions incorporate information on higher moments and determine if higher moments are important when choosing optimal portfolios, we perform two kinds of asset allocation exercises. In the first one we consider an unconditional investment, i.e. the moments of the distribution are constant, and in the second one (see next Section) the moments of the distribution are time varying and, therefore, investors will take different investment options at each date. Moreover, we will analyze two different optimization problems in each case, either short sales and borrowing are allowed, so that no constraints are considered on the portfolio weights, or we assume that the investor has strong constraints on the portfolio strategy, so that portfolio weights are positive and sum to one. The latter case implies that short sales are not possible and that the entire wealth must be invested in risky assets.

In order to analyze the problem we consider the data base consisting on the five series of market indexes that we have used in Chapter 1. As we have seen in Section 1.4, the data base consists on weekly returns (from Wednesday to Wednesday) for dollar denominated stock indexes for the main geographical areas: North America, Japan, Europe, Emerging Markets and Eastern Europe Emerging Markets, represented by the Standard and Poor's 500 Index (S&P), the Nikkei-225 Stock Average (NKI), the Dow Jones EURO STOXX (STX), MSCI Emerging Markets Index (EM) and the MSCI Eastern Europe Emerging Market Index (EME). This data set consists of total logarithmic return indexes from January 4 of 1995 to March 23 of 2005, making in total 519 observations[8].

To estimate the moments necessary to calculate the optimal portfolios, we fit the static VCB-MCFD model to this data base. In Table 1.21 of Chapter 1 we report estimated parameters for the five series. Although we have already commented this estimation in the first Chapter we would like to emphasize here that for the five series we find that the a_3 parameter, which incorporates kurtosis, are significantly different from zero, indicating that the unconditional distribution of returns presents a non-gaussian behavior.

In order to obtain the optimal portfolios and analyze if higher moments are important we proceed as follows. First, we approximate the utility function using a Taylor expansion up to order i, with $i = 2$ (second order approximation), $i = 3$ (third order) and $i = 4$ (fourth

---

8 Consult Section 1.4 for more descriptive details on these time series.

order). This corresponds to cases where we incorporate information on volatility ($i = 2$), volatility and skewness ($i = 3$), and volatility, skewness and kurtosis ($i = 4$). Then, to obtain the optimal portfolio weights we solve the maximization problem of the approximated utility CRRA function (Equation 2.5) or CARA function (Equation 2.6), considering that the riskless asset annual rate is 2%. For every set of weights we calculate the first four moments of the optimal portfolio: mean and standard deviation (annualized), skewness and kurtosis (both standardized) to analyze if different approximations drive substantially different portfolio moments. In addition, to analyze the importance of higher order moments we analyze different indicators, as the Sharpe Ratio, the percentage invested in risky assets (only for the unrestricted optimizations) and the opportunity cost (Simaan 1993), defined as the percent the investor would be just willing to pay out of the portfolio for the privilege of choosing the true Expected Utility maximizing portfolio rather than being confined to the mean-variance "second best". The Sharpe Ratio is designed to measure the risk-adjusted performance and it is calculated by subtracting the risk-free rate from the rate of return of the portfolio and dividing the result by the standard deviation associated to portfolio returns:

$$Sharpe = \frac{\mu_p - r_f}{\sigma_P} \tag{2.7}$$

As it can be seen from its definition, the Sharpe Ratio considers only gaussian characteristics as it associates risk to volatility and, therefore, it might not be well suited to evaluate investments in a non gaussian world and we do not expect this ratio to change significantly between different Taylor approximations. In contrast, as we are interested in comparing if higher order moments are of practical interest compared to the traditional mean-variance approach, *the opportunity cost* becomes a more appropriate measure. In order to calculate the opportunity cost we suppose that the most optimal portfolio is the one obtained with the fourth order approximation. For instance, we could take as the best optimal portfolio the one obtained by maximizing the whole expected utility function of Equation 2.2 but, as we have mentioned before, this task is unmanageable for five assets. Consequently, we will take as the best approximation to the maximization of expected utility the fourth order approximation. Therefore, the opportunity cost, OC, is implicitly defined as:

$$E_t[U(1+\mu_p^*)] = E_t[U(1+\mu_p+OC)] \tag{2.8}$$

where μ_p^* is the return of the fourth order optimal portfolio and μ_p is the return corresponding to the suboptimal one (i.e. the second or third order approximation).

Tables 2.1 to 2.4 report results for optimally selected portfolios for several values of the risk aversion parameter (γ and η ranges between 0.05 and 0.9 which covers values considered in the literature)[9]. Table 2.1 corresponds to the case CRRA function without constraints on weights while Table 2.2 is associated to the case where the investor is forced to invest in the risky assets only under the no-shortsale constraint. Tables 2.3-2.4 report the respective results for the CARA function. The results of Tables 2.1 to 2.4 allow us to derive the following results:

[9] Which range to use for the risk aversion parameters γ and η is actually an open question as it depends on the investors' initial wealth. For example, Rabin 2000 and Rabin and Thaler 2001 argue that, under expected utility theory, the levels of risk aversion we observe over modest stakes would imply absurdly high levels of risk aversion over large stakes. Therefore, we can observe in the literature values for the risk aversion parameters ranging from single digit values to four digit values (Schechter 2005). Nonetheless, Holt and Laury 2002 finds average coefficients of approximately 0.4 when defining utility over gains, not wealth. In this work we take this position and suppose an initial wealth W_0 equal to one and therefore we define the utility in terms of percent gains. As a consequence, $\eta = \gamma$, i.e. relative and absolute risk aversion coincide, and the reasonable range lies between 0.05 and 0.9. Considering the optimal portfolios we obtain with this range (see Tables 2.1 to 2.4) we can conclude that it is sufficiently adequate to model observed risk aversion patterns.

Table 2.1. Results for the optimization under the CRRA utility hypothesis without investment restrictions. We report the optimal weights approximating the utility function using a Taylor expansion up to order i, with $i = 2$ (second order), $i = 3$ (third order), $i = 4$ (fourth order). For every set of weights, ω_i, we calculate the first four moments of the optimal portfolio: mean, μ_p and standard deviation (annualized), σ_p, skewness, s_p^3, and kurtosis κ_p^4. In addition, we report the Sharpe Ratio (SR) as defined by Equation 2.7, the opportunity cost as defined by Equation 2.8 and the percentage invested in risky assets (TW).

	Portfolio Weights					Portfolio Moments				Portfolio Statistics		
γ	ω_1	ω_2	ω_3	ω_4	ω_5	μ_p	σ_p	s_p^3	κ_p^4	**SR**	**OC**	**TW**
						Second order Approximation						
0.05	1.2611	-0.7864	0.3047	-0.6289	0.1699	19.7095	23.6828	-0.1588	3.8033	0.1037	0.5558	0.3203
0.1	0.4978	-0.3103	0.1212	-0.2485	0.0666	8.9937	9.3526	-0.1593	3.8052	0.1037	0.0467	0.1268
0.2	0.2262	-0.1409	0.0551	-0.1131	0.0302	5.1771	4.2488	-0.1590	3.8046	0.1037	0.0038	0.0575
0.4	0.1082	-0.0674	0.0263	-0.0541	0.0144	3.5202	2.0329	-0.1590	3.8045	0.1037	0.0003	0.0275
0.6	0.0708	-0.0443	0.0177	-0.0355	0.0093	2.9975	1.3340	-0.1604	3.8078	0.1037	0.0000	0.0180
0.8	0.0526	-0.0329	0.0131	-0.0264	0.0070	2.7407	0.9905	-0.1598	3.8061	0.1037	0.0000	0.0134
0.9	0.0468	-0.0292	0.0115	-0.0234	0.0062	2.6578	0.8797	-0.1595	3.8054	0.1037	0.0000	0.0119
						Third order Approximation						
0.05	1.0718	-0.6890	0.2104	-0.6776	0.1643	17.0763	20.4789	-0.0725	3.6176	0.1021	0.3399	0.0799
0.1	0.4589	-0.2943	0.0942	-0.2478	0.0649	8.4286	8.6273	-0.1179	3.6884	0.1033	0.0289	0.0760
0.2	0.2168	-0.1372	0.0472	-0.1121	0.0297	5.0330	4.0604	-0.1358	3.7346	0.1036	0.0024	0.0444
0.4	0.1054	-0.0663	0.0241	-0.0537	0.0143	3.4778	1.9770	-0.1455	3.7627	0.1037	0.0001	0.0238
0.6	0.0694	-0.0440	0.0163	-0.0353	0.0096	2.9787	1.3091	-0.1484	3.7683	0.1037	0.0000	0.0161
0.8	0.0513	-0.0326	0.0129	-0.0263	0.0069	2.7280	0.9736	-0.1533	3.7828	0.1037	0.0000	0.0123
0.9	0.0456	-0.0289	0.0114	-0.0234	0.0062	2.6468	0.8651	-0.1535	3.7831	0.1037	0.0000	0.0109
						Fourth order Approximation						
0.05	0.4199	-0.2785	0.0923	-0.2229	0.0606	7.9856	8.0288	-0.1219	3.6886	0.1034	-	0.0713
0.1	0.2760	-0.1809	0.0609	-0.1450	0.0395	5.9182	5.2512	-0.1272	3.7036	0.1035	-	0.0505
0.2	0.1682	-0.1089	0.0380	-0.0872	0.0237	4.3807	3.1876	-0.1351	3.7263	0.1036	-	0.0338
0.4	0.0942	-0.0600	0.0219	-0.0482	0.0129	3.3277	1.7763	-0.1437	3.7541	0.1036	-	0.0208
0.6	0.0649	-0.0412	0.0154	-0.0330	0.0088	2.9154	1.2244	-0.1479	3.7671	0.1037	-	0.0149
0.8	0.0490	-0.0315	0.0119	-0.0250	0.0069	2.6963	0.9315	-0.1485	3.7639	0.1037	-	0.0113
0.9	0.0439	-0.0279	0.0105	-0.0224	0.0061	2.6209	0.8305	-0.1490	3.7681	0.1037	-	0.0102

Table 2.2. Results for the optimization with the CRRA utility function with investment restrictions. The legend is the same as in Table 2.1

	Portfolio Weights					Portfolio Moments				Portfolio Statistics	
γ	ω_1	ω_2	ω_3	ω_4	ω_5	μ_p	σ_p	s_p^3	κ_p^4	**SR**	**OC**
						Second order Approximation					
0.05	0.9915	0.0000	0.0000	0.0000	0.0085	9.3295	17.0374	-0.3027	5.0226	0.0597	0.0309
0.1	0.9439	0.0000	0.0000	0.0000	0.0561	9.1579	16.8328	-0.2997	4.9442	0.0590	0.0442
0.2	0.8982	0.0100	0.0158	0.0000	0.0761	8.9164	16.6913	-0.3021	4.8909	0.0575	0.0516
0.4	0.8144	0.0886	0.0000	0.0270	0.0700	7.6050	16.1152	-0.2756	4.7090	0.0482	0.0239
0.6	0.7921	0.1033	0.0000	0.0374	0.0671	7.3183	16.0146	-0.2708	4.6634	0.0461	0.0170
0.8	0.7903	0.1045	0.0000	0.0383	0.0669	7.2946	16.0067	-0.2704	4.6596	0.0459	0.0146
0.9	0.7928	0.1029	0.0000	0.0372	0.0672	7.3264	16.0173	-0.2710	4.6647	0.0461	0.0140
						Third order Approximation					
0.05	0.9638	0.0000	0.0000	0.0000	0.0360	9.2297	16.9042	-0.3007	4.9786	0.0593	0.0238
0.1	0.9314	0.0000	0.0000	0.0000	0.0686	9.1128	16.7986	-0.2993	4.9225	0.0587	0.0417
0.2	0.8047	0.1267	0.0000	0.0000	0.0685	7.2704	16.0118	-0.2596	4.6605	0.0456	0.0166
0.4	0.7217	0.2142	0.0000	0.0000	0.0641	6.0150	15.7350	-0.2187	4.4371	0.0354	0.0053
0.6	0.6958	0.2423	0.0000	0.0000	0.0619	5.6138	15.6940	-0.2037	4.3660	0.0319	0.0044
0.8	0.6840	0.2557	0.0000	0.0000	0.0603	5.4245	15.6831	-0.1963	4.3337	0.0303	0.0047
0.9	0.6805	0.2599	0.0000	0.0000	0.0596	5.3664	15.6810	-0.1940	4.3240	0.0298	0.0050
						Fourth order Approximation					
0.05	0.7360	0.1485	0.0000	0.0421	0.0735	6.6021	15.8095	-0.2529	4.5349	0.0404	-
0.1	0.6756	0.1939	0.0000	0.0654	0.0650	5.7860	15.6460	-0.2333	4.3968	0.0336	-
0.2	0.6502	0.2126	0.0000	0.0755	0.0616	5.4459	15.6019	-0.2248	4.3406	0.0306	-
0.4	0.6439	0.2167	0.0000	0.0784	0.0609	5.3657	15.5934	-0.2229	4.3276	0.0299	-
0.6	0.6471	0.2138	0.0000	0.0776	0.0614	5.4123	15.5980	-0.2244	4.3351	0.0303	-
0.8	0.6525	0.2092	0.0000	0.0759	0.0623	5.4898	15.6063	-0.2266	4.3478	0.0310	-
0.9	0.6556	0.2066	0.0000	0.0749	0.0628	5.5336	15.6113	-0.2278	4.3550	0.0314	-

Table 2.3. Results for the optimization with the CARA utility function without investment restrictions. The legend is the same as in Table 2.1

	Portfolio Weights					Portfolio Moments				Portfolio Statistics		
η	ω_1	ω_2	ω_3	ω_4	ω_5	μ_p	σ_p	s_p^3	κ_p^4	**SR**	**OC**	**TW**
						Second order Approximation						
0.05	0.8763	-0.4613	0.2024	-0.3713	0.1159	13.5932	15.6945	-0.2279	4.1205	0.1024	0.0000	0.3621
0.1	0.4427	-0.2295	0.0992	-0.1876	0.0576	7.8100	7.8671	-0.2275	4.1264	0.1024	0.0000	0.1825
0.2	0.2225	-0.1161	0.0483	-0.0935	0.0283	4.9136	3.9403	-0.2227	4.1059	0.1025	0.0000	0.0896
0.4	0.1075	-0.0586	0.0254	-0.0442	0.0140	3.4381	1.9455	-0.2257	4.0987	0.1025	0.0000	0.0441
0.6	0.0718	-0.0391	0.0168	-0.0294	0.0094	2.9594	1.2979	-0.2255	4.0989	0.1025	0.0000	0.0295
0.8	0.0538	-0.0292	0.0127	-0.0223	0.0070	2.7195	0.9734	-0.2260	4.1015	0.1025	0.0000	0.0221
0.9	0.0485	-0.0257	0.0110	-0.0205	0.0063	2.6417	0.8677	-0.2237	4.1015	0.1026	0.0000	0.0196
						Third order Approximation						
0.05	0.8681	-0.4577	0.1933	-0.3634	0.1119	13.4339	15.4657	-0.2240	4.1063	0.1025	0.0001	0.3522
0.1	0.4348	-0.2294	0.0967	-0.1812	0.0554	7.7246	7.7434	-0.2241	4.1069	0.1025	0.0000	0.1764
0.2	0.2185	-0.1148	0.0482	-0.0917	0.0282	4.8726	3.8852	-0.2235	4.1056	0.1025	0.0000	0.0884
0.4	0.1101	-0.0569	0.0234	-0.0463	0.0141	3.4355	1.9419	-0.2228	4.1102	0.1025	0.0000	0.0444
0.6	0.0700	-0.0395	0.0173	-0.0309	0.0107	2.9574	1.2920	-0.2179	4.0449	0.1028	0.0000	0.0276
0.8	0.0550	-0.0283	0.0118	-0.0235	0.0072	2.7182	0.9715	-0.2230	4.1111	0.1025	0.0000	0.0222
0.9	0.0498	-0.0252	0.0098	-0.0206	0.0063	2.6400	0.8662	-0.2208	4.1140	0.1025	0.0000	0.0200
						Fourth order Approximation						
0.05	0.8619	-0.4553	0.1917	-0.3597	0.1104	13.3543	15.3563	-0.2236	4.1042	0.1025	-	0.3490
0.1	0.4413	-0.2261	0.0892	-0.1836	0.0563	7.7099	7.7254	-0.2208	4.1087	0.1025	-	0.1771
0.2	0.2178	-0.1136	0.0472	-0.0918	0.0281	4.8525	3.8578	-0.2226	4.1053	0.1025	-	0.0878
0.4	0.1096	-0.0566	0.0229	-0.0460	0.0141	3.4263	1.9293	-0.2217	4.1072	0.1025	-	0.0441
0.6	0.0731	-0.0377	0.0152	-0.0306	0.0094	2.9504	1.2858	-0.2218	4.1089	0.1025	-	0.0294
0.8	0.0549	-0.0284	0.0117	-0.0233	0.0072	2.7162	0.9686	-0.2216	4.1046	0.1025	-	0.0220
0.9	0.0480	-0.0253	0.0107	-0.0201	0.0062	2.6328	0.8558	-0.2236	4.1041	0.1025	-	0.0194

Table 2.4. Results for the optimization with the CARA utility function with investment restrictions. The legend is the same as in Table 2.1

	Portfolio Weights					**Portfolio Moments**				**Portfolio Statistics**	
η	ω_1	ω_2	ω_3	ω_4	ω_5	μ_p	σ_p	s_p^3	κ_p^4	**SR**	**OC**
					Second order Approximation						
0.05	0.9675	0.0000	0.0000	0.0000	0.0325	9.2429	16.9196	-0.3009	4.9848	0.0594	0.0000
0.1	0.9193	0.0000	0.0094	0.0000	0.0714	9.0883	16.7827	-0.3022	4.9157	0.0586	0.0000
0.2	0.7975	0.0998	0.0000	0.0349	0.0678	7.3877	16.0381	-0.2720	4.6746	0.0586	0.0000
0.4	0.7156	0.1539	0.0000	0.0733	0.0572	6.3340	15.7414	-0.2527	4.4994	0.0382	0.0140
0.6	0.7010	0.1635	0.0000	0.0801	0.0553	6.1461	15.7023	-0.2490	4.4677	0.0366	0.0288
0.8	0.7034	0.1619	0.0000	0.0790	0.0556	6.1768	15.7084	-0.2496	4.4729	0.0369	0.0392
0.9	0.7074	0.1593	0.0000	0.0771	0.0561	6.2288	15.7190	-0.2506	4.4817	0.0373	0.0424
					Third order Approximation						
0.05	0.9657	0.0000	0.0000	0.0000	0.0343	9.2364	16.9119	-0.3008	4.9818	0.0593	0.0000
0.1	0.9261	0.0000	0.0000	0.0000	0.0739	9.0938	16.7866	-0.2993	4.9134	0.0586	0.0000
0.2	0.7791	0.1262	0.0000	0.0273	0.0674	7.0650	15.9364	-0.2611	4.6222	0.0441	0.0000
0.4	0.6899	0.2036	0.0000	0.0483	0.0582	5.8061	15.6577	-0.2266	4.4042	0.0337	0.0045
0.6	0.6656	0.2353	0.0000	0.0420	0.0571	5.3994	15.6229	-0.2095	4.3327	0.0302	0.0084
0.8	0.6567	0.2547	0.0000	0.0309	0.0577	5.2033	15.6260	-0.1982	4.2980	0.0284	0.0106
0.9	0.6547	0.2622	0.0000	0.0249	0.0582	5.1404	15.6325	-0.1937	4.2868	0.0279	0.0113
					Fourth order Approximation						
0.05	0.9645	0.0000	0.0000	0.0000	0.0355	9.2322	16.9070	-0.3007	4.9798	0.0593	-
0.1	0.9235	0.0000	0.0000	0.0000	0.0765	9.0841	16.7811	-0.2993	4.9089	0.0585	-
0.2	0.7436	0.1486	0.0000	0.0431	0.0647	6.6235	15.8127	-0.2526	4.5457	0.0405	-
0.4	0.6306	0.2364	0.0000	0.0791	0.0539	5.0997	15.5748	-0.2123	4.2866	0.0276	-
0.6	0.5949	0.2645	0.0000	0.0891	0.0515	4.6207	15.5611	-0.1981	4.2122	0.0234	-
0.8	0.5820	0.2732	0.0000	0.0938	0.0509	4.4579	15.5622	-0.1938	4.1885	0.0219	-
0.9	0.5794	0.2743	0.0000	0.0953	0.0510	4.4305	15.5622	-0.1934	4.1846	0.0217	-

1. As the risk aversion parameter increases we obtain, as expected, that for all approximation orders, in restricted or unrestricted optimizations, the investor opts for a portfolio with lower expected mean and lower expected standard deviation, so that the resulting portfolio becomes more diversified. In the same way, we find that in the unrestricted cases the exposition to risky assets decreases as the risk aversion parameter increases.

2. The gain in terms of opportunity cost of taking higher order terms in the utility expansion is relevant, specially in the unrestricted cases where it can be as high as 55% for the second order approximation in the CRRA function with γ equal to 0.05. The weights in this case are (1.2611, -0.7864, 0.3047, -0.6289) for the second order approximation and (0.4199, -0.2785, 0.0923, -0.2229, 0.0606) for the fourth order, which are substantially different.

3. On the other hand, we find that in some cases, in CARA without restrictions, the opportunity cost is almost zero, while for the CRRA and CARA with restrictions we find a cost of opportunity of around 5%. For example, the weights for the CRRA function with restricted optimization and γ equal to 0.05 are (0.8982, 0.0100, 0.0158, 0.0000, 0.0761) for the second order approximation and (0.6502, 0.2126, 0.0000, 0.0755, 0.0616) for the fourth order, which are also substantially different.

4. In contrast to previous research (Jondeau and Rockinger 2005, Y. Aït-Sahalia and Brandt 2001) we do not find that for small values of the parameter of risk aversion, higher moments do not matter. In particular, we find both cases: in the CRRA with

restrictions we find that for $\gamma = 2$ the OC is equal to 3%, while in the CARA it is 0%. Therefore, we do not find an increasing relationship between OC and risk-aversion.

5. The cost of opportunity is always higher for the second order approximation compared to the third order one.

6. Comparing how portfolio moments are modified when higher moments are introduced also provides interesting results. In general we can observe that the values of kurtosis are in general higher and skewness more negative for the second order approximation than for the fourth order. In addition, when the Taylor expansion is performed up to order 2, portfolio variance decreases when risk aversion increases. However, when a concern for higher moments is allowed, this is not necessarily the case. For large values of risk aversion ($\gamma > 0.6$) a very slight increase in the variance is admitted to obtain an increase in skewness and a decrease in kurtosis.

7. In the unrestricted cases the Sharpe Ratio does not depend on the approximation order nor on the risk aversion. However, in the restricted cases we find that in general it gets smaller with increasing risk aversion or approximation order. The dependence on risk aversion can be explained due to the fact that as we get more risk averse our investment set gets smaller and, therefore, asset allocation will be less optimal, while the dependence on order approximation can be explained given that the Sharpe Ratio only measures risk as volatility. However, if we care about higher-order moments, (reducing the importance on volatility), we will be constructing a less optimal in a Sharpe Ratio sense portfolio.

8. Incorporating higher order terms in the expansion induces an "effective increase" in the risk-aversion coefficient, as portfolios within higher approximations, maintaining all other variables constant, have a lower expected mean and variance compared to the lower order ones. This is so because considering skewness or kurtosis are taking into account that the investor will be aware of higher order moments, and he will be willing to trade some expected return to increase expected skewness or lower expected kurtosis. In fact, for restricted portfolios, the expected kurtosis coefficient is lower for higher order approximations and the skewness coefficient is bigger in the third order approximation than in the second order.

Summarizing, we find that in some cases *but not always* higher order moments are relevant in the asset allocation process and that these cases depend on the specific utility function used.

2.1.5 Conditional Investment Under Non-normality

We now turn to the optimal allocation when returns have time-varying moments. For each week of the sample, we use the dynamic VCB-MCFD model described in Section 2.1.2 to predict the first four moments and co-moments of market returns. As mentioned before, with this model the conditional mean is assumed to follow an AR(1) model, while the conditional variances are given by a GARCH(1,1) model. Conditional correlations vary over time according to a DCC model and the non-gaussian behaviour is captured through a VCB-MCFD.

In Table 1.28 we report the estimations for the set of parameters of the DCC model for these time series. First, we notice that all series except the E-STOXX present a non-zero AR(1) dynamic for the mean, being the coefficient positive for the S&P500 and Nikkei-225 and negative for the Emerging Markets Index and, that all series present volatility clustering and persistence effects, as can be seen from the GARCH parameters, p_i and q_i. Second, the DCC parameters, which measure the dependence of the correlation matrix on past returns and past correlations, are also significant at a 99% confidence interval[10]. And third, the coefficients $a_{i,3}$ of the Multivariate Cornish-Fisher Density, which capture the kurtosis, are significant for all series, meaning that the conditional distributions after accounting for the first and second order dynamics still present fat-tailedness, while the coefficients $a_{i,2}$, which are associated to the skewness are significant for all series except for the Nikkei-225.

With these estimations we maximize the approximated expected utility function for the two alternative approximations (CARA and CRRA) given by Equations 2.5-2.6 and the optimal weights for each approximation will be obtained. In addition, we also calculate for each week of the sample the same indicators as in the previous Section[11]. As we have one value of these statistics for each time period we report the mean values of each indicator. In addition, we present the opportunity cost of using a sub-optimal forecasting model, namely, the unconditional framework analyzed in the latter Section compared to the one using conditional estimates. In this way, we test both shortcomings that can make the investor take bad decisions: the quadratic approximation and the hypothesis of i.i.d. returns. In par-

[10] These results are in agreement with those found in the literature (Engle and Sheppard 2001, Bollerslev 1986 and Jondeau and Rockinger 2005).

[11] We do not report the Sharpe Ratio given that it does not add more information to the one obtained in the unconditional investment Section.

ticular, from the different combinations that we can consider, we analyze the differences of optimal allocations between a conditional and an unconditional investment framework using the fourth order approximation.

In Tables 2.5-2.6 we report the results for optimally selected portfolios for several values of the risk aversion parameter. Table 2.5 corresponds to the CARA function without constraints on weights where we consider that the riskless asset rate is 2% (annualized), while Table 2.6 is associated to the case where the investor is forced to invest in the risky assets only under the no-shortsale constraint. In addition, Tables 2.7 and 2.8 contain the results for the CRRA utility function in the unconstrained and constrained cases, respectively. In this case, from the results found in Tables 2.5-2.8 we can derive the following conclusions:

1. As in the unconditional setting as the risk aversion increases the investor opts for a portfolio with lower expected mean and lower expected standard deviation. We also find that in the unrestricted case the mean and variance values are in general higher than in the conditional case, given that with reallocations the agent with no restrictions is able to capture more aggressively its views about markets.

2. In this case, the gain in terms of opportunity cost of taking higher order terms in the utility expansion is relevant, specially in the unrestricted cases, where it can be as high as 60.78% for the second order approximation in the CRRA function with γ equal to 0.05. The mean weights in this case are (2.5489, -1.5112, 0.6796, -1.2452, 0.41) for

Table 2.5. Results for the optimization with the CARA utility function without investment restrictions and time-varying moments. Legend is the same as in Table 2.1. The opportunity cost of using an unconditional framework appears in the OC column of the fourth order approximation.

