

Multifactor Stochastic Variance Models in Risk Management: Maximum Entropy Approach and Lévy Processes

Alexander Levin and Alexander Tchernitser*

Bank of Montreal, Toronto[†]

August 2001

Abstract

This paper investigates a class of multifactor non-normal models for Market Risk Management, and, specifically, for Value-at-Risk (VaR) calculations, with stochastic variance (SV) driven by Lévy processes. Relevant statistical and dynamic properties for the risk factors are discussed. A short review of the Market Risk Management requirements and stochastic models for VaR is presented.

In the case of one asset, a broad class of pure jump Generalized Gamma processes for the SV is derived from the Maximum Entropy principle. The corresponding family of Lévy processes for the risk factors (RF) possesses skewed leptokurtic marginal distributions with a wide range of heavy tails, from exponential and sub-exponential (stretched exponential) to polynomial. The introduced extended Generalized Gamma Variance family is a two shape parameter class of conditionally normal symmetric distributions (there is the third shape parameter in the case of non-zero skewness) with the SV represented as an arbitrary power (positive, zero or negative) of a gamma distribution. It includes normal, Variance Gamma (Generalized Laplace), Student t, and Weibull Variance Mixture distributions as special cases. Ornstein-Uhlenbeck type processes for the SV driven by positive Lévy noise and the corresponding term structure of the RF kurtosis and quantiles are considered for the purpose of modelling non-linear dependence in the asset returns.

A general framework for constructing multidimensional conditionally Gaussian stochastic processes with the correlated multivariate stochastic variances that follow Lévy processes is considered. This methodology allows for different shape and tail behavior of the marginal RF and linear sub-portfolio distributions, exact fit into the RF correlation structure, and proper non-linear scaling of VaR for different holding periods. Presented empirical evidence for different markets confirms a good agreement between the model and historical RF distributions. Effective numerical calibration and Monte Carlo simulation procedures are developed.

1 Review of Market Risk Models

1.1 Market Risk Management and Value-at-Risk

Market Risk Management deals with the risk of potential portfolio losses due to adverse changes in the price of financial instruments caused by stochastic fluctuations of the market variables [58], [17], [57], [32]. There are many types of general market and specific risk factors (RF) with different distributional properties and stochastic behavior in the foreign exchange, interest rate, commodity and equity markets. Market variables include, for example, stock prices, equity indices, spot foreign exchange rates, commodity prices, as well as

*Mail address: Market Risk, Bank of Montreal, First Canadian Place, 3rd Floor, Toronto, Ontario M5X 1A1, Canada.
E-mail address: Alex.Levin@bmo.com, Alexander.Tchernitser@bmo.com

[†]The views expressed in this Working Paper are those of the authors and not necessarily of the Bank of Montreal

complex aggregate structures: interest rate curves, commodity futures price curves, credit spread curves, implied volatility surfaces (e.g., European option implied volatility as a function of strike and maturity) or "cubes" (e.g., swaption implied volatility as a function of underlying swap tenor, swaption maturity and strike). Also, there are such "wild" and "exotic" market variables as, for example, electricity prices and interest rate or foreign exchange rate cross-correlations (the changes of latter variables effect the spread and cross-currency option prices).

Proper modelling of the multivariate future RF distributions is important for financial institutions for the purpose of accurate estimation of the market risk, identification of the risk concentration, developing of trading and hedging strategies, portfolio optimization, consistent measurement of the risk adjusted performance for different units (Risk Adjusted Return On Capital (RAROC) and Capital-at-Risk methodologies), setting up the trading limits, calculating of the regulatory capital [17], back-testing of the market risk models required by regulators [18]. Many financial institutions need to consistently estimate market risk for large portfolios and sub-portfolios (aggregation levels) that comprise hundreds of thousands of instruments dependent on thousands of risk factors in all markets. These portfolios usually include sub-portfolios of options, which magnify and non-linearly transform deviations of the underlyings. Modern Market Risk Management is interested in comprehensive modelling of the multidimensional risk factor stochastic processes and marginal distributions for different time horizons rather than static multivariate distributions for some fixed holding period. This interest comes from the requirements to capture liquidity risk for many instrument types with varying liquidation periods (see [32]), estimate intraday risk for some frequently rebalanced positions, consistently evaluate VaR for one-day and ten-day time horizons prescribed by BIS documents [18], [17] for back-testing and regulatory capital calculations respectively, and actively dynamically manage risk. This problem points out on the importance of adequate modelling of a non-linear dependence in the underlying returns observed in the market to capture a proper VaR term profile.

