Polynomial sample transmutations of the Normal distribution with an application to option valuation*

Unai Ansejo[1] Aitor Bergara[2] Antoni Vaello-Sebastià[3†]

[1] *Itzarri*

[2] *University of the Basque Country / DIPC / CSIC*

[3] *University of the Balearic Islands*

Abstract

We present a new distribution that consists on a polynomial expansion of the Gaussian quantiles, which nests the well known Cornish-Fisher expansion. Using third-order polynomials we obtain analytical expressions for the distribution and density functions, also yielding a skweness-kurtosis region much wider than the usually considered Gram-Charlier expansion. We discuss three different estimation methodologies using stock index data. We also apply our density for option pricing and hedging obtaining closed-form formulae. Finally, we conduct an empirical application with S&P 500 options data, showing that our option pricing model outperforms other models based on semi–nonparametric densities of similar order.

Keywords: Option Pricing; Cornish-Fisher Expansions; Semi-Parametric Expansions; Skewness; Kurtosis.

JEL Code: G13, G14, C13, C14, C16

*Authors appreciate interesting and valuable discussions with Alfonso Novales, Angel León, Gonzalo Rubio and Javier Mencia. Authors are also grateful to Angel León for providing the options database. Unai Ansejo acknowledges the financial support from the Fundación Ramón Areces. Antoni Vaello-Sebastià acknowledges the financial support from the Spanish Ministry for Science and Innovation through the grant SEJ2007-67895-C04-03.

†Corresponding author: Univ. Illes Balears, Dept. Economía de la Empresa. Crtra. Valldemossa, km. 7.5, Palma de Mallorca, 07122, Illes Balears, Spain. Tel. +34 971 17 2024. e-mail: antoni.vaello@uib.es

1 Introduction

The Black and Scholes (1973) option pricing formula (BS henceforth) has a widespread acceptance among practitioners and academics. However, it is well known that the model is not accurate enough for both deep in-the-money and out-of-the money options. This phenomenon is commonly referred as volatility skew or smile. These misspricing patterns are known to be a result of one of the overidealized assumptions used to derive the formula, namely, the hypothesis of normality for the distribution of log-prices. Actually, daily and weekly returns of the great majority of variables in financial markets present a positive excess of kurtosis (Bouchaud and Potters 2000). This fact implies that extreme events in market variables are much more frequent compared to predictions based on a Gaussian distribution. Usually, some observations in daily series are over 10 standard deviations of the mean. A Gaussian distribution would assign to this phenomenon a probability of order 10^{-23}, while in many cases the observed empirical probability can be of the order of 10^{-4}.

An early improvement to overcome this pricing bias is the semi–nonparametric approximation initially proposed by Jarrow and Rudd (1982). Specifically, they derive an option pricing formula considering an Edgeworth expansion for the log-normal probability density function of stock prices. This approach has been slightly modified in Corrado and Su (1996a) and Corrado and Su (1997b), which have considered a Gram-Charlier expansion to model the distribution of stock log-returns. These semi–nonparametric approximations rely on a moment expansion of the BS formula to account for non-normal skewness and kurtosis. Nevertheless, both approximations share the drawback of yielding negative density function values for certain parameter ranges. Jondeau and Rockinger (2001) propose the restriction of the parameter-range to ensure positive density functions. Additionally, León et al. (2009) have recently proposed the application of semi–nonparametric distributions, considered to be an expansion of any density function. Although the above mentioned approximations keep the analytical flexibility of the Edgeworth-Gram-Charlier moment expansion and solve the negativeness problem, also inherit the strong constraint to describe high degrees of kurtosis and skewness. Actually, León et al. (2009) have demonstrated that in the absence of skewness the maximum kurtosis these expansions can reach is eight, which turns out to be very restrictive.

In this article we propose another expansion-based distribution function. Specifically, we introduce a "sample transmutation" of the Normal distribution through a cubic polynomial. The essence of distribution transmutations is well explained in Shaw and Buckley (2007). Following this work, a "sample transmutation" consists on a deterministic function that transmutes samples of a "base" distribution (the Gaussian in our case) to another distribution. An early example of a polynomial "sample transmutation" of the Normal distribution is the Cornish-Fisher expansion (CFE henceforth), initially proposed by Cornish and Fisher (1937) as an approximation method to estimate quantiles for distributions from known moments.[1] We explain the similarities between both approaches and this is why we call our distribution as Cornish-Fisher distribution (CFD henceforth).

In the empirical applications, we concentrate only on cubic polynomials because bearing the same analytical tractability characteristic of the Edgeworth based density functions, it extends their modeling flexibility in terms of the covered range of skewness-kurtosis possibilities. Moreover, the inverse of a cubic polynomial is available in closed-form. Thus, we can compute the density function analytically, without the need of using numerical methods to invert the transmutation. Moreover, we also provide closed-form formulas for the moments of the CFD. Additionally, we present three different methods for

[1] This expansion has been applied in finance, for example, in delta-gamma approximations for VaR calculation of portfolios containing options, to approximate the percentile of the profit and loss distribution in terms of the sensitivities of the options to underlying movements (see Hull 2004, Ch. 16).

estimating the CFD parameters in a model where returns are identically distributed as a CFD. We also analyze the implications for risk management by comparison of CFE and CFD quantiles in the left tail of the distribution. We obtain significant differences which suppose biased capital requirements when using the Value-at-Risk.

We derive an option pricing model that accounts for skewness and kurtosis, obtaining closed-form expressions for both the price of plain vanilla options and the corresponding "Greeks". Finally, we present an empirical application with S&P 500 options data. The empirical results demonstrate that, first, the fitted implied density function sometimes gives values for the kurtosis coefficient as high as 20 suggesting that Gram-Charlier-based methods are too restrictive in terms of skewness-kurtosis coverage and, second, that it also performs better than other models that allow for asymmetries and positive excess of kurtosis.

The rest of the article is structured as follows. In Section 2, we analyze the statistical properties of the CFD. We discuss in Section 3 the implications for risk management. Section 4 is devoted to the estimation of the CFD parameters. Section 6 deals with the derivation of the option valuation formula and its "Greeks". The empirical application of the option pricing model is conducted in Section 7. Finally, Section 8 presents the conclusions.

2 Statistical characterization of Cornish-Fisher Distributions

In order to introduce our distribution, we initially present the CFE, as defined by Cornish and Fisher (1937). Then we show the relation of CFE with polynomial "sample transmutations". The last part of this section deals with the statistical properties of the new distribution.

2.1 Cornish-Fisher Expansions

A CFE 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. 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. Then, the first terms of the CFE become:

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

where m, σ and κ_i stand for the mean, standard deviation and the *i-th* order cumulant of the variable R respectively. Equation (1) shows the quantile R as a polynomial expansion of the gaussian quantile X. Regrouping the different terms of equation (1) we can appreciate the polynomial coefficients:

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

Formally, the CFE can be understood as a polynomial expansion of the quantile R in terms of the quantile X, where the parameters a_i depend on the cumulants of the distribution F:

$$R = a_0 + a_1X + a_2X^2 + a_3X^3 + \ldots \tag{3}$$

Actually, in order to obtain the CFE, as it is presented in many good references on statistics (for

instance, Johnson and Kotz 1972), one has to re-group the infinite series of equation (3) by a criteria motivated by the central limit theorem. 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 to infinity then R becomes a Gaussian variable. In terms of this variable n the expansion can be written as:

$$R = X + \sum_{k=1}^{\infty} n^{-k/2} \zeta_k(X) \tag{4}$$

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.

Although CFEs are usually applied to approximate theoretically determined distributions (with known moments), they can be used to model random variables as they are directly related to the Edgeworth form of the distribution.[2] Edgeworth (and Gram-Charlier) distributions belong to a family of distributions which provide an explicit relationship between the quantiles of the distribution and the moments or cumulants (see Kendall et al. 1994). In practice, it is unusual to use moments higher than the fourth one when fitting an Edgeworth (or Gram-Charlier) expansion. This is mainly because the possibility of negative density values becomes more probable as higher terms are added, but also because empirical estimation of higher moments is usually highly inaccurate.

With the aim of obtaining a semi–nonparametric distribution, we truncate equation (3) up to order m and consider the coefficients a_i as the fitting parameters of the distribution. With this parametrization, equation (3) can be rewritten as:

$$R = \sum_{i=0}^{m} a_i X^i \equiv Q_m(X) \tag{5}$$

Following Shaw and Buckley (2007), equation (5) establish the variable R as polynomial "sample transmutation" of the variable X. Formally, given a base distribution $\Phi(x)$ (the Gaussian in our case) and another distribution $F(x)$, a sample transmutation mapping T_s (the polynomial in our distribution) is defined by the identity

$$F^{-1}(U) = T_s\left(\Phi^{-1}(U)\right), \qquad \text{i.e.,} \quad T_s(z) = F^{-1}\left(\Phi(z)\right) = Q_F\left(\Phi(z)\right) \tag{6}$$

where $0 \leq U \leq 1$. That is, given the quantile function of the base distribution, Φ, we can obtain the quantiles of the new distribution, Q_F, applying the function T_s.

In other words, instead of using the first terms of the CFE to model the distribution of returns as a function of m, σ, κ_3 and κ_4, we will consider a direct parametrization of the first coefficients, $\{a_i\}_{i=0}^{m}$. Therefore, since the CFE is essentially a polynomial sample transmutation (see equation (2)), from now a variable R that is a m-th order polynomial sample transmutation of the Gaussian variable will be referred as a m-th order *Cornish-Fisher Distribution* (CFD_m) or a m-th order *Cornish-Fisher density* (CFd_m).

Basically, this new parametrization can be understood as a summation of the series made in a different and more efficient order. Comparing equations (2) and (5) we can find the expressions for the coefficients a_i corresponding to a third-order CFE. The main point here is that if we take more terms in the expansion of equation (4) we would find that higher powers of 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

[2] Edgeworth and Gram-Charlier distributions have been implemented in very different fields to model financial returns. 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-Blancard et al. (2001) and Jurczenko et al. (2002).

$\{a_i\}_{i=0}^3$. Thus, taking more terms in the above expansion is required to improve the accuracy of the first coefficients. However, if we parametrize 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.

Other appealing feature of the CFD is the ease of simulation. The following procedure can be considered. First, we simulate standard Gaussian variables X, and second, we apply equation (5), $R = Q_m(X)$, to each simulated observation. The so obtained variable R becomes Cornish-Fisher distributed. In most cases, the function $Q(X)$ is not analytical, because it involves the inverse of the considered distribution function, $Q(X) = F^{-1}[\Phi(X)]$. In these situations the calculation of $Q(X)$ will add more time process to the already "expensive" Monte Carlo method. For instance, Hull and White (1998) use a density function of a mixture of two gaussians which presents the disadvantage of being a transcendent equation, so that it is not possible to obtain the inverse function analytically. Nonetheless, this is not the case of the CFD, given that the function $Q(X)$ is directly modeled as a third-order polynomial.

2.2 Statistical properties of Cornish-Fisher Distributions

Considering that X is the standard Gaussian distribution with distribution function $\Phi(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X} e^{-\frac{1}{2}t^2} dt$, the distribution function of a Cornish-Fisher variable, R, defined by equation (5), will be denoted by $CFD_m(R)$ and can be expressed in the following way:

$$CFD_m(R) = \Phi\left[Q_m^{-1}(R)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Q_m^{-1}(R)} e^{-\frac{1}{2}t^2} dt, \tag{7}$$

where Q_m is the m-th order polynomial and Q_m^{-1} is the inverse function of Q_m.[3] Furthermore, derivating the later expression with respect to R, one can easily find that the density function of a Cornish-Fisher variable, denoted by $CFd_m(R)$, is given by:

$$CFd_m(R) = \frac{d\left[Q_m^{-1}(R)\right]}{dR} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[Q_m^{-1}(R)\right]^2} \tag{8}$$

This work concentrates on a third-order polynomial, since it is sufficiently appropriate to fit experimental data, and it is the first non-trivial approximation useful for financial applications, as we will see in Section 4.[4] On the other hand, the analyticity of the inverse of a third-order polynomial is also specially interesting, since it becomes the basic ingredient for the distribution and density functions (equations (7) and (8)). The explicit form of the third-order CFD function is given in Appendix A. Although these formulas may seem something cumbersome, they are fully analytical and they do not require of numeric procedures like Newton-Raphson to obtain $Q^{-1}(R)$ and its derivative, and they can be easily storaged in a program routine or spreadsheet.

The parameters $\{a_i\}_{i=0}^3$ of the CFD_3 have to be restricted in order to ensure that the density is properly defined. Considering the expression for CFd in equation (8) we can see that to 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 (this also holds for

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

[4] As we are interested in modeling financial returns, we can not consider quadratic polynomials. 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.

any even polynomial). In the following Proposition 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 Propositions are presented in Appendix B.

Proposition 1 *Let $CFd_3(R)$ be a third-order Cornish-Fisher density defined by equation (8), with coefficients $\{a_i\}_{i=0}^3$, then the sufficient and necessary conditions on the coefficients to guarantee the existence of the $CFd_3(R)$ are:*

$$a_3 > 0 \quad , \quad a_1 > 0 \quad , \quad -\sqrt{3a_3a_1} < a_2 < \sqrt{3a_1a_3}. \tag{9}$$

Therefore, a third-order CFD becomes completely defined by equations (5), (8) and (9). Notice that third-order CFDs nest the well known Gaussian distribution ($a_3 = a_2 = 0$), the χ^2 distribution ($a_1 = a_3 = 0$) and the non-central χ^2 ($a_3 = 0$).

Proposition 2 *Let $CFd_m(R)$ be a m-th order Cornish-Fisher density defined by equation (8) for a random variable R, then non-centered r-th order moments, μ'_r, are given by:*

$$\mu'_r = \mathbb{E}\left[R^r\right] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^r e^{\frac{1}{2}J^2}\Bigg|_{J=0} \tag{10}$$

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

It is straightforward to derive the first four non-centered moments of a third-order CFD:

$$\begin{aligned}
\mu'_1 &= a_2 + a_0 \\
\mu'_2 &= 15a_3^2 + a_0^2 + 2a_2a_0 + 6a_3a_1 + a_1^2 + 3a_2^2 \\
\mu'_3 &= 9a_2a_1^2 + 15a_2^3 + a_0^3 + 45a_3^2a_0 + 3a_1^2a_0 + 315a_3^2a_2 + 9a_2^2a_0 + 18a_3a_1a_0 + 90a_3a_2a_1 + 3a_2a_0^2 \\
\mu'_4 &= 105a_2^4 + 60a_0a_2^3 + 18a_0^2a_2^2 + 4a_0^3a_2 + 3a_1^4 + 10\,395a_3^4 + 6a_0^2a_1^2 + 1260a_1a_2^2a_3 + 90a_0^2a_3^2 + \\
&\quad 36a_0a_1^2a_2 + 3780a_1a_3^3 + 36a_0^2a_1a_3 + 60a_1^3a_3 + 90a_2^2a_1^2 + 630a_1^2a_3^2 + 1260a_0a_2a_3^2 + \\
&\quad 5670a_2^2a_3^2 + 360a_0a_1a_2a_3 + a_0^4
\end{aligned}$$

According to these expressions, the centered moments, μ_r, are given by:

$$\begin{aligned}
\mu_1 &= 0 \\
\mu_2 &= 6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2 \\
\mu_3 &= 72a_3a_2a_1 + 8a_2^3 + 270a_3^2a_2 + 6a_2a_1^2 \\
\mu_4 &= 10395a_3^4 + 60a_2^4 + 3a_1^4 + 60a_3a_1^3 + 3780a_3^3a_1 + 936a_3a_2^2a_1 + 4500a_3^2a_2^2 + 630a_3^2a_1^2 + 60a_2^2a_1^2
\end{aligned}$$

Thus, skewness and kurtosis coefficients become:

$$\xi = \frac{72a_3a_2a_1 + 8a_2^3 + 270a_3^2a_2 + 6a_2a_1^2}{\left(6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2\right)^{3/2}} \tag{11}$$

$$\kappa = \frac{10395a_3^4 + 60a_2^4 + 3a_1^4 + 60a_3a_1^3 + 3780a_3^3a_1 + 936a_3a_2^2a_1 + 4500a_3^2a_2^2 + 630a_3^2a_1^2 + 60a_2^2a_1^2}{\left(6a_3a_1 + 15a_3^2 + 2a_2^2 + a_1^2\right)^2} \tag{12}$$

The above equations can be used to characterize the standardized third-order CFD.

Proposition 3 *Let $CFd_3(R)$ be a third-order Cornish-Fisher density function defined by equation (8) for the random variable R, with coefficients $\{a_i\}_{i=0}^3$. Then, one can define a standardized variable R' with zero mean and unit variance imposing*

$$a_0 = -a_2 \qquad , \qquad a_1 = \sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3 \tag{13}$$

with the following conditions on a_2 and a_3 to guarantee the existence of the $CFd_3(R)$:

$$0 \;<\; a_3 < \frac{1}{\sqrt{15}} \tag{14}$$

$$-\sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} \;<\; a_2 < \sqrt{3a_3\left(\sqrt{21a_3^2+1} - 6a_3\right)} \tag{15}$$

Therefore, the standardized third-order CFD is completely defined by equations (5), (8), (13), (14) and (15). The left panel on Figure 1 shows the validity region for (a_2, a_3) pairs according to equations (14) and (15). On the other hand, the right panel on Figure 1 exhibits the corresponding region of skewness and kurtosis. Although the allowed parameter range is bounded, with just two shape parameters, a_2 and a_3, we are able to capture a kurtosis as high as 45 and a skewness up to a value of ± 4. We also plot the equivalent expanded region for two different semi–nonparametric distributions of the same order, the Gram-Charlier distribution of Jondeau and Rockinger (2001), and the semi–nonparametric one of León et al. (2009), as well as the boundary limit for any distribution.[5] According to this Figure, the extended region in the skewness-kurtosis plane within the CFD model is much wider than for the Gram-Charlier distribution or the semi–nonparametric one. Therefore, the CFD becomes much more flexible than other semi–nonparametric approximations.