	Portfolio Weights					Portfolio Moments				Portfolio Statistics	
γ	$\bar{\omega}_1$	$\bar{\omega}_2$	$\bar{\omega}_3$	$\bar{\omega}_4$	$\bar{\omega}_5$	$\bar{\mu}_p$	$\bar{\sigma}_p$	$\bar{s}_p^3$	$\bar{\kappa}_p^4$	**OC**	**TW**
						Second order Approximation					
0.05	1.5311	-0.6290	1.5455	-1.1539	0.3019	86.0913	37.7800	-0.1368	3.6398	0.0452	1.5956
0.1	0.7666	-0.3152	0.7829	-0.5800	0.1508	44.2910	18.9663	-0.1367	3.6393	0.0239	0.8051
0.2	0.3822	-0.1578	0.3887	-0.2879	0.0753	23.0889	9.4567	-0.1367	3.6377	0.0115	0.4004
0.4	0.1899	-0.0794	0.1955	-0.1449	0.0381	12.5543	4.7242	-0.1348	3.6359	0.0059	0.1992
0.6	0.1260	-0.0526	0.1264	-0.0952	0.0249	8.9274	3.1016	-0.1340	3.6323	0.0034	0.1295
0.8	0.0940	-0.0390	0.0940	-0.0714	0.0188	7.1817	2.3179	-0.1331	3.6335	0.0025	0.0962
0.9	0.0843	-0.0347	0.0826	-0.0632	0.0168	6.0355	1.9037	-0.1277	3.6283	0.0021	0.0858
						Third order Approximation					
0.05	1.4645	-0.6225	1.5216	-1.1831	0.3005	84.8517	37.2239	-0.1265	3.6346	0.0174	1.4810
0.1	0.7331	-0.3116	0.7656	-0.5939	0.1503	43.5661	18.6537	-0.1264	3.6349	0.0087	0.7435
0.2	0.3669	-0.1553	0.3815	-0.2969	0.0755	22.7810	9.3246	-0.1263	3.6345	0.0043	0.3716
0.4	0.1831	-0.0777	0.1906	-0.1484	0.0377	12.3847	4.6616	-0.1262	3.6354	0.0028	0.1853
0.6	0.1223	-0.0520	0.1263	-0.0985	0.0251	8.9146	3.1045	-0.1256	3.6354	0.0021	0.1232
0.8	0.0916	-0.0391	0.0945	-0.0735	-0.0735	7.1747	2.3240	-0.1258	3.6352	0.0016	0.0921
0.9	0.0811	-0.0345	0.0825	-0.0649	0.0168	6.5558	2.0525	-0.1253	3.6338	0.0337	0.0810
						Fourth order Approximation					
0.05	1.3953	-0.5906	1.3036	-1.0058	0.2721	74.7933	34.3337	-0.1275	3.6291	0.5494	1.3746
0.1	0.6976	-0.2948	0.6490	-0.5017	0.1360	38.3539	17.1375	-0.1276	3.6294	0.2747	0.6861
0.2	0.3508	-0.1472	0.3232	-0.2512	0.0681	20.2039	8.5852	-0.1279	3.6298	0.1373	0.3437
0.4	0.1749	-0.0736	0.1612	-0.1254	0.0341	11.0823	4.2815	-0.1275	3.6286	0.0687	0.1712
0.6	0.1163	-0.0492	0.1077	-0.0834	0.0226	8.0540	2.8553	0.1271	3.6295	0.0458	0.1141
0.8	0.0878	-0.0369	0.0809	-0.0629	0.0170	6.5493	2.1458	-0.1278	3.6288	0.0343	0.0860
0.9	0.0779	-0.0328	0.0715	-0.0556	0.0151	6.5958	2.0627	-0.1332	3.6335	0.0305	0.0761

Table 2.6. Results for the optimization with the CARA utility function with investment restrictions and time-varying moments. The legend is the same as in Table 2.1

	Portfolio Weights					Portfolio Moments				Portfolio Statistics
γ	$\bar{\omega}_1$	$\bar{\omega}_2$	$\bar{\omega}_3$	$\bar{\omega}_4$	$\bar{\omega}_5$	$\bar{\mu}_p$	$\bar{\sigma}_p$	$\bar{s}_p^3$	$\bar{\kappa}_p^4$	**OC**
					Second order Approximation					
0.05	0.5469	0.0061	0.3530	0.0299	0.0641	12.7993	12.5990	-0.2992	4.2554	0.0000
0.1	0.5468	0.0061	0.3530	0.0301	0.0640	11.7600	11.5475	-0.3028	4.0743	0.0001
0.2	0.5464	0.0056	0.3525	0.0318	0.0637	10.7250	10.9000	-0.3080	3.9576	0.0002
0.4	0.5420	0.0040	0.3531	0.0373	0.0636	9.8212	10.5751	-0.3117	3.9084	0.0019
0.6	0.5174	0.0019	0.3579	0.0569	0.0659	9.4467	10.4835	-0.3120	3.8957	0.0050
0.8	0.4743	0.0020	0.3531	0.0879	0.0827	9.3097	10.4547	-0.3119	3.8919	0.0083
0.9	0.4056	0.0030	0.3526	0.1212	0.1176	9.2893	10.4506	-0.3119	3.8914	0.0098
					Third order Approximation					
0.05	0.4818	0.0129	0.4059	0.0251	0.0743	12.7786	12.5744	-0.2988	4.2514	0.0000
0.1	0.4881	0.0121	0.4008	0.0258	0.0732	11.7176	11.5150	-0.3014	4.0671	0.0000
0.2	0.5014	0.0099	0.3894	0.0281	0.0711	10.6738	10.8780	-0.3042	3.9499	0.0001
0.4	0.5136	0.0065	0.3770	0.0345	0.0684	9.6894	10.5466	-0.3009	3.8917	0.0008
0.6	0.5076	0.0023	0.3673	0.0551	0.0676	9.2528	10.4556	-0.2949	3.8709	0.0021
0.8	0.4719	0.0021	0.3571	0.0861	0.0828	9.0776	10.4310	-0.2894	3.8603	0.0038
0.9	0.4062	0.0030	0.3534	0.1203	0.1171	9.0404	10.4289	-0.2869	3.8567	0.0047
					Fourth order Approximation					
0.05	0.5010	0.0189	0.3880	0.0190	0.0731	12.7625	12.5559	-0.2988	4.2485	0.0750
0.1	0.5028	0.0177	0.3873	0.0197	0.0726	11.6762	11.4838	-0.3014	4.0602	0.0685
0.2	0.5092	0.0138	0.3838	0.0221	0.0710	10.5714	10.8349	-0.3044	3.9407	0.1158
0.4	0.5183	0.0087	0.3755	0.0290	0.0685	9.3542	10.4711	-0.3006	3.8764	0.2068
0.6	0.5107	0.0025	0.3678	0.0520	0.0670	8.7462	10.3710	-0.2950	3.8524	0.2870
0.8	0.4737	0.0021	0.3579	0.0845	0.0818	8.4533	10.3381	-0.2912	3.8402	0.3387
0.9	0.4072	0.0030	0.3535	0.1198	0.1164	8.3741	10.3310	-0.2900	3.8366	0.3530

Table 2.7. Results for the optimization with the CRRA utility function without investment restrictions and time-varying moments. The legend is the same as in Table 2.1

	Portfolio Weights					Portfolio Moments				Portfolio Statistics	
γ	$\bar{\omega}_1$	$\bar{\omega}_2$	$\bar{\omega}_3$	$\bar{\omega}_4$	$\bar{\omega}_5$	$\bar{\mu}_p$	$\bar{\sigma}_p$	$\bar{s}_p^3$	$\bar{\kappa}_p^4$	**OC**	**TW**
					Second order Approximation						
0.05	2.5489	-1.5112	0.6796	-1.2452	0.4100	45.4893	22.7446	–0.1289	3.7654	0.6078	0.8821
0.1	1.0553	-0.5610	0.3068	-0.4393	0.1902	23.5980	11.7994	-0.1276	3.7666	0.0889	0.5520
0.2	0.4865	-0.2771	0.1984	-0.2014	0.0930	20.3764	10.3282	-0.1272	3.7644	0.0133	0.2994
0.4	0.2936	-0.1313	0.1410	-0.1074	0.0995	18.5664	9.0679	-0.1245	3.7635	0.0023	0.2954
0.6	0.2005	-0.0586	0.0802	-0.0434	0.0629	13.906	4.6888	-0.1178	3.7574	0.0012	0.2416
0.8	0.1718	-0.0273	0.0123	-0.0363	0.0345	9.5876	2.5589	-0.1124	3.7563	0.0008	0.1550
0.9	0.1560	-0.0128	-0.0345	-0.0345	0.0122	6.7116	1.6697	-0.1156	3.7589	0.0002	0.0864
					Third order Approximation						
0.05	2.2003	-1.3156	0.4301	-1.3182	0.3345	20.2543	11.8876	–0.0733	3.5923	0.5328	0.3311
0.1	0.9730	-0.5591	0.1946	-0.4807	0.1602	18.2920	8.3990	-0.0945	3.5954	0.0875	0.2880
0.2	0.4416	-0.2016	0.0946	-0.1697	0.0907	17.3034	8.3106	-0.0976	3.5925	0.0103	0.2556
0.4	0.2438	-0.1337	0.0794	-0.0081	0.0567	12.501	7.4423	-0.1072	3.5923	0.0021	0.2381
0.6	0.1577	-0.0231	0.0505	0.0043	0.0216	11.900	4.6353	-0.1087	3.5953	0.0001	0.2110
0.8	0.1255	-0.0231	0.0260	-0.0134	0.0103	7.5233	2.3116	-0.1012	3.5922	0.0000	0.1253
0.9	0.1357	-0.0226	0.018	-0.0058	0.0032	5.7138	1.6000	-0.1053	3.5900	0.0000	0.0585
					Fourth order Approximation						
0.05	0.8866	-0.676	0.2766	-0.3880	0.1334	15.0056	7.5546	-0.1234	3.5187	0.4587	0.2326
0.1	0.6148	-0.433	0.1768	-0.2472	0.1054	13.5770	7.4156	-0.1265	3.5188	0.3407	0.2168
0.2	0.3381	-0.2111	0.0879	-0.1076	0.0790	12.1112	6.9002	-0.1265	3.5152	0.1142	0.1863
0.4	0.2595	-0.1943	0.1356	-0.0583	0.0328	10.5533	5.2298	-0.1232	3.5111	0.0836	0.1753
0.6	0.1889	-0.1843	0.1128	-0.0442	0.0718	7.2254	3.4478	-0.1334	3.5120	0.0526	0.145
0.8	0.1241	-0.1843	0.1161	0.0042	0.0825	6.5899	2.3333	-0.1365	3.5274	0.0336	0.115
0.9	0.0419	-0.1755	0.1047	0.0053	0.0401	5.2340	1.4098	-0.1322	3.5634	0.0334	0.0165

Table 2.8. Results for the optimization with the CRRA utility function with investment restrictions and time-varying moments. The legend is the same as in Table 2.1

	Portfolio Weights					Portfolio Moments				Portfolio Statistics
γ	$\bar{\omega}_1$	$\bar{\omega}_2$	$\bar{\omega}_3$	$\bar{\omega}_4$	$\bar{\omega}_5$	$\bar{\mu}_p$	$\bar{\sigma}_p$	$\bar{s}_p^3$	$\bar{\kappa}_p^4$	**OC**
						Second order Approximation				
0.05	0.3758	0.0036	0.3456	0.1315	0.1435	13.3093	13.2503	-0.2997	4.3314	0.0609
0.1	0.4452	0.0023	0.3501	0.1041	0.0984	12.3044	12.0243	-0.3020	4.1570	0.0417
0.2	0.4969	0.0018	0.3592	0.0689	0.0732	11.2393	11.1792	-0.3051	4.0080	0.0213
0.4	0.5246	0.0021	0.3590	0.0498	0.0644	10.4709	10.7887	-0.3096	3.9413	0.0123
0.6	0.5341	0.0026	0.3561	0.0431	0.0641	10.1277	10.6656	-0.3110	3.9239	0.0087
0.8	0.5390	0.0029	0.3532	0.0406	0.0643	9.9705	10.6166	-0.3117	3.9176	0.0068
0.9	0.5400	0.0029	0.3525	0.0402	0.0644	9.9332	10.6057	-0.3119	3.9163	0.0062
						Third order Approximation				
0.05	0.3855	0.0039	0.3584	0.1206	0.1316	12.9772	12.8143	-0.2943	4.2769	0.0424
0.1	0.4340	0.0031	0.3758	0.0904	0.0967	11.9885	11.7517	-0.2924	4.1000	0.0273
0.2	0.4594	0.0035	0.3997	0.0596	0.0777	10.9894	11.0561	-0.2899	3.9680	0.0143
0.4	0.4620	0.0065	0.4142	0.0413	0.0759	10.2357	10.7353	-0.2858	3.9036	0.0088
0.6	0.4580	0.0095	0.4196	0.0353	0.0776	9.8984	10.6378	-0.2818	3.8820	0.0066
0.8	0.4515	0.0111	0.4251	0.0331	0.0792	9.7357	10.6023	-0.2784	3.8719	0.0057
0.9	0.4479	0.0118	0.4279	0.0326	0.0799	9.6946	10.5962	-0.2769	3.8688	0.0055
						Fourth order Approximation				
0.05	0.4800	0.0028	0.3825	0.0598	0.0749	11.1657	11.1630	-0.2970	3.9767	0.2321
0.1	0.4995	0.0047	0.3866	0.0409	0.0684	10.3962	10.7810	-0.2982	3.9132	0.1215
0.2	0.5112	0.0081	0.3814	0.0290	0.0702	9.7261	10.5636	-0.2979	3.8783	0.0891
0.4	0.5135	0.0112	0.3793	0.0237	0.0722	9.2093	10.4470	-0.2965	3.8594	0.0645
0.6	0.5142	0.0119	0.3782	0.0228	0.0729	9.0750	10.4230	-0.2961	3.8548	0.0345
0.8	0.5150	0.0120	0.3768	0.0229	0.0732	9.0450	10.4183	-0.2961	3.8537	0.0321
0.9	0.5154	0.0119	0.3762	0.0232	0.0733	9.0508	10.4194	-0.2963	3.8539	0.0311

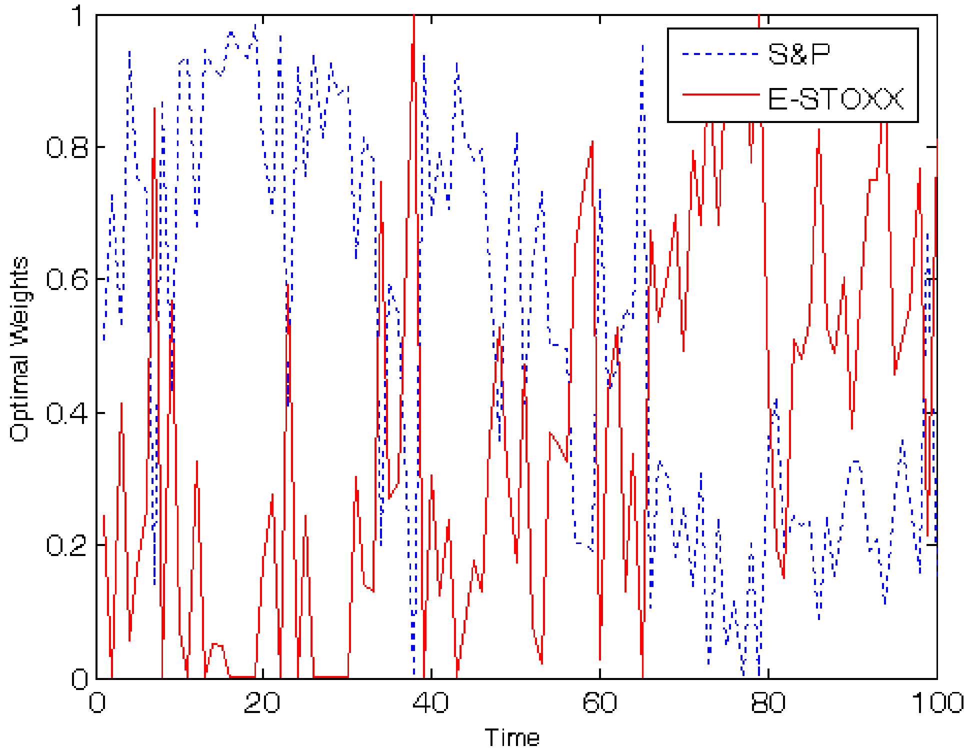


Fig. 2.1. Optimal weights for the S&P and the E-STOXX in the CARA conditional setting for the first 100 days with $\gamma = 0.4$.

the second order approximation and (0.8866 -0.676 0.2766 -0.3880 0.1334) for the fourth order, which are substantially different.

3. The cost of opportunity of ignoring the time varying first moments is also relevant. The opportunity cost can be as high as 35.30% in the restricted case (CARA) for $\gamma = 0.9$ and as high as 54.90% in the restricted case with $\gamma = 0.05$. In Figure 2.1 we present the time evolution of the optimal weights for the S&P Index and the E-STOXX that shows that lines are certainly not constant as it should in the unconditional setting, explaining, therefore, the high opportunity costs found.

In addition to these conclusions, most of the results of the previous Section can also be applied in the conditional framework. Therefore, we can summarize that the opportunity cost of ignoring first and second order dynamics is significant.

2.1.6 Conclusions

In this Section we have presented a dynamic correlation model with Multivariate Cornish-Fisher distributed innovations in order to analyze the impact of higher order and time-varying moments on the asset allocation problem. We have found that considering higher moments is in general, *but not always*, important in portfolio selection. The multivariate normal distribution is not an appropriate probability model for portfolio returns, primarily because it fails to allow for higher moments, for example, skewness, co-skewness, kurtosis and co-kurtosis. We also demonstrate that the Multivariate Cornish-Fisher Density is able to capture these higher moments and becomes flexible enough to allow for skewness and co-skewness and, at the same time, accommodates heavy tails while capturing first and second order dynamics. On the other hand, we have obtained that including time varying moments is crucial for asset allocation, given that ignoring changes in investment moments can result in big costs of opportunity.

While we believe that we have made progress on an important issue in portfolio selection, there are at least three limitations to our approach. First, we handle with parameters which are by definition uncertain and subject to variations. Second, our exercise is an 'in-sample' portfolio selection and we have not applied yet our method to out-of sample portfolio allocation. Finally, the portfolio choice problem we have analyzed only includes the dynamics in the first two moments, while an autoregressive behavior could also be imposed to the parameters a_2 and a_3 to model the dynamics on higher moments[12] to investigate if ignoring higher order dynamics results in costs of opportunity for the investor. In future research we expect to make progress on these limitations considering works like Rubio, Serna, and León 2006 and Jondeau and Rockinger 2005, and we are also interested in incorporating parameter uncertainty in a Bayesian framework, as in Harvey, Liechty, Liechty, and Müller 2005, focusing on prospective moments rather than future moments. Finally, the calculation of the cost of opportunity with respect to the true utility maximization and the relationship between the parameter γ and the parameters of a model with parameters for variance, skewness and kurtosis aversion would be of interest.

[12] Several attempts have been made in this direction but the estimation procedure of the coefficients in the model becomes very unstable. Basically, the problem is that the coefficients a_2 and a_3 are bounded as can be seen in Figure 1.1 and usually the autoregressive behaviour moves the solution to this boundary. One way of solving these problem would be to choose an appropriate transformation of the coefficients that would expand this region to the whole real plane. Nevertheless, when doing so, the problem would be to find a reasonable model for the new coefficients.

2.2 VaR Calculation

2.2.1 Introduction

Value at Risk (VaR) is the standard tool for measuring market risks (Jorion 2000). VaR is an estimate, with a predefined confidence interval, of the quantity that can be lost due to a position during a concrete time interval in standard market conditions. The most typical horizons used in practice are those ranging from one day to one month or one year. The regulatory environment and the need for controlling risk in the financial community have provided incentives for banks to develop proprietary risk measurement models, as the VaR. Among other advantages, VaR provides a quantitative and synthetic measure of risk that allows to take into account various kinds of cross-dependence between asset returns, fat-tail and non-normality effects, arising from the presence of financial options or default risk, for example.

The estimation of this risk measure is related to quantile estimation and tail analysis. Fully parametric approaches are widely used by practitioners (see e.g. JP Morgan Riskmetrics documentation), and most often based on the assumption of multivariate normality of assets or factors returns. These parametric approaches are rather restrictive, as they generally imply misspecification of the tails, and VaR underestimation, since real distributions for the evolution of asset prices are besides asymmetric, notably leptokurtic, as we have shown in Section 1.5. Other parametric approaches which include heavy tails are modeled by means of jumps in prices or stochastic volatilities (GARCH) (Duffie and Pan 1997, Duffie and Pan 2001), or using extreme value densities (Neftci 2000, Longin 2000, Embrechts, Klüppelberg, and Mikosch 1997). Fully non-parametric approaches have also been proposed and consist in determining the empirical quantile, the historical VaR or a smoothed version of it (Harrel and Davis 1982, Falk 1984, Falk 1985, Jorion 1996 and Ridder 1997). Additionally, semi-parametric approaches have been developed, which are mainly based either on an extreme value approximation for the tails (Bassi, Embrechts, and Kafetzaki 1997, Embrechts, Resnick, and Samorodnitsky 1998) or local likelihood methods (Gourieroux and Jasiak 1999).

In this Section, given the fitting-quality of the Univariate and Multivariate Cornish-Fisher density functions to market data found in Chapter 1, we provide two new VaR estimation methods for portfolios based on this density: one analytical and another one simulation-based. Considering the exchange rates database both approaches will be compared with standard market models via a Backtesting. It is interesting to point out that the Backtesting methodology can be seen as an out-of-sample test for the CFD hypothesis. The outline of this Section is as follows: first, we will describe briefly the traditional approaches for calculating the VaR, then we will analyze a one asset VaR calculation using the CFD density and, finally, we will study the two VaR estimation methods for portfolios.

2.2.2 Traditional approaches to VaR

First, we present the standard methods to the VaR calculation of a portfolio, namely, (i) the delta-normal one which assumes that risk factor returns are multivariate normally distributed and that the change in portfolio value is linearly dependent on all risk factor returns, (ii) the historical simulation, which assumes that asset returns in the future will have the same distribution as they had in the past (historical market data) and (iii) Montecarlo simulation, where future asset returns are randomly simulated.

Delta-normal VaR

A very popular approach introduced by J.P Morgan is the one based on the hypothesis that asset returns (equity, prices of zero coupon bonds, exchange rates, options and futures) follow a multivariate normal distribution. The value of the Profit and Loss (P&L) function, P, has the following form:

$$P = \sum_{i=1}^{n} \omega_i R_i \tag{2.9}$$

where R_i is the daily return of the i-th asset, and ω_i $\{i = 1, ..., n\}$ are constants that determine the weight of each asset inside the portfolio. When considering that market variables follow a multivariate normal distribution the value of the portfolio will follow a normal distribution with mean, μ_P, and variance, σ_P, given by:

$$\begin{aligned} \mu_P &= \sum_{i=1}^{n} \omega_i \mu_i \\ \sigma_P^2 &= \sum_{i,j=1}^{n} \omega_i \omega_j \sigma_i \sigma_j \rho_{ij} \end{aligned}$$

with σ_i being the daily volatility of the i-th market variable, μ_i the daily mean and ρ_{ij} the correlation between R_i and R_j. Choosing a confidence interval and a temporary horizon, the VaR can be easily calculated from σ_P. For example, the VaR with a confidence interval of α and with a temporal horizon of T days is given by the expression:

$$1 - \alpha = \int_{-\infty}^{\infty} \theta(R - \mathrm{VaR}_\alpha) \frac{1}{\sqrt{2\pi\sigma_P^2}} e^{-\frac{1}{2}\left(\frac{R-\mu_P}{\sigma_P}\right)^2} dR$$

where $\theta(x)$ is the Heaviside step function[13]. Carrying out the integral one obtains:

$$\text{VaR}_\alpha = \mu_P - \Phi(\alpha)\sigma_P\sqrt{T} \tag{2.10}$$

where $\Phi\left(x\right)$ is the standard normal distribution function. The benefits of the delta-normal model are the use of a compact and maintainable data set (only a variance-covariance data base), which can often be bought from third parties, and the speed of calculation using optimized linear algebra libraries. Drawbacks include the hypothesis that the portfolio is composed of assets whose delta is linear, and the assumption of a normal distribution of asset returns.

Historical-VaR

Historical simulation is the simplest and most transparent method of calculation. This involves running the current portfolio across a set of historical price changes to yield a distribution of changes in portfolio value, and computing a percentile (the VaR). In other words, it is supposed that the weights of the different assets of the current portfolio do not vary and we simulate scenarios equal to the historical changes of our database. The benefits of this method are its simplicity to implement and the fact that it does not assume a normal distribution of asset returns. Drawbacks are the requirement for a large market database and the computationally intensive calculation. Additionally, the number of simulations is evidently limited to the number of available days so that it is very difficult to carry out a sensibility analysis, as for example, the inclusion of assets in the portfolio for which

---

[13] The Heaviside step function is defined as:

$$\theta(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{cases}$$

historical data do not exist. Finally, with this model we assume that the historical distribution is constant over time, obviating dynamic features as conditional heteroskedasticity or changing correlations.

Montecarlo-VaR

In the Montecarlo-VaR we generate a random scenario of market movements using some market model (generally multivariate normality) and revaluate the portfolio under each simulated market scenario. Afterwards, we compute the P&L and sort the resulting P&L to give us the simulated returns distribution for the portfolio. Finally, the VaR at a particular confidence level is calculated using the percentile function. For example, if we computed 5000 simulations, our estimate of the 95% percentile would correspond to the 250th largest loss. Monte Carlo simulation is generally used to compute VaR for portfolios containing securities with non-linear returns (e.g. options) since the computational effort required is exhaustive. Note that for portfolios without these complicated securities, such as a portfolio of stocks, the delta-normal method is perfectly suitable and should probably be used instead if returns follow a multivariate normal distribution. This method implies a high time consumption given that the value of the portfolio must be recalculated in each simulation (i.e. the value of each option must be recalculated using, for example, a Black-Scholes model). In order to reduce another commonly used method consists on approaching the relationship between the P&L, P, and the value of the market variables, R_i, by means of second-order Taylor expansion:

$$P = \sum_{i=1}^{n} \delta_i R_i + \frac{1}{2} \sum_{i,j=1}^{n} \gamma_{ij} R_i R_j + O\left[R_i^3\right] \tag{2.11}$$

where δ_i is the delta (linear sensitivity) of the portfolio to return R_i and γ_{ij} are the second-order derivatives (or gammas) of the portfolio to R_i and R_j. This latter expansion can be used to carry out Montecarlo simulations if we know parameters δ_i and γ_{ij}: we generate a random scenario for R_i and calculate the simulated P&L, P, using Equation 2.11.

2.2.3 One Asset VaR using the Cornish-Fisher Density

VaR estimation for one asset[14] considering that returns follow a m-th Cornish-Fisher density function, $cf_m(R)$, is simply reduced to the calculation of the following integral:

$$z_\alpha = \int_{-\infty}^{\infty} \theta(R - \text{VaR}_\alpha) cf_m(R) dR$$

where z_α determines the confidence level with which we want to calculate the VaR_α. In Appendix B we demonstrate the following result that gives us an analytic formula to calculate the VaR of an asset when the returns follow a Cornish-Fisher distribution:

Proposition 11 *Let R be a m-th order CFD variable. Then, the VaR at a confidence level* $1 - z$ *is given by:*

$$\mathit{VaR} = \Phi^{-1}[Q_m(z)] \tag{2.12}$$

where $\Phi(x)$ *is the standard normal distribution function,* $Q_m(x)$ *is the* m*-th order polynomial and* Φ^{-1} *its respective inverse.*

Therefore, in order to calculate the VaR under the Cornish-Fisher assumption we only need to estimate the parameters of the CFD distribution, a_3, a_2 a_1 and a_0. As we have seen in Section 1.5, this can be done using a static or a dynamic framework but we notice that for

[14] Although this Section could seem to be very restrictive given that only one asset is involved, this asset can be in fact a whole portfolio.

both cases the formula is exactly the same. Given the superiority of in-sample goodness of fit, in the illustration of VaR calculation that follows we have considered the more complete dynamic framework using the CFD-GARCH model presented in Section 1.5.

As an illustrative example we have developed an integrated graphic interface in the MatRisk application, initially implemented by Suárez and Carrillo 2003 in a MATLAB environment. In Figure 2.2 we show the final interface of the application with an example corresponding to the CHF/USD exchange rate. The central graph presents a histogram with a graph in red of the third-order Cornish-Fisher density function and in green a line that indicates the VaR at the confidence level that has been selected in the lower displacement bar. In the square bellow the resulting VaR using the third-order CFD model (CFD-VaR), the sample (historical) VaR and the Expected Shortfall or Conditional VaR[15] resulting from the model are also shown. On the right part the p-value of the Kolmogorov-Smirnov statistic and the parameter estimates are reported. Finally, we present the QQ-Plot of the data against the normal distribution and the fit of a third-order polynomial. In addition, in Figure 2.3 we have compared these results with those obtained calculating the VaR by means of the fit of a normal distribution (normal-VaR). It is evident that in this case the fitting quality is poorer that in the case of the third-order CFD, as measured by the KS statistic.

In Table 2.9 we compare the differences in the VaR value between the normal-VaR, historical-VaR and CFD-VaR considering different confidence levels. It can be observed

[15] The Expected Shortfall or Conditional VaR of a sample is defined as:

$$\mathrm{ES}_\alpha = E\left[X \mid X > \mathrm{VaR}_\alpha\right]$$

The Expected Shortfall is a measure of the mean loss that will have a portfolio, given that this loss is bigger than a certain limit. It has been proposed by several authors like an alternative risk measure with the desirable properties of subadditivity and coherence (Artzner et al. 1999). All calculations performed in this Section can be easily adapted to this risk measure.

	$\mathbf{90\%}$	$\mathbf{95\%}$	$\mathbf{99\%}$
Historical-VaR	-0.8675	-1.2067	-2.1240
CFD-VaR	-0.8903	-1.2260	-2.0247
Normal-VaR	-0.9739	-1.2518	-1.7732

Table 2.9. VaR comparison for the CHF/USD exchange rate. Historical-VaR, CFD-VaR and Normal-VaR, denote the Historical simulation, the CFD and the delta-normal estimates for the VaR, respectively.

that the normal-VaR for the percentiles corresponding to 90% and 95% are bigger than the CFD-VaR, while for a percentile of 99% it is notably smaller. This is so because when calculating the normal-VaR very little importance is given to the most extreme losses, since this one is not capturing appropriately the asymmetry nor the kurtosis of the distribution. In this way, under the assumption of normality, the VaR is overestimated for elevated percentile values, while it is undervalued for the lowest percentile values, which correspond to the most extreme events. On the other hand, it is interesting to point out that the calculated CFD-VaR is much more similar to the historical-VaR. It is interesting to highlight that the differences between the CFD-VaR and the normal-VaR with respect to the estimates of the historical-VaR at a 1% confidence level are 4.68% and 16%, respectively.

2.2.4 Backtesting

Backtesting is a methodology used in the supervision of the quality of VaR calculations and consists on keeping track of the realized losses and gains, checking that for a given confidence level they are not bigger than the calculated VaR for that confidence level. Therefore, the most straightforward way to backtest a VaR calculation is to plot the P&L against predicted VaRs and monitor the number of excessions. For example, if you have a 1-day 95% confidence VaR, you should expect 5% upside and downside excessions over time. If ac-

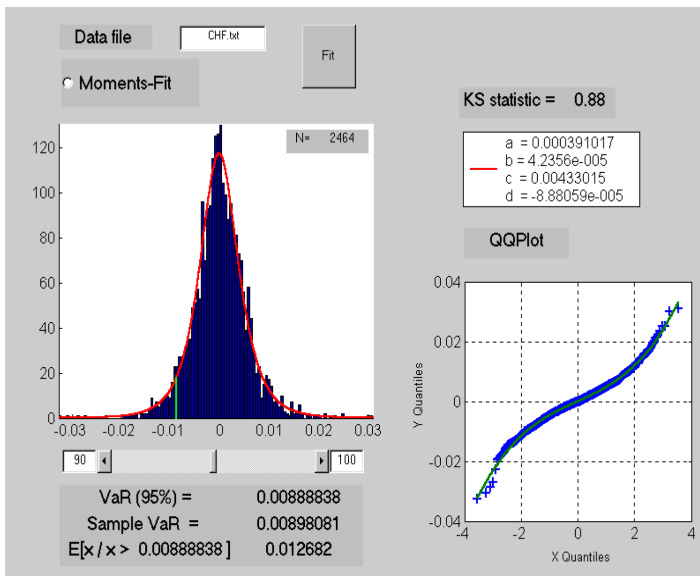


Fig. 2.2. Graphical interface to perform VaR calculations using a third-order CFD model. On screen it can be seen the fit for the CHF/USD exchange rate. The central figure shows an histogram with a red curve representing the fitted CFD. A green line indicates the confidence level percentage for the VaR that corresponds to the one that has been chosen in the bar. On the lower figure the numerical VaR values using the CFD model, the sample-VaR and the Expected Shortfall calculated using the model are shown. On the right side appears the p-value of the Kolmogorov-Smirnov statistic and the values of the estimated parameters from the model. Finally, on the lower-right corner the sample QQ-Plot with the cubic polynomial fit is shown.