Along with the RF volatilities (standard deviations of daily changes) and correlations combined with the portfolio sensitivities (Greeks, [49]), the most widely accepted methodology for measuring market risk is the Value-at-Risk approach. The VaR can be defined as the worst possible loss in the portfolio value over a given holding period (1 or 10 days) at the 99% confidence level ([57], [32]). Essentially, a mathematical model for VaR consists of two main parts: 1) modelling of proper multivariate risk factor distributions (processes) for the required time horizons; 2) evaluation of the portfolio (linear instruments, options and other derivatives) changes for the risk factor scenarios to produce a portfolio distribution. The evaluation part can be based on a full revaluation for the prices of instruments or partial revaluation methodologies (for example, Delta-Gamma-Vega approximation, [49]). Regulators also require complementing the VaR analysis with stress testing (scenarios for crashes, extreme movements in the market, etc.). Traditional methods of the VaR calculation are analytical (variance-covariance) method [58], historical simulation (combined with some bootstrapping procedures or other non-parametric methods, [32]), and parametric Monte Carlo simulation approach (see [33]). Primarily developed for the "normal" market conditions (multivariate Gaussian distribution for the risk factors), the variance-covariance method can be applied only for linear portfolios. The variance-covariance method can be extended from multivariate normal to the non-normal elliptical RF distributions (see Section 3.3). VaR for option portfolios is usually calculated based on simulation approaches. In this article, we concentrate on the parametric modelling of the RF distributions based on the Monte Carlo simulation procedures given an appropriate portfolio valuation methodology.

There are some market risk measures other than VaR closely related to the tails of the RF probability distributions, for example, Expected Shortfall (see [82]). The Expected Shortfall is defined as an average loss calculated from the losses that exceed VaR. The Expected Shortfall, as a conditional mathematical expectation, is an example of so-called coherent risk measures (see [7]) that, contrary to VaR, possess a natural subadditivity property (total risk of entire portfolio should be less or equal to a sum of risks of all sub-portfolios). In some cases, Expected Shortfall reflects the market risk better than VaR (it gives an answer to the question, what is the average of the worst case losses that occur at the corresponding confidence level). This market risk measure is more sensitive to the tail behavior than VaR. In general, it is wrong to say that only tails of the underlying RF distributions are important for the VaR or other risk measures. For example, a left tail for the portfolio of some barrier options or even European near at-the-money options may mostly depend on the central part of the underlying distribution. Therefore, it is a necessity to accurately model all parts of the RF distributions, including peaks at the origin and tails.

Due to short time horizons utilized in Market Risk Management (1 - 10 business days, contrary to Credit Risk Management with usual time horizons of years [32], [34]), the market risk factors are defined as daily log-returns, relative or absolute changes in the underlying prices, rates or implied volatilities, rather than these underlyings themselves. Such long-term effects as mean-reversion in the interest rate, commodity price, and implied volatility dynamics (with characteristic times 1 - 20 years) are not taken into account in the VaR modelling. Most of financial variables are positive (although, spreads and interest rate differentials may be negative). Except some rare situations (e.g., Japanese interest rates), daily changes for the underlyings are much less than 100% of the notional values, and, therefore, there is no need to apply any positive transformations to the market variables, like exponential or square transformations. Heuristically, this means that in most cases one can use "linear" RF simulation models for the VaR calculation.