To better understand the behaviour of the skewness and kurtosis, on one hand, Figure 2 exhibits in the left graphic the skewness surface of the standardized CFD as a function of (a_2, a_3). The right graphic shows level curves of the skewness surface for different values of a_3. An interesting finding is that, given a value of a_3, ξ is linear with a_2. Moreover, the skewness lines do not have intercept since all of them share the origin, and the slope of the skewness lines depends on a_3: the larger a_3, the higher the slope. On the other hand, Figure 3 displays the kurtosis surface (left graphic) showing its convexity and the "spoon" shape. Moreover, κ is symmetric respect to a_2, that is, $\kappa(a_2, a_3) = \kappa(-a_2, a_3)$. The right-side graphic shows the level curves for different values of a_2. We observe that given a_2, κ is a quadratic function of a_3, and the larger the absolute value of a_2, the higher the slope of the parabola.

Potential users of the CFD usually will need to find which parameters (a_2, a_3) match certain values of (ξ, κ), that is, just the inverse of equations (11) and (12). Although we cannot obtain the analytical inverse of those equations we propose two solutions.

First, Table 1 shows the tabulation of (a_2, a_3) of the standardized third-order CFD for certain values for ξ and κ in the horizontal and vertical axes respectively. Every element (i, j) of this table is the pair $\binom{a_2}{a_3}$ that implies the values for the kurtosis and skewness of κ_i and ξ_j. For instance, the pair (a_2, a_3) which generates the values (ξ, κ) = (2, 24) is (0.1455, 0.1719). To obtain the parameters for non tabulated values of ξ and κ we can use a linear interpolation between the nearest values. For example, for (ξ, κ) = (1.75, 19) we obtain by interpolation (a_2, a_3) = (0.1390, 0.1509). The true skewness and kurtosis for these parameters are (1.764, 19.094) which are a good approximation to the initially desired values. For negative values of ξ, we just have to find the parameters as if ξ were positive, and then change the sign of a_2.

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

Secondly, we can also proceed numerically minimizing the following distance:

$$\min_{a_2,\, a_3} \left(\xi(a_2,\, a_3) - \bar{\xi}\right)^2 + \left(\kappa(a_2,\, a_3) - \bar{\kappa}\right)^2 \tag{16}$$

where $\xi(a_2,\, a_3)$ and $\kappa(a_2,\, a_3)$ stand for equations (11) and (12) respectively and $\bar{\xi}$ and $\bar{\kappa}$ are the desired levels of skewness and kurtosis. This optimization converges quickly. For instance, to compute all cases in Table 1 ($18 \times 10 = 180$ optimizations) with a tolerance of 10^{-5} a common laptop only takes 1.92 seconds.[6]

The next proposition characterizes the unimodality of the CFD, which seems a serious restriction of the CFD respect to other expansions. However, as we will see in the next sections, the empirical performance of the CFD is superior when fitting return data and the CFD option pricing model performs better than other approaches.

Proposition 4 *Third-order Cornish-Fisher Densities are unimodal.*

Figure 4 presents some possible shapes for standardized third-order CFds showing their flexibility to describe different degrees of skewness and kurtosis. Figure 5 shows the detail of the tails of the distribution in a logarithmic scale. The solid line represents the relative frequencies (histogram) of daily returns of the YEN/USD exchange rate for the period 01/04/1988 - 08/15/1997. It can be observed that the CFD (dashed line)[7] presents an almost linear behaviour in the tails, as it corresponds to an exponential distribution, and the rate of decrease is much lower than in the Gaussian approximation (parabolic dotted line), so that a much higher weight can be assigned to the tails.

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.

Proposition 5 *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 $\{a'_i\}_{i=0}^m$ given by $a'_i = \sigma a_i$ and $a'_0 = \sigma a_0 + m$.*

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:

$$R = \sigma a_3 X^3 + \sigma a_2 X^2 + \sigma \left(\sqrt{1 - 6a_3^2 - 3a_2^2} - 3a_3\right) X + m - \sigma a_2 \tag{17}$$

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.

2.3 Relation of CFD with QQ-Plots

One can interpret equation (5) as a percentile-percentile relation between a fictitious Gaussian variable, X, and the non-normal variable, R, that we want to describe. We can consider that the value of the variable X that we are fixing is the one that corresponds to a certain percentile α of the distribution R. In this way, equation (5) relates percentiles of R with percentiles of the normal distribution. Therefore, in order to estimate the parameters of the function $Q(X)$ in equation (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.

[6] We have considered all (ξ, κ) combinations despite some of them do not guarantee well-behaved densities.
[7] Section 4.1 provides three different methods to estimate the parameters of the CFD.

As it is well-known, 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 6 we present the QQ-Plot (left panel) for a standardized daily return series of the YEN/USD exchange rate for the period 04/01/1988-15/08/1997 versus a standard normal distribution. Crosses represent experimental data and the solid line corresponds to the fit of a third-order polynomial by least squares. As can be observed, it is remarkable the non-linear shape of the QQ-Plot that discards gaussianity and the high fitting quality obtained with a third-order polynomial. The right-side graphic exhibits the corresponding histogram, along with the fitting of a third-order CFD and the Gaussian distribution. It becomes clear that the resulting non-linear shape of the QQ-Plot derives in a leptokurtic distribution which is more peaked than the Gaussian and exhibits heavier tails. It seems obvious that the CFD function becomes more adequate than the Gaussian one for the USD/YEN exchange rate.

We also can interpret equation (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, implicitly supposes a certain parametric distribution function. Therefore, one can interpret the variable transformation as an alternative form of defining distribution functions, as pointed out by Johnson (1949) or Kendall et al. (1994).[8]

As it is shown in Figure 6, the QQ-Plot of financial series shows a clear deviation from the identity function, as corresponds to a Gaussian variables. Actually, our starting point of using polynomials to fit 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. Nonetheless, as it is presented in Section 4.2, third-order polynomials are enough to describe highly non-linearly behaved financial variables. This characteristic is of special interest, since the transformation based on a series expansion of the QQ-Plot allows us to make a non-perturbative approximation of the distribution function that we want to model.

3 Implications for risk management

The Basel II agreement accepts the *Value at Risk*, VaR, as a measure to calculate the capital requirements of financial institutions. Therefore, VaR has became one of the most important measures in risk management for both practitioners and academics. The VaR is the maximum loss we will expect with probability $1-\alpha$ over a certain time period Δt. Formally, $VaR(1-\alpha,\, \Delta t) = F^{-1}_{\Delta t}(\alpha)$, where $F^{-1}_{\Delta t}$ is the inverse of the distribution function of returns for a Δt period. In other words, $VaR(1-\alpha,\, \Delta t)$ is the α-quantile of the distribution F. Cornish-Fisher expansions have been widely used to account for the presence of skewness and kurtosis in the distribution of returns. Nonetheless, CFE are an approximation and important biases may arise due to the truncation of the expansion series.

In Table 2 we show four quantiles on the left-tail of the standardized third-order CFE and CFD. We consider different levels of kurtosis for the symmetric (left-hand panel) and a negative skewed distribution (right-hand panel), as is observed in many financial series. The quantiles have been computed by simulation, that is, we have simulated 500,000 random i.i.d. standardized Gaussian numbers and we have applied equations (1) and (5) to get the CFE and CFD samples respectively. Then, we obtain

[8] Something similar happens when one defines a density function $f(x)$ in terms of its characteristic function $g(k) = \mathbb{E}(e^{ikx})$.

the quantiles empirically using the 500,000 observations. The last row shows the quantiles for the standardized Gaussian distribution for comparison. As expected, the CFE and CFD quantiles are higher (in absolute value) than the Gaussian ones. Moreover, the CFE provides higher quantiles (in absolute value) than CFD. The differences between CFE and CFD quantiles enlarges with κ and $(1-\alpha)$. This suggests that the CFE provides a more conservative VaR since the associated quantile is larger (in absolute value). For instance, for $\xi = 0$, $\kappa = 20$ and $\alpha = 0.01$, the CFE provides a quantile of -6.20, while the corresponding CFD quantile is -3.04. This result supposes a capital requirement under CFE two times larger than under CFD (!).

These differences are due to the lack of accuracy of the third-order CFE. Columns $\hat{\xi}$ and $\hat{\kappa}$ exhibit the sample skewness and kurtosis computed using the 500,000 simulations for each case. We appreciate how the sample skewness and kurtosis under CFE is much larger than the desired ones, but it is not the case for the CFD sample. Using the previous example, the theoretical kurtosis is 20 but the CFE sample yields $\hat{\kappa} = 84.62$, while it becomes 19.67 for the CFD. To summarize, the CFD is an alternative to CFE to compute the VaR since it is a distribution and does not involve large quantile errors due to the truncation of the expansion.

4 CFD estimation

In this Section, we discuss three different estimation methods to fit the parameters of the CFD, introduced in Section 2. Next, we will consider the CFD model to analyze the in–sample performance of the CFD against other distributions using data of five stock indexes.

4.1 Methods of Estimation

We analyze three different methodologies to fit the coefficients in the transformation function $Q_i(X) = a_{3,i}X^3 + a_{2,i}X^2 + a_{1,i}X + a_{0,i}$ from equation (5):

1. QQ-estimates: The least-squares fitting of the function Q_i arising from the QQ-Plot.
2. MM-estimates: Fit the first four moments of the theoretical distribution.
3. ML-estimates: Choose the parameters that maximizes the logarithm of the likelihood function of the distribution CFD_3.

The first one is the most direct, or computationally less expensive one among the three methodologies because it only involves a least-squares algorithm mainly based in matrix calculations. The fact that the density function, $CFd_3(R)$ 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 (MM). This method involves a non-linear process of optimization, because the moments are functions of the coefficients and the inverse is required to estimate the coefficients as a function of the moments associated to the distribution. It is relevant to point out that in all cases considered the convergence has always been fast when we use the QQ-estimates as the initial guess.

Finally, the maximum-likelihood method (ML) would not be, in principle, well adapted in our case, since coefficients of the third-order polynomial enter in the density function in a highly non-linear way (see Appendix A), which greatly complicates the numerical optimization. However, we have experienced that using adequate starting parameters for the non-linear optimization algorithm, namely the QQ-estimates, we have been able to achieve the global maxima for the log-likelihood functions easily and quickly.

On the other hand, function $Q_i(X)$ has to be invertible and restrictions shown in equation (9) must hold. Given that historical data present a positive excess of kurtosis, these conditions are naturally satisfied when using the QQ-estimates. For the MM and ML methods it is necessary to impose explicitly the restrictions in the optimization procedures. However, using the starting estimates of the QQ-method and given the high flexibility of the third-order CFD, we always get inner solutions. In contrast, Jondeau and Rockinger (2001) obtain many frontier solutions using Gram-Charlier densities, which are not so flexible.

These algorithms have been tested using Monte Carlo experiments. We have considered the fit of CFDs to data generated with both a CFD and a Gaussian mixture. Furthermore, in the latter case we distinguish the situation where parameters are in or out of the restricted domain showed in Figure 1. We consider $N = 100$ series of length $T = 2000$ of standardized CFD data. According to these experiments, in general, the three algorithms are well behaved and the estimations of the QQ and ML are sensibly better than those corresponding to the MM. Besides that, we find that ML-estimate is the most efficient one when the true distribution is CFD. In addition, in the second experiment we have found that the estimation errors and dispersions are very similar in the whole region of permitted values.[9]

4.2 Descriptive data analysis

To illustrate the estimation of the CFD we consider a database of 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 on 519 observations per series ranging from January 4 of 1995 to March 23 of 2005.

As a preliminary investigation of the data, Table 3 presents a summary of the most important univariate statistics. This table shows the first four moments and their corresponding standard errors, in parenthesis, computed with the Generalized Moments Method (GMM) proposed by Bekaert and Harvey (1997).[10] Since the normality hypothesis is crucial to our analysis, we also report information of three well-known tests. First we consider the Jarque-Bera (JB) statistic proposed by Bera and Jarque (1982), which analyzes whether skewness and excess kurtosis are jointly zero. This test is 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.

[9] More details are available from authors upon request.

[10] The mean, σ^2, ξ and κ are jointly estimated using an exactly identified GMM system with four orthogonality conditions. The variance-covariance matrix of the parameters is heteroskedasticity consistent and corrects for serial correlation using a Bartlett kernel with an optimal band as in Andrews (1991).

Since the corresponding p-values for all series and for all statistics are lower than 0.01, we do not report them in Table 3 for shorten.

Parameter estimates for the mean are only significant for the S&P and the STX, and the most volatile indexes are those corresponding to Emerging Markets. The JB 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 coefficients are 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. Summarizing, the JB, Wald and KS tests indicate that the gaussianity hypothesis cannot be accepted for the index data, with p-values close to zero.

4.3 Estimation Results

In this section we use the three estimation methods defined above to estimate the parameters under the hypothesis that series described in Section 4.2 are independent CFD distributed.[11] We test the goodness of the fit using two 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 test we take advantage of the property that if R is a CFD variable, then the variable X, defined as $X = Q^{-1}(R)$, should be normal. Therefore, we can apply on these fictitious X variables any of the usual normality tests to analyze if returns are CFD distributed. In particular, we use the JB test.

Table 4 shows the QQ (Panel A), MM (Panel B) and ML (Panel C) estimates and the corresponding standard errors (in parenthesis). They have been calculated using a bootstrap method with 1000 simulations for the QQ and MM-estimates, and using the Hessian matrix evaluated at the ML estimates for the ML method. The p-values of the goodness of fit tests are displayed in brackets. First, it is interesting to notice 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) behaviour of financial series. Using the KS-statistic at a 5% significance level, the null hypothesis that the data are CFD distributed cannot be rejected for any case. Comparing across the different estimation methods, 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 ones. Similar results can be observed within the JB-statistic. According to this test, we find a rejection of the CFD assumption in none of the series for the three estimation methods. However, the p-values of the JB-statistic suggest that the QQ-estimation method gets the best results.

According to the results presented above, it is easy to conclude that the ML method is the most flexible one in finding good fits and that the assumption of the CFD 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.

Next, we compare the goodness of fit of the CFD with other distributions. Table 5 presents the values of the log-likelihood function for the indexes database considering four different distributions: Gaussian, Johnson U-type distribution (Johnson 1949), the third-order CFE (see equation 2) and the CFD. We have compared the CFD 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 the standard market model, so that in our comparison it might be considered as the first order approximation. The Johnson distribution, as well as the third-order CFD, is a four-parameters distribution and is also very flexible, allowing for

[11] We have also used the CFD to describe daily returns of twelve exchange rates. The findings are rather the same than for stock indexes, and they are available from authors upon request.

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 CFE, we have considered the expansion as a distribution[12] and compare the results with the CFD. Table 5 also reports both the Akaike and Bayesian criteria, which penalize for an increase in complexity through the inclusion of more parameters. The main finding is that the alternative distributions present a lower (similar for NKI series) log-likelihood value.[13] Moreover, both Akaike and Bayesian criteria establish that CFD is preferred to explain all series, except for EM series, where the Johnson distribution is the best for both criteria, although the differences are very small.

5 A dynamic model for financial series

5.1 The model

In this section we present a GARCH-type model with CFD innovations allowing for time varying skewness and kurtosis. Specifically, the most general process we assume for the log-returns of daily financial series is

$$
\begin{aligned}
r_t &= C + y_t && (18)\\
y_t &= \sigma_t x_t \qquad \text{where} \qquad x_t \sim CFD(a_{2,\,t}, a_{3,\,t}) && (19)\\
\sigma_t^2 &= \beta_0 + \beta_1 \sigma_{t-1}^2 + \beta_2 \left(y_{t-1} + \beta_3 \sigma_{t-1}\right)^2 && (20)\\
\tilde{a}_{2,t} &= \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \mathbb{1}_{\{y_{t-1}<0\}} y_{t-1} && (21)\\
\tilde{a}_{3,t} &= \eta_0 + \eta_1 y_{t-1} + \eta_2 \mathbb{1}_{\{y_{t-1}<0\}} y_{t-1} && (22)\\
a_{2,t} &= \mathcal{G}(\tilde{a}_{2,t}) && (23)\\
a_{3,t} &= \mathcal{G}(\tilde{a}_{3,t}) && (24)
\end{aligned}
$$

We initially pre-estimate an ARMA(1,1) process on the return series to remove the predictable part of the returns. Then, we use the residuals of this previous estimation to fit our model, and this why in equation (18) we model the returns as a constant, C, plus an innovation, y_t, which is CFD distributed. We decompose the innovation as the product of σ_t, which stands for the contemporaneus volatility, and $x_t(a_{2,t},\, a_{3,t})$, which is distributed as a standardized CFD, according to equations (5), (8), (13), (14) and (15).