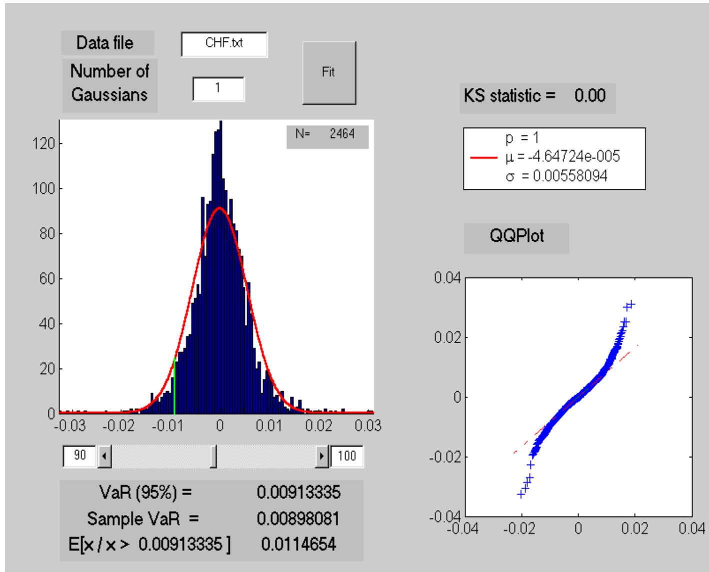


Fig. 2.3. VaR estimation using a gaussian fit for the CHF/USD exchange rate. The quality-of-fit obtained is remarkably inferior to the one obtained considering a third-order CFD, as can be observed from the Kolmogorov-Smirnov statistic.

tual excessions are much larger or smaller this is an indication of an inaccuracy in the calculation models. A VaR measure in which the losses were always smaller would be a too conservative measure and one of the consequences would be, for example, the obligation of keeping too much capital as provision. Equivalently, if losses above the VaR were more frequent than in reality, we would be undervaluing the risk and it could imply losses for which we are not appropriately prepared. This methodology, can be seen as an out-of-sample test for the CFD hypothesis as we are evaluating the capacity of predicting a given percentile of the empirical distribution.

In Table 2.10 we present the results of the Backtesting for the twelve exchange rates, using the fit of a gaussian as well as the Cornish-Fisher density function and the historical time series. We perform one day VaR calculations with a confidence interval of 99% for

	Normal-VaR	Historical-VaR	CFD-VaR	Excessions Range
AUD	9	6	9	1-11
BEF	8	6	5	1-11
CHF	7	5	5	1-11
DEM	12	7	5	1-11
DKK	10	7	6	1-11
ESP	7	5	4	1-11
FRF	8	5	6	1-11
GBP	10	8	4	1-11
ITL	8	5	11	1-11
JPY	7	8	4	1-11
NGL	6	5	4	1-11
SEK	7	4	6	1-11
TOTAL	99	71	69	50-70

Table 2.10. Backtesting results for the twelve exchange rates using the gaussian fit as well as a fit using the third-order CFD and the historical-VaR. We present the number of excessions for the VaR at a 99% level of confidence throughout 500 days. A good model should give around 60 excessions in total.

500 successive days for the twelve exchange rates. In each calculation we use the 200 previous days for the fitting of the distribution[16]. Since we are calculating the VaR with a confidence level of 99%, for each exchange rate we expect the number of losses exceeding the VaR to be a 1% of the number of days considered in the Backtesting. Since we carry out 500 measures of the VaR, the number of excessions must be around 5, i.e., the number of excessions will be inside the confidence intervals considered in Peña 2002.

As we have already mentioned, the basic result indicates that the normal-VaR underestimates the historical-VaR and that the CFD-VaR fits better to the historical-VaR. As an example, in Figures 2.4 we show the graph of the Backtesting for the last exchange rate, the SEK/USD exchange rate.

[16] The time window used to fit the parameters of the model usually depends on the criterion of the agent. The correct election of this parameter is essential to set good predictions in VaR calculation. We have carried out the Backtesting considering both a 200 days time window and a 500 one and, although VaR calculations using the Cornish-Fisher function were also closer to sample-VaR compared to the normal case, with the normal-VaR better results were obtained in the Backtesting.

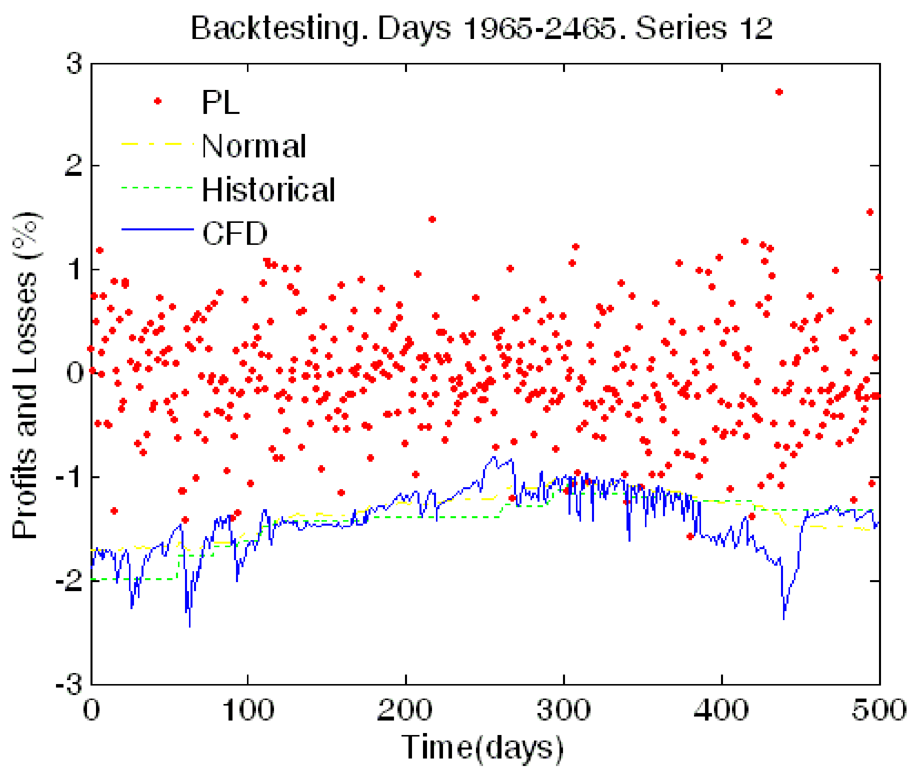


Fig. 2.4. Backtesting for a one day time horizon VaR at a 99% level of confidence throughout 500 consecutive days for the SEK/USD exchange rate series. The normal-VaR, historical-VaR and CFD-VaR are shown. Dots represent realized losses or profits the next day after the calculation. A good model should give a number of excessions around 5.

2.2.5 Portfolio VaR using the Multivariate Cornish-Fisher Density

Methodologies

VaR calculation for a portfolio that is formed by assets that follow a Multivariate Cornish-Fisher distribution (in both of its forms, the Copula Based and the Variance-Covariance Based, see Section 1.3) does not have a direct analytical approach. The problem is that Multivariate Cornish-Fisher distributions are not closed under convolution and, in principle, we ignore the general Profit and Loss distribution of a portfolio[17]. Nevertheless, we provide two different approximative alternatives to carry out the VaR calculation of a portfolio whose assets follow a MCFD: the first one is based on Montecarlo simulations and the other one is based on the calculations of the moments of the Profit and Loss distribution.

The first possibility consists on fitting the parameters of the MCFD, and use one of the simulation algorithms described in Section 1.3 to simulate CB-MCFD or VCB-MCFD variables. If one wants to calculate the VaR for a period of T days, paths should be simulated with T points. After generating enough scenarios, the VaR$_\alpha$ is obtained from the percentile z_α of the simulated distribution function of P&L of the simulated portfolio.

The second possibility consists on applying the Cornish-Fisher Expansion (Johnson and Kotz 1972b, see Appendix A.1 for details) for the calculation of the percentile using the moments of the distribution of the Profit and Loss function. In Section 1.3 we have calculated the m-th moment of the change in the value of a portfolio given by the equation

[17] This is not true for the VCB-MCFD model as we have shown in Chapter 1. However, although been closed under summation, we do not know the specific form of the univariate (marginal) distributions of a VCB-MCFD and, therefore, a completely analytical VaR calculation is also not possible for this distribution.

2.9 both under the Copula-Based MCFD and under the Variance-Covariance-Based MCFD models. Therefore, using the Cornish-Fisher Expansion on the P&L variable, we can obtain an estimate of the VaR applying the following equation:

$$\text{VaR} = \mu + \sigma z_R$$

where μ corresponds to the mean of the distribution, σ is the standard deviation and z_R is given by:

$$z_R = z_X + \frac{1}{6}\left(z_X^2 - 1\right)\xi + \frac{1}{24}(z_X^3 - 3z_X)\kappa - \frac{1}{36}(2z_x^3 - 5z_x)\xi^2 \tag{2.13}$$

where z_X is the corresponding standard gaussian percentile, ξ is the skewness and κ is the kurtosis. If we want to obtain the VaR for a longer time horizon and we have only fitted the distribution for daily data in this second approach it is necessary to make additional hypothesis on the time-scaling properties of moments. In the gaussian approach it is well-known that the variance scales as $\sqrt{t}$ and it is not necessary to make any hypothesis on the scale properties of the skewness ξ or on the excess kurtosis κ, since these are equal to zero. However, in this more general context the escalation problem is more complicated. It is known (Bouchaud and Potters 2000) that in a uncorrelated series the centered moments must have an evolution proportional to $\sqrt{t}$. This implies that skewness must follow a $1/\sqrt{t}$ law and the kurtosis $1/t$ law. Therefore, for distant times the skewness, the kurtosis and, of course, the superior order moments tend to zero, which constitutes a manifestation of the central limit theorem. However, the existence of autocorrelation, either lineal or of superior order, implies some slower convergence rates towards normality.

As in the previous Subsection, we can use either a static or a dynamic framework to estimate the parameters of the MCFD. Given the multivariate in-sample estimation results of Chapter 1, we use the dynamic Variance-Covariance-Based MCFD in the rest of this Section to estimate the VaR of a portfolio.

Application of both methodologies

We carry out the calculation of the VaR for a portfolio formed by a monetary unit expressed in US Dollars of each one of the following foreign currencies: Japanese yen, Dutch pound and Swedish crown[18]. We calculate the VaR to one day and for the last day available in our data, 08/15/1997, in four different ways: the first of them considering the normal framework (normal-VaR), the second one using the sample percentile of the historical distribution of price changes of the portfolio (historical-VaR), the third and fourth ones following the two methodologies proposed in the previous Section. The third one using the Cornish-Fisher Expansion (CFD_{ana}-VaR) and in the fourth one we use the Montecarlo simulation in the way that we detail next (CFD_{Mon}-VaR). First, we carry out a fit of the dynamic VCB-MCFD function considering all the data of the sample (2645 data, between 04/07/1988 and 15/08/1997) using the algorithm described in Chapter 1[19]. Afterwards we simulate 10000 paths of VCB-MCFD variables using the simulation procedure described in Section 1.3.2. From these simulated evolutions we calculate the evolutions of the value

[18] In this work we will not analyze the impact of including options for the VaR calculations under a multivariate CFD. Nonetheless, we would like to outline a procedure to include these assets in the model. In this case, we can evaluate the VaR using Montecarlo simulation and considering the quadratic relation between the Profit and Loss function and the market variables given by Equation 2.11. As a matter of fact, we simulate the variables R_i and then use this Equation to simulate the portfolio.

[19] The estimation results are available from the authors upon request.

	90%	**95%**	**99%**
Normal-VaR	-1.9242	-2.4659	-3.4822
Historical-VaR	-1.8286	-2.4478	-3.7399
$\mathbf{MCFD_{ana}}$-VaR	-1.7176	-2.4148	-4.1811
$\mathbf{MCFD_{Mon}}$-VaR	-1.7891	-2.4120	-3.7995

Table 2.11. VaR for a portfolio formed by one monetary unit countervalued in dollars for each of the following currencies: Japanese yen, Dutch pound and Swedish crown, supposing normality, historical-VaR and the VaR using the VCB-MCFD model, with the analytical approximation and by means of the Montecarlo simulation.

of the portfolio in each path. Finally, on the simulated histogram of changes of the value of the portfolio we calculate the VaR as the percentile to the 10, 5 and 1%.

We show VaR results obtained using the different methods in Table 2.11. The results of the CFD_{Mon}-VaR are very similar to the obtained ones for the VaR of an asset in the previous Section. The calculation of the CFD_{ana}-VaR gives a value something superior for the percentile to 1% that the CFD_{Mon}-VaR. This is so because the CFD-VaR works with up to the fourth order of the distribution in this approach. Therefore, the CFD_{ana}-VaR will always tend to overestimate the VaR for leptokurtic distributions.

Backtesting

In order to check the goodness-of-fit of each method we carry out VaR calculations for a one day time horizon and with a confidence level of 99% for 500 successive days for four portfolios. These four portfolios are formed by one unit of each one of the three following foreign currencies: portfolio 1 (Australian dollar, Belgian franc and Swiss franc), portfolio 2 (German mark, Danish crowns and Spanish peseta), portfolio 3 (French franc,

Portfolio\Excessions	Normal	Historical	MCFD_{ana}	MCFD_{Mon}	Excessions Range
1	3	2	5	5	1-11
2	9	6	7	5	1-11
3	5	2	7	6	1-11
4	5	5	6	5	1-11
Total	22	15	25	21	10-30

Table 2.12. Excessions in the estimation of the normal-VaR, of the sample-VaR, of the VaR using the Cornish-Fisher expansion approximation and of the Montecarlo-VaR. A good measure for the VaR should give a number of excessions included within the specified range in the last column.

English pound and Italian lira) and portfolio 4 (Japanese yen, Dutch pound and Swedish crown)[20].

The procedure used for the simulations is the following one: first, we carry out the VCB-MCFD fitting using the Maximum Likelihood method as explained in Section 1.5, taking for this fit the 200 data previous to the moment of the VaR calculation, period that approximately corresponds to one year data[21]. Later on we simulate 10000 evolution paths for each one of the assets. Using these simulations of the market variables we calculate the evolutions of the portfolio value in each one of the paths and with this simulated histogram of the variations of the value of the portfolio we obtain the VaR as the 1% percentile.

For the calculation of the VaR by means of the Cornish-Fisher Expansion approach we use the same fitting that the one used for the Montecarlo simulation and by means of the coefficients, $a_{3,i}$, $a_{2,i}$, $a_{1,i}$ and $a_{0,i}$, and the variance-covariance matrix we obtain the moments of the P&L distribution. Finally, we use the Cornish-Fisher approach to evaluate the percentile and to estimate the VaR.

[20] Selection of the component assets of each portfolio is completely random.

[21] Estimation results are available from the authors upon request.

In Figure 2.5 we present the Backtesting for the CFD-VaR, the VaR calculation supposing normality (also with a 99% confidence level), the historical VaR and the realized Profit and Losses values for each of the days. As it can be seen, at a 99% confidence the CFD-VaR is in general higher than the normal-VaR, which does not take into account skewness and excess kurtosis. In some intervals the CFD-VaR is smaller, because the dynamic model allows to capture periods with lower volatility. The calculated CFD-VaR is systematically higher than the one derived considering the normal framework, since the excess kurtosis coefficient is always positive. On the other hand, it is reliable the little variability that the historical-VaR presents, being almost constant for some stretches and presenting small jumps due to extreme events in profit or losses. This strong dependence on big market movements is one of the reasons for which the methodology of the historical-VaR is usually criticized. CFD-VaR calculations are the ones closest to the historical-VaR, but they present a stronger variability due to the aleatory sampling that we are considering.

Since we are calculating the VaR with a confidence level of 99% it is to be expected that for each portfolio the number of losses exceeding the VaR to be more or less of 1%. In Table 2.12 we present the excessions in the measure of the normal-VaR, historical-VaR, $\mathrm{CFD}_{\mathrm{ana}}$VaR and $\mathrm{CFD}_{\mathrm{Mon}}$-VaR and also the acceptance range of excessions for the VaR. One can observe that although the number of excessions of the normal-VaR is within the allowed range, the excessions of the $\mathrm{CFD}_{\mathrm{Mon}}$-VaR is much closer to the historical-VaR and the theoretical value of twenty. The $\mathrm{CFD}_{\mathrm{ana}}$-VaR, as we have already mentioned, it overvalues the theoretical value although it lies within the admitted range. Results are therefore very similar to the ones obtained for a single asset.

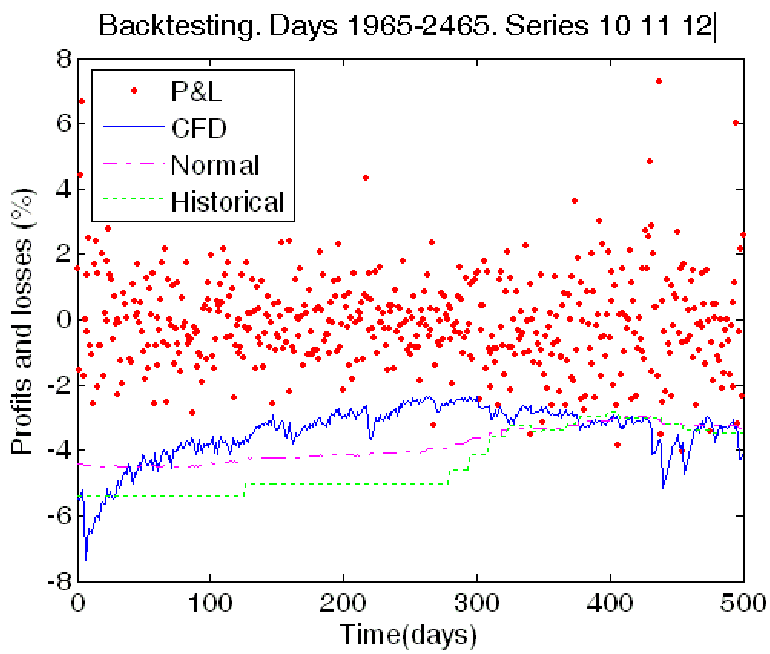


Fig. 2.5. Backtesting for the fourth portfolio formed by the assets JPY, NGL and SEK. These are one day VaR calculations with a 99% confidence interval for 500 consecutive days. The following estimations are shown: normal VaR, historical-VaR and the VaR obtained by means VCB-MCFD model with the Montecarlo-VaR approximation using 10000 scenarios. For the fits we use the 200 prior days to the calculation day. Dots are the realized losses or gains. The optimal number of excessions is equal to 5.

2.2.6 Conclusions

In this Section we have proposed two ways of calculating the Value at Risk for a portfolio formed by assets following a Multivariate CFD. In particular, we have analyzed a semi-analytical procedure based on the Cornish-Fisher Expansion of the Profit and Loss function and another one simulation based. We have tested the quality of these VaR meassures using a Backtesting and we have found that both models tend to outperform the traditional normal or delta-VaR. We have also developed a graphic interface to apply easily these meassures.

In further research it would be interesting to compare these two methods with other market models which incorporate dynamics and heavy tails, like the ones proposed in Chapter 1.

2.3 Option Valuation

In this Section we will generalize the Black and Scholes option valuation formula (Black and Scholes 1973) to include underlyings whose returns can be characterized by a third-order Cornish-Fisher distribution function. We also obtain analytical formula for the hedging parameters, which do not show the anomalies present in other semi-parametric option pricing models (e.g. Corrado and Su 1997b and Jarrow and Rudd 1982). In addition, we compare in and out-sample estimations of option prices, using Spanish options data, and find that our model clearly out-performs the standard Black-Scholes model.

2.3.1 Introduction

As it is well known, the Black and Scholes 1973 option pricing formula was an authentic breakthrough in finance and even nowadays it is commonly used to value derivative securities. In spite of its widespread acceptance among practitioners and academics, the model has a well known drawback of pricing inconsistently deep in-the-money and deep out-of-the-money options. This phenomenon is referred by professionals as volatility *skew* or *smile*. These mispricing patterns are known to be a result of one of the overidealized assumptions used to derive the formula, namely, the hypothesis of normality with constant volatility for the distribution of log prices.

One of the most recognized solutions to this pricing bias is the semi-parametric methodology proposed first by Jarrow and Rudd 1982 and slightly modified by Corrado and Su 1996a and Corrado and Su 1997b. These methodologies are based on a moment expansion of the Black-Scholes formula to account for non-normal skewness and kurtosis in prices and returns, respectively. Jarrow and Rudd derived an option pricing formula from an Edgeworth expansion of the log-normal probability density function to model the distribution of stock prices, while Corrado and Su used a Gram-Charlier series expansion of the normal probability function to model the distribution of stock log returns. Nevertheless, both approaches, being polynomial approximations, share the theoretical drawback of yielding negative density function values for certain parameter ranges of the implied risk-neutral density function. In addition, these models present anomalies in the hedging parameters, as we will see. Jondeau and Rockinger 2001 proposed as a solution to this problems: the restriction of the parameter space so that the density remains positive, while recently, León, Mencía, and Sentana 2005 analyze the use of the semi-nonparametric distribution for option pricing. These two last approximations solve the negativeness problem of the implied risk-neutral density function while keeping the analytical flexibility of the Edgeworth-Gram-Charlier moment expansion but, on the other hand, they inherit the lack of flexibility for capturing high degrees of kurtosis and skewness, characteristic of these approximations. In Figure 1.2 of Chapter one, we show that in the absence of skewness the maximal kurtosis that can be achieved using this expansions is eight which could be very restrictive to model implied density functions. Moreover, it also has to be pointed out that these approximations derive multi-modal implied densities functions for certain parameter values, which should be treated very carefully as it would imply the uncommon feature that investors find more than one return value as expectable.

In this Section we propose the use of the Cornish-Fisher distribution function to characterize the prices of the underlying in order to develop a model for option valuation. As we will see, this model bears the same analytical tractability characteristic of the Jarrow and Rudd 1982, Corrado and Su 1996a, Jondeau and Rockinger 2001 or León, Mencía, and Sentana 2005 models while it extends their modeling flexibility in terms of covered range of skewness-kurtosis. We will derive closed form expressions for the price of plain vanilla options and its "Greeks" under the assumption that at maturity underlyings follow a third-order Cornish-Fisher distribution. Nonetheless, it is interesting to note that the Cornish-Fisher model does not present certain "anomalies" in the volatility smile and the Greeks, found in the Corrado-Su model. Additionally, an empirical application will be also presented using Spanish options data. Performing in-sample and out-of-sample quality of fit test, the empirical results demonstrate that besides solving the negative density values, properly capturing the risk-neutral kurtosis coefficients as high as twenty, this model highly outperforms the Black-Scholes model. In contrast, comparing the Cornish-Fisher option pricing model with the Corrado and Su 1996a model, we find similar results between them at the cost of allowing negative density values.

2.3.2 European Option Valuation

In this Section we will derive a closed-form pricing formula for European options for an underlying characterized by a Cornish-Fisher density defined in Equation 1.9. As it is well known, the Black-Scholes model is based on the hypothesis that asset prices, S_t, follow a geometric Brownian motion under the risk-neutral probability Q given by:

$$S_T = S_t e^{\left(r-\frac{\sigma^2}{2}\right)\tau+\sigma W_T} \tag{2.14}$$

where r is the risk-free rate, σ is the underlying's volatility, S_t is the initial price of the underlying, τ is the time to maturity and W_T is a standard Brownian motion. The basic hypothesis of the model is that log-returns defined as $R_{t-s} = \ln(S_t/S_{t-s})$ follows a normal distribution, $N(r, \sigma\sqrt{t-s})$. In general, we can obtain the price of an European call option as the expected value of the payment under the risk-neutral probability:

$$C_t = e^{-r\tau} E_Q[(S_T - K)^+] \tag{2.15}$$

The computation of this formula for the geometric Brownian motion gives as a result the well known Black-Scholes formula:

$$\begin{aligned} C_{BS} &= S_t \Phi(d) - K e^{-r\tau} N(d - \sigma\sqrt{\tau}) \\ d &= \frac{\ln(S_0/K) + (r + \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}} \end{aligned}$$

Equation 2.15 is of general use for any other distribution governing S_T. In the model that we propose in this work we make the hypothesis that prices differences can be expanded through a Cornish-Fisher density function which, as we have already mentioned in the previous Chapter, arises naturally as a series expansion of the percentiles of the true distribution in terms of the percentiles of the gaussian distribution. Therefore, the first order term in the expansion of the price S_T will be normally distributed, instead of log-normal as is the case in the Black-Scholes framework, and hence, the first order approximation of the price of a European call using the CFD will not exactly coincide with the Black-Scholes formula, as it is the case of approaches based on an Edgeworth or Gram-Charlier expansion like Jarrow and Rudd 1982 or Corrado and Su 1996a. It is interesting to point out that we do not consider the expansion of the log-returns, R_t, in terms of CFD, as would be more natural, because doing so prices, S_t, would follow a distribution with divergent mean (i.e. $E\left[e^{R_t}\right] = \infty$) and, therefore, its application to option pricing would be useless[22]. Nevertheless, this method regarded as an approximation is theoretically equally valid as any other expansion based methods and, moreover, it assures unimodality and avoids negative density values.

Due to the above mentioned reasons, the first order approximation of the CFD expansion will be an arithmetic Brownian motion rather than a geometric one. Bouchaud and Potters 2000 analyze pricing differences between a geometric Brownian motion and an arithmetic Brownian motion based[23]. It is important to remark that, although the arithmetic Brownian motion includes the possibility of negative prices, in our framework this feature has to be considered as an approximation nuisance; a very small probability as 10^{-20} is in practice virtually equal to zero. Moreover, this theoretical drawback allows us to obtain a positive definite expansion based density function which permits rather high kurtosis and skewness levels[24].

---

[22] Let us consider the expression:

$$E\left[e^{R}\right] = \int_{-\infty}^{\infty} e^{R} cf_m\left(R\right) dR$$

where $cf_m(R)$ is the CFd given by Equation 1.5. Carrying out the variable change, $R = Q_m(X)$, we would obtain the following integral:

$$\int_{-\infty}^{\infty} e^{\sum_{i=1}^{m} a_i X^i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2} dX$$

which clearly diverges for $m \geq 3$ if a_3 is greater than zero.

[23] They find that with a daily volatility of 1% and a zero interest rate, the relative difference between the geometric and arithmetic Black-Scholes price is almost zero for at and in-the-money options but can be as high as 30% for options 20% out-of-the money. Nevertheless, this relative difference is very small when we consider absolute differences.

[24] This problem could be solved considering a *truncated* CFD which would assign a zero probability for the occurrence of negative prices and a CFD distribution for positive prices. Nonetheless, with this distribution valuation formulas are much more complicated and the pricing differences between the truncated and the untruncated are almost negligible for reasonable parameter sets.

The basic model in this framework, the arithmetic Brownian motion, can be written as:

$$S_T = S_t\left(1 + r\tau\right) + S_t \sigma W_T \tag{2.16}$$

The following Proposition, which is demonstrated in Appendix B, gives the price of an European call option under the hypothesis that the underlying follows an arithmetic Brownian motion.

Proposition 12 *Let r be the risk-free interest rate, K the strike of the option, S_t the initial price of the underlying, τ the time to maturity and σ the volatility of the process. In absence of arbitrage opportunities the call option, whose underlying follows an arithmetic Brownian motion, is given by:*

$$\begin{aligned} C &= \frac{1}{1+r\tau}\left[\left(S_t\left(1+r\tau\right)-K\right)\Phi(-d)+S_t\sigma\sqrt{\tau}\phi(d)\right] \qquad (2.17) \\ d &= \frac{K - S_t(1+r\tau)}{S_t\sigma\sqrt{\tau}} \end{aligned}$$

where $\Phi\left(x\right)$ is the distribution function of a standard gaussian variable and $\phi(x)$ is its corresponding density.

In this work we assume that the model for the asset price, S_T, under the risk neutral measure Q is given by:

$$S_T = S_t(1 + r\tau) + S_t\sigma\sqrt{\tau}z^* \tag{2.18}$$

where z^* is a variable following a standardized CFD. Note that we already have selected the drift of the process in such a way that the martingale restriction holds, i.e. $E_Q(S_T) = S_t(1 + r\tau)$[25]. Under the assumption of a CFD distribution we can obtain a more general

[25] The martingale restriction states that under the risk-neutral probability the underlying must have an ex-

formula for the option price since both skewness and excess kurtosis different from zero are possible under this distribution.

Proposition 13 *Let r be the risk-free interest rate, K the strike of the option, S_t the initial price of the underlying, τ the time to maturity and σ the volatility of the process. In absence of arbitrage opportunities the call option, C^{CFD}, whose underlying follows a third-order CFD given by Equation* 1.9 *is:*

$$\begin{aligned} C^{CFD} &= \frac{1}{1+r\tau}\left\{ \begin{array}{c} (S_t(1+r\tau)-K)\,\Phi\left(-d\right)+a_1 S_t\sigma\sqrt{\tau}\phi(d) \\ S_t\sigma\sqrt{\tau}\phi(d)\left(a_3\left(d^2+1\right)+a_2 d\right) \end{array} \right\} \qquad (2.19) \\ d &= Q_\tau^{-1}(K) \\ Q_\tau(x) &= S_t(1+r\tau)+\sigma\sqrt{\tau}S_t\left(a_3x^3+a_2x^2+\left(\sqrt{1-6a_3^2-2a_2^2}-3a_3\right)x-a_2\right) \end{aligned}$$

where $\Phi\left(x\right)$ is the distribution function of a standard gaussian variable, $\phi(x)$ is its corresponding density and $Q_\tau^{-1}(x)$ is the inverse of the third-order polynomial $Q_\tau(x)$.

Note that the formula under the third-order CFD (Equation 2.19) becomes Equation 2.17 when $a_3 = a_2 = 0$, since we obtain the normal distribution for S_T[26]. Although not reported, it is straightforward to generalize the later proposition to allow for a general m-th order CFD distribution. Nevertheless, with a third-order polynomial we already are able to capture the non-normality features of the implicit density function.

Next, we analyze the absolute valuation differences between the Black-Scholes model (1973) and the CFD model calculating call options prices with a strike of $K = 10$, times to maturity of one month considering that the annual volatility is 40% and the risk free in-

pected return equal to the risk-free rate. In these risk neutral worlds, options can be priced as discounted expected payoffs (Hull 2004).