1.2 Statistical Properties of the Market Risk Factors

There is extensive empirical evidence that historical daily return distributions for different underlyings in the foreign exchange, interest rate, commodity, and equity markets have high peaks, "fat" tails (excess kurtosis, Fig. 1 and 2) and skewness (right graph on Fig. 2) contrary to the normal distribution (see, for example, [76], [39], [33], [89], [16], [96], [25], [30]). Also, it is well known that the volatility of these financial variables varies stochastically with clustering [24] (see Fig. 3). These distributional properties have significant impact on Risk Management, specifically on VaR. A standard methodology usually used for the VaR calculation ([58]) exploits a multivariate normal distribution as a proxy for the RF distributions. The standard model corresponds to stable market conditions when one can neglect large jumps of the underlyings and volatility fluctuations. This results in underestimating of the actual VaR by the standard methodology and breaching the backtesting. A comprehensive RF simulation model should additionally capture the following important features observed in the market:

- different distributional shapes for different risk factors and markets (for example, short interest rates have much heavier tails, higher peaks and kurtosis than long term rates even for the same interest rate curve, Fig. 1; some commodity price distributions deviate more from normal than others);

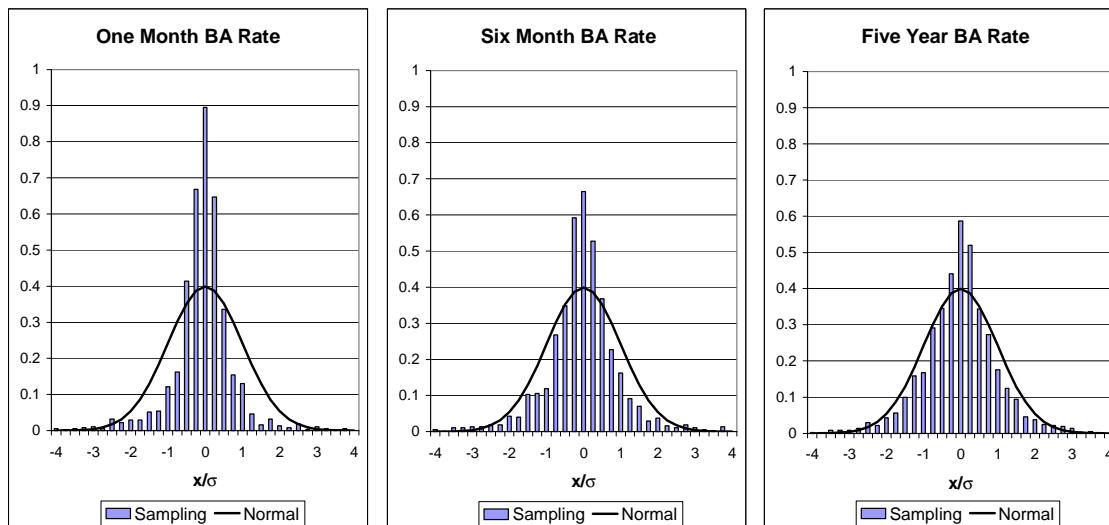


Figure 1: Variety of distributional shapes for CAD BA interest rate daily returns

- anomalously small normalization effect for large diversified portfolios contrary to the one predicted by the Central Limit Theorem (for example, S&P 500 Industrial Index or TSE 300 Index (Fig. 2), viewed as large portfolios of stocks, have markedly non-normal distributions with kurtosis about ten). This phenomenon points to a non-linear dependence between different risk factors (see also [38]);