For the variance process in equation (20), we consider the NGARCH(1,1) specification proposed by Engle and Ng (1993). This model nests the common GARCH(1,1) developed by Bollerslev (1986) when $\beta_3 = 0$. Moreover, the NGARCH also allows the well known leverage effect when $\beta_3 < 0$, that is, a negative relationship between returns and volatility.

In order to model the dynamics of the skewness and kurtosis, equations (21) and (22) stand for the dynamics of the unconstrained CFD parameters $\tilde{a}_{2,t}$ and $\tilde{a}_{3,t}$. We can observe that the current value of these parameters depend on the past innovation. The function $\mathbb{1}_{\{\ldots\}}$ is an indicator function that equals one if the expression in brackets is true, then it allows for a asymmetric effects in the skewness and kurtosis (through a_2 and a_3) depending on the sign of y_{t-1}. After that, we obtain the correct (bounded) parameters, $(a_{2,t}, a_{3,t})$ through a logistic mapping denoted by the function $\mathcal{G}(\ldots)$ in equations (23) and

[12] See the comments on the relationship between the CFE and the CFD in Section 2.

[13] Since CFD nests the Gaussian distribution when $a_2 = a_3 = 0$, we have also compared both distributions using the likelihood ratio test obtaining that CFD is always preferred.

(24). We consider a two-steps procedure. Initially, we map $\tilde{a}_{3,t}$ into its correct domain[14]. Secondly, given the mapped value of $a_{3,t}$ we compute the bounds for $a_{2,t}$ using equation (15). Finally, we map $\tilde{a}_{2,t}$ into the allowed domain to obtain $a_{2,t}$.

5.2 Estimation results

Besides the model described by equations (18) to (24) we also consider five compiting models all of them nested in the most general one. We consider two variance specifications, GARCH and NGARCH, to measure the relevance of the leverage effect. We also introduce the classical gaussian innovations. Finally, we restrict equations (21) and (22) to asses the relevance of asymmetries in the conditional skewness and kurtosis. Table 6 summarizes the different models considered. We estimate all models by QMLE and the standard errors have been calculated using the Hessian matrix.

For the empirical estimation we consider the last decade of daily returns of the stock indexes S&P 500 and EuroStoxx 50. Figure 7 exhibits the time series of both indexes. Specifically, our sample goes from 03/17/2000 to 03/20/2010. This period covers the ".com" and the subprime crises located at the beginig and at the end of the sample respectively, while the period between both crises is characterized by a lower volatility and a positive trending. In Table 7 we summarize the main statistics of both series. Both series show similar standard deviations and the maximum and the minimum daily return in both samples are also quite similar. Nonetheless, the S&P 500 is more leptokurtic and the asymmety (negative) is higher. As expected, the Jarque-Bera statistic rejects the hypothesis of normality in both series.

Tables 8 and 9 show the results of estimating the six candidate models. The last three rows correspond to the Akaike (AIC), Schwarz (SIC) and Bayesian (BIC) criteria to measure the goodness of fit of the models. In general, the estimates are quite significant (estimates without markers). Comparing Model 1 against Model 2, we observe that the leverage effect is a relevant feature. For both series Model 2 yields a much higher value of log-likelihood function, and it is preferred to model 1 for the three measures. When we compare Model 1 and Model 3, we observe similar results, that is, a GARCH model with non-Gaussian innovations is preferred to the normal-GARCH.

It is also relevant to compare Model 2 against Model 3, that is: what is preferable, to use a better specification of the variance equation considering the leverage effect (Model 2), or to use a simple GARCH model but with non-normal innovations. According with our results, Model 2 is preferred to Model 3. Model 2 yields a higher value of the likelihood function. Moreover, note that Model 2 is more parsimonius than Model 3, therefore the three goodness of fit measures are clearly favourable to the Model 2 against Model 3. Model 4 considers jointly the leverage effect and non-normal innovations with time-varying skewness and kurtosis outperforms the more parsimonious especifications despite the increase in the number of parameter. Finally, to introduce an asymmetric effect in the dynamics of the CFD parameters, Model 6, supposes an improvement according to AIC, but not for SIC and BIC, which penalyze stronger the number of parameters.

An interesting finding is that the estimated parameters of the variance equation do not change widely when we consider different dynamics for the innovations. Thus, a good strategy to estimate the most sophisticated models is to estimate first the simple GARCH or NGARCH, and to use these estimates as the starting values to estimate more general models. For those cases where the optimization does not converge, we could restrict the model assuming the variance parameter estimates of the simplest models.

[14]See proposition 3 and equations (13) to (15) for more details.

From the estimation of the different GARCH models we obtain the daily series of $a_{2,t}$ and $a_{3,t}$. From these series we compute the daily conditional skewness and kurtosis using equations (11) and (12). Figure 8 displays the skewness and kurtosis time series for the S&P 500 sample using Model 6. The first result is that the skewness is always negative and the kurtosis can change widely. For the "crisis" periods it can take values around 40. It seems that there is relationship between skewnees and kurtosis: the higher the skewness (in absolute value), the higher the kurtosis...

6 Option valuation

In this Section we derive closed-form formulae for the price and the *greeks* of European-style options with underlying characterized by a CFD. We also compare the behaviour of the CFD model with the model proposed by Corrado and Su (1996a).

6.1 European Option Valuation

As it is well known, the BS model is based on the hypothesis that asset prices, S_t, follow a geometric Brownian motion under the risk-neutral probability $\mathbb{Q}$ given by:

$$S_T = S_t e^{\left(r-\frac{\sigma^2}{2}\right)\tau+\sigma(W_T-W_t)} \tag{25}$$

where r is the risk-free rate, σ is the volatility of the underlying, S_t is the initial price, 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 $R_{t-s} = \ln(S_t/S_{t-s})$, follow a normal distribution, $N(r, \sigma\sqrt{t-s})$. In general, we can obtain the price of an European call option as the discounted expected value of the payment under the risk-neutral probability:

$$C_t = e^{-r\tau}\mathbb{E}_{\mathbb{Q}}[(S_T - K)^+] \tag{26}$$

Considering a geometric Brownian motion, the application of this formula yields the well known BS formula. Nonetheless, equation (26) is of general use for any other distribution governing S_T. In this work we make the hypothesis that prices differences can be expanded through a CFD. Therefore, the first order term in the expansion of the price S_T will be normally distributed, instead of log-normal as in the BS framework, and hence, the first order approximation of the price of a European call using the CFD will not exactly coincide with the BS formula, as it is the case of approaches based on an Edgeworth or Gram-Charlier expansion (Jarrow and Rudd 1982 or Corrado and Su 1996a).

Notice that we do not consider the expansion of the log-returns, R_t, in terms of CFD, as would be more natural, because in such a case prices, S_t, would follow a distribution with divergent mean (*i.e.* $\mathbb{E}\left[e^{R_t}\right] = \infty$) and, therefore, its application to option pricing would be useless.[15] Nevertheless, this method, regarded as an approximation, is theoretically equally valid as any other expansion based method and, moreover, it avoids negative density values.

Due to the above mentioned reasons, the first order approximation of the CFD expansion considered in this work will be the process

$$S_T = S_t(1 + r\tau) + S_t\sigma(W_T - W_t) \tag{27}$$

[15] Let us consider the expression: $\mathbb{E}\left[e^R\right] = \int_{-\infty}^{\infty} e^R CFd_m(R)\,dR$, where $CFd_m(R)$ is given by equation (8). 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.

rather than the geometric one.[16] It is important to remark that, although the process in equation (27) includes the possibility of negative prices, in our framework this feature has to be considered as an approximation nuisance; a very small probability of 10^{-20} is in practice virtually equal to zero. Moreover, this theoretical drawback allows us to obtain a positive definite density function which permits rather high kurtosis and skewness levels.[17]

Let r be the risk-free interest rate, K the exercise price of the option, S_t the initial price of the underlying, τ the time to maturity and σ the volatility of the process. It is easy to show that in the absence of arbitrage opportunities the price of a call option, whose underlying follows equation (27), is given by:

$$C = \frac{1}{1+r\tau}\left[\left(S_t\left(1+r\tau\right)-K\right)\Phi(-d)+S_t\sigma\sqrt{\tau}\phi(d)\right] \tag{28}$$

where $d = \left(K - S_t(1+r\tau)\right)/\left(S_t\sigma\sqrt{\tau}\right)$ and $\Phi\left(x\right)$ and $\phi(x)$ stand for the distribution and density functions of the standard Gaussian variable respectively.

This article assumes that the model for the asset price, S_T, under the risk neutral measure $\mathbb{Q}$ is given by:

$$S_T = S_t(1+r\tau) + S_t\sigma\sqrt{\tau}z^* \tag{29}$$

where z^* is a variable following a standardized CFD. Note that the the martingale restriction holds, *i.e.* $\mathbb{E}_{\mathbb{Q}}(S_T) = S_t(1+r\tau)$. Under the assumption of a CFD distribution we can obtain a more general formula for the option price, since both skewness and excess of kurtosis different from zero are possible under this distribution.

Proposition 6 *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 the absence of arbitrage opportunities the call option, C^{CFD}, whose underlying follows a third-order CFD, given by equation (8), is:*

$$\begin{aligned} C^{CFD} &= \frac{1}{1+r\tau}\left\{\begin{array}{c}\left(S_t(1+r\tau)-K\right)\Phi\left(-d\right)+\\ \sigma\sqrt{\tau}S_t\phi(d)\left(a_3\left(d^2-2\right)+a_2 d+\sqrt{1-6a_3^2-2a_2^2}\right)\end{array}\right\} \\ Q_\tau(x) &= S_t(1+r\tau)+S_t\sigma\sqrt{\tau}\left(a_3x^3+a_2x^2+\left(\sqrt{1-6a_3^2-2a_2^2}-3a_3\right)x-a_2\right) \end{aligned} \tag{30}$$

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 equation (30) becomes equation (28) when $a_3 = a_2 = 0$, as we would restore the normal distribution for S_T.[18] Although not reported here, 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 are already able to capture most non-normal features.

We consider a hypothetical call option with $K = 10$, one month to maturity, $\sigma = 0.40$ and $r = 0.05$, to analyze the valuation differences between the Black-Scholes model and the CFD approximation. We select different degrees of ξ and κ. In particular, we choose parameters $(a_2,\, a_3)$ equal to $(0,\, 0.092)$,

[16] Bouchaud and Potters (2000) analyze pricing differences between both approximations. They find that with a daily volatility of 1% and a zero interest rate, the relative difference between the geometric and arithmetic BS 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 difference is very small when we consider absolute differences.

[17] 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. However, in this case valuation formulas are more complicated and the pricing differences are almost negligible for reasonable parameter sets.

[18] The inverse of the polynomial $Q_\tau^{-1}(K)$ is equal to $\left(K - S_0(1+r\tau)/S_0\sigma\sqrt{\tau}\right)$ when $a_3 = a_2 = 0$.

(−0.11, 0.079) and (0.11, 0.079), which correspond to (ξ, κ) 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. Figure 9 shows the differences between the CFD and BS call prices. The presence of positive (negative) skewness makes the upper (lower) tail of the price density thicker under the CFD model, and this feature induces an increase in the probability of a large drop (raise) in prices, which is responsible for the relative lower (higher) prices given by the BS formula for deeply out-of or in-the money call options.

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 for European options based on Gram-Charlier distributions, which also includes the effects of skewness and kurtosis. Thus, we discuss it briefly and consider it later in Section 7. The pricing formula of Corrado and Su (1996a) for a European call, C_{CS}, is[19]:

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

It is trivial to see that if the implied density function of log-returns is Gaussian (*i.e.* $\xi = 0$ and $\kappa = 3$), this expression becomes the BS formula.

In Figures 10 and 11 we analyze the differences between the Corrado and Su (CS henceforth) and the Cornish-Fisher valuation formula for different values of ξ and κ. These figures compare the BS implied volatilities obtained with both models for the European option described above. Interestingly, we can observe significant pricing differences although the first four moments of the implicit density function for both models are the same. Figure 10 shows that the CS model gives higher implied volatility values for deep out-the-money (otm) and deep in-the-money (itm) call options: with the underlying equal to 5 and $\xi = 0$ the CS model implies a volatility of 57%, meanwhile for the CFD the implied volatility is 42%. On the other hand, for at-the-money (atm) and near itm options the CFD model gives higher implied volatility values than CS. Analyzing the absolute and relative differences we observe that the highest deviations are found for otm options for relative differences and in atm options for absolute differences. The observed differences allow us to conclude that the implied density of the CS presents heavier tails on the negative part than the CFD model.

It is also interesting to note that the CS model gives a hump shaped volatility pattern for deep otm options, which the CFD model does not exhibit, displaying a maximum implied volatility for an underlying value around 6. This "anomaly" is not observed in real volatility smile representations (Jondeau and Rockinger 2000, e.g.) and it can be explained because of the incorrect negative density values included in the CS model.

[19]Brown and Robinson (2002) found an error in the formula initially proposed by Corrado and Su (1996a). Along this article we have considered the corrected option pricing formula.

6.2 Hedging Parameters

In this subsection we characterize the hedging parameters (the *Greeks*, as they are more commonly known) under the CFD option pricing model. One of the most important is the *Delta*, Δ, which 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 CFD model are rather cumbersome, and we will just show here the expressions for the *Delta* and the *Gamma*:

$$\Delta_{CFD} = \phi(d)\sqrt{\tau}\sigma \left(\begin{array}{c} \Phi(-d) + \phi(d)\left(\left(S_t\left(1+r\tau\right)-K\right)d_K\right) + \\ \left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) - S_t d d_K\left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) + \\ S_t\left(a_2 d_k + 2a_3 d d_K\right) \end{array} \right)$$

$$\Gamma_{CFD} = \phi(d)\sqrt{\tau}\sigma \left(\begin{array}{c} \phi(d)\left(S_t\left(1+r\tau\right)-K\right)\left(d d_K^2 - d_{KK}\right) + \\ S_t\sigma\left(a_2 d_k + 2a_3 d d_K\right) - 2\phi\left(d\right)d_K\left(1+r\tau\right) + \\ -2 d d_K\left(a_1 + a_2 d + a_3\left(1+d^2\right)\right) \\ S_t\left(a_1 + a_2 d + a_3\left(1+d^2\right)\left(d^2 d_K^2 - d_K^2 - d d_{KK}\right)\right) + \\ 2\left(a_2 d_K + 2a_3 d d_K\right) + S_t\left(a_2 d_{KK} + 2a_3\left(d_K^2 + d d_{KK}\right)\right) \end{array} \right)$$

where :

$$d = \tilde{Q}_\tau^{-1}(K/S_t)$$

$$d_K = \left.\frac{-K}{S_t^2}\frac{\partial \tilde{Q}_\tau^{-1}(x)}{\partial x}\right|_{x=K/S_t}$$

$$d_{KK} = \left.\frac{K^2}{S_t^4}\frac{\partial^2 \tilde{Q}_\tau^{-1}(x)}{\partial x^2}\right|_{x=K/S_t} - 2\left.\frac{K}{S_t^3}\frac{\partial \tilde{Q}_\tau^{-1}(x)}{\partial x}\right|_{x=K/S_t}$$

$$\tilde{Q}_\tau(x) = \sigma\sqrt{\tau}\left(a_3 x^3 + a_2 x^2 + \left(\sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3\right)x - a_2\right) + \left(1 + r\tau\right)$$

When the European call market price is given by the CS formula, the *Delta* and *Gamma* of a call options can be written (Backus et al. 1997) as:

$$\Delta_{CS} \simeq \Phi(d) + \phi(d)\left\{\frac{\gamma_1}{6}\left(d^2 - 3d\sigma\sqrt{\tau} + 2\sigma^2\tau - 1\right) + \right.$$
$$\left. + \frac{\gamma_2}{24}\left(-d^3 + 4d^2\sigma\sqrt{\tau} + 3d - 3d\sigma^2\tau - 4\sigma\sqrt{\tau}\right)\right\}$$

$$\Gamma_{CS} \simeq \frac{\phi(d)}{S_0\sigma\sqrt{\tau}}\left\{1 + \frac{\gamma_1}{6}\left(-d^3 + 3d^2\sigma\sqrt{\tau} - 2d\sigma^2\tau + 3d - 3\sigma\sqrt{\tau}\right)\right.$$
$$\left. + \frac{\gamma_2}{24}\left(d^4 - 4d^3\sigma\sqrt{\tau} + 3d^2\sigma^2\tau - 6d^2 + 12d\sigma\sqrt{\tau} - 3\sigma^2\tau + 3\right)\right\}$$

$$d = \frac{\ln\left(S_0/K\right) + \left(r + \sigma^2/2\right)\tau}{\sigma\sqrt{t}}$$

In Figures 12 and 13 we compare the *Delta* and *Gamma* functions derived from the CFD and CS models, considering the same parameters used in Figures 9 and 10. We appreciate that for both deep itm and otm options the CS model gives higher *Delta* values than the CFD, and vice versa for atm options. These differences are significant as they can be as high as 20%. On the other hand, calculated *Gamma* values also show substantial differences: atm options present similar *Gamma* values for both models, but for deep otm or itm options the CFD model derives more reliable results given that the CS model

presents negative values, which are theoretically "surprising" for call options (see Hull 2004). Again, this anomaly (the *Delta* function presents a local maxima) is associated to the negative density values and through the fact that the *Delta* and *Gamma* expressions for the CS model are only approximations and cannot be calculated analytically. Therefore, before turning to the empirical comparison between both models, we can conclude that the CFD option pricing model is theoretically more coherent than the CS model as it does not derive "dubious" hedging values for deep otm or itm options.