[26] It is important to point out that the inverse of the polynomial $Q_\tau^{-1}(K)$ is equal to $(K - S_0(1+r\tau)/S_0\sigma\sqrt{\tau})$ in the case of $a_3 = a_2 = 0$.

terest rate is equal to 5%. Considering different parameters a_2 and a_3, we select different degrees of skewness and kurtosis. In particular, we choose parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness and kurtosis coefficients of (0,8), (-1,8) and (1,8), respectively. Therefore, we are considering the influence of positively and negatively skewed implied density functions with fat tails. In Figure 2.6 we show the differences between the CFD and Black-Scholes call prices. The presence of positive (negative) skewness makes the upper (lower) tail of the price density fatter under the CFD model, and this produces an increase in the probability of a large drop (raise) in prices, which is responsible for the relative lower (higher) prices given by the Black-Scholes formula for deeply out-of or in-the money call options. We also have applied the CFD model to obtain implied volatility smiles, which are represented in Figure 2.7. In this Figure we can notice the flexibility to obtain different smile patterns.

Corrado and Su (Corrado and Su 1996a, Corrado and Su 1996b, Corrado and Su 1997a, and Corrado and Su 1997b) have developed a valuation formula, based on Gram-Charlier distributions (Section 1.2), for european options, which does also include the contribution of skewness and kurtosis to the price of the option. As this model is also based on a semi-parametric density we will discuss it briefly and we will use it later in the empirical Section for comparison porpouses. The valuation formula of a European call of Corrado

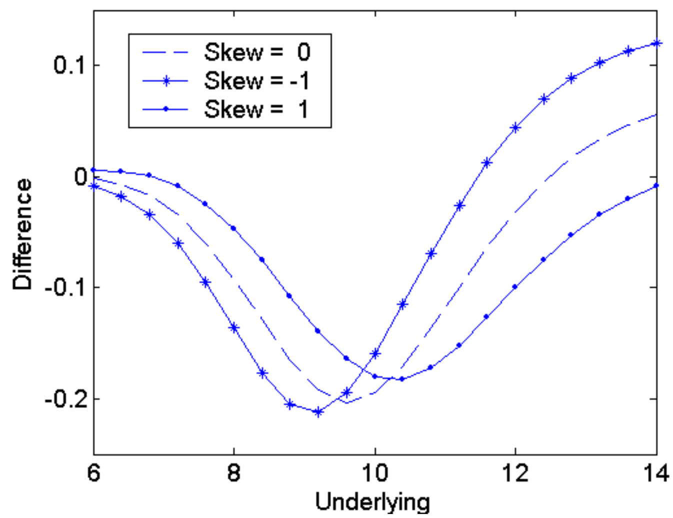


Fig. 2.6. Differences between CFD and geometric Black-Scholes call prices. Call options have a strike of $K = 10$, time to maturity of one month, the annual volatility is 40% and the risk free interest rate is equal to 5%. We assume parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness and kurtosis coefficients of (0,8), (-1,8) and (1,8), respectively.

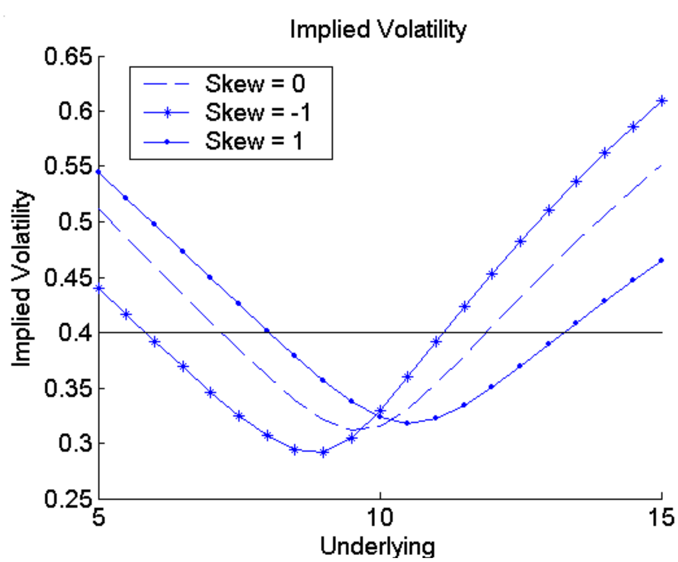


Fig. 2.7. Implied volatilities assuming that the true implied distribution is CFD. The parameters used are the same as in Figure 2.6.

and Su, C_{CS}, is:

$$
\begin{aligned}
C_{CS} &= C_{BS} + \mu_3 Q_3 + (\mu_4 - 3) Q_4 \\
C_{BS} &= S_0 \Phi(d) - K e^{-r\tau} N(d - \sigma\sqrt{\tau}) \\
Q_3 &= \frac{1}{6} S_0 \sigma\sqrt{\tau} \left(\left(2\sigma\sqrt{\tau} - d \right) \phi(d) - \sigma^2 \tau \Phi(d) \right) \\
Q_4 &= \frac{1}{24} S_0 \sigma\sqrt{\tau} \left(\left(d^2 - 1 - 3\sigma\sqrt{\tau} \left(d - \sigma\sqrt{\tau} \right) \right) \phi(d) + \sigma^3 \tau^{3/2} \Phi(d) \right) \\
d &= \frac{\ln(S_0/K) + (r + \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}
\end{aligned}
$$

where μ_3 and μ_4 are the skewness and kurtosis coefficients of the implied density function. It is trivial to see that if the implied density function of log-returns is gaussian (i.e. $\mu_3 = 0$ and $\mu_4 = 3$), this valuation formula converges to the Black-Scholes formula. As pointed out in Section 1.2, this distribution presents the theoretical drawback of yielding negative density values for certain parameter ranges[27].

In Figures 2.8 and 2.9 we analyze the differences between the Corrado and Su and the Cornish-Fisher valuation formula for different degrees of asymmetry and kurtosis. In particular, in Figure 2.8 we compare the implied volatility with both models calculating the implied volatility of call options with a strike of $K = 10$ and times to maturity of one month while in Figure 2.9 we analyze the absolute and relative valuation differences between both models. In these Figures we suppose that the annual volatility is 40%, the risk free interest rate equal to 5% and we consider different parameters in both approximations: a_2 and a_3 for

[27] In Figure 1.2 we can observe the permitted range of the Gram-Charlier distribution (which is the density which underlies in the Corrado and Su model). The permitted range is defined as the limit value of the parameters, so that the density function remains positive. We can observe that only kurtosis coefficients of eight are achievable with this limitation and in the empirical Section we will find that the implied kurtosis coefficients that we find in the Spanish options markets are usually higher. Moreover, Jondeau and Rockinger 2000, using French Franc/German Mark European type exchange rate options data, find that the implied density function of the Corrado and Su model derives negative density values.

the CF model and μ_3 and μ_4 for the CS. In particular, as in the previous Figures, we assume parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness, μ_3, and kurtosis, μ_4, coefficients in the CS model of (0,8), (-1,8) and (1,8), respectively. As a first conclusion we can observe that although the first four moments of the implicit density function for both models are equal, significant pricing differences can be found. Figure 2.8 shows that the CS model gives higher implied volatility values for deep out-of-the-money and deep in-the-money calls: with the underlying equal to 5 and the zero skew case the CS model implies a 57% volatility and the CF a 42% volatility. On the other hand, for at-the-money and in near in-the-money options the CF model gives higher implied volatility values than the CS. Analyzing the absolute and relative valuation differences we can find that the biggest differences are found in out-of-the-money options for relative differences and in at-the-money options for absolute differences. These findings allow as to conclude that the implied density of the CS presents heavier tails on the negative part than the CF model. It is also interesting to note that the CS model gives a hump shaped volatility pattern for deep out-of-the-money options, which the CF model does not exhibit, displaying a maximum implied volatility for an underlying value of near 6. This "anomaly" can be explained because of the negative density values included in the CS model and is not observed in real volatility smile representations (e.g. Jondeau and Rockinger 2000).

2.3.3 Hedging Parameters

In this subsection we consider the CFD model to calculate the hedging parameters (the *Greeks*, as they are more commonly known). The most important of them, the *Delta*, Δ,

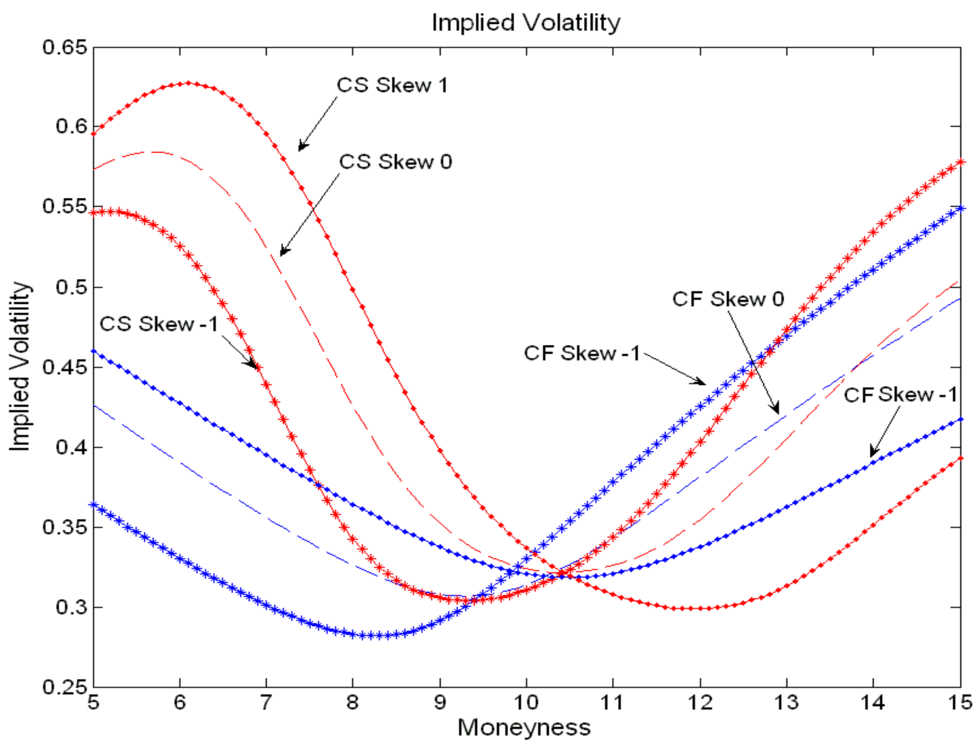


Fig. 2.8. Comparison of the implied volatility of both models for call options with a strike of $K = 10$ and times to maturity of one month. The annual volatility is supposed to be 40%, the risk free interest rate is equal to 5% and we consider different parameters a_2 and a_3 for the CF model and μ_3 and μ_4 for the CS. In particular, as in the previous Figures we assume parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness, μ_3, and kurtosis, μ_4, coefficients in the CD model of (0,8), (-1,8) and (1,8), respectively. The blue lines represent the CF volatilities and the red lines the CS ones.

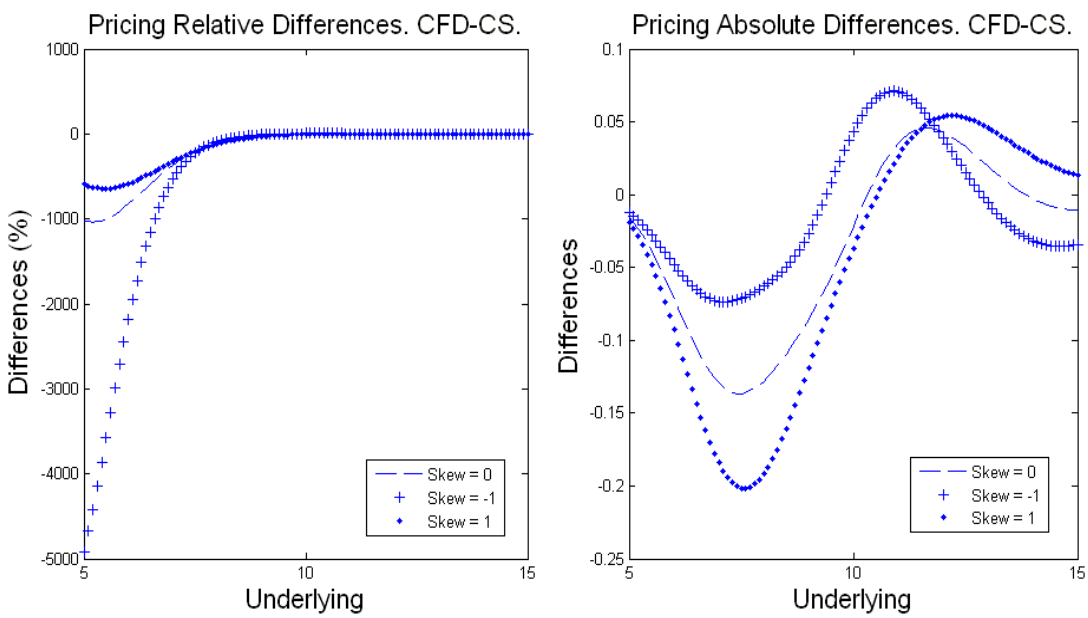


Fig. 2.9. Analysis of the relative and absolute differences of european call prices calculated using the CFD and the CS model. The set of parameters is the same as the one used in Figure 2.8.

determines the sensitivity of the option price to the underlying asset price. *Gamma*, Γ, is defined as the sensitivity of the Delta with respect to the underlying asset price and the *Vega* measures the dependence to the option price to volatility shifts.

Although completely analytical, the expressions for the Greeks in the CF model are rather cumbersome. Therefore, we will just report here the Delta and the Gamma as examples:

$$
\begin{aligned}
\Delta_{CF} &= \phi(d)\sqrt{\tau}\sigma \left(\begin{array}{c} \Phi(-d) + \phi(d)\left(\left(S_t\left(1+r\tau\right) - K\right) d_K\right) + \\ \left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) - S_t d d_K \left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) + \\ S_t\left(a_2 d_k + 2a_3 d d_K\right) \end{array} \right) \\
\Gamma_{CF} &= \phi(d)\sqrt{\tau}\sigma \left(\begin{array}{c} \phi(d)\left(S_t\left(1+r\tau\right) - K\right)\left(d d_K^2 - d_{KK}\right) + \\ S_t\sigma\left(a_2 d_k + 2a_3 d d_K\right) - 2\phi\left(d\right) d_K\left(1+r\tau\right) + \\ -2 d d_K\left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) \\ S_t\left(a_1 + a_2 d + a_3\left(1+d^2\right)\left(d^2 d_K^2 - d_K^2 - d d_{KK}\right)\right) + \\ 2\left(a_2 d_K + 2a_3 d d_K\right) + S_t\left(a_2 d_{KK} + 2a_3\left(d_K^2 + d d_{KK}\right)\right) \end{array} \right) \\
\text{with} &: \\
d &= \tilde{Q}_\tau^{-1}(K/S_t) \\
d_K &= \frac{-K}{S_t^2}\frac{\partial \tilde{Q}_\tau^{-1}(x)}{\partial x}\bigg|_{x=K/S_t} \\
d_{KK} &= \frac{K^2}{S_t^4}\frac{\partial^2 \tilde{Q}_\tau^{-1}(x)}{\partial x^2}\bigg|_{x=K/S_t} - 2\frac{K}{S_t^3}\frac{\partial \tilde{Q}_\tau^{-1}(x)}{\partial x}\bigg|_{x=K/S_t} \\
\tilde{Q}_\tau(x) &= \sigma\sqrt{\tau}\left(a_3 x^3 + a_2 x^2 + \left(\sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3\right) x - a_2\right) + (1 + r\tau)
\end{aligned}
$$

These expressions can be readily obtained derivating the expression for the price of a call option under the CFD hypothesis (Equation 2.19) with respect to S_t, once (Delta) or twice (Gamma), bearing in mind that the coefficient $d = Q_\tau^{-1}(K)$ also depends on S_t and defining $\tilde{Q}_\tau(x) = Q_\tau(x)/S_t$. In Table of Figures 2.13 we illustrate the shape of the CFD Greeks for the same parameters used in Figures 2.6 and 2.7 as well as the Black-Scholes counterparts. The most interesting case, as it is the one often found in practice, is the one with negative skew. As we can see, when investors find negative return values more plausible than positive ones, and the underlying is deep out-of-the-money the delta function is closer to zero, but when the underlying gets at-the-money the increase of the Delta is much more sharpen than in the Black-Scholes case. This feature is corroborated by the Gamma function, as we can see how the Gamma values when the option is near at-the-money for the CFD are much higher than for the BS. This phenomenon is crucial for hedgers which hedge a long position in an option with a position equal to the Delta in the underlying, because if they do not consider negative skew and heavy tails they can suffer from open positions. Even if they take care of it and model heavy tails, hedging an option in a non-gaussian world is intrinsically more difficult as the Delta function is more sharpen.

When the European call market price is given by the Corrado and Su (Corrado and Su 1996a, Corrado and Su 1997b) formula, the *Delta* and *Gamma* of a call options can be written (Backus, Foresi, Li, and Wu 1997) as:

$$
\begin{aligned}
\Delta_{CS} &\simeq \Phi(d) + \phi(d)\left\{\frac{\gamma_1}{6}\left(d^2 - 3d\sigma\sqrt{\tau} + 2\sigma^2\tau - 1\right)\right. \\
&\quad \left. + \frac{\gamma_2}{24}\left(-d^3 + 4d^2\sigma\sqrt{t} + 3d - 3d\sigma^2\tau - 4\sigma\sqrt{\tau}\right)\right\} \\
\Gamma_{CS} &\simeq \frac{\phi(d)}{S_0\sigma\sqrt{\tau}}\left\{1 + \frac{\gamma_1}{6}\left(-d^3 + 3d^2\sigma\sqrt{\tau} - 2d\sigma^2\tau + 3d - 3\sigma\sqrt{\tau}\right)\right. \\
&\quad \left. + \frac{\gamma_2}{24}\left(d^4 - 4d^3\sigma\sqrt{\tau} + 3d^2\sigma^2\tau - 6d^2 + 12d\sigma\sqrt{\tau} - 3\sigma^2\tau + 3\right)\right\} \\
d &= \frac{\ln\left(S_0/K\right) + \left(r + \sigma^2/2\right)\tau}{\sigma\sqrt{t}}
\end{aligned}
$$

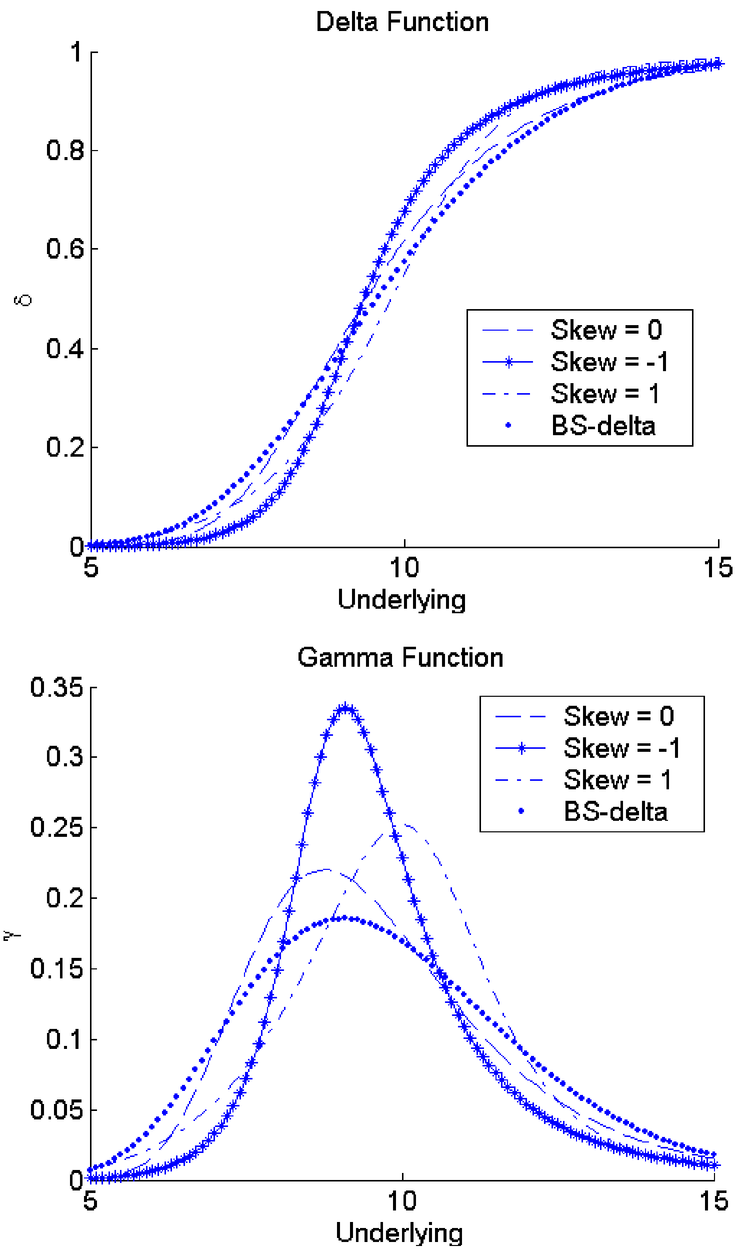


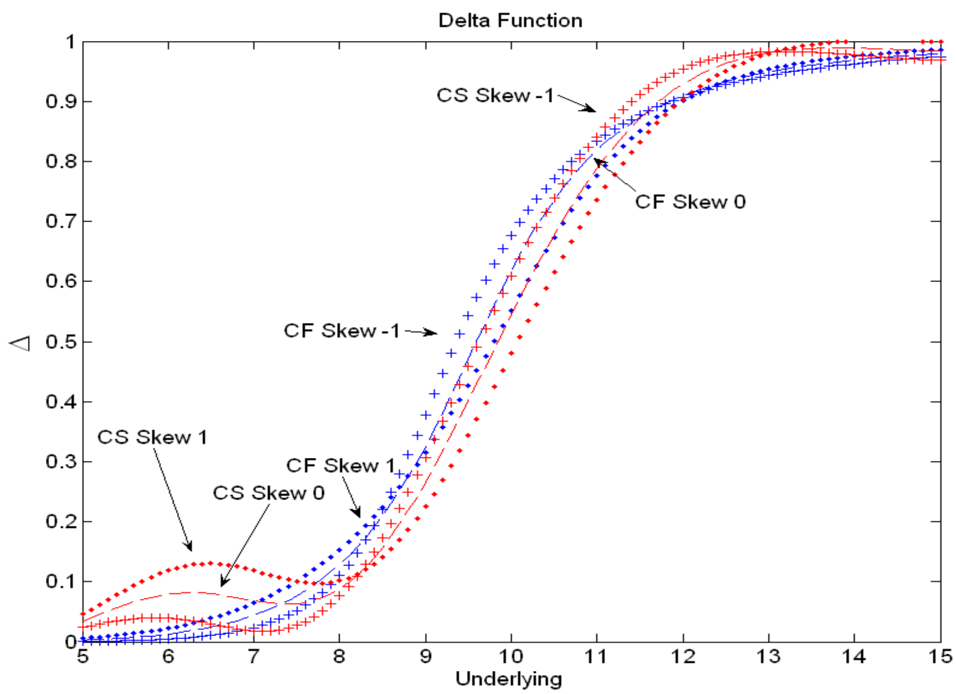
Table 2.13. Delta and Gamma functions assuming that the true implied distribution is CFD. Call options have a strike of K=10, time to maturity of one month, the annual volatility is 40% and the risk free interest rate is equal to 5%. We assume parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness and kurtosis coefficients of (0,8), (-1,8) and (1,8), respectively.

In Table of Figures **??** and **??** we compare the Delta and Gamma functions derived from the CF and CS models for similar degrees of asymmetry and kurtosis, considering the same parameters used in Figures 2.6, 2.7 and 2.8. In Table **??** we present the Delta and Gamma values and in Table **??** the relative and absolute differences between both models. From the Delta function it can be observed that for both deep in and out-of-the-money options the CS model gives higher Delta values than the CF, and viceversa for at-the-money options. These differences are significant as they can be as high as 0.1 in absolute value or 20% in relative one. On the other hand, the Gamma functions present also substantial differences: at-the-money options present similar Gamma values for both models but for deep out or in-the-money options the CF model derives more reliable results given that the CS model presents negative values, which are theoretically unsound for call options (Hull 2004). Again, this anomaly (the Delta function presents a local maxima) can be explained through the negative density values and through the fact that de Delta and Gamma expressions for the CF model are only approximations and cannot be calculated analytically. Therefore, before turning to the empirical comparison between both models, we can conclude that the CF option pricing model is theoretically more coherent than the CS model as it does not derive "dubious" hedging values for deep out or in-the-money options.

2.3.4 Empirical Performance of CFD Option Pricing

Data Description and Estimation Procedure

This analysis on the empirical performance will be based on the Spanish Futures and Options market (MEFF) for the most actively traded options contracts, namely, options on



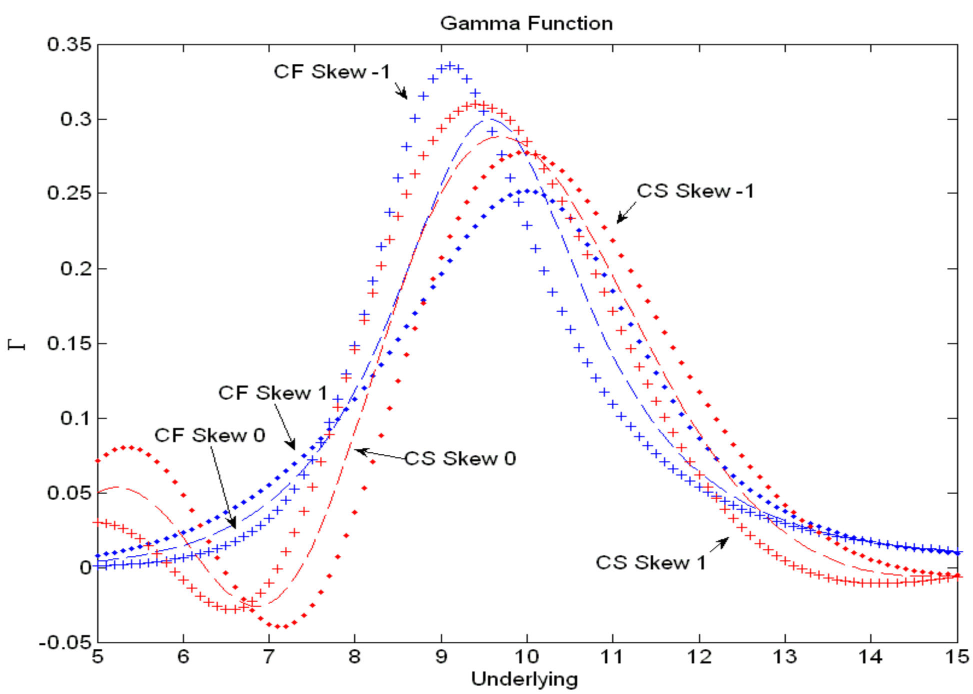


Table 2.14. Comparison of the Delta and Gamma functions between the CF and CS models. Call options have a strike of K=10, time to maturity of one month, the annual volatility is 40% and the risk free interest rate is equal to 5%. We assume parameters a_2 and a_3 equal to (0.092,0), (0.079 -0.11) and (0.079 0.11) which correspond to skewness, μ_3, and kurtosis, μ_4, coefficients of (0,8), (-1,8) and (1,8), respectively. The blue lines represent the CF volatilities and the red lines the CS ones.

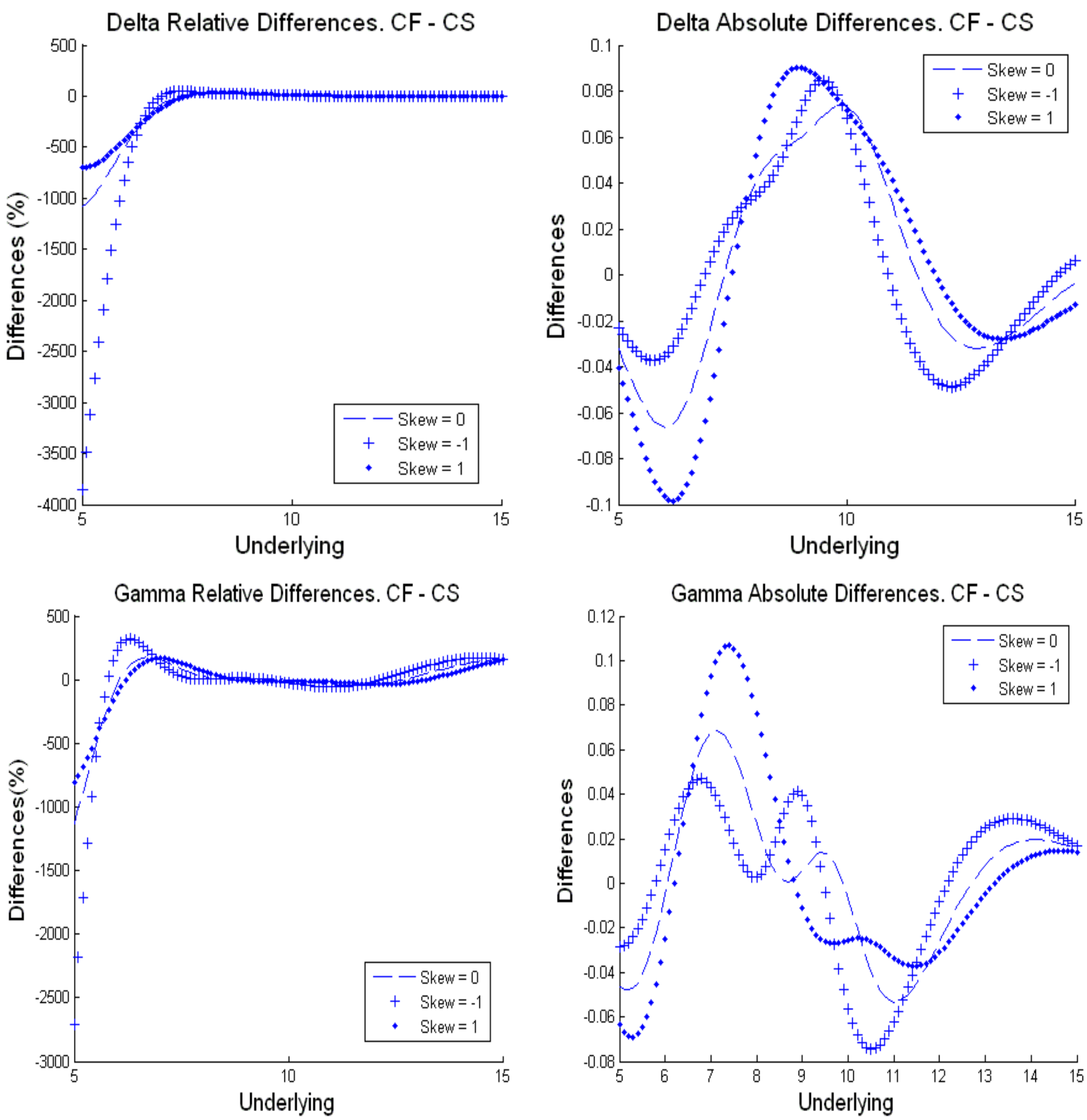


Table 2.15. Analysis of the relative and absolute differences of Delta and Gamma functions using the CFD and the CS model. The set of parameters is the same as the one used in the Table of Figures **??**.

futures over the IBEX-35 index. The Spanish Ibex-35 index is a value-weighted index comprising the 35 most liquid Spanish Stocks traded in the continuous auction market system. The option contract that we consider is a cash settled European option with trading over three nearest consecutive months and the other three months of the March-June-September-December cycle. The expiration day is the third Friday of the contract month. Prices are quoted with a minimum price change of one index point, and the exercise prices are given by 50 index point intervals. The database considered in Ferreira, Gago, León, and Rubio 2005[28], contains all call and put options on the Ibex-35 Index futures traded daily on MEFF during the period February 1996 to November 1998 and liquidity is concentrated on the nearest expiration contract. We restrict our attention to all options transacted from 11:00 to 16:45 in order to avoid the well known intra day cycle. Thus, every trade recorded during this window is used in the estimation. For this sample, about 90% of the crossing transactions took place with these contracts. In each trade, we collect its transaction option price, the simultaneous future price (F) as measured by its bid-ask spread average, the exercise price (X), the expiration date and the annualized repo T-bill rates with approximately the same maturity as the option. Those calls and put prices that violate arbitrage bounds and also the options with a market value of less than 5 Spanish pesetas are eliminated[29].

The final daily sample is 37807 options (22099 calls and 15708 puts). Moneyness is defined as the ratio of the exercise price to the futures price. A call (put) option is said

---

28 We are very thankful to the authors for making available the data.

29 The upper bound for an European call option is given by $C \leq S_0$ where C is the price of the option and S_0 is the initial price of the underlying. On the other hand, the lower bound for the price of an European call on a non-dividend paying stock (as is the case for the futures contract) is given by $C \geq S_0 - Ke^{rT}$ (Hull 2004).

to be out-of-the-money (in-the-money) when $K/F > 1.015$, at-the-money when $1.015 > X/F > 0.985$ and in-the-money (out-of-the-money) when $0.985 > X/F$.

We compare the performance of the CFD option valuation with the standard option valuation for the European call on futures of Black 1976 (B76) and with the model of Corrado and Su (CS, Corrado and Su 1996a). We carry out both in-sample and out-of-sample analysis with a one day time horizon. Our analysis is based on the following sample statistics: the mean error (ME), the mean absolute error (MAE) and the root of the mean squared errors (RMSE). All errors will be measured in relative value in order to weight each option equally[30]. The parameters of each model, the volatility σ for the Black76 model, the volatility σ and the parameters a_2 and a_3 for the CFD model and the volatility σ, skewness, ξ , and kurtosis coefficients, κ, for the CS model, are estimated for each of the 601 days of in the sample-period and have been computed from the cross-section of options prices. The implicit estimator of date t is defined as the minimizer of the mean of the squared pricing errors for the options traded that day, that is,

$$\hat{\theta}_t = \arg\min_{\theta} \frac{1}{n_t} \sum_{i=1}^{n_t} [c_i(\theta) - c_i]^2$$

where $c_i(\theta)$ is the theoretical option price, c_i, denotes the market option price and n_t the total number of options at date t. Refer to Appendix C.1.5 for details on the optimization procedure.

30 In these calculations we remove the dates where the relative mean error is bigger than 1000%, as in these cases the true value of the option was close to zero (we only found 6 dates).

Optimization Results

Next, we analyze the results of the optimizations. In Table 2.16 we present the results of the three models (B76, CFD and CS) analyzing separately the fitting for calls and puts, and for options out, at or in-the money. Table 2.16 shows that the CFD clearly beats B76 since both MAE and RMSE are lower under CFD for the whole sample as well as for every Moneyness level. With respect to the ME statistics we conclude that while the B76 model clearly fails to valuate calls the CFD works really well, as the ME for Calls (Puts) is 0.65 (0.27) for the CFD model while it is 10.45 (-0.84) for the B76 model. On the other hand, comparing the CS model with the CFD, we observe that both models perform similarly. For instance, mean errors are 0.65% (CFD) and 0.10% (CS) for the calls and 0.27% (CFD) and 0.22% (CS) for the puts and the mean absolute errors are quite similar 5.22% (CFD) versus 5.44% (CS) for the calls and 2.28% (CFD) versus 2.29% (CS).

Finally, given the parameters used to calculate the in-sample results for day t we use them to predict prices at date $t + 1$. We calculate the statistics ME, MAE and RMSE for each date and the results for this out-of-sample analysis are shown in Table 2.17. Now, although the CFD model remains almost unbiased (the pricing error for both total calls and puts is less than 1% in absolute value), the dispersion statistics (MAE and RMSE) now give values that are closer to the B76 model ones. Nevertheless, both dispersion statistics remain smaller for every moneyness level and, therefore, we can conclude that the CFD model outperforms the Black-Scholes model also in an out-of-sample sense. Comparing the CS model and the CFD model we observe again that the out-of-sample performance of both models remains very similar: the mean errors are 0.77% (CFD) and -0.25% (CS) for

	Calls (B76)			Calls(CFD)			Calls(CS)		
	ME	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**
OTM	18.50	20.11	30.34	0.68	7.10	11.62	-0.42	7.51	13.31
ATM	2.10	4.34	6.35	0.64	3.34	4.99	0.71	3.36	5.03
ITM	-2.54	3.04	3.85	0.51	1.73	2.50	0.47	1.72	2.49
Total	10.45	12.55	22.39	0.65	5.22	9.04	0.10	5.44	10.20

	Puts (B76)			Puts(CFD)			Puts(CS)		
	ME	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**
ITM	17.45	22.99	42.49	6.18	15.13	28.83	5.73	15.08	28.87
ATM	-0.19	3.46	5.62	-0.06	2.79	4.69	-0.07	2.78	4.69
OTM	-2.75	2.84	3.28	-0.03	0.91	1.48	-0.06	0.92	1.49
Total	-0.84	4.07	10.39	0.27	2.28	7.11	0.22	2.29	7.14

Table 2.16. Daily in-sample performance of alternative option pricing model (Black 1976, (B76), Cornish-Fisher (CFD) and Corrado-Su 1996 (CS) . ME, MAE and RMSE stand for Mean Errors, Mean Absolute Errors and Root Mean Squared Errors, while OTM, ATM and ITM denote out-of-the money, at-the-money and in-the-money options.

the calls and -0.68% (CFD) and -0.81% (CS) for the puts. In addition, we can observe that the dispersion statistics are also very similar between these models.

On the other hand, it is interesting to analyze the evolution of the implied volatility for both models and the evolution of the kurtosis coefficient for the CFD model. In Figure 2.10 we can observe that both volatilities, the one implied by the Black-Scholes model and the one implied by the CFD are very similar. However, the evolution of the CFD volatility is slightly sharper. Figure 2.11 shows the evolution of the kurtosis coefficient and we can see that some days the implied distribution of agents is much more leptokurtic than other Edgeworth based models can present. Therefore, this fact corroborates the statement that Edgeworth based distributions are too limited when capturing high degrees of kurtosis, that are often found in reality to be greater than the typical limit of eight corresponding the CS model.

	Calls (B76)			Calls(CFD)			Calls(CS)		
	ME	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**
OTM	18.02	24.32	36.65	0.22	15.23	23.83	-1.76	16.00	24.05
ATM	2.77	6.76	10.00	1.44	6.06	8.63	1.50	6.04	8.55
ITM	-2.36	3.38	4.72	0.88	2.95	4.12	0.89	2.95	4.06
Total	10.50	15.80	27.38	0.77	10.78	18.19	-0.25	11.11	18.32

	Puts (B76)			Puts(CFD)			Puts(CS)		
	ME	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**	**ME**	**MAE**	**RMSE**
ITM	-8.46	24.71	43.34	-3.83	21.23	31.92	-4.85	21.32	31.97
ATM	-1.41	5.59	9.11	-1.09	5.17	7.88	-1.11	5.11	7.80
OTM	–3.18	3.45	4.69	-0.26	1.93	3.04	-0.30	1.94	3.00
Total	-1.98	5.27	11.68	-0.68	4.07	8.85	-0.81	4.03	8.84

Table 2.17. Daily out-of-sample performance of alternative option pricing. The legend is the same as in Table 2.16.

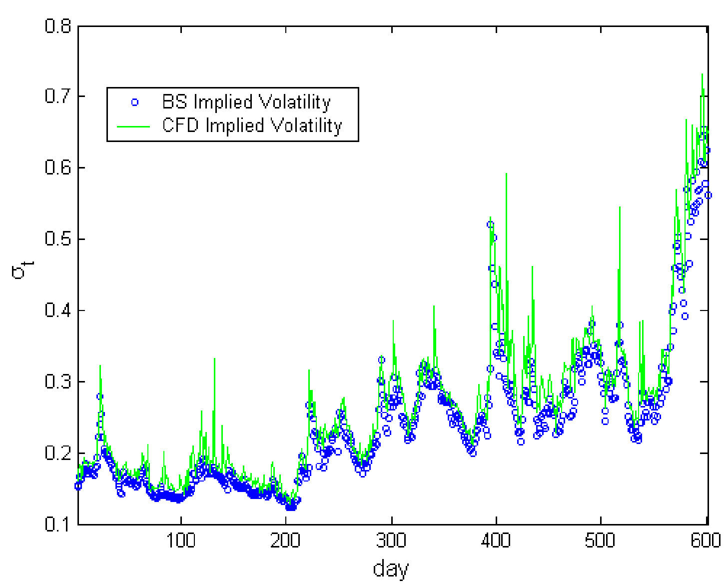


Fig. 2.10. Time evolution of the implied volatility for the Black-Scholes model and the Cornish-Fisher model

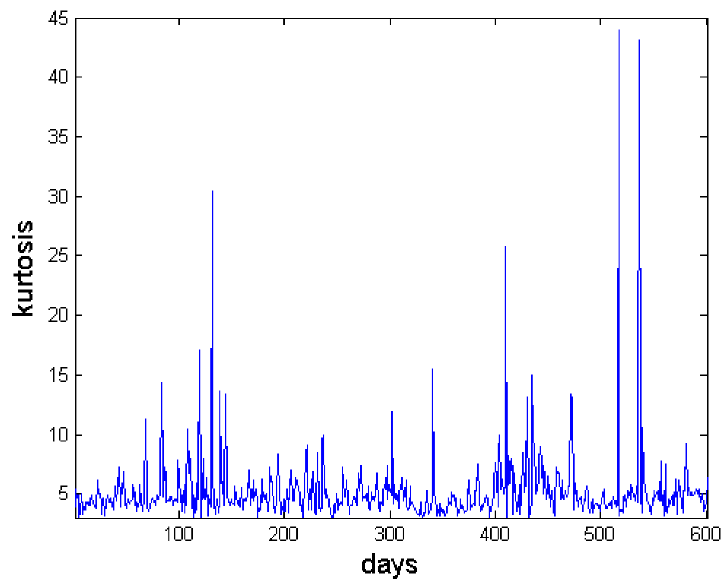


Fig. 2.11. Time evolution of the kurtosis implied by the Cornish-Fisher model.

Therefore, we can conclude that the CFD model, both in-sample and out-of-sample, clearly out-performs the classical Black-Scholes model, which does not include heavy tails, and that the CFD model gives similar prediction results compared to the Corrado and Su model which does include asymmetry and heavy tails. Nevertheless, it is interesting to point out that the implied densities of the Corrado and Su model present negative values for some days in the sample. As an example, in Figure 2.12 we present the implied density functions derived from the Corrado-Su (CS), Cornish-Fisher (CFD) and Black-Scholes (BS) models for the dates 06/06/1996 and 08/09/1997, considering options with three months to expiration. The implied BS density, obviously, does not present heavy tails or asymmetry, while the CS and CF do show negative asymmetry and a considerable left heavy tail. These findings are in agreement with the results in the literature (e.g. Jondeau and Rockinger 2000)

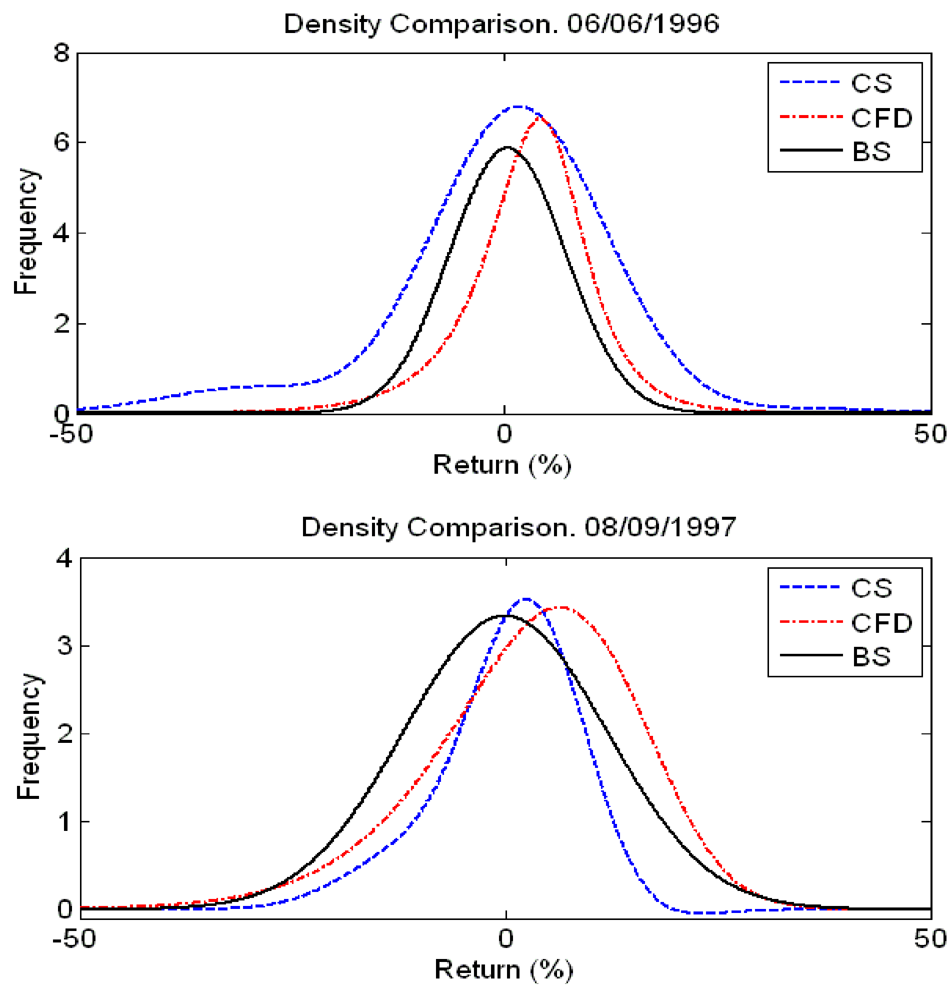


Fig. 2.12. Comparison of the implied density functions derived from the Corrado-Su (CS), Cornish-Fisher (CFD) and Black-Scholes (BS) models for the dates 06/06/1996 and 08/09/1997.

and demonstrate that options trading agents assign a higher probability to large losses than to large wins. In addition, in the upper graph we note that the fitted CS density is bimodal and in the lower one we note that it derives negative density values, fact that could be expected given that in Figure 2.11 we observe kurtosis coefficient much bigger than eight. Therefore, although the fitting quality results are similar between the CS and the CFD, the CS model possess a theoretical drawback that the CFD model does not.

2.3.5 Conclusions

In this Section we have derived closed-form option price and hedging parameters formulas assuming that prices differences follow a third-order CFD under the risk-neutral measure. In this way, we have proposed a new semi-parametric generalization of the Black and Scholes option valuation formula (Black and Scholes 1973) to include underlyings whose returns can be characterized by high degrees of skewness and kurtosis. In addition, we have shown that hedging parameters in the CFD option pricing model do not present the same anomalies as negative Gamma values for deep out or in-the-money options like in the CS model, although being both from the semi-parametric class.

In the empirical application to Spanish options data we have evaluated the performance of our pricing formulas using the Black-Scholes model as a benchmark, finding that the CFD model beats the Black-Scholes whether in-sample or out-of-sample. We have also compared the CS and CF model finding that the in-sample and out-of-sample performance of both models remains very similar. In addition, we have shown that the CS presents negative and bimodal implied density values while the CFD model does not.

Therefore, given first that both models present the same predictive power and second, that the CF does not present the anomaly in the hedging parameters, we can conclude that the CF model is preferable than the CS model within the semi-parametric class of distributions for valuing options.

Comparison of the option pricing performance and of the modeling flexibility between the Cornish-Fisher distribution and other option pricing model that also allow for heavy-tailedness but do not belong to the semi-parametric class as a non-parametric method based on a mixture of log–normal densities, a parametric approach of Malz which assumes a jump-diffusion for the underlying process (Malz 1996), or Heston's approach assuming a stochastic volatility model (Heston 1993), could be of interest and this is left for future research.

Appendix A

Third-order Cornish-Fisher Density

The third-order Cornish-Fisher Density with parameters a_3, a_2, a_1 and a_0 is given by:

$$cf_3(R) = \frac{d\left[Q^{-1}(R)\right]}{dR} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[Q^{-1}(R)\right]^2} \tag{A.1}$$

where $Q^{-1}(R)$ is the inverse of the third-order polynomial, $a_3X^3 + a_2X^2 + a_1X + a_0 = R$, and $d\left[Q^{-1}(R)\right]/dR$ is the first derivative of this inverse function with respect to R. Both the inverse function of a third-order polynomial and its first derivative can be calculated analytically and are given by:

$$\begin{aligned} Q^{-1}(R) &= -\frac{a_2}{3a_3} + \sqrt[3]{\sqrt{B_1}R + a_3B_2 + \sqrt{B_1R^2 + B_2R + B_3}} - \\ &\quad \frac{1}{3}\frac{a_1/a_3 - a_2^2/3a_3^2}{\sqrt[3]{\sqrt{B_1}R + a_3B_2 + \sqrt{B_1R^2 + B_2R + B_3}}} \end{aligned} \tag{A.2}$$

and:

$$\begin{aligned} \frac{d\left[Q^{-1}(R)\right]}{dR} &= \frac{\sqrt{B_1} + \frac{\left(\frac{1}{2}B_2 + RB_1\right)}{\sqrt{B_3 + RB_2 + R^2B_1}}}{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}} \times \\ &\quad \left(\begin{array}{c} \frac{1}{3}\sqrt[3]{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}} + \\ \frac{1}{27a_3^2}\frac{3a_3a_1 - a_2^2}{\sqrt[3]{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}}} \end{array} \right) \end{aligned} \tag{A.3}$$

where the coefficients B_1, B_2 and B_3 depend on the coefficients a_0, a_1, a_2 and a_3 and are given by:

$$
\begin{aligned}
B_1 &= \frac{1}{4a_3^2} \\
B_2 &= \frac{a_2 a_1}{6a_3^3} - \frac{a_0}{2a_3^2} - \frac{a_2^3}{27a_3^4} \\
B_3 &= -\frac{a_2 a_1 a_0}{6a_3^3} + \frac{a_0^2}{4a_3^2} + \frac{a_1^3}{27a_3^3} + \frac{a_2^3 a_0}{27a_3^4} - \frac{a_2^2 a_1^2}{108a_3^4}
\end{aligned}
$$

A.1 Other distributions related to the Cornish-Fisher Distribution

We present here some univariate distributions somehow related to the CFD function and the skewed student-t (Hansen 1994) that we use for comparison purposes. On the following, we will just outline their definitions and main properties, for a more extensive presentation see Kendall, Stuart, and Ord 1994.

Gram-Charlier Series of Type A and Edgeworth Series (1879, 1904)

Instead of expanding the QQ-Plot function, it is proposed that a general density function $f(x)$ can be expanded in a series of derivatives of the standard gaussian density function $\phi(x)$. Then:

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x)\phi(x) \tag{A.4}$$

where $H_j(x)$ is the j-th order Hermite polynomial (Abramowitz and Stegun 1964) and c_j are coefficients given by:

$$c_r = \frac{1}{r!}\left\{\mu'_r - \frac{r^2}{2 \cdot 1!}\mu'_{r-2} + \frac{r^4}{2^2 \cdot 2!}\mu'_{r-4} - \dots\right\}$$

where μ'_r are the non-central moments of $f(x)$. It is important to note here, that although Edgeworth and Gram-Charlier densities have the same formal definition (Equation A.4), given that for practical purposes it is necessary to take a finite number of terms the finite sum of, this density is different in both cases. For example, in the Gram-Charlier case we find the following formal expansion:

$$f(x) = \phi(x)\{1 + \frac{1}{2}(\mu_2 - 1) H_2 + \frac{1}{6}\mu_3 H_3 + \frac{1}{24}(\mu_4 - 6\mu_2 + 3) H_4 + ...\}$$

where μ_r are the central moments of $f(x)$. Therefore, considering the first four empirical moments of the data, we can fit the distribution up to fourth order. This expansion has been frequently used in finance, for example, by Jarrow and Rudd 1982 in option pricing theory. However, a major drawback of this approximation is that it provides negative probabilities for high kurtosis, for example, as pointed out in Kendall, Stuart, and Ord 1994. Jondeau and Rockinger 2001 calculate the permitted parameter range in order to keep the density always positive, and find that it becomes very restrictive.

Functional transformations to normality: Johnson distributions (1949)

Johnson considers a general transformation to normality of the type:

$$R = \gamma + \delta g\left(\frac{X - \mu}{\lambda}\right)$$

where μ, λ, γ and δ are parameters at choice and g is some convenient function. In particular, he considers three types of systems:

1. The R_L or lognormal system: $g(y) = log(y)$.

2. The R_B (bounded range) system: $g(y) = log(y/1 - y)$.

3. The R_U (unbounded range) system: $g(y) = sinh^{-1}(y)$.

As financial returns are unbounded, we would be more interested in the third one. R_U distributions are unimodal and its mode lies between the median, $sinh(-\gamma/\delta)$, and zero. This distribution is positively or negatively skewed according to γ being negative or positive. Besides that, it is interesting to note that a R_U variable, presents an hyperbolic sinusoidal form QQ-Plot.

Skewed student-t: Hansen (1994)

The skewed student-t density $t(x)$ with parameters ν and λ is defined as:

$$t(x) = bc\left(1 + \frac{\zeta}{\nu - 2}\right)^{-\frac{\nu+1}{2}}$$

where:

$$\begin{aligned}
\zeta &= \begin{cases} (bx + a)/(1 - \lambda) & \text{if} \quad x < -a/b \\ (bx + a)/(1 + \lambda) & \text{if} \quad x > -a/b \end{cases} \\
a &= 4\lambda c \frac{\nu - 2}{\nu - 1}, \qquad\qquad b = 1 + 3\lambda^2 - a^2 \\
c &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\,(\nu - 2)}\Gamma\left(\frac{\nu}{2}\right)}
\end{aligned}$$

with $\Gamma(x)$ being the Gamma function. Parameters a and b are required to center and scale the asymmetric distribution. The probability density is well-defined if $\nu > 2$ and $-1 < \lambda < 1$. If $\lambda = 0$ the density reduces to the standard t distribution. If in addition $\nu > 4$ the density will have positive and finite excess kurtosis.

In addition, the multivariate Skewed student-t density is defined as (Jondeau and Rockinger 2003):

$$Mt(x) = \prod_{i=1}^{n} b_i c_i \left(1 + \frac{\zeta_i}{\nu_i - 2}\right)^{-\frac{\nu_i+1}{2}}$$

where b_i, c_i, ζ_i and ν_i have the same expression as above. With this multivariate Skewed student-t, the variables x_i are independent and each follows a univariate skewed student-t distribution.

Semi non-parametric distributions (2005)

The *semi non-parametric* distribution is defined as a variable $z = a + bx$ where x is distributed with the following density:

$$f(x) = \frac{\phi(x)}{\nu'\nu}\left(\sum_{i=0}^{m} \nu_i H_i(x)\right)^2$$

where ν is an $m+1$ dimensional real vector (i.e. $\nu \in R^{m+1}$), $\phi(x)$ denotes the probability density function of a standard normal distribution, and $H_i(x)$ is the normalized Hermite polynomial of order i. The density function of z is therefore given by:

$$g(z) = \frac{1}{b}\frac{\phi((z-a)/b)}{\nu'\nu}\left(\sum_{i=0}^{m} \nu_i H_i((z-a)/b)\right)^2$$

In this context, a is the location parameter, b as a scale parameter and the parameters ν are the form parameters. It is interesting to note that this density is basically a squared Gram-Charlier density and, therefore, it cannot be negative.

Appendix B

Proofs

Lemma 1

Let $cf_3(R)$ be the third-order Cornish-Fisher density function defined by Equation 1.9 *for a random variable R, with coefficients a_i, ($i = 0, 1, 2, 3$), then the sufficient and necessary conditions on the coefficients to guarantee the existence of the CFD are*

$$a_3 > 0 \quad , \quad a_1 > 0 \quad , \quad -\sqrt{3a_3a_1} < a_2 < \sqrt{3a_1a_3}$$

Proof. In order to guarantee that a third-order polynomial $Q(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ is invertible it is enough and necessary to impose that it is a strictly increasing function. In consequence, the condition of positive derivative must hold for every point x:

$$Q'(x) = 3a_3x^2 + 2a_2x + a_1 > 0$$

This equation represents a parabola that must be positive for every point x, and is equivalent to impose the following conditions: existence of a unique minimum and a positive function value in this minimum. The first one implies that it must exist a unique solution, x_m, to $Q''(x_m) = 6a_3x_m + 2a_2 = 0$, which is always verified, and that $Q'''(x_m) > 0$, which gives

$a_3 > 0$. The second conditions implies:

$$\begin{aligned} Q'(x_m) &= 3a_3x_m^2 + 2a_2x_m + a_1 \\ &= a_1 - \frac{1}{3}\frac{a_2^2}{a_3} > 0, \end{aligned}$$

which implies the following condition:

$$\begin{aligned} -\sqrt{3a_3a_1} &< a_2 < \sqrt{3a_1a_3} \\ a_1 &> 0 \end{aligned}$$

The result of this Lemma could be easily generalized to a fifth order polynomial. ■

Proposition 2

Let $cf_m(R)$ be the m-th order Cornish-Fisher density function defined by Equation 1.9 for the random variable R, then non-central r-th order moments, μ_r, are given by:

$$\mu_r' = E\left[R^r\right] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^r e^{\frac{1}{2}J^2}\Bigg|_{J=0}$$

where $Q\left(\frac{\partial}{\partial J}\right) = \sum_{i=1}^m a_i \frac{\partial^i}{\partial J^i}$ is a differential operator.

Proof. Consider the Fourier transform of the aleatory variable R:

$$\hat{P}(k) = E\left[e^{ikR}\right]$$

where $E\left[\cdot\right]$ denotes its expected value. We perform the variable change $R = Q(X)$:

$$\hat{P}(k) = E\left[e^{ikQ(X)}\right] \Leftrightarrow$$

$$\hat{P}(k) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + ikQ(X)} dX$$

Expanding the exponential in power series we obtain:

$$\hat{P}(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \int \frac{1}{\sqrt{2\pi}} Q^m(X) e^{-\frac{1}{2}X^2} dX$$

Using this expression we obtain the moments of the variable R can be obtained through the following integral:

$$E\left[R^m\right] = \int \frac{1}{\sqrt{2\pi}} Q^m(X) e^{-\frac{1}{2}X^2} dX \tag{B.1}$$

Next, we define the functional generator $\hat{P}(J)$ as:

$$\hat{P}(J) = \int \frac{1}{\sqrt{2\pi}} Q^m(X) e^{-\frac{1}{2}X^2 + JX} dX$$

where the variable J is introduced in order to calculate the integrals. If $Q(X)$ is a polynomial it is easy to demonstrate (through direct derivation) that Equation [B.1] reduces to:

$$E\left[R^m\right] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^m \cdot \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + JX} dX \Bigg|_{J=0} \tag{B.2}$$

The integral [B.2] can be carried out analytically:

$$\int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + JX} dX = e^{\frac{1}{2}J^2}$$

obtaining, finally, the following expression:

$$E\left[R^m\right] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^m \cdot e^{\frac{1}{2}J^2} \Bigg|_{J=0}$$

■

Lemma 3

Let $cf_3(R)$ be the third-order Cornish-Fisher density function defined by 1.9 *for the random variable R, with coefficients a_i, ($i = 0, 1, 2, 3$). Then, one can construct a standardized* *variable R with zero mean and unit variance imposing*

$$a_0 = -a_2 \quad , \ a_1 = \sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3$$

with the following conditions on a_2 and a_3 to guarantee the existence of the $cf_3(R)$:

$$0 \ < \ a_3 < \frac{1}{\sqrt{15}}$$
$$-\sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} \ < \ a_2 < \sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)}$$

Proof. For the first result one just has to impose $\mu_1' = 0$ and $\mu_2' = 1$ in Equations 1.11. In order to prove the second one we need to replace the conditions $a_0 = -a_2$ and $a_1 = \sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3$ in the inequation $-\sqrt{3a_3a_1} < a_2 < \sqrt{3a_1a_3}$, which guarantees the existence of the CFD. As a result, one finds:

$$-\sqrt{3a_3\left(\sqrt{1 - 2a_2^2 - 6a_3^2} - 3a_3\right)} < a_2 < \sqrt{3a_3\left(\sqrt{1 - 2a_2^2 - 6a_3^2} - 3a_3\right)}$$

Solving this equation for a_2 one obtains that this is equivalent to:

$$-\sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} < a_2 < \sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)}$$

Imposing that $\sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)}$ has to be a real number and that $a_3 > 0$, we obtain the last condition:

$$0 < a_3 < \frac{1}{\sqrt{15}}$$

■

Lemma 4

Let R we a m-th order CFD distributed variable with parameters $\{a_i\}_{i=1}^m$. Consider the variable $Z = m + \sigma R$, then, the new variable Z is also distributed as a CFD with parameters $\{a'_i\}_{i=1}^m$ given by

$$\begin{aligned} a'_i &= \sigma a_i \qquad\qquad i = 1, ..., m \\ a_0 &= \sigma a_0 + m \end{aligned}$$

Proof. To prove this Lemma we only have to consider the definition of a Cornish-Fisher distributed variable. In Equation 1.5 we see that a CF variable is given by:

$$R = \sum_{i=0}^{m} a_i X^i$$

where X is a standard gaussian variable. Therefore, the variable Z given by:

$$Z = m + \sigma R = m + \sum_{i=0}^{m} \sigma a_i X^i = m + \sigma a_0 + \sum_{i=1}^{m} \sigma a_i X^i,$$

which can be again rewritten as:

$$Z = a'_0 + \sum_{i=1}^{m} a'_i X^i$$

where:

$$\begin{aligned} a'_i &= \sigma a_i \qquad\qquad i = 1, ..., m \\ a_0 &= \sigma a_0 + m \end{aligned}$$

■

Proposition 5

Third-order Cornish-Fisher Densities are unimodal.