7 Empirical Performance of CFD Option Pricing

7.1 Data Description and Estimation Procedure

The analysis of the empirical performance of the CFD option pricing model is based on the S&P 500 index options database of Dumas et al. (1998). Options in this database are of European-style and expire on the third Friday of each month. Option prices were collected every Wednesday between 2:45 p.m. and 3:45 p.m. from June 1988 to December 1993, which includes a total number of 292 weekly spaced days. We have considered the bid-ask mid price for estimation purposes. The risk-less interest rate is approximated by the T-bill rate implied by the average bid and ask discounts reported in the *Wall Street Journal*. This database accounts for the presence of dividends providing the implied forward prices, computed as the current stock price minus the present value of dividends times the interest accrued until maturity, *i.e.* $F_{t,T} = (S_t - \hat{D}_t)e^{r\tau}$. The final sample contains 28417 option prices (15889 calls and 12528 puts) after removing options that do not verify arbitrage restrictions.

We compare the performance of the CFD option valuation model with respect to the standard option valuation for European call on futures of Black (1976), B76, the (corrected) model of Corrado and Su, CS, and a practitioner-style method, P-BS, which combines the BS formula with a deterministic volatility function that is assumed to be a quadratic function of moneyness.[20] Specifically, we assume that

$$\sigma(x) = \lambda_0 + \lambda_1(x - \lambda_2)^2 \tag{31}$$

with $x = K/F_{t,T}$ and $\lambda_0 > 0$, $\lambda_1 \geq 0$ in order to ensure positivity. To illustrate this methodology, in Figure 14 we show the daily estimated volatility functions using option data from 08/28/1991 and the BS implied volatilities (crosses). Moreover, we split the sample into two groups depending on the option maturity as it is explained in the next paragraph. First, it is interesting to observe the well known smile/smirk shape of the estimated volatility functions. Second, the volatility smile is more pronounced for short-term options, according to previous articles, such as Bakshi et al. (1997).

We have carried out both an in–sample and an out–of–sample analysis. The parameters of each model (σ for B76, $\{\lambda_i\}_{i=0}^2$ for P-BS, $\{\sigma, \xi, \kappa\}$ for CS and $\{\sigma, a_2, a_3\}$ for CFD) are calibrated for each of the 292 days of the sample and have been computed from the cross-section of options prices. We have also analyzed the model performances for different subsamples. Specifically, we consider i) call and put options separately, and ii) we split the option database between short-term and long-term options. Short-term options are those with a time to maturity lower than 61 days, while long-term options have an expiration larger than 60 days. The implicit estimator for date t is defined as the minimizer of the mean of the squared pricing errors for the options traded that day, that is,

$$\hat{\theta}_t = \arg\min_{\theta} \frac{1}{n_t} \sum_{i=1}^{n_t} [c_i(\theta) - c_i]^2 \tag{32}$$

[20] This method has been studied widely in Dumas et al. (1998).

where $c_i(\theta)$ is the theoretical option price, c_i, denotes the market option price and n_t is the total number of options at date t. For calibration reasons, we discard those days where the number of observations (contracts) is lower than six, that is, two times the number of parameters to be calibrated. This rule only applies for the subsamples of long-term options and specially for the puts. Moreover, for the most unfavorable case it only supposes the discard of less than the 5% of the sample days.

7.2 Option Pricing Results

Table 10 displays the in–sample results of the four option pricing models (B76, P-BS, CS and CFD). We report the daily average for the 292 sample days of the root mean squared error (RMSE) and mean absolute error (MAE) as the performance measures. Bold numbers represents the model with the best performance. The main result is that CFD produces the lowest RMSE and MAE in all cases. We can also appreciate that the Practitioner-BS model performs better than CS. Figure 15 exhibits the scatter plot for the option implied pairs (ξ, κ) under both the CS and the CFD option pricing models for the whole sample. This figure also displays the bounds for the allowed (ξ, κ) values for the underlying distributions (see Figure 1). We observe that the CS pairs (circles) equal the bound for non negativity for a large number of sample days, while the CFD model yields higher values for both the implied skewness (in absolut value) and kurtosis for a great deal of days (dots). This finding corroborates the statement that Edgeworth based distributions are too restricted to capture high degrees of skewness and kurtosis implicit in option prices.[21]

Given the in–sample calibrated parameters, we use them to test the models out–of–sample. We use the parameters calibrated at date t to value options at time $t + \Delta t$. We would like to remark that this exercise supposes a one week out–of–sample analysis, since our option data is only sampled on Wednesdays ($\Delta t = 7$ days). The results for the out–of–sample analysis are shown in Table 11. Again, the main result is that the CFD model outperforms the rest of the models for the whole option sample and all subsamples considered, yielding the lowest error measures. Nonetheless, in contrast with the in–sample results, we observe that now the CS model generates lower RMSE and MAE than P-BS in the majority of the situations.[22]

It is also interesting to analyze the evolution of the option implied moments. Figure 16 displays the implicit volatility (top), skewness (medium) and kurtosis (bottom) for each sample day for the models B76, CS and CFD. The implicit volatilities are quite similar across models, although the evolution of the CFD volatility is slightly sharper. The CFD implied skewness, in general, becomes more negative than CS one, and the option implied CFD distributions are more leptokurtic than the Gram-Charlier ones.

Summarizing, we can state that the CFD model generates lower option pricing errors than B76, P-BS and CS, in–sample and one week out–of–sample. Moreover, the CFD model is much more flexible than CS to capture risk-neutral moments implied in the option prices.

7.3 Hedging Performance

In this section we conduct an analysis of the hedging performance of the different option pricing models. We use the approach of Dumas et al. (1998), that is, we assume that the hedge portfolio is continuously

[21] 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.

[22] Besides of the sample average, we also have considered the median of the 292 estimated RMSEs and MAEs, which is a more robust statistic. In this case, the results hold in and out–of–sample. Tables similar to Tables 10 and 11, but using the median, are available upon request.

rebalanced. Therefore, the hedge portfolio error is defined as

$$\epsilon_t = \Delta C_{real,t} - \Delta C_{model,t} \tag{33}$$

where $\Delta C_{real,t}$ is the observed (real) variation in the option price from day t until day $t+7$ and $\Delta C_{model,t}$ is the corresponding variation using the theoretical value. The intuition behind equation (33) is that when the hedge is continuously rebalanced, the hedging error is simply equal to the time increments in the valuation errors.[23]

Table 12 shows the root mean squared hedging error (panel A) and the mean absolute hedging error (panel B) for the four option pricing models considered and differentiating by option maturity and type of option like in Tables 10 and 11. The best performance is again in bold. We observe that the model with the lowest hedging errors is CFD for the two performance measures and for all subsamples considered.

8 Conclusions

Cubic polynomial base transmutations of the Gaussian distribution are shown to be very efficient to model densities which even strongly deviate from the normal case. In this paper we analyze the parametric properties of the expansion and allows us to present the Cornish-Fisher distribution. We determine the moments and the domain of skewness and kurtosis over which the CFD is valid. This domain is considerably wider than the corrections to the Edgeworth density proposed by Jondeau and Rockinger (2001) or León et al. (2009).

The new distribution is applied to the estimation of densities in an independent and identically distributed framework, and three different estimation methods have been proposed: maximum likelihood, the moments method and a quantile based method. We find that maximum likelihood provides the best estimates in terms of the Kolmogorov-Smirnov test, and the quantile-based method is a good choice for starting values in optimization procedures.

We provide closed-form option price and hedging parameters formulae assuming that price differences follow a third-order CFD under the risk-neutral measure. In this way, we have proposed a new semi–nonparametric generalization of the Black and Scholes formula 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 anomalies of the CS model (negative *Gamma* values for deep out or in-the-money options), although both are semi–nonparametric.

We have evaluated the performance of the CFD option pricing model using a S&P 500 option database. Both the in–sample and out–of–sample results confirm that CFD option pricing model outperforms the CS approach and a practitioner–style method. Moreover, the CFD option pricing model also yields lower hedging errors. These results hold when separating the option database between short and long maturity options. We also find that the Gram-Charlier density is very restrictive to capture the option implied moments. Therefore, given that, i) CFD option pricing model performs better than CS "in" and out–of–sample, ii) CFD does not present the anomaly in the hedging parameters, and iii) CFD generates lower hedging errors, we can conclude that the CFD model is preferable than the CS model within the semi–nonparametric class of distributions for valuing options.

Comparison of the option pricing performance and of the modeling flexibility between the CF distribution and other option pricing models, that also allow for heavy-tailedness but do not belong to

[23] See Dumas et al. (1998) for more details about equation (33).

the semi–nonparametric class, as a non-parametric method based on a mixture of log–normal densities, the parametric approach of Malz (1996) which assumes a jump-diffusion for the underlying process, or the stochastic volatility model of Heston (1993), could be of interest and is left for future research. Moreover, to develop a GARCH model with CFD innovations allowing for time-varying skewness and kurtosis is a promising application of the results of this article.

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:

$$CFd_3(R) = \frac{d\left[Q^{-1}(R)\right]}{dR}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[Q^{-1}(R)\right]^2} \tag{34}$$

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 of a third-order polynomial and its first derivative can be calculated analytically:

$$\begin{aligned} Q^{-1}(R) &= -\frac{a_2}{3a_3} + \sqrt[3]{\sqrt{B_1}R + a_3B_2 + \sqrt{B_1R^2 + B_2R + B_3}} - \\ &\quad \frac{1}{3}\frac{a_1/a_3 - a_2^2/3a_3^2}{\sqrt[3]{\sqrt{B_1}R + a_3B_2 + \sqrt{B_1R^2 + B_2R + B_3}}} \end{aligned} \tag{35}$$

and:

$$\begin{aligned} \frac{d\left[Q^{-1}(R)\right]}{dR} &= \frac{\sqrt{B_1} + \frac{\left(\frac{1}{2}B_2 + RB_1\right)}{\sqrt{B_3 + RB_2 + R^2B_1}}}{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}} \times \\ &\quad \left(\begin{array}{c} \frac{1}{3}\sqrt[3]{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}} + \\ \frac{1}{27a_3^2}\frac{3a_3a_1 - a_2^2}{\sqrt[3]{a_3B_2 + R\sqrt{B_1} + \sqrt{B_3 + RB_2 + R^2B_1}}} \end{array} \right) \end{aligned} \tag{36}$$

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

$$\begin{aligned} B_1 &= \frac{1}{4a_3^2} \\ B_2 &= \frac{a_2a_1}{6a_3^3} - \frac{a_0}{2a_3^2} - \frac{a_2^3}{27a_3^4} \\ B_3 &= -\frac{a_2a_1a_0}{6a_3^3} + \frac{a_0^2}{4a_3^2} + \frac{a_1^3}{27a_3^3} + \frac{a_2^3a_0}{27a_3^4} - \frac{a_2^2a_1^2}{108a_3^4} \end{aligned} \tag{37}$$

B Proofs

B.1 Proposition 1

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: $Q'(x_m) = 3a_3x_m^2 + 2a_2x_m + a_1 = a_1 - \frac{1}{3}\frac{a_2^2}{a_3} > 0$, which implies the following conditions: $-\sqrt{3a_3a_1} < a_2 < \sqrt{3a_1a_3}$ and $a_1 > 0$. The result of this Proposition could be easily generalized to a fifth order polynomial. ■

B.2 Proposition 2

Proof. Consider the Fourier transform of the random variable R:

$$\hat{P}(k) = \mathbb{E}\left[e^{ikR}\right]$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator. After making the following variable transformation $R = Q(X)$:

$$\hat{P}(k) \quad = \quad E\left[e^{ikQ(X)}\right] \Leftrightarrow \hat{P}(k) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + ikQ(X)} dX$$

Expanding the exponential in power series we obtain:

$$\hat{P}(k) = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \int \frac{1}{\sqrt{2\pi}} Q(X) e^{-\frac{1}{2}X^2} dX \tag{38}$$

Using this expression we obtain the moments of the variable R can be expressed by the following integral:

$$\mathbb{E}[P^m] = \int \frac{1}{\sqrt{2\pi}} Q(X) e^{-\frac{1}{2}X^2} dX \tag{39}$$

Next, we define the functional generator $\hat{P}(k, J)$ as:

$$\hat{P}(k, J) = \int \frac{1}{\sqrt{2\pi}} Q(X) e^{-\frac{1}{2}X^2 + JX} dX \tag{40}$$

where the variable J is included in order to calculate the integrals. If $Q(X)$ is a polynomial it is easy to demonstrate that equation (39) reduces to:

$$\mathbb{E}[P^m] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^m \cdot \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + JX} dX \Bigg|_{J=0} \tag{41}$$

The integral (41) can be carried out analytically:

$$\int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}X^2 + JX} dX = e^{\frac{1}{2}J^2} \tag{42}$$

obtaining, finally, the following expression:

$$\mathbb{E}[P^m] = \left[Q\left(\frac{\partial}{\partial J}\right)\right]^m \cdot e^{\frac{1}{2}J^2} \Bigg|_{J=0} \tag{43}$$

■

B.3 Proposition 3

Proof. For the first result one just has to impose $\mu_1' = 0$ and $\mu_2' = 1$ in equations (11). In order to prove the second one we need to replace the conditions $a_0 = -a_2$ and $a_1 = \sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3$ in the inequation $-\sqrt{3a_3 a_1} < a_2 < \sqrt{3a_1 a_3}$, which guarantees the existence of the CFD. As a result, one

finds:

$$-\sqrt{3a_3\left(\sqrt{1-2a_2^2-6a_3^2}-3a_3\right)} < a_2 < \sqrt{3a_3\left(\sqrt{1-2a_2^2-6a_3^2}-3a_3\right)} \tag{44}$$

Solving this equation for a_2 one obtains that this is equivalent to:

$$-\sqrt{3a_3\left(\sqrt{21a_3^2+1}-6a_3\right)} < a_2 < \sqrt{3a_3\left(\sqrt{21a_3^2+1}-6a_3\right)} \tag{45}$$

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:

$$0 < a_3 < \frac{1}{\sqrt{15}} \tag{46}$$

■

B.4 Proposition 4

Proof. To prove this Proposition we will consider the expression of the density function of a third-order Cornish-Fisher distributed variable:

$$CFd_3(R) = \frac{d\left[Q_3^{-1}(R)\right]}{dR}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[Q_3^{-1}(R)\right]^2} \tag{47}$$

Derivating this expression with respect to R and making $d\left(CFd_3(R)\right)/dR = 0$, we find the condition for a maximum of the density function, R_m,:

$$\begin{aligned}\frac{d\left(CFd_3(R)\right)}{dR} &= \left[\frac{d^2\left[Q_3^{-1}(R)\right]}{dR^2} - \left(\frac{d\left[Q_3^{-1}(R)\right]}{dR}\right)^2 Q_3^{-1}(R)\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left[Q_3^{-1}(R)\right]^2} = 0 \\ &\Rightarrow \frac{d^2\left[Q_3^{-1}(R_m)\right]}{dR^2} - \left(\frac{d\left[Q_3^{-1}(R_m)\right]}{dR}\right)^2 Q_3^{-1}(R_m) = 0 \end{aligned} \tag{48}$$

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 $R \to \pm\infty$ we have $Q_3^{-1}(R) \to \pm\infty$ and $Q_3^{-1}(R)'' \to 0$. Therefore, we have three 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 (48) cannot hold. When $R < R_c$ then $Q_3^{-1}(R) < 0$ and $Q_3^{-1}(R)'' > 0$ and also equation (48) cannot hold. Thus, 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 (48) cannot hold. If $R < R_i < R_c$ then $Q_3^{-1}(R) < 0$ and $Q_3^{-1}(R)'' > 0$ and equation (48) 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 (48) 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 (48) can only hold

once.

iii) Finally, the third case is very similar to the second one. ■

B.5 Proposition 5

Proof. We only have to consider the definition of a Cornish-Fisher distributed variable to prove this proposition. In Equation 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 $a'_i = \sigma a_i$ and $a'_0 = \sigma a_0 + m$. ■

B.6 Proposition 6

Proof. We consider that the underlying S_T at the time of maturity T can be approximated through:

$$S_T = S_t(1 + r\tau) + S_t\sigma\sqrt{\tau}z^* \tag{49}$$

where z^* is a standardized third-order CFD variable with parameters a_2 and a_3. In the absence of arbitrage opportunities the price of a European call option should be:

$$C = \frac{1}{1 + r\tau}\mathbb{E}\left[(S_T - K)^+ | F_t\right] \tag{50}$$

Therefore, we have to evaluate the following integral

$$C = \frac{1}{1 + r\tau}\int_K^\infty (S_T - K)\, CFd_3(S_T)dS_T \tag{51}$$

where $cf_3(S_T)$ denotes the density function of the third-order CFD variable S_T. In order to evaluate this integral, we consider the 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 CFD with coefficients a_2 and a_3

$$\begin{aligned} C &= \frac{1}{1 + r\tau}\int_K^\infty \left(S_t(1 + r\tau) + S_t\sigma\sqrt{\tau}\tilde{Q}_\tau(x) - K\right) CFd_3(S_T)dS_T \\ &= \frac{1}{1 + r\tau}\left\{(S_t(1 + r\tau) - K)\int_K^\infty CFd_3(S_T)dS_T + S_t\sigma\sqrt{\tau}\int_K^\infty \tilde{Q}_\tau(x)CFd_3(S_T)dS_T\right\} \end{aligned}$$

We can calculate the first integral using the above mentioned variable change, $S_T = Q_\tau(x)$ and the relation that holds in any percentile-to-percentile transformation, namely, $CFd_3(S_T)dS_T = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$, obtaining:

$$C = \frac{1}{1 + r\tau}\left\{(S_t(1 + r\tau) - K)\left(1 - \Phi(Q_\tau^{-1}(K)\right) + S_t\sigma\sqrt{\tau}\underbrace{\int_{Q_\tau^{-1}(K)}^\infty \tilde{Q}_\tau(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx}\right\} \tag{52}$$

where $\Phi(x)$ represents the standard normal distribution function. Next, we carry out the marked integral, denoting $\tilde{Q}_\tau(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ and $d = Q_\tau^{-1}(K)$:

$$\begin{aligned}
\int_d^\infty \tilde{Q}_\tau(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx &= \int_d^\infty \left(a_3x^3 + a_2x^2 + a_1x + a_0\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx \\
&= \int_d^\infty \left(a_3x^3 + a_2x^2 + a_1x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx + a_0\int_d^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx \\
&= \int_d^\infty \left(a_3x^3 + a_2x^2 + a_1x\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx + a_0\left(1 - \Phi\left(d\right)\right) \\
&= a_2(1 - \Phi(d)) + \phi(d)\left(a_3d^2 + a_2d + a_1 + a_3\right) + a_0\left(1 - \Phi\left(d\right)\right) \\
&= \phi(d)\left(a_3\left(d^2 - 2\right) + a_2d + \sqrt{1 - 6a_3^2 - 2a_2^2}\right) \qquad (53)
\end{aligned}$$

where $\phi(x)$ denotes the standard normal density function and we have to bear in mind that for a standardized third-order CF variable the following relations hold: $a_0 = -a_2$ and $a_1 = \sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3$. Finally, we obtain the required result substituting equation (53) in equation (52):

$$\begin{aligned}
C &= \frac{1}{1 + r\tau}\left\{\begin{array}{c}\left(S_t(1 + r\tau) - K\right)\Phi\left(-d\right) + \\ \sigma\sqrt{\tau}S_t\phi(d)\left(a_3\left(d^2 - 2\right) + a_2d + \sqrt{1 - 6a_3^2 - 2a_2^2}\right)\end{array}\right\} \\
Q_\tau(x) &= S_t(1 + r\tau) + S_t\sigma\sqrt{\tau}\left(a_3x^3 + a_2x^2 + \left(\sqrt{1 - 6a_3^2 - 2a_2^2} - 3a_3\right)x - a_2\right)
\end{aligned}$$

■

References

Andrews, D. W.: 1991, Heterokedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817–858.

Backus, D., Foresi, S., Li, K. and Wu, L.: 1997, Accounting for biases in Black-Scholes. Working Paper, Stern School of Business, 40 pages.

Bakshi, G., Cao, C. and Chen, Z.: 1997, Empirical performance of alternative option pricing models, *Journal of Finance* **52**(5), 2003–2049.

Bekaert, G. and Harvey, C. R.: 1997, Emerging equity market volatility, *Journal of Financial Economics* **43**, 29–77.

Bera, A. and Jarque, C.: 1982, Model specification tests: A simultaneous approach, *Journal of Econometrics* **20**, 59–82.

Black, F.: 1976, The pricing of commodity contracts, *Journal of Financial Economics* **3**, 167–79.

Black, F. and Scholes, M.: 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.

Bouchaud, J. and Potters, M.: 2000, *Theory of Financial Risk*, Cambridge University Press, Cambridge.

Brown, C. A. and Robinson, D. M.: 2002, Skewness and kurtosis implied by option prices: A correction, *Journal of financial research* **25**(2), 279–282.

Capelle-Blancard, G., Jurczenko, E. and Maillet, B.: 2001, The approximate option pricing model: Performances and dynamic properties, *Journal of Multinational Financial Management* **11**, 427–444.

Cornish, E. and Fisher, R.: 1937, Moments and cumulants in the specification of distributions, *Review of International Statistics* **5**, 307–337.

Corrado, C. and Su, T.: 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 Su, T.: 1996b, S&P 500 index option tests of Jarrow and Rudd's approximate option valuation formula, *Journal of Futures Markets* **6**, 611–629.

Corrado, C. and Su, T.: 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 Su, T.: 1997b, Implied volatility skews and stock return skewness and kurtosis implied by stock option prices, *The European Journal of Finance* **3**, 73–85.

Dumas, B., Fleming, J. and Whaley, R. E.: 1998, Implied volatility functions: Empirical tests, *Journal of Finance* **53**(6), 2059–2106.

Engle, R. and Ng, V. K.: 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* **48**, 1749–1778.

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.

Hull, J. C.: 2004, *Options, Futures, and Other Derivatives*, Prentice Hall, Toronto, Canada.

Hull, J. and White, A.: 1998, Value at risk when daily changes in market variables are not normally distributed, *Journal of derivatives* **5**(3), 9–19.

Jarrow, R. and Rudd, A.: 1982, Approximate option valuation for arbitrary stochastic processes, *Journal of Financial Economics* **10**, 347–369.

Johnson, N.: 1949, Systems of frecuency curves generated by methods of translation, *Biometrika* **36**(1/2), 149–176.

Johnson, S. and Kotz, S.: 1972, *Distributions in Statistics: Continuous Univariate Distributions. Volume One.*, John Wiley and Sons, New York.

Jondeau, E. and Rockinger, M.: 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 Rockinger, M.: 2001, Gram-Charlier densities, *Journal of Economic Dynamics and Control* **25**, 1457–1483.

Jurczenko, E., Maillet, B. and Negrea, B.: 2002, Multi-moment option pricing models: A general comparison (part 1). Working paper, Univesity of Paris I Panthteon-Sorbonne.

Kendall, M., Stuart, A. and Ord, J.: 1994, *Kendall's advanced theory of statistics*, Vol. 1, Oxford University Press, New York, N.Y.

León, A., Mencía, J. and Sentana, E.: 2009, Parametric properties of semi-nonparametric distributions, with an application to option valuation, *Journal of Business and Economic Statistics* **27**(2), 176–192.

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.

Shaw, W. T. and Buckley, I. R. C.: 2007, The alchemy of probability distributions: Beyond Gram-Charlier & Cornish-Fisher expansions, and Skew-Normal or Kurtotic-Normal distributions., *Working Paper*. Available at: `http://library.wolfram.com/ infocenter/Articles/ 6670/alchemy.pdf` .

Table 1: CFD parameters tabulated by skewness and kurtosis

$\kappa\backslash\xi$	0	0.25	0.5	1	1.5	2	2.5	3	3.5	4
3.5	0 0.0177	0.0381 0.0152	0.0802 0.0066	—	—	—	—	—	—	—
4	0 0.0314	0.0351 0.0294	0.0729 0.0229	—	—	—	—	—	—	—
6	0 0.0679	0.0288 0.0668	0.0585 0.0635	0.1274 0.0476	—	—	—	—	—	—
8	0 0.0921	0.0255 0.0913	0.0515 0.0891	0.1083 0.0790	0.1831 0.0561	—	—	—	—	—
10	0 0.1107	0.0234 0.1101	0.0471 0.1084	0.0976 0.1009	0.1574 0.0857	0.2517 0.0514	—	—	—	—
12	0 0.1261	0.0219 0.1256	0.0440 0.1242	0.0903 0.1182	0.1426 0.1066	0.2112 0.0849	—	—	—	—
14	0 0.1393	0.0207 0.1389	0.0415 0.1377	0.0848 0.1327	0.1325 0.1233	0.1901 0.1070	0.2814 0.0733	—	—	—
16	0 0.1511	0.0197 0.1507	0.0396 0.1497	0.0806 0.1453	0.1248 0.1374	0.1761 0.1242	0.2451 0.1009	—	—	—
18	0 0.1616	0.0189 0.1613	0.0380 0.1604	0.0771 0.1565	0.1188 0.1496	0.1657 0.1385	0.2242 0.1203	0.3267 0.0814	—	—
20	0 0.1713	0.0183 0.1710	0.0366 0.1702	0.0742 0.1667	0.1139 0.1606	0.1577 0.1509	0.2097 0.1359	0.2842 0.1091	—	—
22	0 0.1802	0.0177 0.1800	0.0354 0.1792	0.0717 0.1761	0.1097 0.1705	0.1511 0.1619	0.1987 0.1490	0.2608 0.1281	—	—
24	0 0.1886	0.0172 0.1883	0.0344 0.1876	0.0695 0.1847	0.1061 0.1797	0.1455 0.1719	0.1899 0.1606	0.2446 0.1431	0.3336 0.1095	—
26	0 0.1964	0.0167 0.1961	0.0335 0.1955	0.0676 0.1928	0.1030 0.1882	0.1408 0.1811	0.1827 0.1709	0.2324 0.1559	0.3024 0.1305	—
28	0 0.2038	0.0163 0.2036	0.0327 0.2029	0.0659 0.2004	0.1002 0.1961	0.1366 0.1896	0.1765 0.1804	0.2226 0.1670	0.2827 0.1463	—
30	0 0.2108	0.0159 0.2106	0.0319 0.2100	0.0643 0.2077	0.0978 0.2036	0.1330 0.1976	0.1712 0.1891	0.2145 0.1771	0.2682 0.1593	0.3576 0.1253
35	0 0.2269	0.0152 0.2268	0.0304 0.2263	0.0611 0.2242	0.0926 0.2207	0.1255 0.2156	0.1605 0.2085	0.1989 0.1988	0.2434 0.1854	0.3010 0.1651
40	0 0.2416	0.0145 0.2414	0.0291 0.2410	0.0585 0.2392	0.0885 0.2361	0.1197 0.2316	0.1523 0.2254	0.1875 0.2172	0.2268 0.2063	0.2736 0.1911
45	0 0.2551	0.0140 0.2550	0.0280 0.2546	0.0563 0.2529	0.0852 0.2501	0.1148 0.2461	0.1458 0.2406	0.1787 0.2335	0.2146 0.2241	0.2555 0.2117

This table shows the pairs $\binom{a_2}{a_3}$ of the standardized third-order CFD that yield the different values of skewness, ξ, and kurtosis, κ, in the horizontal and vertical axes respectively. For negative values of ξ, find the parameters as if ξ were positive, and then change the sign of a_2.

Table 2: CFD and CFE quantiles comparison

		$\xi = 0$						$\xi = -0.5$					
κ		3%	2%	1%	0.5%	$\hat{\xi}$	$\hat{\kappa}$	3%	2%	1%	0.5%	$\hat{\xi}$	$\hat{\kappa}$
5	CFE	-1.96	-2.25	-2.78	-3.37	-0.02	7.11	-2.16	-2.49	-3.06	-3.69	-0.71	6.54
	CFD	-1.91	-2.16	-2.59	-3.03	-0.01	5.02	-2.08	-2.36	-2.84	-3.35	-0.50	4.96
10	CFE	-2.20	-2.81	-3.95	-5.33	-0.02	34.76	-2.35	-2.99	-4.23	-5.65	-1.00	34.68
	CFD	-1.93	-2.27	-2.86	-3.52	0.00	9.55	-2.03	-2.39	-3.05	-3.76	-0.47	9.75
20	CFE	-2.57	-3.78	-6.20	-9.26	0.14	84.62	-2.79	-4.05	-6.53	-9.58	-1.01	106.64
	CFD	-1.87	-2.28	-3.04	-3.92	0.02	19.67	-1.97	-2.41	-3.22	-4.14	-0.55	22.13
	Gauss	*-1.88*	*-2.05*	*-2.33*	*-2.58*	*0*	*3*						

This table shows four different quantiles on the left tail using the Cornish-Fisher expansion (CFE) and the Cornish-Fisher distribution (CFD) for different degrees of kurtosis, κ, and skewness, ξ. The quantiles have been obtained numerically from a sample of 500,000 simulations. $\hat{\xi}$ and $\hat{\kappa}$ columns stand for the sample skewness and kurtosis computed with the 500,000 observations. The last row exhibits the corresponding quantiles for the standardized Gaussian distribution.

Table 3: Descriptive univariate statistics on returns.

	Mean ($\times 10^2$)	σ^2	ξ	κ	JB	Wald	KS
S&P	0.002 (0.097)	5.615 (0.644)	-0.085 (0.185)	4.509 (0.406)	49	14.67	0.343
NKI	-0.099 (0.128)	9.344 (0.750)	-0.147 (0.146)	3.928 (0.308)	21	9.10	0.184
STX	0.150 (0.128)	8.672 (1.404)	-0.414 (0.244)	5.997 (0.762)	209	28.74	0.447
EM	0.027 (0.149)	8.114 (0.854)	-0.718 (0.229)	5.098 (0.774)	140	9.98	0.424
EME	0.110 (0.216)	18.832 (3.001)	-0.443 (0.200)	5.604 (0.679)	164	29.26	0.292

Mean, σ^2, ξ and κ denote the mean, variance, skewness, and kurtosis of returns, respectively. Estimates are not annualized and standard errors, in parenthesis, are computed using the GMM-based procedure proposed by Bekaert and Harvey (1997). JB, Wald and KS denote 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. The corresponding p-values are lower than 0.01 for all statistics and all series, this is why we do not report them.

Table 4: Cornish-Fisher Density estimates.

	$\hat{a}_3$	$\hat{a}_2$	$\hat{a}_1$	$\hat{a}_0$	KS-Test	JB-Test
Panel A: Quantile-Quantile (QQ) estimates						
S&P	0.1218 (0.030)	-0.0406 (0.061)	1.9906 (0.097)	0.2205 (0.099)	0.035 [0.55]	0.145 [0.93]
NKI	0.1103 (0.029)	0.0547 (0.065)	2.7170 (0.133)	-0.1541 (0.129)	0.025 [0.90]	0.027 [0.98]
STX	0.2288 (0.048)	-0.1930 (0.098)	2.2029 (0.138)	0.3428 (0.118)	0.034 [0.39]	0.407 [0.81]
EM	0.1493 (0.036)	-0.2617 (0.077)	2.3624 (0.129)	0.2881 (0.121)	0.026 [0.87]	0.092 [0.95]
EME	0.3391 (0.068)	-0.2828 (0.126)	3.2338 (0.231)	0.3928 (0.172)	0.041 [0.35]	0.373 [0.83]
Panel B: Moments Method (MM) estimates						
S&P	0.1008 (0.025)	-0.0264 (0.067)	2.0541 (0.098)	0.2064 (0.105)	0.038 [0.43]	0.379 [0.82]
NKI	0.0883 (0.022)	0.0632 (0.066)	2.7832 (0.126)	-0.1628 (0.132)	0.023 [0.92]	0.235 [0.88]
STX	0.1914 (0.045)	-0.1419 (0.116)	2.3264 (0.149)	0.2921 (0.130)	0.044 [0.24]	1.520 [0.46]
EM	0.1188 (0.031)	-0.2703 (0.083)	2.4512 (0.137)	0.2974 (0.122)	0.030 [0.72]	0.446 [0.80]
EME	0.2533 (0.068)	-0.2316 (0.162)	3.5225 (0.242)	0.3423 (0.196)	0.054 [0.08]	4.043 [0.13]
Panel C: Maximum Likelihood (ML) estimates						
S&P	0.1217 (0.032)	-0.0926 (0.056)	1.9908 (0.094)	0.2654 (0.095)	0.026 [0.86]	0.316 [0.85]
NKI	0.1127 (0.038)	0.0239 (0.071)	2.7063 (0.127)	-0.1271 (0.132)	0.022 [0.96]	0.112 [0.94]
STX	0.2200 (0.042)	-0.3220 (0.084)	2.2358 (0.106)	0.4475 (0.108)	0.028 [0.77]	1.153 [0.56]
EM	0.1376 (0.038)	-0.2239 (0.063)	2.3839 (0.111)	0.2560 (0.115)	0.020 [0.97]	0.114 [0.94]
EME	0.4829 (0.091)	-0.3516 (0.125)	2.8760 (0.180)	0.4610 (0.145)	0.021 [0.97]	1.143 [0.56]

This table shows the Quantile-Quantile, Moments-Method and Maximum Likelihood estimates of the parameters of the CFD using stock indexes series. The corresponding standard errors are displayed in parenthesis. KS-Test denotes the Kolmogorov-Smirnov statistic for the CFD null hypothesis and JB-Test stands for the Jarque-Bera statistic that tests the normality of the fictitious normal variables defined by equation (5). The corresponding p-values are showed in brackets.

Table 5: Density Goodness of fit

	Log-Likelihoods				Akaike Criteria				Bayesian Criteria			
	Gaus.	John.	CFE	CFD	Gaus.	John.	CFE	CFD	Gaus.	John.	CFE	CFD
S&P	-1183	-1171	-1171	-1170	2375	2350	2351	2349	2392	2367	2368	2366
NKI	-1315	-1309	-1309	-1309	2639	2626	2626	2625	2656	2643	2643	2642
STX	-1296	-1263	-1268	-1260	2600	2533	2545	2529	2617	2550	2562	2546
EM	-1279	-1256	-1258	-1256	2566	2520	2525	2521	2583	2537	2542	2538
EME	-1497	-1459	-1454	-1454	3003	2926	2917	2916	3020	2943	2934	2933

This table presents three different model selection criteria, the Log-Likelihood, the Akaike and the Bayesian criteria, for the estimation of five stock indexes. We consider four alternative models: the Gaussian (Gaus.), the Johnson distribution (John.), the Cornish-Fisher Expansion (CFE) and the Cornish-Fisher Distribution (CFD).