Proof. To prove this Proposition we will consider the expression of the density function of a third-order Cornish-Fisher distributed variable:

$$cf_3(R) = \frac{d\left[Q_3^{-1}(R)\right]}{dR}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[Q_3^{-1}(R)\right]^2}$$

Derivating this expression with respect to R and making $d\left(cf_3(R)\right)/dR = 0$, we find the condition for a maximum of the density function, R_m,:

$$\begin{aligned}\frac{d\left(cf_3(R)\right)}{dR} &= \left[\frac{d^2\left[Q_3^{-1}(R)\right]}{dR^2} - \left(\frac{d\left[Q_3^{-1}(R)\right]}{dR}\right)^2 Q_3^{-1}(R)\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[Q_3^{-1}(R)\right]^2} = 0 \\ &\Rightarrow \frac{d^2\left[Q_3^{-1}(R_m)\right]}{dR^2} - \left(\frac{d\left[Q_3^{-1}(R_m)\right]}{dR}\right)^2 Q_3^{-1}(R_m) = 0 \qquad \text{(B.3)}\end{aligned}$$

In order to demonstrate the unimodality of the CF_3 we have to prove that we have only one R_m. The function $Q_3^{-1}(R)$ is necessarily strictly increasing for the existence of the density and, therefore, as Q_3^{-1} is the inverse of a third-order polynomial, it can only have one inflexity point, R_i, where $d^2\left[Q_3^{-1}(R)\right]/dR^2$ is equal to zero, and one cross point with the x-axis, R_c, where $Q_3^{-1}(R)$ is equal to zero. For the shake of ease of readiness for this demonstration we will denote $d^2\left[Q_3^{-1}(R)\right]/dR^2$ by $Q_3^{-1}(R)''$. In the limits where $R \rightarrow \pm\infty$ we have $Q_3^{-1}(R) \rightarrow \pm\infty$ and $Q_3^{-1}(R)'' \rightarrow 0$. We have therefore three possible cases: i) $R_c = R_i$, ii) $R_c > R_i$, and iii) $R_c < R_i$. If we demonstrate that there is only one R_m for each case we will prove the proposition.

i) $R_c = R_i$. When $R_c = R_i$, then R_m coincides with them, $R_m = R_c = R_i$. As we have that $Q_3^{-1}(R) > 0$ and $Q_3^{-1}(R)'' < 0$, Equation B.3 cannot hold. When $R < R_c$ then

$Q_3^{-1}(R) < 0$ and $Q_3^{-1}(R)'' > 0$ and also Equation B.3 cannot hold. Therefore we have only a maximum R_m.

ii) $R_c > R_i$. If $R > R_c > R_i$ then $Q_3^{-1}(R) > 0$ and $Q_3^{-1}(R)'' < 0$ and Equation B.3 cannot hold. If $R < R_i < R_c$ then $Q_3^{-1}(R) < 0$ and $Q_3^{-1}(R)'' > 0$ and Equation B.3 cannot hold. When $R_c > R > R_i$ then $Q_3^{-1}(R) < 0$ and $Q_3^{-1}(R)'' < 0$ and, therefore, there exists at least one point R_m where Equation B.3 holds. But there can only be one R_m because the function $Q_3^{-1}(R)$ is strictly increasing, $Q_3^{-1}(R)' > 0$, and $Q_3^{-1}(R)''$ is strictly decreasing and, therefore, Equation B.3 can only hold once.

ii) Finally, the third case is very similar to the second. ■

Proposition 6

Let $(R_i)_{i=1}^n$ be the daily returns of n assets that follow a Copula-Based Multivariate Cornish-Fisher Density CB-MCFD$_m$ and κ be the normal rank correlation *matrix of these assets. The Profit and Loss (P&L) distribution of a portfolio with weights $(\omega_i)_{i=1}^n$ corresponding to these assets, constrained to $\sum_{i=1}^n \omega_i = 1$, is given by:*

$$P = \sum_{i=1}^{n} \omega_i R_i$$

Then, the k-order non-centered moments of the aleatory variable P are given by:

$$E\left[P^k\right] = \left[\sum_i \omega_i Q_i(\frac{\partial}{\partial J_i})\right]^k \cdot e^{\frac{1}{2}J\kappa J^t}\Bigg|_{J=0}$$

where $J = (J_i)_{i=1}^n$ is an auxiliary n-dimensional vector.

Proof. Consider the Fourier transform of the aleatory variable P:

$$\hat{P}(k) = E\left[e^{ik\sum_{i=1}^{n}\omega_i R_i}\right]$$

where $E\left[\cdot\right]$ denotes the expected value. We perform the variable change $R = Q(X)$:

$$\hat{P}(k) = E\left[e^{ik\sum_{i=1}^{n}\omega_i Q_i(X_i)}\right] \Leftrightarrow$$

$$\hat{P}(k) = \int e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t + ik\sum_{i=1}^{n}\omega_i Q_i(X_i)} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}}$$

Expanding the exponential in power series we obtain:

$$\hat{P}(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \int \left[\sum_i \omega_i Q_i(X_i)\right]^m e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}},$$

and the moments of the variable R can be obtained by the following integral:

$$E\left[P^m\right] = \int \left[\sum_i \omega_i Q_i(X_i)\right]^m e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}} \tag{B.4}$$

Next, we define the functional generator $\hat{P}(J)$ as:

$$\hat{P}(J) = \int e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t + JX^t} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}}$$

where the vector $J = (J_1, ..., J_n)$ is introduced as a fictitious vector that will allow to calculate the integrals. If $Q(X)$ is a polynomial it is easy to demonstrate that Equation [B.4] reduces to:

$$E\left[P^m\right] = \left[\sum_i \omega_i Q_i\left(\frac{\partial}{\partial J_i}\right)\right]^m \cdot \int e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t + JX^t} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}}\Bigg|_{J=0} \tag{B.5}$$

The integral [B.5] can be carried out analytically:

$$\int e^{-\frac{1}{2}X\kappa^{-1}(R_i,R_j)X^t + JX^t} \frac{dX}{\sqrt{(2\pi)^n \det\left[\kappa\left(R_i, R_j\right)\right]}} = e^{\frac{1}{2}J\kappa(R_i,R_j)J^t}.$$

Finally, we obtain the following result:

$$E\left[P^m\right] = \left[\sum_i \omega_i Q_i \left(\frac{\partial}{\partial J_i}\right)\right]^m \cdot e^{\frac{1}{2}J\kappa(R_i,R_j)J^t}\Bigg|_{J=0}$$

■

Corollary 14 *If the portfolio consists on two assets which follow a CFD, characterized by the coefficients a_1, b_1, c_1, d_1 and a_2, b_2, c_2, d_2 respectively, with normal rank correlation κ, and weights w_1 and w_2 the variance of the total return $w_1R_1 + w_2R_2$ can be calculated using the above Proposition and is equal to:*

$$\begin{aligned}\sigma^2\left[r_w\right] &= w_1^2\left(6a_1c_1 + 15a_1^2 + 2b_1^2 + c_1^2\right) + w_2^2\left(6a_2c_2 + 15a_2^2 + 2b_2^2 + c_2^2\right) + \\ & \quad 2w_1w_2\left(6a_1a_2\kappa^3 + 2b_1b_2\kappa^2 + 3c_1a_2\kappa + c_1c_2\kappa + 9a_1a_2\kappa + 3a_1c_2\kappa\right)\end{aligned} \tag{B.6}$$

and the kurtosis is given by:

$$\kappa\left[r_w\right] = \frac{E\left[r_w - \mu[r_w]\right]^4}{\sigma^4\left[r_w\right]} \tag{B.7}$$

where: $E\left[r_w - \mu[r_w]\right]^4 = 10\,395a_1^4w_1^4 + 60b_1^4w_1^4 + 10\,395a_2^4w_2^4 + 3c_1^4w_1^4 + 60b_2^4w_2^4 + 3c_2^4w_2^4 + 60a_1c_1^3w_1^4 + 3780a_1^3c_1w_1^4 + 60a_2c_2^3w_2^4 + 3780a_2^3c_2w_2^4 + 936a_1b_1^2c_1w_1^4 + 936a_2b_2^2c_2w_2^4 + 11\,340\kappa a_1a_2^3w_1w_2^3 + 11\,340\kappa a_1^3a_2w_1^3w_2 + 36\kappa a_1c_2^3w_1w_2^3 + 36\kappa a_2c_1^3w_1^3w_2 + 3780\kappa a_1^3c_2w_1^3w_2 + 3780\kappa a_2^3c_1w_1w_2^3 + 12\kappa c_1c_2^3w_1w_2^3 + 12\kappa c_1^3c_2w_1^3w_2 + 2808\kappa a_1a_2b_1^2w_1^3w_2 + 2808\kappa a_1a_2b_2^2w_1 w_2^3 + 540\kappa a_1a_2c_1^2w_1^3w_2 + 3780\kappa a_1^2a_2c_1w_1^3w_2 + 540\kappa a_1a_2c_2^2w_1w_2^3 + 3780\kappa a_1a_2^2c_2w_1w_2^3 + 936\kappa a_1b_1^2c_2w_1^3w_2 + 360\kappa a_2b_1^2c_1w_1^3w_2 + 360\kappa a_1b_2^2c_2w_1w_2^3 + 180\kappa a_1c_1^2c_2w_1^3w_2 + 936\kappa a_2b_2^2c_1w_1w_2^3 + 1260\kappa a_1^2c_1c_2w_1^3w_2 + 180\kappa a_2c_1c_2^2w_1w_2^3 + 1260\kappa a_2^2c_1c_2w_1w_2^3 + 120\kappa b_1^2c_1c_2w_1^3w_2 + 120\kappa b_2^2c_1c_2w_1w_2^3 + 216a_1a_2c_1c_2w_1^2w_2^2 + 3456\kappa a_1a_2b_1b_2w_1^2w_2^2 + 576\kappa a_1b_1b_2c_2w_1^2w_2^2 + 576\kappa a_2b_1b_2c_1w_1^2w_2^2 + 1872\kappa^2a_1b_1b_2c_1w_1^3w_2 + 96\kappa b_1b_2c_1c_2w_1^2w_2^2 + 1872\kappa^2a_2b_1b_2c_2w_1w_2^3 + 4500a_1^2b_1^2w_1^4 + 630a_1^2c_1^2w_1^4 +$

$60b_1^2c_1^2w_1^4+4500a_2^2b_2^2w_2^4+630a_2^2c_2^2w_2^4+60b_2^2c_2^2w_2^4+30\,240\kappa^3a_1a_2^3w_1w_2^3+30\,240\kappa^3a_1^3a_2w_1^3w_2+$
$240\kappa^2b_1b_2^3w_1w_2^3+240\kappa^2b_1^3b_2w_1^3w_2+24\kappa^3a_1c_2^3w_1w_2^3+24\kappa^3a_2c_1^3w_1^3w_2+540a_1a_2^2c_1w_1^2w_2^2+$
$72a_1b_2^2c_1w_1^2w_2^2+540a_1^2a_2c_2w_1^2w_2^2+36a_1c_1c_2^2w_1^2w_2^2+72a_2b_1^2c_2w_1^2w_2^2+36a_2c_1^2c_2w_1^2w_2^2+$
$9000\kappa^2a_1^2b_1b_2w_1^3w_2+6192\kappa^3a_1a_2b_1^2w_1^3w_2+9000\kappa^2a_2^2b_1b_2w_1w_2^3+6192\kappa^3a_1a_2b_2^2w_1w_2^3+$
$720\kappa^3a_1a_2c_1^2w_1^3w_2+7560\kappa^3a_1^2a_2c_1w_1^3w_2+120\kappa^2b_1b_2c_1^2w_1^3w_2+720\kappa^3a_1a_2c_2^2w_1w_2^3+7560\kappa^3a_1a_2^2c_2w_1w_2^3+$
$576\kappa^3a_2b_1^2c_1w_1^3w_2+120\kappa^2b_1b_2c_2^2w_1w_2^3+576\kappa^3a_1b_2^2c_2w_1w_2^3+11\,664\kappa^3a_1a_2b_1b_2w_1^2w_2^2+$
$1728\kappa^2a_1a_2c_1c_2w_1^2w_2^2+1296\kappa^3a_1b_1b_2c_2w_1^2w_2^2+1296\kappa^3a_2b_1b_2c_1w_1^2w_2^2+2880\kappa^5a_1a_2b_1b_2w_1^2w_2^2+$
$576\kappa^4a_1a_2c_1c_2w_1^2w_2^2+144\kappa^3b_1b_2c_1c_2w_1^2w_2^2+1350a_1^2a_2^2w_1^2w_2^2+180a_1^2b_2^2w_1^2w_2^2+180a_2^2b_1^2w_1^2w_2^2+$
$90a_1^2c_2^2w_1^2w_2^2+90a_2^2c_1^2w_1^2w_2^2+24b_1^2b_2^2w_1^2w_2^2+12b_1^2c_2^2w_1^2w_2^2+12b_2^2c_1^2w_1^2w_2^2+6c_1^2c_2^2w_1^2w_2^2+$
$6480\kappa^2a_1a_2^2c_1w_1^2w_2^2+576\kappa^2a_1b_2^2c_1w_1^2w_2^2+6480\kappa^2a_1^2a_2c_2w_1^2w_2^2+144\kappa^2a_1c_1c_2^2w_1^2w_2^2+576\kappa^2a_2b_1^2c_2$
$w_1^2w_2^2+4320\kappa^4a_1a_2^2c_1w_1^2w_2^2+144\kappa^2a_2c_1^2c_2w_1^2w_2^2+288\kappa^4a_1b_2^2c_1w_1^2w_2^2+4320\kappa^4a_1^2a_2c_2w_1^2w_2^2+$
$288\kappa^4a_2b_1^2c_2w_1^2w_2^2+24\,300\kappa^2a_1^2a_2^2w_1^2w_2^2+2160\kappa^2a_1^2b_2^2w_1^2w_2^2+2160\kappa^2a_2^2b_1^2w_1^2w_2^2+540\kappa^2a_1^2c_2^2w_1^2w_2^2+$
$540\kappa^2a_2^2c_1^2w_1^2w_2^2+192\kappa^2b_1^2b_2^2w_1^2w_2^2+32\,400\kappa^4a_1^2a_2^2w_1^2w_2^2+48\kappa^2b_1^2c_2^2w_1^2w_2^2+48\kappa^2b_2^2c_1^2w_1^2w_2^2+$
$2160\kappa^4a_1^2b_2^2w_1^2w_2^2+2160\kappa^4a_2^2b_1^2w_1^2w_2^2+12\kappa^2c_1^2c_2^2w_1^2w_2^2+144\kappa^4b_1^2b_2^2w_1^2w_2^2+4320\kappa^6a_1^2a_2^2w_1^2w_2^2.$

Lemma 7

Let ρ_{ij} *be the linear correlation coefficient and* κ_{ij} *the* normal rank correlation *between the variables* R_i *and* R_j *which follow a CB-MCFD*$_3$*, and* $a_{i,j}$ *the* i*-th coefficient in Equation* 1.5 *corresponding to the asset* j *determining the transformation. The relation between both correlation coefficients is given by:*

$$\rho_{ij}=\frac{6a_{3,i}a_{3,j}\kappa_{ij}^3+2a_{2,i}a_{2,j}\kappa_{ij}^2+\left(3a_{1,i}a_{3,j}+a_{1,i}a_{1,j}+9a_{3,i}a_{3,j}+3a_{3,i}a_{1,j}\right)\kappa_{ij}}{\sqrt{\left(6a_{3,i}a_{1,i}+15a_{2,i}^2+2a_{2,i}^2+a_{1,i}^2\right)\left(6a_{3,j}a_{1,j}+15a_{3,j}^2+2a_{2,j}^2+a_{1,j}^2\right)}}$$

Proof. The demonstration of this Lemma requires the result in Equation B.11 for the variance of a portfolio formed by two assets R_i and R_j with weights w_i and w_j:

$$\begin{aligned}\sigma^2(w_iR_i + w_jR_j) \;=\; & w_i^2\left(6a_{3,i}a_{1,i} + 15a_{2,i}^2 + 2a_{2,i}^2 + a_{1,i}^2\right) + \\ & w_j^2\left(6a_{3,j}a_{1,j} + 15a_{3,j}^2 + 2a_{2,j}^2 + a_{1,j}^2\right) + \\ & 2w_iw_j\left(\begin{array}{c} 6a_{3,i}a_{3,j}\kappa_{ij}^3 + 2a_{2,i}a_{2,j}\kappa_{ij}^2 + \\ (3a_{1,i}a_{3,j} + a_{1,i}a_{1,j} + 9a_{3,i}a_{3,j} + 3a_{3,i}a_{1,j})\,\kappa_{ij}\end{array}\right)\end{aligned}$$

Comparing this expression with the general formula for the variance of a sum of aleatory variables:

$$\sigma^2(w_iR_i + w_jR_j) = w_i^2\sigma^2(R_i) + w_j^2\sigma^2(R_j) + 2w_iw_j\text{cov}\,(R_i, R_j)$$

and having in mind that the Pearson correlation coefficient, ρ_{ij}, is defined as:

$$\rho_{ij} = \frac{\text{cov}\,(R_i, R_j)}{\sigma(R_i)\sigma(R_j)}$$

we finally obtain the desired result. ■

Lemma 8

Let CB-MCFD$_3$ be a third-order Copula-Based Multivariate Cornish-Fisher density defined by Equation 1.19 *for the random vector of variables R_i, with coefficients a_i, $i = 0, 1, 2, 3$. Then one can define a new vector of variables R_i' with zero mean and unitary variance-covariance matrix imposing κ to be the identity matrix and:*

$$a_{i,0} = -a_{i,2} \quad , \; a_{i,1} = \sqrt{1 - 6a_{i,3}^2 - 3a_{i,2}^2} - 3a_{i,3} \tag{B.8}$$

Besides that, $a_{i,2}$ and $a_{i,3}$ need to satisfy the following conditions in order to guarantee the existence of the distribution:

$$\begin{aligned} 0 &< a_{i,3} < \frac{1}{\sqrt{15}} \\ -\sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2+1}-6a_{i,3}\right)} &< a_{i,2} < \sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2+1}-6a_{i,3}\right)} \end{aligned}$$

Proof. This proof is almost identical to the one of Lemma 3. In order to guarantee the existence of the CB-MCFD$_3$ we must impose the existence of all the univariate distributions, which are by definition CF distributed. Therefore we must have the same conditions as in Lemma 3 for each variable R_i:

$$\begin{aligned} -\sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2+1}-6a_3\right)} &< a_{i,2} < \sqrt{3a_{i,3}\left(\sqrt{21a_{i,3}^2+1}-6a_{i,3}\right)} \\ 0 &< a_{i,3} < \frac{1}{\sqrt{15}}. \end{aligned}$$

■

Proposition 9

Let $(R_i)_{i=1}^n$ be variables that follow a VCB-MCFD with parameters of the distribution given by $a_{3,i}$, $a_{2,i}$, m and Σ. Then, any variable W defined as a sum of variables R_i is also a VCB-MCFD variable.

Proof. We will prove this proposition for the bivariate case, $n = 2$, but the generalization to n dimensions is straightforward. Let $W = R_1 + R_2$. From Equation 1.27 we have that the variables R_i can be written in terms of the z_1 and z_2 which follow an I-MCFD:

$$\begin{aligned} R_1 &= \Sigma_{11}^{1/2} \cdot z_1 + \Sigma_{12}^{1/2} \cdot z_2 + m_1 \\ R_2 &= \Sigma_{21}^{1/2} \cdot z_1 + \Sigma_{22}^{1/2} \cdot z_2 + m_2 \end{aligned}$$

Therefore, the variable W is given by:

$$\begin{aligned} W &= R_1 + R_2 \\ &= \Sigma_{11}^{1/2} \cdot z_1 + \Sigma_{12}^{1/2} \cdot z_2 + m_1 + \Sigma_{21}^{1/2} \cdot z_1 + \Sigma_{22}^{1/2} \cdot z_2 + m_2 \\ &= \left(\Sigma_{11}^{1/2} + \Sigma_{21}^{1/2}\right) \cdot z_1 + \left(\Sigma_{12}^{1/2} + \Sigma_{22}^{1/2}\right) \cdot z_2 + m_1 + m_2 \\ &= \tilde{\Sigma}_{11}^{1/2} \cdot z_1 + \tilde{\Sigma}_{12}^{1/2} \cdot z_2 + m \end{aligned}$$

Last Equation shows that the variable W can be also written as a linear combination of z's and, therefore, by definition the variable W must have the same marginal distribution as the R_i, i.e. W is a VCB-MCFD variable. ■

Proposition 10

Let $(R_i)_{i=1}^n$ *be variables that follow a VCB-MCFD with parameters of the distribution given by* $a_{3,i}$*,* $a_{2,i}$*,* m *and* Σ*. Then, the mean vector* M_1*, and the second, third and fourth centered multivariate moments, given by the variance-covariance matrix* M_2*, the skewness* *tensor M_3 and the kurtosis tensor, M_4, respectively, are given by:*

$$\begin{aligned}
M_{1,i} &= m_i \\
M_{2,ij} &= \Sigma_{ij} \\
M_{3,ijk} &= \sum_{r=1}^{n} w_{ir} w_{jr} w_{kr} \mu_{3,r} \\
M_{4,ijkl} &= \sum_{r=1}^{n} w_{ir} w_{jr} w_{kr} w_{lr} \mu_{4,r} \\
&\quad + \sum_{r=1}^{n} \sum_{\substack{s=1 \\ s \neq r}}^{n} \begin{pmatrix} w_{ir} w_{jr} w_{ks} w_{ls} + w_{ir} w_{js} w_{kr} w_{ls} \\ + w_{ir} w_{js} w_{ks} w_{lr} \end{pmatrix}
\end{aligned}$$

where $\mu_{3,r}$ and $\mu_{4,r}$ are the third and fourth order centered univariate moments of the Cornish-Fisher Density for the variable r given by Equations 1.11*, and w_{ij} denotes the ij-element of the Cholesky decomposition of the covariance matrix, i.e, $\Sigma^{1/2} = (w_{ij})$, $i, j = 1, ..., n$.*

Proof. We denote $\Sigma_{ij} = (w_{ij})$, $i, j = 1, ..., n$ the Cholesky decomposition of the covariance matrix of returns, Σ. The first two equalities are straightforward and are already demonstrated in the main text. Using tensor notations and denoting $\otimes$ the Kronecker product, we define the (n, n^2) co-skewness matrix as:

$$M_3 = E\left[(R-\mu)(R-\mu)' \otimes (R-\mu)'\right] = \{s_{ijk}\}$$

and the (n, n^3) co-kurtosis matrix as:

$$M_4 = E\left[(R-\mu)(R-\mu)' \otimes (R-\mu)' \otimes (R-\mu)'\right] = \{c_{ijkl}\}$$

The (i, j, k) component of the third central moment, s_{ijk}, is given by:

$$
\begin{aligned}
s_{ijk} &= E\left[(R-\mu)(R-\mu)' \otimes (R-\mu)'\right]_{ijk} \\
&= E\left[(R-\mu)_i (R-\mu)_j (R-\mu)_k\right] \\
&= E\left[\left(\Sigma^{1/2} z\right)_i \left(\Sigma^{1/2} z\right)_j \left(\Sigma^{1/2} z\right)_k\right] \\
&= E\left[\left(\sum_r w_{ir} z_r\right)\left(\sum_s w_{js} z_s\right)\left(\sum_t w_{kt} z_t\right)\right] \\
&= E\left[\sum_{r,s,t} w_{ir} w_{js} w_{kt} z_r z_s z_t\right] = \sum_{r,s,t} w_{ir} w_{js} w_{kt} E\left[z_r z_s z_t\right] \qquad \text{(B.9)}
\end{aligned}
$$

By definition, the variables $z_r z_s z_t$ are independent and therefore $E\left[z_r z_s z_t\right] = E\left[z_r\right] E\left[z_s\right] E\left[z_t\right]$ which is 0 by definition when one of the z's is different from the others. Therefore the only terms different from zero in Equation B.9 are those with $r = j = k$:

$$
s_{ijk} = \sum_r w_{ir} w_{jr} w_{kr} E\left[z_r^3\right] = \sum_r w_{ir} w_{jr} w_{kr} \mu_{3,r}
$$

where $\mu_{3,r}$ is the third moment of a standard (zero mean and unit variance) CFD for asset r and is given by Equations 1.11. For the fourth moment the demonstration is similar:

$$
\begin{aligned}
c_{ijkl} &= E\left[(R-\mu)(R-\mu)' \otimes (R-\mu)' \otimes (R-\mu)'\right]_{ijkl} \\
&= E\left[(R-\mu)_i (R-\mu)_j (R-\mu)_k (R-\mu)_l\right] \\
&= E\left[\left(\Sigma^{1/2} z\right)_i \left(\Sigma^{1/2} z\right)_j \left(\Sigma^{1/2} z\right)_k \left(\Sigma^{1/2} z\right)_l\right] \\
&= E\left[\left(\sum_r w_{ir} z_r\right)\left(\sum_s w_{js} z_s\right)\left(\sum_t w_{kt} z_t\right)\left(\sum_u w_{lu} z_u\right)\right] \\
&= E\left[\sum_{r,s,t,u} w_{ir} w_{js} w_{ks} w_{lu} z_r z_s z_t z_u\right] \\
&= \sum_{r,s,t,u} w_{ir} w_{js} w_{ks} w_{lu} E\left[z_r z_s z_t z_u\right] \qquad \text{(B.10)}
\end{aligned}
$$

In this case we have two sets of non-zero terms in the last sum, the $E\left[z_r^4\right]$ and $E\left[z_r^2 z_s^2\right]$. The other combinations which involve odd powers of z, as $E[z_r z_s^3] = E[z_r]E\left[z_s^3\right]$, are zero given the independence of the z's. The terms with $E\left[z_r^4\right]$ are simply equal to the fourth moment of a standard CFD for asset r, $\mu_{4,r}$, given by Equations 1.11. The terms $E\left[z_r^2 z_s^2\right]$ are equal to $E\left[z_r^2\right]E\left[z_s^2\right]$ given the independence of the z's and, therefore, are equal to 1, because the z's have unit variance. Therefore, the above sum from Equation B.10 is given by:

$$
\begin{aligned}
c_{ijkl} &= \sum_{r,s,t,u} w_{ir} w_{js} w_{ks} w_{lu} E\left[z_r z_s z_t z_u\right] \\
&= \sum_{r} w_{ir} w_{jr} w_{kr} w_{lr} \mu_{4,r} + \\
&\quad \sum_{r} \sum_{\substack{s \\ s \neq r}} \left(\begin{array}{c} w_{ir} w_{jr} w_{ks} w_{ls} + w_{ir} w_{js} w_{kr} w_{ls} \\ + w_{ir} w_{js} w_{ks} w_{lr} \end{array} \right)
\end{aligned}
$$

■

Proposition 11

Let R we a m-th order CFD variable, then the VaR calculated at a $1 - z$ confidence level is given by:

$$VaR = \Phi^{-1}(Q_m(z))$$

where Φ is the standard normal distribution function, $Q_m(z)$ is the m-th order polynomial and Φ^{-1} its respective inverse.

Proof. The definition of VaR for a m-th order Cornish-Fisher distributed variable with a fixed time horizon and percentile z, is given by the implicit following Equation:

$$z = \int_{-\infty}^{\infty} \theta(R - \text{VaR}) cf_m(R) dR$$

Using the expression $R = Q(X)$, we obtain:

$$z = \int_{-\infty}^{\infty} \theta(Q_m(X) - \text{VaR}) cf_m(Q_m(X)) \frac{\partial Q_m(X)}{\partial X} dX$$

In order to calculate $cf_m(Q_m(X))$, we use the definition of the density in Equation A.1:

$$cf_m(R) = \frac{1}{\sqrt{2\pi}} \frac{d\left[Q_m^{-1}(R)\right]}{dR} e^{-\frac{1}{2}\left[Q^{-1}(R)\right]^2}$$

then:

$$cf_m(Q_m(X)) = \frac{1}{\sqrt{2\pi}} \frac{d\left[Q_m^{-1}(Q_m(X))\right]}{dR} e^{-\frac{1}{2}\left[Q_m^{-1}(Q_m(X))\right]^2} = \frac{1}{\sqrt{2\pi}} \frac{dX}{dR} e^{-\frac{1}{2}X^2} \Leftrightarrow$$

$$cf_m(Q_m(X)) = \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{dR}{dX}} e^{-\frac{1}{2}X^2} = \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{dQ_m(X)}{dX}} e^{-\frac{1}{2}X^2}$$

Replacing this result in the latter equation, we obtain:

$$z = \int_{-\infty}^{\infty} \theta(Q_m(X) - \text{VaR}) \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{dQ_m(X)}{dX}} e^{-\frac{1}{2}X^2} \frac{\partial Q_m(X)}{\partial X} dX =$$

$$\int_{-\infty}^{\infty} \theta(Q_m(X) - \text{VaR}) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2} dX =$$

$$\int_{-\infty}^{Q_m^{-1}(\text{VaR})} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2} dX \Leftrightarrow$$

$$z = \Phi\left[Q_m^{-1}(\text{VaR})\right] - \Phi(-\infty) = \Phi\left[Q_m^{-1}(\text{VaR})\right]$$

Solving this equation for VaR we finish the demonstration. ■

Proposition 12

Let r be the risk-free interest rate, K the strike of the option, S_t the initial price of the underlying, T the maturity date, $\tau = T - t$ the time to maturity and σ the volatility of the process. In absence of arbitrage opportunities the call option, whose underlying follows an arithmetic Brownian motion, is given by:

$$\begin{aligned} C &= \frac{1}{1+r\tau}\left[(S_t(1+r\tau)-K)\,\Phi(-d)+S_t\sigma\sqrt{\tau}\phi(d)\right] \\ d &= \frac{K-S_t(1+r\tau)}{S_t\sigma\sqrt{\tau}} \end{aligned}$$

where $\Phi(x)$ is the distribution function of a standard gaussian variable and $\phi(x)$ is its corresponding density.