Table 6: Model summary

	Innovation		
	Gaussian	CFD ($\gamma_2 = \eta_2 = 0$)	CFD-asymm.
GARCH ($\beta_3 = 0$)	Model 1	Model 3	Model 5
NGARCH	Model 2	Model 4	Model 6

This table summarizes the six models considered in this article according to equations (18) to (24). Model 6 is the most general model and nests the other ones.

Table 7: S&P 500 and EURO STOXX descriptives

	mean	median	min.	max.	std.	skewness	kurtosis	JB-test
S&P 500	-0.0001	-0.0001	-0.0947	0.1096	0.2198	-0.1178	11.3164	7521.8
Euro Stoxx	-0.0002	0.0005	-0.0825	0.0996	0.2338	-0.0662	7.5801	2223.7

This table shows the main statistics for the daily series of the log-returns of the S&P 500 and the EuroStoxx indexes for the period 03/17/2000 to 03/20/2010.

Table 8: S&P 500 GARCH models estimates (pre-filtered).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$C(\times 10^3)$	5.6777	-0.1918°	0.4008^{*}	0.1489°	0.3857^{*}	0.0999°
	(0.7173)	(1.0802)	(0.1730)	(0.1658)	(0.1725)	(0.1287)
$\beta_0(\times 10^6)$	0.1124°	1.5287	0.7186	1.0881	0.6599^{*}	1.0649
	(0.5701)	(0.5549)	(0.2828)	(0.2338)	(0.3161)	(0.2568)
β_1	0.9051	0.8620	0.9186	0.8697	0.9195	0.8673^{*}
	(0.0105)	(0.0142)	(0.0088)	(0.0143)	(0.0106)	(0.0150)
β_2	0.0745	0.0464^{*}	0.0814	0.0471	0.0805	0.0491^{*}
	(0.0089)	(0.0081)	(0.0097)	(0.0078)	(0.0111)	(0.0038)
β_3	—	−1.3471	—	−1.2836	—	−1.2769
		(0.2684)		(0.1896)		(0.0637)
γ_0	—	—	−0.3286	−0.4188	-0.2124^{**}	−0.2938
			(0.0695)	(0.0829)	(0.1103)	(0.0514)
γ_1	—	—	-8.2893^{*}	-9.9554^{**}	-25.3031^{*}	−26.1919
			(3.8505)	(5.1635)	(11.2827)	(6.0073)
γ_2	—	—	—	—	28.2362°	31.9380
					(18.6944)	(11.3521)
η_0	—	—	−1.6845	−2.2844	−1.7691	−2.4553
			(0.2051)	(0.2499)	(0.2126)	(0.1657)
η_1	—	—	55.7481	61.5990	52.3229	71.0286
			(6.0639)	(5.9193)	(6.4698)	(0.9468)
η_2	—	—	—	—	-44.3677^{**}	−81.7170
					(23.4164)	(15.0613)
$\mathcal{L}(\hat{\theta})$	8211.3	8236.5	8230.3	8284.1	8230.8	8287.6
			Goodness of fit			
AIC	−16415	−16463	−16445	−16550	−16442	−16553
SIC	−16391	−16434	−16398	−16497	−16383	−16489
BIC	−6.2849	−6.3012	−6.2875	−6.3257	−6.2818	−6.3223

This table shows the estimation of six GARCH models for the daily log-returns of the EuroStoxx 50 series from 03/17/2000 to 03/20/2010. The returns have been previously filtered with an ARMA(1,1) structure. The last three rows stand for the Akaike, Schwarz and Bayesian criteria to measure the goodness of fit. Estimates without markers are significant at 1%. (*) and (**) denote parameters significant at 5% and 10% respectively, while (°) denotes not significant parameters.

Table 9: EuroStoxx 50 GARCH models estimates (filtered).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$C(\times 10^3)$	0.7701	0.3382°	0.7045	0.3128°	0.6648	0.3170°
	(0.2058)	(1.4612)	(0.1933)	(0.1881)	(0.1936)	(0.1886)
$\beta_0(\times 10^6)$	1.5562	2.1484	1.0715	1.8899	1.0655	1.8723
	(0.3179)	(0.8875)	(0.3791)	(0.3111)	(0.3785)	(0.3126)
β_1	0.8990	0.8430	0.9010	0.8505	0.9018	0.8516
	(0.0087)	(0.0156)	(0.0101)	(0.0139)	(0.0103)	(0.0141)
β_2	0.0944	0.0544	0.0994	0.0576	0.0982	0.0573
	(0.0081)	(0.0089)	(0.0112)	(0.0093)	(0.0115)	(0.0095)
β_3	—	−1.2775	—	−1.1926	—	−1.1893
		(0.2613)		(0.1995)		(0.2029)
γ_0	—	—	−0.5240	−0.7266	−0.7196	−0.7951
			(0.1141)	(0.1834)	(0.1342)	(0.1708)
γ_1	—	—	22.9211	26.6281	30.7324	29.1474
			(6.6412)	(8.0110)	(6.8044)	(7.8742)
γ_2	—	—	—	—	−58.6326	−56.2688
					(15.6929)	(20.8892)
η_0	—	—	−2.4733	−3.3316	−2.7191	−3.3676
			(0.2804)	(0.4508)	(0.3122)	(0.4785)
η_1	—	—	84.0231	93.2125	95.2238	93.5462
			(13.4471)	(20.2862)	(12.4667)	(25.5264)
η_2	—	—	—	—	−103.8716	−110.9625*
					(29.0813)	(52.1395)
$\mathcal{L}(\hat{\theta})$	7686.6	7760.8	7734.1	7791.4	7737.0	7792.1
			Goodness of fit			
AIC	−15365	−15512	−15452	−15565	−15454	15562
SIC	−15342	−15482	−15405	−15512	−15396	15498
BIC	−6.0354	−6.0906	−6.0603	−6.1024	−6.0565	−6.0967

This table shows the estimation of six GARCH models for the daily log-returns of the EuroStoxx 50 series from 03/17/2000 to 03/20/2010. The returns have been previously filtered with an ARMA(1,1) structure. The last three rows stand for the Akaike, Schwarz and Bayesian criteria to measure the goodness of fit. Estimates without markers are significant at 1%. (*) and (**) denote parameters significant at 5% and 10% respectively, while (°) denotes not significant parameters.

Table 10: In Sample Option Pricing Results

	All maturities			Short maturities			Long maturities		
	All	Calls	Puts	All	Calls	Puts	All	Calls	Puts
				Panel A: RMSE					
B76	1.269	1.228	1.216	0.697	0.694	0.580	1.515	1.442	1.488
P-BS	0.664	0.606	0.567	0.336	0.280	0.225	0.638	0.545	0.509
CS	0.683	0.661	0.627	0.385	0.349	0.258	0.715	0.605	0.654
CFD	**0.587**	**0.514**	**0.515**	**0.327**	**0.265**	**0.206**	**0.563**	**0.455**	**0.452**
				Panel B: MAE					
B76	1.042	1.025	0.993	0.587	0.602	0.489	1.316	1.263	1.303
P-BS	0.525	0.483	0.448	0.267	0.216	0.171	0.508	0.433	0.404
CS	0.546	0.495	0.507	0.308	0.280	0.206	0.581	0.494	0.541
CFD	**0.456**	**0.402**	**0.404**	**0.260**	**0.206**	**0.158**	**0.444**	**0.361**	**0.361**

This table shows the in–sample average of both the root mean squared error (panel A) and the mean absolute error (panel B) computed over the 292 daily estimations for the four option models considered: Black-76 (B76), Practitioner Black-Scholes (P-BS), Corrado and Su (CS) and Cornish-Fisher density (CFD). "Short maturities" columns contain options with a time to maturity lower than 61 days and "Long maturities" columns contain options with a time to maturity larger than 60 days. For each category, the best performance is in bold.

Table 11: Out of Sample Option Pricing Results

	All maturities			Short maturities			Long maturities		
	All	Calls	Puts	All	Calls	Puts	All	Calls	Puts
				Panel A: RMSE					
B76	1.359	1.337	1.368	0.802	0.853	0.745	1.626	1.579	1.683
P-BS	0.911	0.967	0.937	0.588	0.644	0.542	0.958	0.967	0.961
CS	0.849	0.826	0.864	0.567	0.611	0.516	0.935	0.883	0.961
CFD	**0.814**	**0.806**	**0.817**	**0.551**	**0.590**	**0.507**	**0.877**	**0.840**	**0.875**
				Panel B: MAE					
B76	1.100	1.098	1.094	0.661	0.710	0.618	1.396	1.361	1.457
P-BS	0.701	0.741	0.708	0.468	0.521	0.442	0.769	0.786	0.779
CS	0.676	0.673	0.687	0.464	0.517	0.429	0.763	0.739	0.798
CFD	**0.642**	**0.651**	**0.649**	**0.450**	**0.496**	**0.421**	**0.718**	**0.708**	**0.731**

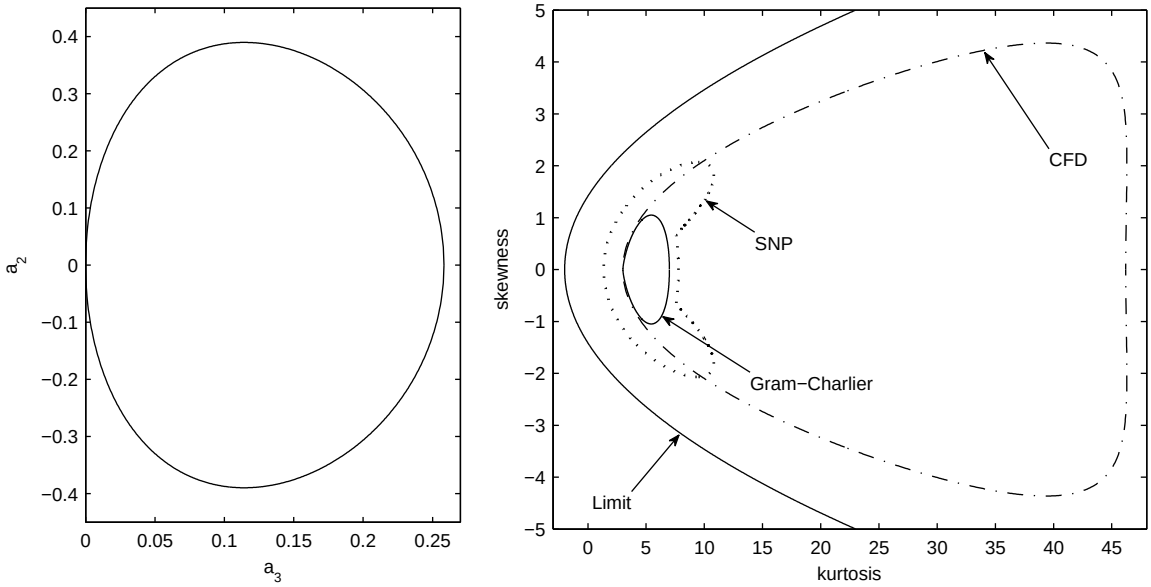
This table shows the one week out–of–sample average of both the root mean squared error (panel A) and the mean absolute error (panel B) computed over the 292 daily estimations for the four option models considered: Black-76 (B76), Practitioner Black-Scholes (P-BS), Corrado and Su (CS) and Cornish-Fisher density (CFD). "Short maturities" columns contain options with a time to maturity lower than 61 days and "Long maturities" columns contain options with a time to maturity larger than 60 days. For each category, the best performance is in bold.

Table 12: Option hedging errors

	All maturities			Short maturities			Long maturities		
	All	Calls	Puts	All	Calls	Puts	All	Calls	Puts
					Panel A: RMSHE				
B76	0.605	0.571	0.566	0.539	0.499	0.445	0.619	0.586	0.565
P-BS	0.526	0.459	0.491	0.407	0.320	0.313	0.511	0.423	0.457
CS	0.536	0.456	0.480	0.435	0.358	0.310	0.534	0.437	0.452
CFD	**0.499**	**0.422**	**0.444**	**0.394**	**0.304**	**0.271**	**0.488**	**0.388**	**0.408**
					Panel B: MAHE				
B76	0.449	0.426	0.408	0.398	0.378	0.318	0.465	0.435	0.410
P-BS	0.394	0.342	0.349	0.301	0.232	0.208	0.387	0.312	0.325
CS	0.400	0.338	0.342	0.324	0.264	0.215	0.400	0.322	0.324
CFD	**0.377**	**0.315**	**0.320**	**0.296**	**0.222**	**0.189**	**0.371**	**0.288**	**0.296**

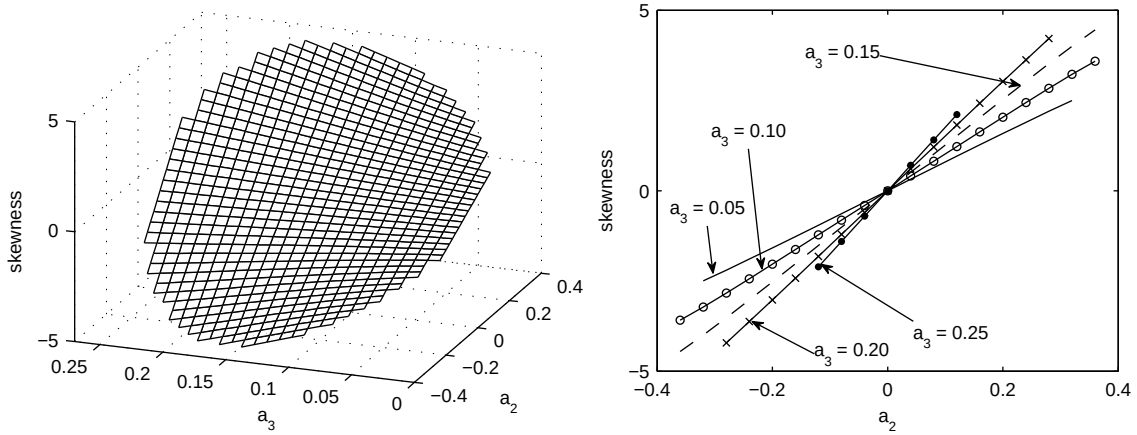
This table shows the average of both the root mean squared hedging errors (panel A) and the mean absolute hedging error (Panel B) for the four option models considered: Black-76 (B76), Practitioner Black-Scholes (P-BS), Corrado and Su (CS) and Cornish-Fisher density (CFD). "Short maturities" columns contain options with a time to maturity lower than 61 days and "Long maturities" columns contain options with a time to maturity larger than 60 days. For each category, the best performance is in bold.

Figure 1: CFD parameters and skewness-kurtosis envelopes



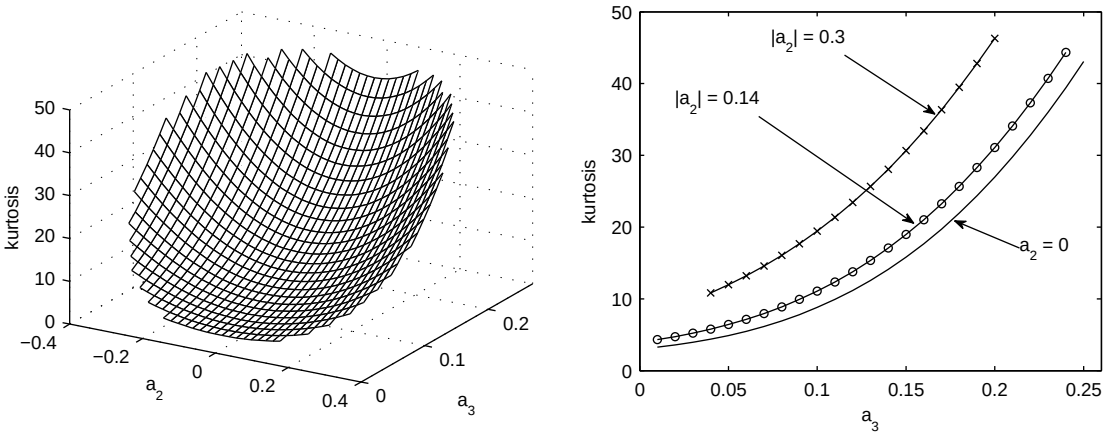
The left panel exhibits the range of validity for the coefficients a_2 and a_3 of a standardized third-order CFD according to equations (14) and (15). The right panel shows the regions of skewness and kurtosis covered by the third-order CFD, the Gram-Charlier distribution used by Jondeau and Rockinger (2001), the semi–nonparametric distribution of León et al. (2009) and the limit for all distributions.

Figure 2: Skewness surface



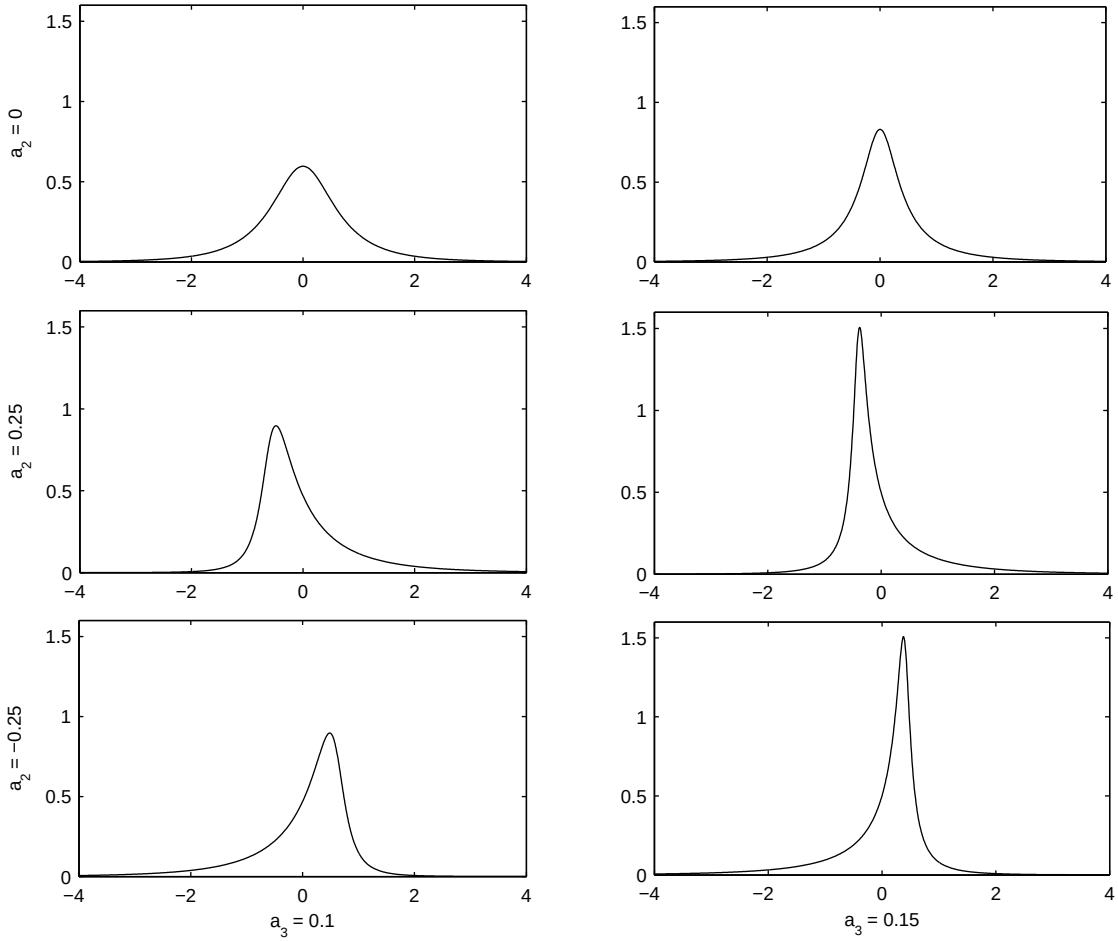
This figure shows the surface of the skewness of a standardized third-order CFD as a function of the parameters of the distribution (left graphic). The right-side graphic displays the skewness as a function of a_2 for different values of a_3 (level curves).

Figure 3: Kurtosis surface



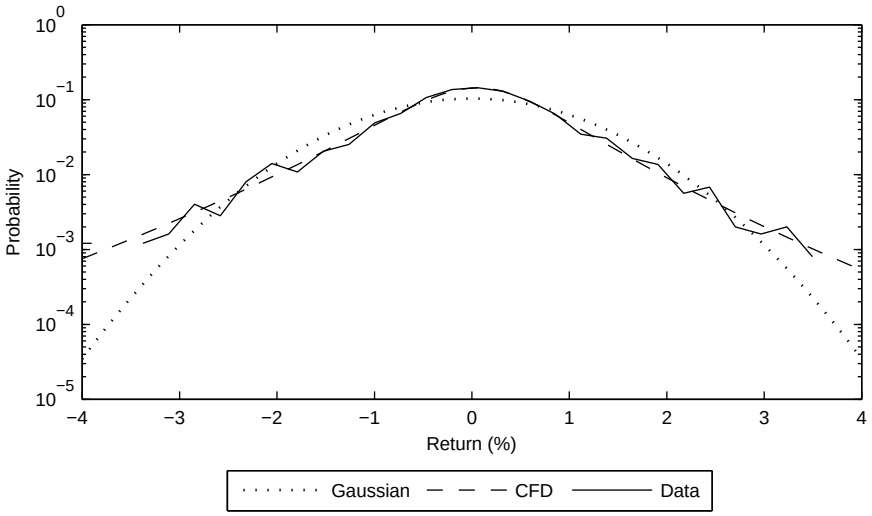
This figure shows the surface of the kurtosis of a standardized third-order CFD as a function of the parameters of the distribution (left graphic). The right-side graphic displays the kurtosis as a function of a_3 for different values of a_2 (level curves).

Figure 4: Third-order Cornish-Fisher densities



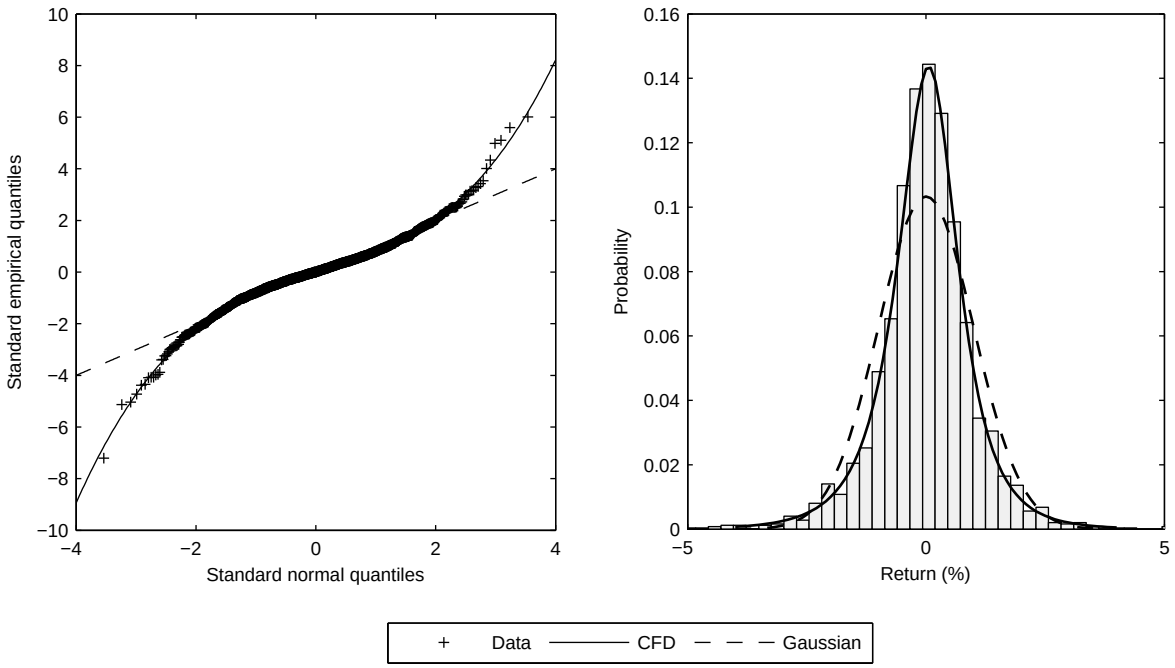
This Figure shows different shapes of the standardized third-order CFD. We consider symmetric (top), positive skewed (medium) and negative skewed (bottom) densities. The right-side graphics are more leptokurtic (higher a_3).

Figure 5: Logarithmic scale densities



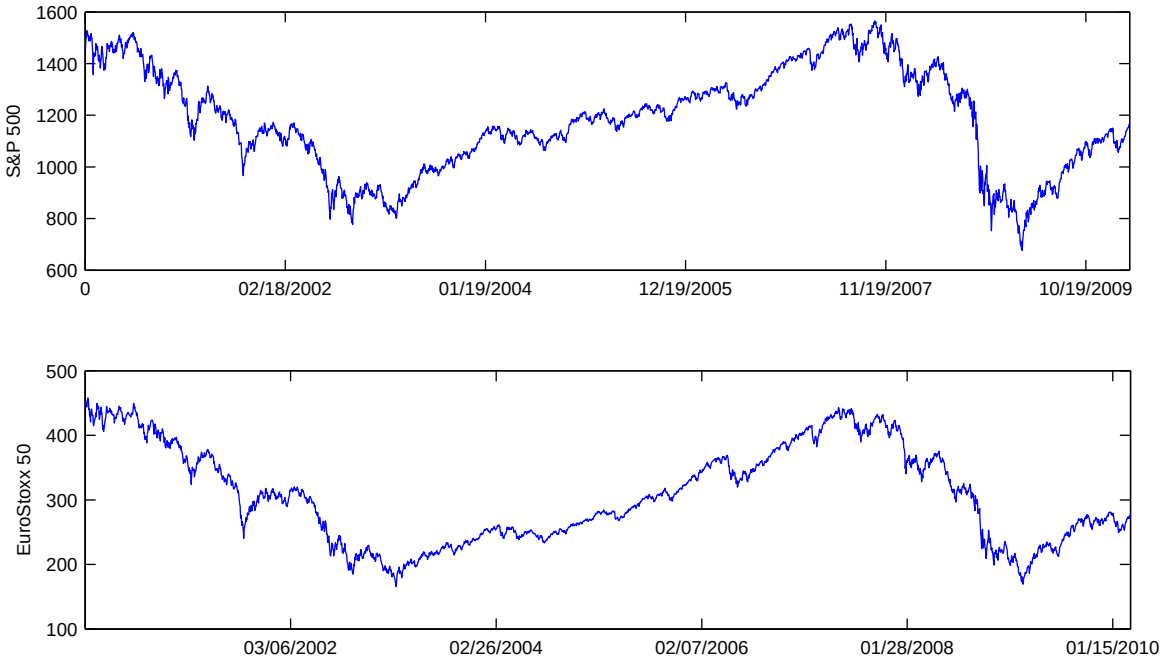
This Figure shows, in logarithmic scale, the fitted third-order Cornish-Fisher Density (dashed line), the fitted Gaussian density (dotted line) and the empirical density of the series (solid line) of daily returns of the YEN/USD exchange rate for the period 01/04/1988–08/15/1997.

Figure 6: QQ-plot and CFD density



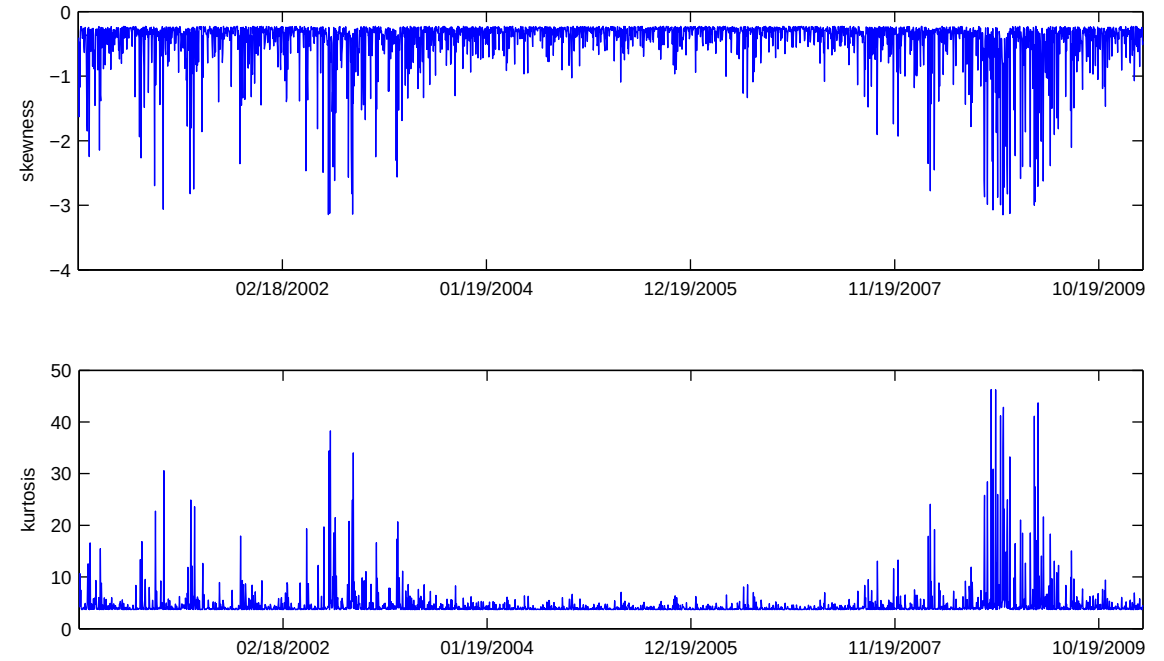
This Figure shows the QQ-plot (left panel) and the histogram (right panel) of the series of daily returns of the YEN/USD exchange rate for the period 01/04/1988–08/15/1997. The solid line represents the CFD approximation estimated by least squares: $R = Q(X) = 0.08379X^3 - 0.02655X^2 + 0.7201X + 0.02628$. The dashed line represents the Gaussian approximation.

Figure 7: S&P 500 and EuroStoxx time series



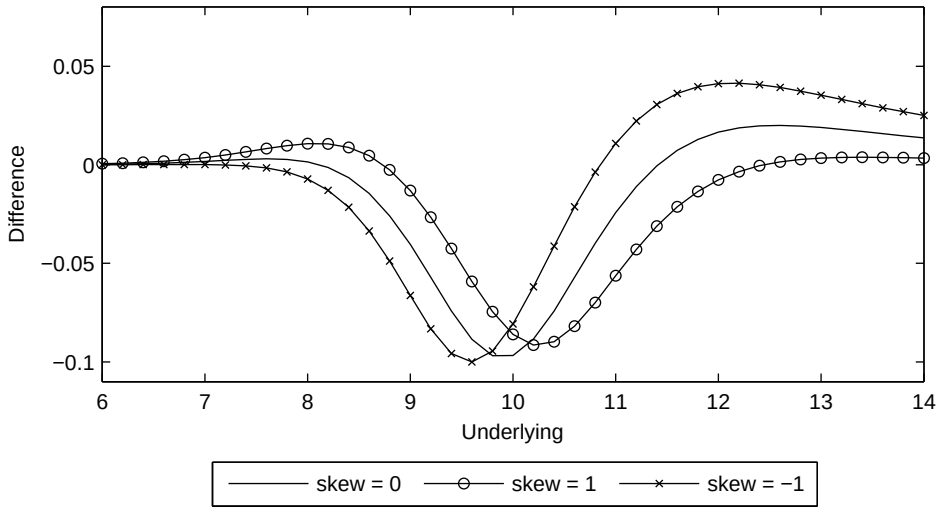
This Figure shows the time series of the S&P 500 and the EuroStoxx 50 from 03/17/2000 to 03/20/2010.

Figure 8: S&P 500 conditional skewness and kurtosis



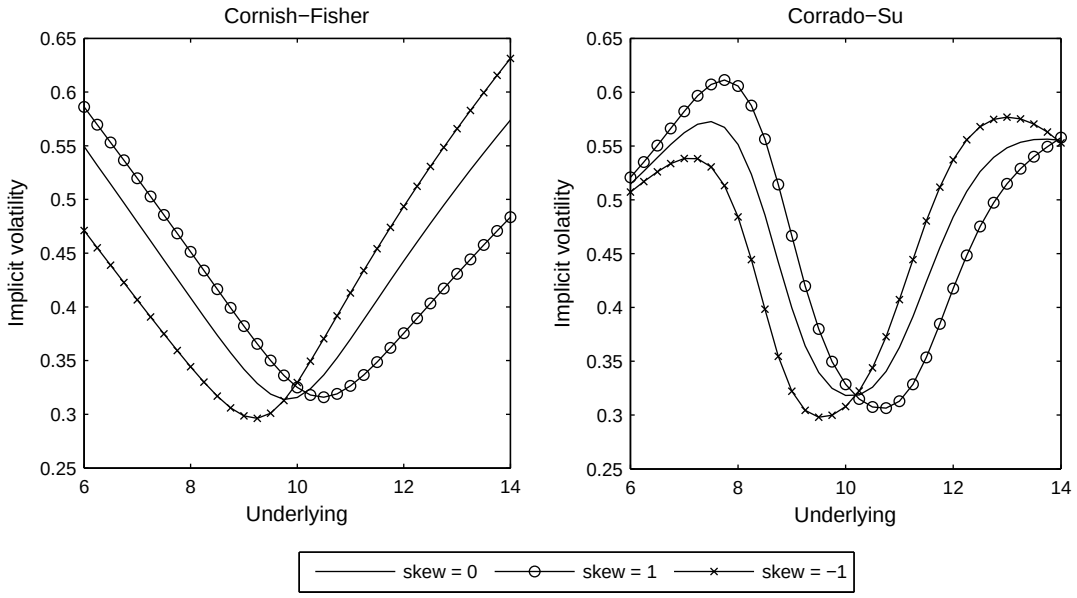
This Figure shows the estimated conditional skewness and kurtosis time series of the S&P 500 daily returns from 03/17/2000 to 03/20/2010.

Figure 9: Differences between CFD and geometric Black-Scholes call prices.