Proof. Under the risk-neutral probability the process of the underlying asset is:

$$S_T = S_t(1+r\tau) + S_t\sigma W_t$$

In absence of arbitrage opportunities the price of a call option should be:

$$C = \frac{1}{1+r\tau}E\left[(S_T-K)^+|F_t\right]$$

where $E[x|F_t]$ denote the expected value of x conditioned to the available information at time t. The later expression is equal to:

$$C_s = \frac{1}{1+r\tau}\int_K^\infty (S_T-K)\,f(S_T)dS_T = \frac{1}{1+r\tau}\int_K^\infty (x-K)\,\frac{1}{\sqrt{2\pi S_t^2\sigma^2\tau}}e^{-\frac{1}{2}\left(\frac{x-S_t(1+r\tau)}{S_t\sigma\sqrt{\tau}}\right)^2}dx \tag{B.11}$$

To evaluate the integral we consider the following variable change:

$$y = \frac{x-S_t(1+r\tau)}{S_t\sigma\sqrt{\tau}} \rightarrow y_k = \frac{K-S_t(1+r\tau)}{S_t\sigma\sqrt{\tau}} \rightarrow dy = \frac{dx}{S_t\sigma\sqrt{\tau}}$$

in Equation B.11:

$$C = \frac{1}{1+r\tau}\left[\left(S_t\left(1+r\tau\right)-K\right)\left(1-\Phi(y_K)\right)+\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_K^2}S_t\sigma\sqrt{\tau}\right]$$

Rearranging terms and undoing the variable change we obtain the required result:

$$C = \frac{1}{1+r\tau}\left[\left(S_t\left(1+r\tau\right)-K\right)\Phi(d)+\frac{S_t\sigma\sqrt{\tau}}{\sqrt{2\pi}}e^{-\frac{1}{2}d^2}\right], \qquad d = \frac{S_t(1+r\tau)-K}{S_t\sigma\sqrt{\tau}}$$

■

Proposition 13

Let r be the risk-free interest rate, K the strike of the option, S_t the initial price of the underlying, τ the time to maturity and σ the volatility of the process. In absence of arbitrage opportunities the call option, C^{CFD}, whose underlying follows a third-order CFD given by Equation 1.9 *is:*

$$\begin{aligned} C^{CFD} &= \frac{1}{1+r\tau}\left\{\begin{array}{c}\left(S_t(1+r\tau)-K\right)\Phi\left(-d\right)+a_1S_t\sigma\sqrt{\tau}\phi(d) \\ S_t\sigma\sqrt{\tau}\phi(d)\left(a_3\left(d^2+1\right)+a_2d\right)\end{array}\right\} \\ d &= Q_\tau^{-1}(K) \\ Q_\tau(x) &= S_t(1+r\tau)+\sigma\sqrt{\tau}S_t\left(a_3x^3+a_2x^2+\left(\sqrt{1-6a_3^2-2a_2^2}-3a_3\right)x-a_2\right)\end{aligned}$$

where $\Phi\left(x\right)$ is the distribution function of a standard gaussian variable, $\phi(x)$ is its corresponding density and $Q_\tau^{-1}(x)$ is the inverse of the third-order polynomial $Q_\tau(x)$.

Proof. We consider that the underlying S_T at the time of maturity T can be approximated through:

$$S_T = S_t(1+r\tau)+S_t\sigma\sqrt{\tau}z^*$$

where z^* is a standarized third-order CF variable with parameters a_2 and a_3. In absence of arbitrage opportunities the price of a call option should be:

$$C = \frac{1}{1+r\tau} E\left[(S_T - K)^+ |F_t\right]$$

Therefore, we have to evaluate the following integral

$$C = \frac{1}{1+r\tau} \int_K^\infty (S_T - K)\, cf_3(S_T) dS_T$$

where $cf_3(S_T)$ is the density of the third-order CFD variable S_T. We make the following variable change, $S_T = S_t(1+r\tau)+S_t\sigma\sqrt{\tau}\tilde{Q}_\tau(x) = Q_\tau(x)$, where $\tilde{Q}_\tau(x)$ is the polynomial corresponding to a standardized third-order Cornish-Fisher density with coefficients a_2 and a_3:

$$\begin{aligned} C &= \frac{1}{1+r\tau} \int_K^\infty \left(S_t(1+r\tau) + S_t\sigma\sqrt{\tau}\tilde{Q}_\tau(x) - K\right) cf_3(S_T) dS_T \\ &= \frac{1}{1+r\tau} \left\{(S_t(1+r\tau) - K) \int_K^\infty cf_3(S_T) dS_T + S_t\sigma\sqrt{\tau} \int_K^\infty \tilde{Q}_\tau(x) cf_3(S_T) dS_T\right\} \end{aligned}$$

Using the property that $cf_3(S_T)dS_T = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$ we calculate the first integral obtaining:

$$C = \frac{1}{1+r\tau} \left\{(S_t(1+r\tau) - K)\left(1 - \Phi(Q_\tau^{-1}(K)\right) + S_t\sigma\sqrt{\tau} \underbrace{\int_{Q_\tau^{-1}(K)}^\infty \tilde{Q}_\tau(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx}\right\} \tag{B.12}$$

Now, we carry out the marked integral, denoting $\tilde{Q}_\tau(x) = a_3x^3 + a_2x^2 + a_1x + a_0$[31]:

$$\int_{Q_\tau^{-1}(K)}^\infty \tilde{Q}_\tau(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \int_{Q_\tau^{-1}(K)}^\infty \left(a_3x^3 + a_2x^2 + a_1x + a_0\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \tag{B.13}$$

[31] It has to be beard in mind that for a standardized third-order CF variable the following relations hold:

$$a_0 = -a_2 \quad , \quad a_1 = \sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3.$$

We consider the variable change in Equation B.13:

$$\int_{Q_\tau^{-1}(K)}^{\infty}\left(a_3x^3+a_2x^2+a_1x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx+a_0\int_{Q_\tau^{-1}(K)}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx=$$
$$\underbrace{\int_{Q_\tau^{-1}(K)}^{\infty}\left(a_3x^3+a_2x^2+a_1x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx}-a_2\left(1-\Phi\left(Q_\tau^{-1}(K)\right)\right)$$

We evaluate the marked expression in the later equation using the variable change, denoting $d=Q_\tau^{-1}(K)$:

$$\int_{d}^{\infty}\left(a_3x^3+a_2x^2+a_1x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx=a_2(1-\Phi(d))+\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}d^2}\left(a_3d^2+a_2d+a_1+a_3\right)$$

Rearranging terms:

$$\int_{K}^{\infty}Q_\tau(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx=a_2\left(1-\Phi(d)\right)+\phi(d)\left(a_3\left(d^2+1\right)+a_2d+a_1\right) \tag{B.14}$$

We obtain the required result substituting B.14 in B.12:

$$\begin{aligned}
C &= \frac{1}{1+r\tau}\left\{\begin{array}{c}\left(S_t(1+r\tau)-K\right)\Phi\left(-d\right)+\\ \sigma\sqrt{\tau}S_t\phi(d)\left(a_3\left(d^2+1\right)+a_2d+\left(\sqrt{1-6a_3^2-2a_2^2}-3a_3\right)\right)\end{array}\right\}\\
d &= Q_\tau^{-1}(K)\\
Q_\tau(x) &= S_t(1+r\tau)+S_t\sigma\sqrt{\tau}\left(a_3x^3+a_2x^2+\left(\sqrt{1-6a_3^2-2a_2^2}-3a_3\right)x-a_2\right)
\end{aligned}$$

■

Appendix C
Algorithms and tests

C.1 Algorithms

In this Appendix, we present the details of the algorithms and optimization procedures we have applied in the realization of this thesis. In particular, we discuss the algorithms used in the estimation of the different models derived from the Cornish-Fisher Density (CFD) and the optimization procedure in the empirical options Section. All computer programs have been written in MATLAB code and are available from the authors upon request. We are very thankful to Kevin K. Sheppard for making available the "UCSD GARCH Toolbox" for estimating univariate and multivariate heteroskedasticity in time series models which we have used and modified in the estimation of the DCC models.

C.1.1 Univariate static CFD

QQ-Estimation

In the Quantile-Quantile method (QQ) we calculate first the quantile-quantile function between the experimental distribution function $S_T(R)$,

$$S_T(R) = \frac{1}{T} \sum_{t \in (R_t \leq R)} R_t$$

where T is the sample size. The number of points, q_i, forming this discrete function $S_T(R)$ to be used in this algorithm is in principle free, but we consider that an appropriate number of points is the one given by the multiples of the inverse of the sample size: $q_i = \left\{\frac{1}{T}, \frac{2}{T}, ..., \frac{T-1}{T}\right\}$. Therefore, we construct the pairs $\left(R_{(i)}, \Phi^{-1}\left(q_i\right)\right)$, where $R_{(i)}$ are the ordered return series elements, q_i are the corresponding experimental quantiles and Φ is the standard normal distribution. Then, we fit a cubic polynomial to these pairs using the Minimum Least Squares algorithm. Interestingly, different versions of this method could be easily incorporated. For example, one could give different weights to different quantiles using the Generalized Minimum Squares algorithm, increasing, therefore, the importance of the tails in the fitting.

Moments Estimation

In this Appendix we describe the algorithm used to obtain estimates for the parameters of the Cornish-Fisher distribution using the Moments Method (MM). A third-order CFD depends on four parameters (a_3, a_2, a_1 and a_0) and at least we must fit the first four empirical moments, given by Equations 1.11, to the theoretical ones. However, as the variable a_0 appears only in the Equation for the mean, we can reduce the estimation to three parameters and get it after fitting the other parameters a_1, a_2 and a_3.

We use the MATLAB macro FMINSEARCH to minimize the objective error function:

$$\min_{a_3,a_2,a_1} \text{err} = \sum_{i=2}^{4} \left(\mu_i\left(a_3, a_2, a_1\right) - \hat{m}^i\right)^2$$

where $\mu_i\left(a_3, a_2, a_1\right)$ is the i-th theoretical centered moment which depends on the parameters a_3, a_2 and a_1 and $\hat{m}^i$ is the sampling centered moment of order i. As initial parameters we use the QQ-estimates described above. If we try to fit the data with very small variances, given that we are handling high powers or very small numbers, numeric problems related to the floating comma precision can arise. Therefore, as the starting point in order to solve this problem, we scale the data to unit variance and fit the corresponding parameters, a'_i of the CFD. Finally, Lemma 4 is used to re-scale the obtained parameters a_i to calculate the real parameters $a_i = \sigma a'_i$, where σ is the sampled standard deviation of the data.

Maximum Likelihood Estimation

In this Appendix we describe the algorithm used to obtain estimates for the parameters of the Cornish-Fisher distribution using the Maximum Likelihood Method (ML). Supposing i.i.d. (independent and identical distributed) CF variables, we have to maximize the following log-likelihood function:

$$\begin{aligned} L_{\text{static-cfd}}(R) &= \sum_{t=1}^{T} \log cf_3(R_t; a_0, a_1, a_2, a_3) \\ &= -\frac{T}{2}\log(2\pi) + \sum_{t=1}^{T}\left(\log \frac{d\left[Q^{-1}(R_t)\right]}{dR} - \frac{1}{2}\left[Q^{-1}(R_t)\right]^2\right) \end{aligned}$$

where $d\left[Q^{-1}(R_i)\right]/dR$ and $Q^{-1}(R_t)$ are given in Equations A.2 and A.3. Considering that the third-order CFD is analytical we can calculate the score function as the derivatives of the log-likelihood with respect to the parameters a_i, therefore, we avoid calculating numeric derivatives in the optimization procedure. Although completely analytical, we do not report here the score function for the shake of conciseness, given their cumbersome expression. Finally, as for the previous method, we use the MATLAB macro FMINSEARCH to maximize the objective error function, and use also the QQ-estimates as initial parameters.

C.1.2 Univariate dynamic CFD

We use the Maximum Likelihood Method to obtain estimations for the GARCH model with Cornish-Fisher distributed innovations. We denote as m_t and σ_t^2 the conditional mean and variance at time t, $E\left[R_t|F_t\right]$ and $E\left[(R-m_t)^2|F_t\right]$, respectively. We leave the specification of the conditional mean and variance free, i.e. any dynamics in the mean and variance can be considered. We have to maximize the log-likelihood function:

$$L_{\text{dynamic-cfd}}\left(R\right) = \sum_{t=1}^{T} \log cf_3(R_t; a_{0,t}, a_{1,t}, a_{2,t}, a_{3,t})$$

Note that with this parametrization of the third-order CFD the four parameters are time varying but only the first two moments of the distribution, m_t and σ_t^2 are really time varying. Therefore, it is more appropriate to translate the dynamics on $a_{0,t}$, $a_{1,t}$, $a_{2,t}$ and $a_{3,t}$ into the dynamics of the parameters of the second parametrization of the CFD, $cf_3(R_t; m_t, \sigma_t^2, a_2, a_3)$, given in Equation 1.15 using the results of Lemma 4:

$$\sum_{t=1}^{T} \log cf_3(R_t; a_{0,t}, a_{1,t}, a_{2,t}, a_{3,t})$$

$$\begin{aligned}
a_{0,t} &= m_t - \sigma_t a_2 \\
a_{1,t} &= \sigma_t\left(\sqrt{1-6a_3^2-3a_2^2}-3a_3\right) \\
a_{2,t} &= \sigma_t a_2 \\
a_{3,t} &= \sigma_t a_3
\end{aligned}$$

For example, for an AR(1) - GARCH(1,1) specification of the mean and variance with Cornish-Fisher distributed innovations:

$$
\begin{aligned}
R_t &= m_t + y_t \\
m_t &= c + \text{AR} \cdot R_{t-1} \\
y_t &= \sigma_t z_t \\
\sigma_t^2 &= w + p y_{t-1}^2 + q \sigma_{t-1}^2 \\
& \quad z_t \sim CF(a_3, a_2),
\end{aligned}
$$

the log-likelihood function to be maximized is given by:

$$
\begin{aligned}
L_{\text{garch-cfd}}(R) &= \sum_{t=1}^{T} \log cf_3(R_t; a_{0,t}, a_{1,t}, a_{2,t}, a_{3,t}) \\
a_{0,t} &= (c + \text{AR} \cdot R_{t-1}) - \left(w + p y_{t-1}^2 + q \sigma_{t-1}^2\right) a_2 \\
a_{1,t} &= \left(w + p y_{t-1}^2 + q \sigma_{t-1}^2\right) \left(\sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3 \right) \\
a_{2,t} &= \left(w + p y_{t-1}^2 + q \sigma_{t-1}^2\right) a_2 \\
a_{3,t} &= \left(w + p y_{t-1}^2 + q \sigma_{t-1}^2\right) a_3
\end{aligned}
$$

C.1.3 Multivariate static CFD

In this Section, we will describe the algorithms that we have applied to estimate the static MCFD densities.

CB-MCFD

In the Copula-Based Multivariate Cornish-Fisher density we have decomposed our estimation algorithm in two parts. First, we estimate the univariate parameters, $a_{0,i}$, $a_{1,i}$,

$a_{2,i}$ and $a_{3,i}$ of each asset i, $(i = 1, ..., n)$, using the above mentioned Likelihood Estimation Method for each asset and, second, we estimate of the normal rank correlation matrix, κ, using the following estimator:

$$\hat{\kappa}_{ij} = \frac{1}{T}\sum_{t=1}^{T} Q_i^{-1}(R_{i,t})Q_j^{-1}(R_{j,t}) = \frac{1}{T}\sum_{t=1}^{T} X_{i,t}X_{j,t}$$

where $R_{i,t}$ is the return of asset i at time t. As we will see next, this estimation procedure is consistent with the Maximum Likelihood optimization algorithm. If we write down the expression for the log-likelihood of an i.i.d. sample of CB-MCFD distributed variables:

$$\begin{aligned} L_{\text{cb-mcfd}}(R) &= -\frac{T}{2}\left[n\log(2\pi) + \log|\kappa|\right] + \\ &\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(\log\left(\frac{\partial\left[Q_i^{-1}(R_{i,t})\right]}{\partial R_i}\right)\right) - \frac{1}{2}\sum_{i,j=1}^{n} Q_i^{-1}(R_{i,t})Q_j^{-1}(R_{j,t})\left(\kappa^{-1}\right)_{ij}\right) \end{aligned}$$

and suppose that the true univariate distribution functions are given by a third-order CFD model, this log-likelihood function can be written in terms of the fictitious variables, $X_{i,t}$, using the variable transformation $X_{i,t} = Q_i^{-1}(R_{i,t})$:

$$L_{\text{cb-mcfd}}(X) = -\frac{T}{2}\left[n\log(2\pi) + \log|\kappa|\right] - \frac{1}{2}\sum_{t=1}^{T}\sum_{i,j=1}^{n} X_{i,t}X_{j,t}\left(\kappa^{-1}\right)_{ij} \ .$$

This log-likelihood corresponds to the one of a centered multivariate normal i.i.d. sample and therefore, the matrix κ can be calculated as a usual correlation matrix with the variables X_i.

VCB-MCFD

The estimation of the parameters with the Variance-Covariance Based Multivariate Cornish-Fisher density model will be done considering the standard Maximum Likelihood Estimator. The parameters of this density (m_i, Σ_{ij} , $a_{i,2}$ and $a_{i,3}$) can be estimated maximizing the log-likelihood function given by:

$$L_{\text{vcb-mcfd}}(R) = -\frac{T}{2}\left(n\log\left(2\pi\right) + \log\left|\Sigma_t\right|\right) + \sum_{t=1}^{T}\sum_{i=1}^{n}\left(\log\left(\frac{\partial\left[Q_i^{-1}(z_{i,t})\right]}{\partial z_i}\right) - \frac{1}{2}\left(Q_i^{-1}(z_{i,t})\right)^2\right)$$

where $z_{i,t}$ is the i-th component of the following vector $z_t = \Sigma^{-1/2}\left(R_t - m_t\right)$ and $\Sigma^{-1/2}$ is the inverse of the Cholesky decomposition of matrix Σ_{ij}. It is straightforward to obtain this expression using the definition of the density of a VCB-MCFD (Equation 1.27). This log-likelihood is maximized as a whole in one step. As initial parameters in the algorithm, we choose $m_i^0 = E\left[R_i\right]$[32] for the mean and $\Sigma_{ij}^0 = E\left[\left(R_i - m_i\right)\left(R_j - m_j\right)\right]$ for the variance-covariance matrix, and the univariate Maximum Likelihood estimates of $a_{2,i}$ and $a_{3,i}$, calculated as explained in the above Section, for $a_{2,i}^0$ and $a_{3,i}^0$.

C.1.4 Multivariate dynamic CFD

In this Section we will present the estimation algorithms for the dynamic MCFD densities.

Dynamic CB-MCFD

Following the procedure of the static case, we also use a two steps Maximum Likelihood estimator for the dynamic CB-MCFD model. In this model, the parameters to be considered are c_i and AR_i for the conditional mean, w_i, p_i and q_i for the univariate GARCH processes, $a_{i,2}$ and $a_{i,3}$ for the skewness and kurtosis parameters of the CFD and ψ_1 and ψ_2 for the DCC model of the normal rank correlation, κ. First, we estimate a set of GARCH processes with CFD distributed innovations following the method described above, i.e. using the second parametrization. If the model is well specified we can transform the vari-

[32] A 0 superscript indicates that the parameter is the initial guess in the optimization procedure.

ables R_t into fictitious normally distributed variables, X_t. So that we finally estimate a standard DCC model (Engle 2002) by using the normal fictitious variables, X_t, with a likelihood given by:

$$L_{\text{dcb-mcfd}}(X) = -\frac{1}{2}\sum_{t=1}^{T}\left((n\log(2\pi) + \log|\kappa_t|) + \sum_{i,j=1}^{n} X_{i,t}X_{j,t}\left(\kappa^{-1}\right)_{ij}\right)$$

However, it has to be pointed out that this algorithm is not equivalent to maximizing the whole likelihood given by:

$$\begin{aligned} L_{\text{cb-mcfd}}(R) &= -\frac{1}{2}\sum_{t=1}^{T}(n\log(2\pi) + \log|\kappa_t|) + \\ &\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\left(\log\left(\frac{\partial\left[Q_i^{-1}(R_{i,t})\right]}{\partial R_i}\right)\right) - \frac{1}{2}\sum_{i,j=1}^{n} Q_i^{-1}(R_{i,t})Q_j^{-1}(R_{j,t})\left(\kappa^{-1}\right)_{ij}\right) \end{aligned}$$

but it is actually much faster and makes this approximation feasible.

Dynamic VCB-MCFD

As in the previous cases, to estimate the dynamic VCB-MCFD density we will use a two steps Maximum Likelihood Estimator. According to this model and as described in Section 1.6.2, the parameters to be considered are c_i and δ_i for the conditional mean, w_i, p_i and q_i for the univariate GARCH processes, $a_{i,2}$ and $a_{i,3}$ for the skewness and kurtosis parameters of the CFD, and ψ_1 and ψ_2 for the DCC model. The likelihood function of a dynamic VCB-MCFD is given by:

$$\begin{aligned} L_{\text{dvcb-mcfd}}(R) &= \sum_{t=1}^{T} -\frac{1}{2}(n\log(2\pi) + \log|\Sigma_t|) + \\ &\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\log\frac{\partial Q_i^{-1}(z_{i,t})}{\partial z_i} - \frac{1}{2}\left(Q_i^{-1}(z_{i,t})\right)^2\right) \end{aligned} \tag{C.1}$$

where $z_{i,t}$ is the i-th component of the vector $z_t = \Sigma_t^{-1/2}(R_t - m_t)$. First we estimate n (one for each asset) univariate ARMA-GARCH models with gaussian innovations to obtain the parameters c_i, δ_i, w_i, p_i and q_i, and, afterwards, with the standardized residuals we estimate the parameters associated to the DCC model together with the $a_{2,i}$ and $a_{3,i}$ parameters. As the likelihood is separable, these two steps procedure is equivalent to the whole maximization of the whole likelihood as we will prove next. We recall here the definitions of D_t, Θ_t from Section 1.6.2 and let $\varepsilon_t = D_t^{-1}(R_t - m_t)$. The innovations, ε_t, have zero conditional means and variances but nonzero covariances. Adding and subtracting $\varepsilon_t'\varepsilon_t$ in Equation C.1 and collecting terms adequately, we can obtain the following expression:

$$\begin{aligned} L_{\text{vcb-mcfd}}(R) &= -\frac{1}{2}\sum_{t=1}^{T}\left(n\log(2\pi) + 2\log|D_t| + R_t' D_t^{-1} D_t^{-1} R_t\right) + \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}\left(\log|\Theta_t| + \varepsilon_t'\varepsilon_t + \sum_{i=1}^{n}\left(-2\log\frac{\partial Q_i^{-1}(\Theta_t^{-1/2}\varepsilon_{t,i})}{\partial z_i} + Q_i^{-1}(\Theta_t^{-1/2}\varepsilon_{t,i})^2\right)\right) \end{aligned}$$

The first part of this Equation is a sum of likelihoods from GARCH processes and the second part contains the dynamic dependence parameters ψ_1 and ψ_2 along with the parameters $a_{2,i}$ and $a_{3,i}$. Therefore, given that the maximum of a sum is the sum of the maximums we can obtain first, the parameters of the univariate ARMA-GARCH models and, afterwards, the parameters associated to the DCC along with $a_{2,i}$ and $a_{3,i}$, improving vastly the convergence speed of the algorithm.

C.1.5 Option Pricing

Here we describe how we implemented the estimation of the parameters for each option pricing model. We first introduce some notation, then we discuss the estimation. We then go on to explain how we estimated parameters in more difficult situations.

The implicit estimator, $\hat{\theta}_t$, of the parameters of each model of date t is defined as the minimizer of the mean of the squared pricing errors for the options traded that day, that is,

$$\hat{\theta}_t = \arg\min_{\theta} \frac{1}{n_t} \sum_{i=1}^{n_t} [c_i(\theta) - c_i]^2$$

where $c_i(\theta)$ is the theoretical option price, c_i, denotes the market option price and n_t the total number of options at date t. For the CS and CF model, the parameter estimation turned out to be difficult. In particular, if parameters need to be obtained in a systematic way such as in the timeseries framework, it becomes necessary to make sure that the algorithm does not diverge. In this case, we restrict parameters in certain intervals and, second, force certain parameters to take values on a grid whereas the other parameters were obtained without restrictions. When a parameter was on a grid we eventually ran an unconstrained estimation using as starting values the estimates obtained over the grid that had a minimum error. In particular, we found that the CS model tended to derive extremely negative values for the skewness, and we obliged the parameter to stay between -2 and 2. And for the CF model, we had to impose the restrictions necessary to keep the parameters a_2 and a_3 within the permitted parameter range shown in Figure 1.2. From the 601 dates available we only found two cases where the CF model derived boundary parameters because of the small number of option prices available these dates.

C.2 Montecarlo Experiments

C.2.1 Comparison of estimators in the static CFD model

In this Section we wish to investigate the properties of the maximum-likelihood, moments and QQ-estimates when Cornish-Fisher Densities (CFD) are used in an attempt to directly obtain higher moments that differ from the ones of the normal distribution. For this purpose we investigate how well CFDs can be fitted to simulated data. We consider the fit of CFDs to data generated with a Cornish-Fisher distribution and to data generated with a mixture of normals. Furthermore, in the latter type of simulation we distinguish the situation where parameters for the simulated data are in or out of the restricted domain given by the Figure 1.1. Once the statistical properties are well understood we turn to the estimation of static and GARCH processes with Cornish-Fisher distributed innovations.

In our first simulation experiment we consider as true data those generated by a random process distributed according to the Cornish-Fisher density, and we fit to those data a CFD with the three-methods. We simulate 100 series of length 2000 of standardized CF variables with zero mean and unit variance. We will retain this type of size for all simulations reported in this work. As we have proven in Section 1.2, we have only two parameters left a_3 and a_2. We have chosen parameters a_3 and a_2 in a manner that we simulate variables with a kurtosis that we have arbitrarily chosen as 4, 8, and 15, and for each value of the kurtosis we have chosen values of skewness that correspond to the 75th, and 95th percentile of the [0, s (k)] segment in Figure 1.2. The first value of kurtosis corresponds to a situation where the tails behave very much like a normal density, and the third value is in the middle of possible kurtosis choices as we can see in Figure 1.2.

We present three different tables with the results of the estimations for each of the methods. In columns 2-5 of Table C.18 we present the various selected parameters (a_3, a_2, a_1, a_0). Our analysis of the results is twofold: first, we want to compare the properties of each estimator and, afterwards, we perform a comparison between the estimates for different skewness-kurtosis coefficients. As the columns 6-9 for the different estimations show, the average of the estimates are in general very good. However, we notice that the parameters a_3 and a_1 derived with the QQ and ML estimations are sensibly better than the ones corresponding to the MM, but the QQ and ML estimates are overall very similar. When turning to the dispersion of the parameter estimates, measured by their standard deviation, we observe that ML estimates are the most efficient according to the theory, given that the theoretical model in this case is also the true one. Comparing the dispersion between the QQ and ML estimates we find that the QQ are in general more efficient. At sight of these results, we should suggest the use of the ML method, but the QQ estimates are very useful also, as for the time of computation is much smaller with this method: for 100 time series, the ML algorithm takes around 50 times longer than the QQ method to reach its solution. Therefore, although ML estimates have better properties, the QQ are interesting as a first guess or as initial parameters in the optimization algorithm.

Comparing the estimations for the different kurtosis and skewness coefficients we do not find a worse behavior near the gaussinity, as in Jondeau and Rockinger 2001, where they analyze the Gram-Charlier density. In general, the estimates errors and dispersions are very similar in the whole region of permitted values.

C.2.2 GARCH + CFD

In Table C.19 we present the results of a set of Montecarlo experiments designed to test the performance of our algorithm. In the first panel we show the parameters used to simulate the data, while in the second and third panels we report the Maximum Likelihood estimates and their standard deviation, respectively. We simulate for each set of parameters 100 samples of length 2000. The chosen parameters derive conditional skewness and kurtosis equal to the values used in the static Montecarlo experiments, while the parameters w, p and q have been given just reasonable values. According to these results, we can deduce that our algorithm works well.

C.3 Tests

In this Appendix we provide a brief description of all the tests used throughout this work.