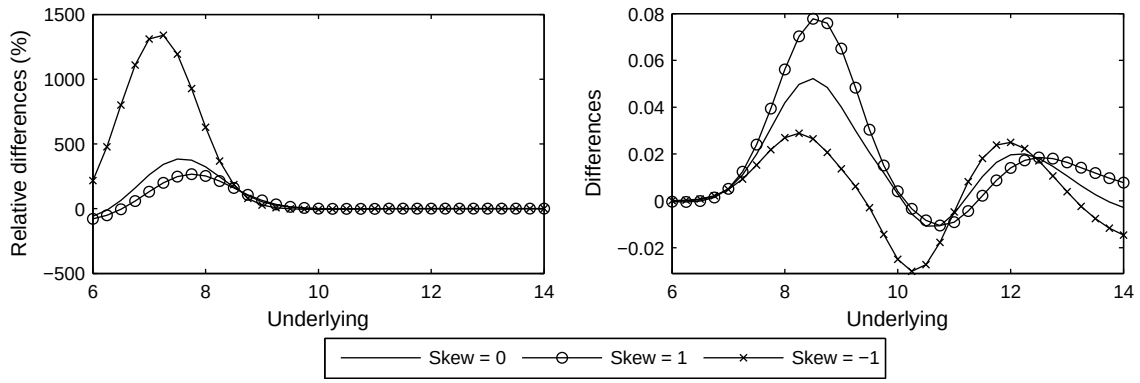
This Figure shows the difference between CFD and geometric Black-Scholes call prices with an strike of $K = 10$, time to maturity of one month, the annual volatility is 40%, and the risk free interest rate equals 5%. We assume parameters (a_2, a_3) equal to $(0, 0.092)$, $(0.11, 0.079)$ and $(-0.11, 0.079)$ which correspond to skewness and kurtosis coefficients of (0,8), (1,8) and (-1,8), respectively.

Figure 10: Black-Scholes implied volatilities.



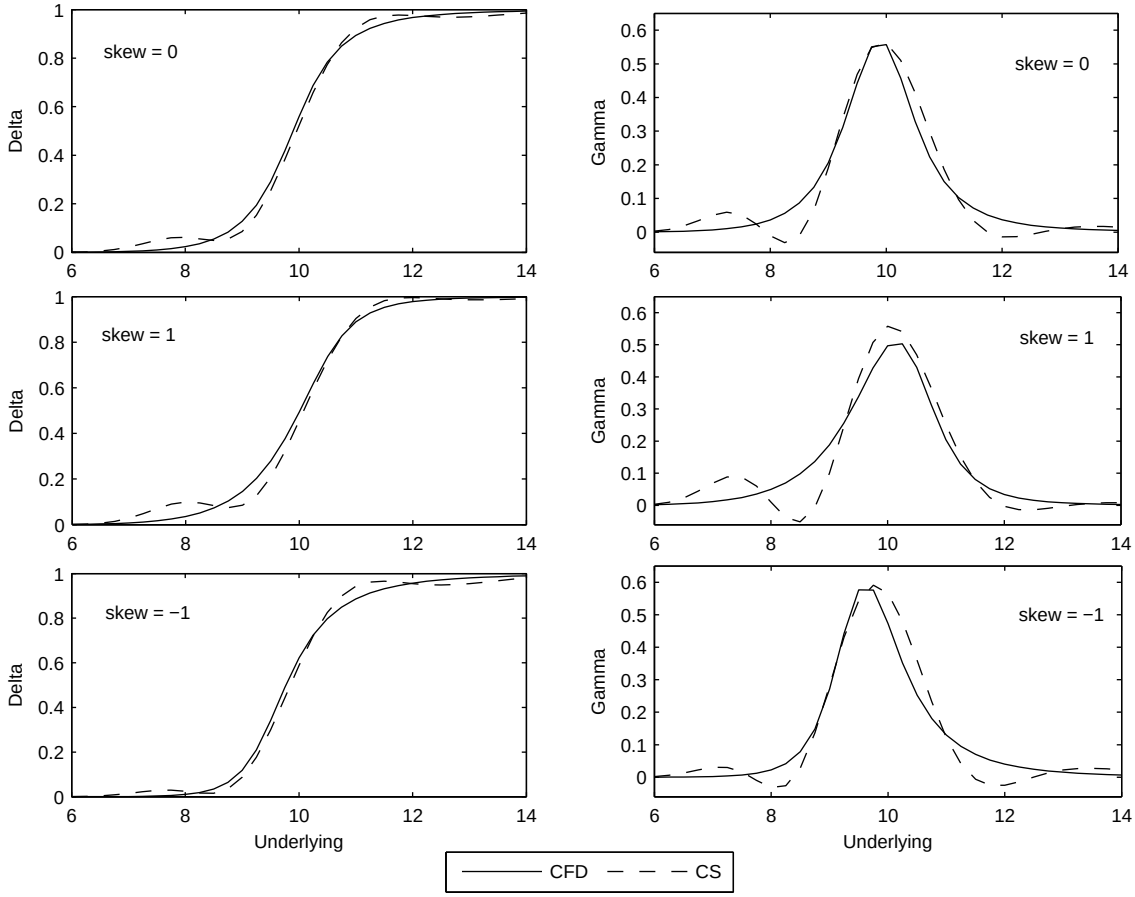
This Figure shows the implied Black-Scholes volatilities for a European-style call option with an strike of $K = 10$, time to maturity of one month, and a risk free interest rate equal to 5% when the "true" option price is driven, either by the CFD model (left panel) and the Corrado and Su model (right panel), with an annual volatility of 40%. We assume parameters (a_2, a_3) equal to $(0, 0.092)$, $(0.11, 0.079)$ and $(-0.11, 0.079)$ which correspond to skewness and kurtosis coefficients of (0,8), (1,8) and (-1,8), respectively.

Figure 11: CFD and Corrado-Su price differences



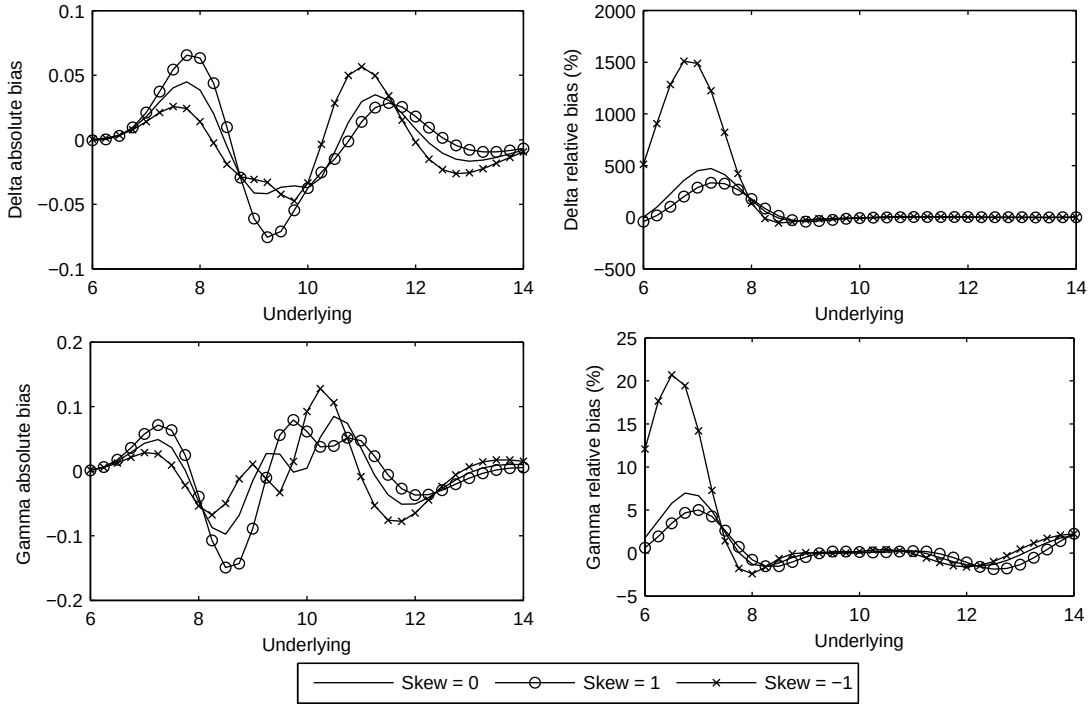
This Figure shows the relative bias (left panel) computed as (CS-CFD)/CFD×100 and the absolute bias (right panel), computed as CS-CFD, for a European-style call option with an strike of $K = 10$, time to maturity of one month, a risk free interest rate equal to 5% and a yearly volatility of 40%. We assume parameters (a_2, a_3) equal to $(0, 0.092)$, $(0.11, 0.079)$ and $(-0.11, 0.079)$ which correspond to skewness and kurtosis coefficients of (0,8), (1,8) and (-1,8), respectively.

Figure 12: CFD and Corrado-Su option *Delta* and *Gamma*



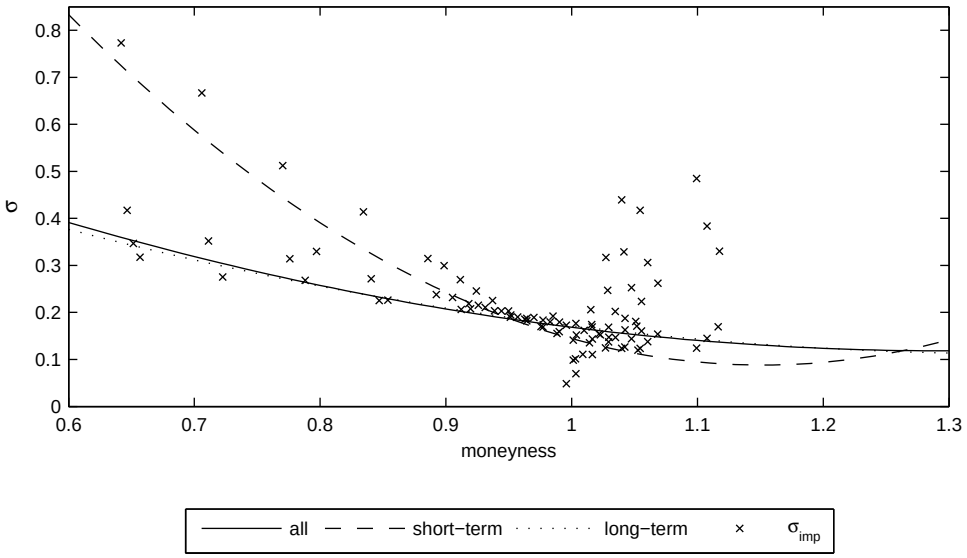
This Figure shows the option *delta* (left panels) and the option *gamma* (right panels) for either the CFD and CS option pricing formulae for a European-style call option with an strike of $K = 10$, time to maturity of one month, a risk free interest rate equal to 5% and a yearly volatility of 40%. We assume parameters (a_2, a_3) equal to $(0, 0.092)$, $(0.11, 0.079)$ and $(-0.11, 0.079)$ which correspond to skewness and kurtosis coefficients of (0,8), (1,8) and (-1,8), respectively.

Figure 13: *Delta* and *Gamma* differences



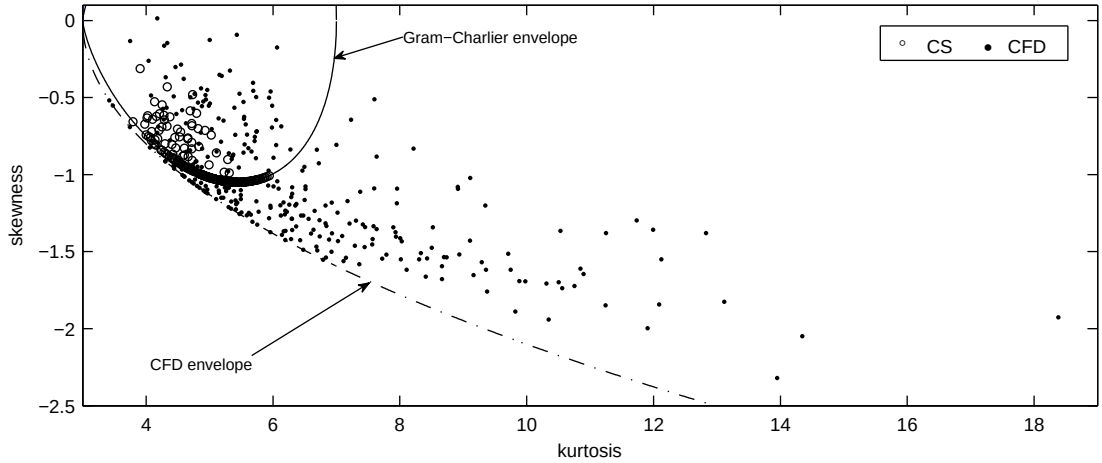
This Figure shows the option delta differences (top) and the gamma differences (bottom) between the CFD and the CS models for a European-style call option with an strike of $K = 10$, time to maturity of one month, a risk free interest rate equal to 5% and a yearly volatility of 40%. We assume parameters (a_2, a_3) equal to $(0, 0.092)$, $(0.11, 0.079)$ and $(-0.11, 0.079)$ which correspond to skewness and kurtosis coefficients of (0,8), (1,8) and (-1,8), respectively. The left panels show the absolute difference measured as CS-CFD, meanwhile the right panels show the relative differences measured as (CS-CFD)/CFD×100.

Figure 14: Option implied volatility functions



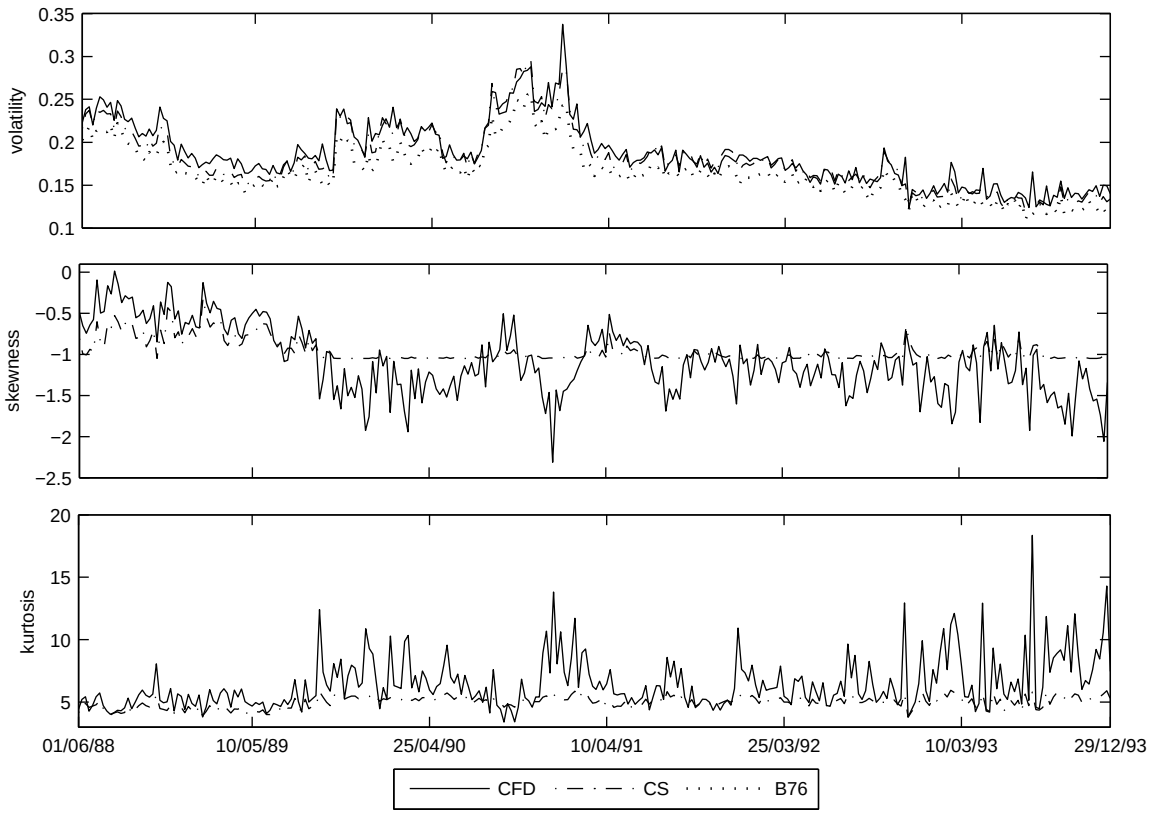
This Figure shows the estimated volatility functions (see equation 31) for S&P 500 options traded on 08/28/1991. We present the estimation for all options (solid line), short-term options (dashed line) and long-term options (dotted line). The crosses are the implied Black-Scholes volatilities.

Figure 15: Option implied skewness and kurtosis and density restrictions



This Figure shows the scatter plot of the pairs *(kurtosis, skewness)* implied in S&P 500 European option prices using two different models: CS and CFD. Moreover, we add the bounds for the allowed kurtosis-skewness values for both the Gram-Charlier and Cornish-Fisher distributions, which underlie the two option pricing models respectively (see Figure 1).

Figure 16: Option implied moments time series



This Figure shows the implied volatility (top), skewness (medium) and kurtosis (bottom) using the prices of S&P 500 European options from June 1988 to December 1993, and for three different option valuation models: CFD (solid line), CS (dashed line) and B76 (dotted line).