C.3.1 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test statistic is used to decide whether a sample comes from a certain distribution or not. In order to compare a sample of size n whose experimental distribution function is given by $S_n(x)$ with a theoretical distribution function $F(x)$, the value the maximum difference between these curves, D_n is calculated:

$$D_n = \max(S_n(x) - F(x)) \qquad \forall x$$

	Theoretical Parameters				QQ estimates				STD of QQ estimates			
	a_3	a_2	a_1	a_0	a_3	a_2	a_1	a_0	a_3	a_2	a_1	a_0
1	0.0302	0.0266	0.9058	-0.0266	0.0300	0.0269	0.9032	-0.0262	0.0072	0.0123	0.0250	0.0223
2	0.0307	0.0209	0.9048	-0.0209	0.0308	0.0216	0.9055	-0.0252	0.0060	0.0115	0.0238	0.0207
3	0.0821	0.0964	0.7236	-0.0964	0.0804	0.0963	0.7254	-0.0967	0.0107	0.0182	0.0276	0.0199
4	0.0861	0.0744	0.7135	-0.0744	0.0873	0.0772	0.7126	-0.0779	0.0123	0.0221	0.0304	0.0206
5	0.1232	0.1663	0.5545	-0.1663	0.1259	0.1623	0.5472	-0.1658	0.0190	0.0270	0.0413	0.0218
6	0.1326	0.1250	0.5313	-0.1250	0.1307	0.1251	0.5316	-0.1289	0.0185	0.0289	0.0370	0.0228
	Theoretical Parameters				**MM estimates**				**STD of MM estimates**			
1	0.0302	0.0266	0.9058	-0.0266	0.0296	0.0269	0.9042	-0.0262	0.0088	0.0137	0.0297	0.0235
2	0.0307	0.0209	0.9048	-0.0209	0.0297	0.0222	0.9086	-0.0258	0.0068	0.0123	0.0260	0.0205
3	0.0821	0.0964	0.7236	-0.0964	0.0737	0.0974	0.7466	-0.0978	0.0150	0.0236	0.0449	0.0248
4	0.0861	0.0744	0.7135	-0.0744	0.0836	0.0777	0.7228	-0.0784	0.0235	0.0304	0.0776	0.0286
5	0.1232	0.1663	0.5545	-0.1663	0.1222	0.1633	0.5494	-0.1669	0.0494	0.0408	0.2049	0.0373
6	0.1326	0.1250	0.5313	-0.1250	0.1209	0.1241	0.5629	-0.1280	0.0323	0.0495	0.0997	0.0440
	Theoretical Parameters				**ML estimates**				**STD of ML estimates**			
1	0.0302	0.0266	0.9058	-0.0266	0.0299	0.0271	0.9033	-0.0264	0.0066	0.0111	0.0245	0.0205
2	0.0307	0.0209	0.9048	-0.0209	0.0310	0.0204	0.9047	-0.0239	0.0060	0.0110	0.0253	0.0217
3	0.0821	0.0964	0.7236	-0.0964	0.0813	0.0959	0.7213	-0.0963	0.0086	0.0129	0.0213	0.0156
4	0.0861	0.0744	0.7135	-0.0744	0.0873	0.0765	0.7122	-0.0773	0.0098	0.0158	0.0229	0.0170
5	0.1232	0.1663	0.5545	-0.1663	0.1239	0.1651	0.5536	-0.1688	0.0126	0.0182	0.0178	0.0141
6	0.1326	0.1250	0.5313	-0.1250	0.1306	0.1245	0.5308	-0.1285	0.0128	0.0169	0.0181	0.0140

Table C.18. This Table presents the results of simulations where a CFD is fitted via the QQ, MM and ML algorithm to a CFD distributed data. We simulate for each set of parameters 100 samples of length 2000. Columns 2-5 present the theoretical parameters chosen so that the corresponding kurtosis is 4, 8, and 15 and for each value of kurtosis we have chosen values of skewness that correspond to the 75th, and 95th percentile of the [0, s (k)] segment in Figure 1.2. Columns 6-9 present averages of the QQ, MM and ML estimates and columns 10-13 present their standard deviations.

Theoretical Parameters

	c	w	p	q	a_3	a_2
1	0.0000	0.0100	0.0500	0.9000	0.0302	0.0266
2	0.0000	0.0100	0.0500	0.9000	0.0307	0.0209
3	0.0000	0.0100	0.0500	0.9000	0.0821	0.0964
4	0.0000	0.0100	0.0500	0.9000	0.0861	0.0744
5	0.0000	0.0100	0.0500	0.9000	0.1232	0.1663
6	0.0000	0.0100	0.0500	0.9000	0.1326	0.1250

Maximum Likelihood Estimates

	c	w	p	q	a_3	a_2
1	0.0013	0.0128	0.0488	0.8862	0.0287	0.0292
2	0.0008	0.0120	0.0526	0.8872	0.0301	0.0209
3	0.0005	0.0110	0.0511	0.8935	0.0813	0.0965
4	0.0008	0.0112	0.0491	0.8947	0.0858	0.0735
5	0.0010	0.0102	0.0469	0.9011	0.1214	0.1696
6	0.0009	0.0106	0.0505	0.8958	0.1313	0.1266

STD of Maximum Likelihood Estimates

	c	w	p	q	a_3	a_2
1	0.0100	0.0178	0.0138	0.0953	0.0056	0.0110
2	0.0095	0.0052	0.0148	0.0364	0.0063	0.0118
3	0.0091	0.0042	0.0124	0.0278	0.0077	0.0138
4	0.0098	0.0047	0.0148	0.0339	0.0074	0.0143
5	0.0089	0.0033	0.0110	0.0223	0.0065	0.0123
6	0.0088	0.0039	0.0142	0.0283	0.0063	0.0134

Table C.19. This Table presents the results of simulations where a GARCH process with CFD innovations is fitted via the ML algorithm to an exactly distributed data set. We simulate for each set of parameters 100 samples of length 2000. The first panel shows the parameters used to simulate the data while in the second and third panels we report the estimates and their standard deviation. The chosen parameters derive conditional skewness and kurtosis equal to the static experiments, while the dynamic parameters have chosen just to be reasonable.

In our case, the proposed theoretical distribution function is the one corresponding to the CFD:

$$CF(S) = \Phi\left[Q^{-1}(S)\right] = \int_{-\infty}^{Q^{-1}(S)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

and the empirical distribution function is calculated as:

$$S_n(x) = \frac{1}{n} \sum_{i \in (x_i \leq x)} x_i \tag{C.2}$$

Under the null hypothesis of correct specification, the values of this statistic follow the Kolmogorov-Smirnov distribution.

C.3.2 Jarque-Bera Test

The Jarque-Bera test is a goodness-of-fit measure of departure from normality based on the sample kurtosis and skewness. The JB test statistic is defined as:

$$JB = \frac{T}{6}\left(\xi^2 + \frac{(\kappa - 3)^2}{4}\right)$$

where ξ is the skewness, κ is the kurtosis, and T is the number of observations. The statistic has an asymptotic χ^2 distribution with two degrees of freedom.

C.3.3 Wald GMM-test

Bekaert and Harvey 1997 test for normality of equity returns based on Hansen 1982 generalized method of moments (GMM). The following system of equations is estimated for each asset i:

$$\begin{aligned} e_{1,i,t} &= R_{i,t} - m_i \\ e_{2,i,t} &= (R_{i,t} - m_i)^2 - \sigma_i^2 \\ e_{3,i,t} &= (R_{i,t} - m_i)^3 \sigma_i^{-3/2} - \xi_i \\ e_{4,i,t} &= (R_{i,t} - m_i)^4 \sigma_i^{-2} - \kappa_i - 3 \end{aligned}$$

where m_i is the mean, σ_i^2 is the variance, ξ_i is the skewness, κ_i is the excess kurtosis, and $e_t = (e_{1it}, e_{2it}, e_{3it}, e_{4it})$ represents the disturbances, with $E[e_{1,i}] = E[e_{2,i}] = E[e_{3,i}] = E[e_{4,i}] = 0$. There are four orthogonality conditions and four parameters implying that the model is exactly identified. The null hypothesis that the coefficients of skewness and excess kurtosis are zero is tested with a Wald test.

The variance-covariance matrix of the parameters is heteroskedasticity consistent and corrects for serial correlation using a Bartlett kernel with an optimal bandwidth as in Andrews 1991. The Wald statistic is asymptotically distributed as a χ^2 distribution.

C.3.4 Ljung-Box statistic

The Ljung-Box (Ljung and Box 1978) test is based on the autocorrelation plot and tests if a number of autocorrelations are zero. It is calculated as:

$$Q(h) = T\,(T+2) \sum_{j=1}^{h} \frac{\rho^2(j)}{T-j}$$

where T is the sample size, $\rho(j)$ is the autocorrelation at lag j, and h is the number of lags being tested. The Ljung-Box statistic is asymptotically distributed as a χ^2 distribution with h degrees of freedom.

C.3.5 LM-test for Heteroskedasticity

Given sample residuals obtained from a curve fit, the LM-test of Engle 1982 tests for the presence of h-th order ARCH effects by regressing the squared residuals on a constant and the lagged values of the previous squared residuals. Under the null hypothesis, the asymptotic test statistic, $T(R^2)$, where T is the number of squared residuals included in the regression and R^2 is the sample multiple correlation coefficient, is asymptotically χ^2 distributed with h degrees of freedom. When testing for ARCH effects, a GARCH(P,Q) process is locally equivalent to an ARCH(P+Q) process.

C.3.6 Shapiro-Wilk (modified - Royston 1982)

The Shapiro-Wilk test (Shapiro and Wilk 1965) tests the null hypothesis that a sample came from a normally distributed population. Let $m' = (m_{1,...,} m_n)$ denote the vector of expected values of standard gaussian order statistics (Kendall, Stuart, and Ord 1994) and Σ_{ij} the covariance matrix. Considering that $R = (R_1, ..., R_n)$ is ordered $R_{(1)} < R_{(2)} < ... < R_{(n)}$, the traditional Shapiro-Wilk statistic is given by:

$$W = \frac{\left[\sum_{i=1}^{n} a_i R_{(i)}\right]^2}{\sum_{i=1}^{n} \left(R_i - \bar{R}\right)^2}$$

where $a' = (a_1, ..., a_n) = m'\Sigma^{-1}\left[\left(m'\Sigma^{-1}\right)\left(\Sigma^{-1}m\right)\right]^{-1/2}$. The test rejects the null hypothesis if W is too small. As this test has been shown to be limited to sample sizes between 3 and 50, we use the Royston 1982 transformation of this statistic which is suitable for sample sizes up to 2000:

$$\begin{aligned} y &= (1-W)^{\lambda} \\ z &= \left(y-m_y\right)/\sigma_y \end{aligned}$$

where λ, m_y and σ_y are parameters tabulated in Royston 1982.

C.3.7 Mardia A and B

Mardia A and B statistics (Mardia 1985) test for multivariate normality focusing on skewness and kurtosis, respectively. They are defined as:

$$\begin{aligned} A &= \frac{1}{6} n b_{1,n} \\ B &= \left(b_{2,n}-\beta_{2,n}\right)/\{8n\left(n+2\right)/T\}^{1/2} \end{aligned}$$

where n is the number of assets, T is the sample size, coefficient $\beta_{2,n}$ is equal to $n\left(n+2\right)$ and:

$$\begin{aligned} b_{2,n} &= E\left[\left(\mathbf{R}-\bar{\mathbf{R}}\right)' \boldsymbol{\Sigma}^{-1}\left(\mathbf{R}-\bar{\mathbf{R}}\right)\right]^2 \\ b_{1,n} &= \sum_{r,s,t}\sum_{r',s',t'} \boldsymbol{\Sigma}_{rr'}^{-1}\boldsymbol{\Sigma}_{ss'}^{-1}\boldsymbol{\Sigma}_{tt'}^{-1} M_{rst}^3 M_{r's't'}^3 \end{aligned}$$

where $\mathbf{R} = \{R_{i,t}\}_{i=1,\ldots,n}^{t=1,\ldots,T}$ is the matrix of returns, $\boldsymbol{\Sigma}$ is the variance-covariance matrix and M_{rst}^3 represents the skewness tensor:

$$M_{rst}^3 = \frac{1}{T}\sum_{i=1}^{n}\left(R_{r,i}-\bar{R}_i\right)\left(R_{s,i}-\bar{R}_i\right)\left(R_{t,i}-\bar{R}_i\right)$$

Mardia A and B statistics are asymptotically distributed as a χ^2 with $n\left(n+1\right)\left(n+2\right)/6$ degrees of freedom and a $N(0,1)$, respectively.

C.3.8 Omnibus

The Omnibus statistic (Doornick and Hansen 1994) tests for multivariate normality. They propose the following statistic:

$$O = Z_1' Z_1 + Z_2' Z_2$$

with $Z_1' = (z_{11}, ..., z_{1n})$ and $Z_2' = (z_{21}, ..., z_{2n})$ where z_{1i} and z_{2j} being a transformation of skewness ξ and kurtosis κ, respectively. This test controls for finite sample properties of skewness and kurtosis.

C.3.9 χ^2-Malevergne and Sornette

This statistic tests for the null hypothesis that the structure of dependencies is given by a gaussian copula. The statistic z^2 is given by:

$$z^2 = \sum_{i,j=1}^{n} \Phi^{-1}(F_i(R_i)) \left(\kappa^{-1}\right)_{ij} \Phi^{-1}(F_j(R_j)) \tag{C.3}$$

where F_i is the marginal distribution function of asset i, Φ^{-1} is the inverse of the gaussian distribution and the matrix κ is the normal rank correlation:

$$\kappa_{ij} = Cov\left[\Phi^{-1}(F_i(R_i)), \Phi^{-1}(F_j(R_j))\right],$$

and follows a χ^2-distribution with n degrees of freedom. Malevergne and Sornette 2003 use the empirical distribution function $S_n(x) = \frac{1}{n}\sum_{i\in(x_i\leq x)} x_i$ to avoid making hypothesis about the marginals.

In our work, we will test for the null hypothesis that innovations are given by a Copula-Based Multivariate CFD (see Section 1.3.1) using $F_i = \Phi\left(Q_i^{-1}\right)$ in Equation C.3.

In our notation we can rewrite this equation more easily in terms of Q_i as:

$$\begin{aligned} z^2 &= \sum_{i,j=1}^{n} Q_i^{-1}(R_i)\left(\kappa^{-1}\right)_{i,j} Q_j^{-1}(R_j) \\ \kappa &= Cov\left[Q_i^{-1}(R_i), Q_j^{-1}(R_j)\right] \end{aligned}$$

Therefore, if our assets are supposed to follow a CB-MCFD, according to this test the z^2 should be distributed as χ^2-distribution with n degrees of freedom.

References

Chapter 1
References

Abramowitz, M. and I. A. Stegun (1964). *Handbook of Mathematical Functions*. USA: National Bureau of Standards.

Acerbi, C. (2002). Spectral representations of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance 26*, 1505–1518.

Acerbi, C. (2004). *Coherent Representations of Subjective Risk Aversion*. New York: John Wiley and Sons.

Adcock, C. (2002). Asset pricing and portfolio selection based on the multivariate skew-student distribution. Paper presented at the Non-linear Asset Pricing Workshop, April, Paris.

Alexander, G. and A. Baptista (2002). A VaR-constrained mean-variance model: Implications for portfolio selection and the basle capital accord. EFA 2001 Barcelona Meetings.

Andrews, D. W. (1991). Heterokedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*, 817–858.

Ang, A. and G. Bekaert (2002). Regime switches in interest rates. *Journal of Business and Economic Statistics 20*(2), 163–172.

Ang, A., J. Chen, and Y. Xing (2002). Downside correlation and expected stock returns. EFA 2002 Berlin Meetings Presented Paper; USC Finance and Business Econ. Working Paper No. 01-25.

Arditti, F. (1967, March). Risk and the required return on equity. *The Journal of Finance 22*(1), 19–36.

Arditti, F. D. and H. Levy (1975). Portfolio efficiency analysis in three moments: The multiperiod case. *The Journal of Finance 30*, 797–809.

Artzner, P., J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Athayde, G. and R. Flôres (2004). Finding a maximum skewness portfolio–a general solution to three moments portfolio choice. *Journal of Economic Dynamics and Control 28*, 1335–1352.

Backus, D., S. Foresi, K. Li, and L. Wu (1997). Accounting for biases in black-scholes. Working Paper, Stern School of Business, 40 pages.

Barone-Adesi, G., P. Gagliardini, and G. Urga (2000). Homogeneity hypothesis in the context of asset pricing models: The quadratic market model. Working Paper. http://ssrn.com/abstract=263215.

Bassi, F., P. Embrechts, and M. Kafetzaki (1997). *Risk Management and Quantile Estimation*, Chapter 5, pp. 100–120. Birkhauser, Boston.

Bawa, V. S. and E. B. Lindenberg (1977). Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics 5*, 189–200.

Bekaert, G. and C. R. Harvey (1997). Emerging equity market volatility. *Journal of Financial Economics 43*, 29–77.

Bera, A. and C. Jarque (1982). Model specification tests: A simultaneous approach. *Journal of Econometrics 20*, 59–82.

Bergara, A., U. Ansejo, and A. Rubia (2006). A semi-parametric cornish-fisher expansion for modelling and forecasting multivariate volatility. Working Paper.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics 3*, 167–79.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–659.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedastic. *Journal of Econometrics 31*, 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics 72*, 498–505.

Bollerslev, T., R. Chou, and K. Kroner (1992). Arch modeling in finance. *Journal of Econometrics 52*, 5–59.

Bouchaud, J. and M. Potters (2000). *Theory of Financial Risk*. Cambridge: Cambridge University Press.

Browne, S. (1999). The risk and rewards of minimizing shortfall probability. *Journal of Portfolio Management 25*, 76–86.

Buckley, I., G. Comezaña, B. Djerroud, and L. Seco (2005). Portfolio optimization when asset returns have the gaussian mixture distribution. Working Paper.

Capelle-Blanchard, G., E. Jurczenko, and B. Maillet (2001). The approximate option pricing model: Performances and dynamic properties. *Journal of Multinational Financial Management 11*, 427–444.

Charlier, C. V. L. (1905). über das fehlergesetz. *Ark. Math. Astr. och Phys 2*, 1–9.

Chen, X., Y. Fan, and A. Patton (2004). Simple tests for models of dependence between multiple financial time series, with applications to u.s. equity returns and exchange rates. Working Paper.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. New York: John Wiley and Sons.

Chunhachinda, P., K. Danpadani, S. Hamid, and A. Prakash (1997). Portfolio selection and skewness: Evidence from the international stock markets. *Journal of banking and Finance 21*, 143–167.

Cornish, E. and R. Fisher (1937). Moments and cumulants in the specification of distributions. *Rev. Int. Statist. Inst. 5*, 307–337.

Corrado, C. and T. Su (1996a). Skewness and kurtosis in s&p 500 index returns implied by option prices. *The Journal of Financial Research 19*(2), 175–193.

Corrado, C. and T. Su (1996b). S&p 500 index option tests of jarrow and ruddŠs approximate option valuation formula. *Journal of Futures Markets 6*, 611–629.

Corrado, C. and T. Su (1997a). Implied volatility skews and stock index skewness and kurtosis implied by s&p 500 index. *Journal of Derivatives 4*, 8–19.

Corrado, C. and T. Su (1997b). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *The European Journal of Finance 3*, 73–85.

Demarta, S. and A. McNeil (2004). The t copula and related copulas. Department of Mathematics Federal Institute of Technology.

Dias, A. and P. Embrechts (2005). Dynamic copula models for multivariate high-frequency data in finance. Working Paper.

Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance 1*, 369–403.

Doornick, J. and H. Hansen (1994). An omnibus test for univariate and multivariate normality. Working Paper, Nuffield College.

Duffie, D. and J. Pan (1997, Spring). An overview of value at risk. *The Journal of Derivatives 4*(2), 7–49.

Duffie, D. and J. Pan (2001). Analytical value-at-risk with jumps and credit risk. *Finance and Stochastics 5*, 155–180.

Edgeworth, F. (1898). On the representation of statistics by mathematical formulae (part i). *Journal of the Royal Statistical Society 61*(4), 670–700.

Edgeworth, F. Y. (1905). The law of error. *Cambridge Philos. Society 20*, 113–141.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.

Embrechts, P., S. Resnick, and G. Samorodnitsky (1998). Living on the edge. *Risk 11*, 96–100.

Engle, R. (1982). Autorregressive conditional heterokedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*, 987–1007.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heterokedasticity models. *Journal of Business and Economic Statistics 20*, 339.

Engle, R. and K. Kroner (1995). Multivariate simultaneous GARCH. *Econometric Theory 11*, 122–150.

Engle, R. F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554.

Falk, M. (1984). Extreme quantile estimation in d-neighborhoods of generalized pareto distributions. *Statistics and Probability Letters 20*, 9–21.

Falk, M. (1985). Asymptotic normality of the kernel quantile estimator. *The Annals of Statistics 13*, 428–433.

Ferreira, E., M. Gago, A. León, and G. Rubio (2005). An empirical comparison of the performance of alternative option pricing models. Working Paper.

Fishburn, P. (1977). Mean-risk analysis with risk associated with below-target returns. *American Economic Review 67*, 116–126.

Flôres, R. G. and G. M. de Athayde (2002). On certain geometric aspects of portfolio optimisation with higher moments. Working Paper. http://epge.fgv.br/portal/arquivo/1302.pdf.

Francis, J. (1975). Skewness and investors' decisions. *The Journal of Quantitative Analysis 10*, 163–172.

Frees, E. and E. Valdez (1998). Understanding relationships using copula. *North American Actuarial Journal 2*, 1–25.

Frey, R., A. McNeil, and M. Nyfeler (2001, October). Copulas and credit models. *RISK*, 11–114.

Gallant, A. and G.Tauchen (1989). Semi-nonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. *Econometrica 57*, 1091–1120.

Gnedenko, B. and A. Kolmogorov (1954). *Limit Distributions for Sums of Independent Random Variables*. Cambridge, Mass: Addison-Wesley.

Gourieroux, C. and J. Jasiak (1999). Truncated local likelihood and nonparametric tail analysis. DP 99 t' CREST.

Guidolin, M. and A. Timmerman (2005). Optimal portfolio choice under regime switching, skew and kurtosis preferences. available at http://ideas.repec.org/p/fip/fedlwp/2005-006.html.

Hansen, B. (1994). Autorregressive conditional density estimation. *International Economic Review 35*.

Hansen, L. P. (1982). Large sample propoerties of generalized methods of moments estimators. *Econometrica 50*, 1029–1054.

Harrel, F. and C. Davis (1982). A new distribution free quantile estimation. *Biometrika 69*, 635–640.

Harvey, C. and A. Siddique (2000). Conditional skewness in asset pricing tests. *Journal of Finance 55*, 1263–1295.

Harvey, C. R., J. C. Liechty, M. W. Liechty, and P. Müller (2005). Portfolio selection with higher moments. Working Paper.

Henessy, D. and H.E.Lapan (2002). The use of archimidean copulas to model portfolio allocations. *Mathematical Finance 12*, 143–154.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*, 327–343.

Hodges, S. (1998). A generalization of the sharpe ratio and its application to valuation bounds and risk measures. FORC Preprint 98/88. University of Warwick.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*, 1644–1655.

Hong, H. (1998). Application of polynomial transformation to normality in structural reliability analysis. *Canadian Journal of Civil Engineering 25*, 241–249.

Hull, J. and A. White (1998). Value at risk when daily changes in market variables are not normalliy distributed. *Journal of derivatives 5*(3), 9–19.

Hull, J. C. (2004). *Options, Futures, and Other Derivatives*. Toronto, Canada: Prentice Hall.

Hung, D. C.-H., M. Shackleton, and X. Xu (2004). CAPM, higher co-moment and factor models of UK stock returns. *Journal of Business Finance and Accounting 31*, 87.

Ingersoll, J. (1975). Multidimensional security pricing. *The Journal of Financial and Quantitative Analysis 10*, 785–798.

Jarrow, R. and A. Rudd (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics 10*, 347–369.

Jean, W. H. (1971). The extension of portfolio analysis to three or more parameters. *The Journal of Financial and Quantitative Analysis 6*, 505–515.

Jean, W. H. (1973). More on multidimensional portfolio analysis. *The Journal of Financial and Quantitative Analysis 8*, 475–490.

Johnson, N. (1949). Systems of frecuency curves generated by methods of translation. *Biometrika 36*(1/2), 149–176.

Johnson, S. and S. Kotz (1972a). *Distributions in Statistics: Continious Multivariate Distributions. Volume Two.* New York: John Wiley and Sons.

Johnson, S. and S. Kotz (1972b). *Distributions in Statistics: Continuous Univariate Distributions. Volume One.* New York: John Wiley and Sons.

Jondeau, E. and M. Rockinger (2000). Reading the smile: The message conveyed by methods which infer risk neutral densities. *Journal of International Money and Finance 19*, 885–915.

Jondeau, E. and M. Rockinger (2001). Gram charlier densities. *Journal of Economic Dynamics and Control 25*, 1457–1483.

Jondeau, E. and M. Rockinger (2002). The allocation of assets under higher moments. Working Paper.

Jondeau, E. and M. Rockinger (2003). Conditional volatility, skewness, and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control 27*, 1699–1737.

Jondeau, E. and M. Rockinger (2005). Conditional asset allocation under non-normality: How costly is the mean-variance criterion? Research Paper Nř132.

Jorion, P. (1996). Risk2: Measuring the risk in value at risk. *Financial Analysts 13*, 47–56.

Jorion, P. (2000). *Value at Risk: The New Benchmark for Controlling Market Risk*. New York: Mc Graw-Hill.

Jurczenko, E. and B.Maillet (2001). *The 3-CAPM: Theoretical Foundations and an Asset Pricing Model Comparison in a Unified Framework*, pp. 239–273. New York: John Wiley.

Jurczenko, E., B. Maillet, and B. Negrea (2002). Multi-moment option pricing models: A general comparison (part 1). Working paper, Univesity of Paris I Panthteon-Sorbonne.

Kendall, M., A. Stuart, and J. Ord (1994). *Kendall's advanced theory of statistics*, Volume 1. New York, N.Y.: Oxford University Press.

Koedijk, K., R. Campbell, and P. Kofman (2002). Increased correlation in bear markets. *Financial Analyist Journal 58*(1), 87–94.

Kraus, A. and R. Litzenberger (1976). Skewness preference and the valuation of risk assets. *Journal of Finance 31*(4), 1085–1100.

Kroll, Y., H. Levy, and H. M. Markowitz (1984). Mean-variance versus direct utility maximization. *The Journal of Finance 39*, 47–61.

León, A., J. Mencía, and E. Sentana (2005). Parametric properties of semi-nonparametric distributions, with an application to option valuation. Working Paper.

Levy, H. and H. M. Markowitz (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review 69*, 308–317.

Levy, J. (1969). A utility function depending on the first three moments. *The Journal of Finance 24*, 715 – 719.

Lintner, J. (1965). The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets. *Review of Economics and Statistics 47*, 13–37.

Ljung, G. M. and G. E. P. Box (1978). On a measure of a lack of fit in time series models. *Biometrika 65*, 297–303.

Lo, A. and C. MacKinlay (1988). Ock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies 1*, 41–66.

Lo, A. and C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3*, 175–206.

Longin, F. and B. Solnik (2001). Extreme correlation of international equity market. *Journal of Finance 56*, 649–679.

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking and Finance 24*, 1097–1130.

Lowenstein, R. (2000). *When Genius Failed. The Rise and Fall of Long-Term Capital Management*. New York: Random House.

Malevergne, M. and D. Sornette (2001). Testing the gaussian copula hypothesis for financial assets dependences. Working Paper.

Malevergne, Y. and D. Sornette (2003). Testing the gaussian copula hypothesis for financial assets dependences. *Quantitative Finance 3*, 231–250.

Malevergne, Y. and D. Sornette (2004). General framework for a portfolio theory with non-gaussian risks and non-linear correlations. arXiv:cond-mat/0103020. v1 1 Mar 2001.

Malz, A. (1996). Using option prices to estimate realignment probabilities in the european monetary system: The case of sterling mark. *Journal of International Money and Finance 15*, 717–748.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business 36*, 294–419.

Mardia, K. (1985). *Mardia's Test of Multinormality*, Volume 5, Chapter 5, pp. 217–221. New York: Wiley.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*, 77–91.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. New York: Wiley.

Markowitz, H. M. (1991). Foundations of portfolio theory. *The Journal of Finance 46*, 469–477.

Neftci, S. N. (2000, Spring). Value at risk calculations, extreme events, and tail estimation. *Journal of Derivatives*, 23–37.

Nelsen, R. B. (1999). *An Introduction to Copulas*. New York: Springer-Verlag.

Neumann, J. V. and . Morgenstern (1953). *Theory of Games and Economic Behavior*. New York: Princeton University Press.

Owen, J. and R. Rabinovitch (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *Journal of Finance 38*, 745–752.

Peña, J. (2002). *La Gestión de Los Riesgos Financieros de Mercado Y Crédito*. España: Prentice Hall.

Prakash, A., C. Chang, and T. Pactwa (2003). Selecting a portfolio with skewness: Recent evidence from US, european, and latin american equity markets. *Journal of Banking and Finance 27*, 1375–1390.

Pratt, J. W. and R. J. Zeckhauser (1987). Proper risk aversion. *Econometrica 55*, 143–154.

Pulley, L. B. (1981). A general mean-variance approximation to expected utility for short holding periods. *The Journal of Financial and Quantitative Analysis 16*, 361–373.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*, 1281–1292.

Rabin, M. and R. H. Thaler (2001). Anomalies: Risk aversion. *Journal of Economic Perspectives 15*, 219–232.

Rachev, S. (2003). *Handbook of Heavy Tailed Distributions in Finance*. New York: Elsevier.

Ridder, T. (1997). Basics of statistical VaR-estimation. Technical report, SGZ-BankRidder, FrankfurtrKarlsruhe.

Rockinger, M. and E. Jondeau (2002). Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics 106*, 119–142.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*, 341–360.

Roy, A. (1952). Safety first and the holding of assets. *Econometrica 20*, 431–439.

Royston, P. (1982). An extension of shapiro and wilk's w test for normality to large samples. *Applied Statistics 31*, 115–124.

Rubinstein, M. E. (1973). The fundamental theorem of parameter-preference security valuation. *The Journal of Financial and Quantitative Analysis 8*, 61–69.

Rubio, G., G. Serna, and A. León (2006). Autoregressive conditional volatility, skewness and kurtosis. Forthcoming in the Quarterly Review of Economics and Finance.

Samuelson, P. A. (1970). The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *The Review of Economic Studies 37*, 537–542.

Schechter, L. (2005). Risk aversion and expected-utility theory: A calibration exercise. Working Paper.

Scott, R. and P. Horvath (1980). On the direction of preference for moments of higher order than the variance. *Journal of Finance 35*(4), 915–919.

Shadwick, W. and C. Keating (2002). A universal performance measure. *Journal of Performance Measurement 6*(3), 59–84.

Shalit, H. and S. Yitzhaki (1984). Mean-gini, porttolio theory, and the pricing of risky assets. *Journal of Finance 39*, 1449–1468.

Shapiro, S. and M. Wilk (1965). An analysis of variance test for normality (complete samples). *Biometrika 52*, 591–611.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business 39*, 119–138.

Simaan, Y. (1993). Portfolio selection and asset pricing- three-parameter framework. *Management Science 39*, 568–577.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leures marges. *Publications de l'Institute de Statistique de L'Université de Paris 8*, 229–231.

Sornette, D., J. Andersen, and P. Simonetti (2000a). ϕ^pnon-linear correlations. *Physics Report 335*, 19–92.

Sornette, D., J. Andersen, and P. Simonetti (2000b). Portfolio theory for "fat tails". *International Journal of Theoretical and Applied Finance 3*, 523–535.

Suárez, A. and S. Carrillo (2003). Computational tools for the analysis of market risk. *Computational Economics 21*(1-2), 153–172.

Tauchen, G. and R. Hussey (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica 59*, 371–396.

Tsiang, S. C. (1972). The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *The American Economic Review 62*, 354–371.

Y. Aït-Sahalia, Y. and M. W. Brandt (2001). Variable selection for portfolio choice. *Journal of Finance 56*, 1297–1351.

Yitzhaki, S. (1982). Stochastic dominance, mean variance and gini's mean difference. *American Economic Review 72*, 178–185.