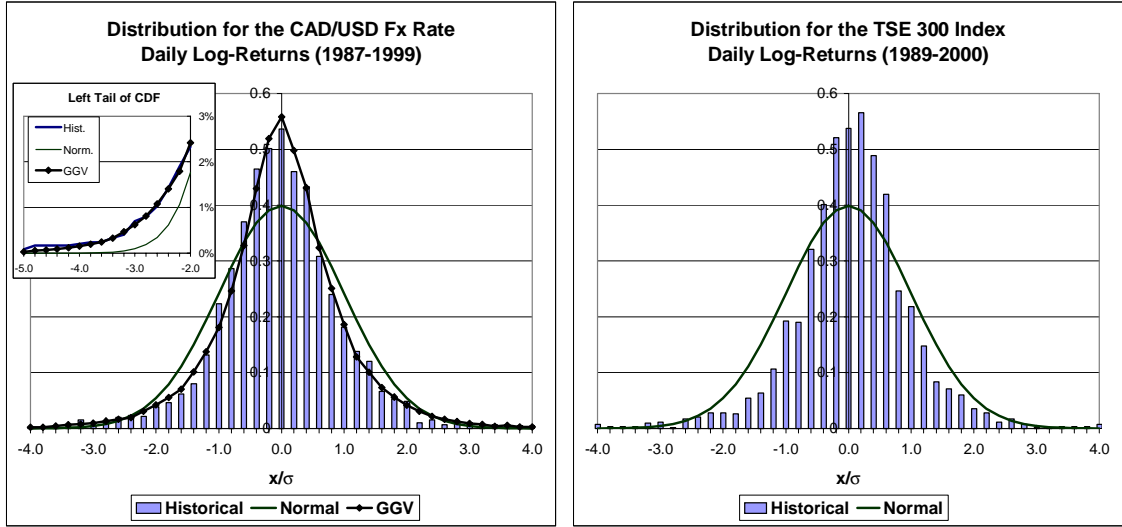


Figure 2: Distributions for the CAD/USD FX and TSE 300 daily log-returns

- normalization of the risk factor distributions for longer holding periods (for example, ten-day return distributions are significantly closer to normal than daily return distributions, on the other hand, intraday change distributions are clearly more distant from normal than daily ones [89], [31], [80]). A decreasing term structure of kurtosis points out to the same effect [33], [25];
- volatility clustering and non-linear time dependence in risk factor returns (for example, statistically significant autocorrelation in squares of virtually uncorrelated daily returns, see top graph on Fig. 3 and Fig. 10 in Section 2.3).

1.3 A Short Review of Stochastic Volatility Models

In this paper we restrict consideration of the SV models to the case of continuous time models. Time series approaches (ARCH, GARCH, etc.) [23], [24] are beyond the scope of the article.

L. Bachelier introduced the normal distribution and Brownian motion in finance in his Ph.D. Thesis [8] more than one hundred years ago. Brownian motion (that corresponds to a standard model for VaR [58]) was rediscovered in finance in [92], and then replaced by a Geometric Brownian motion for modelling of the stock dynamics [100]. Without any doubt, the Black-Scholes-Merton [22] option pricing model has become a main tool in modern finance. Since well-known investigations of Mandelbrot [76], [77] and Fama [39] on stable processes in the market, researchers have developed different approaches for modelling the abnormal behavior of the market variables. Fat-tailed distributions and jumps in the risk factors have been usually modelled by jump-diffusion processes ([86], [87], [20], [65]), processes with diffusion stochastic volatility ([50], [48], [104], [19], [85]), mixtures of normal and other distributions ([33], [95], [2]), and other methods [51], [103]. Also, different types of non-Gaussian Lévy processes were used to describe the dynamics of underlyings (we refer to [21], [41], [71], and [102] for the theory of infinitely divisible distributions and Lévy processes). Stable Paretian models in Finance were considered in [76], [77], [39], [83], [84], [88], [109], [96], and other works (see also [99], [54] and [90] for the theory, simulation and estimation of stable processes). Since pioneering 1973 paper of Clark [29], there have been a lot of research works on subordinated Lévy processes in finance: VG model [74], [73], [72]; Hyperbolic and Generalized Hyperbolic models [9]-[12], [35]-[37] (see also [81], [96]). A fine structure of asset returns from a Lévy process point of view was considered in [28], [46], [45] (CGMY model), [80], [25], [26], [13] (Truncated Lévy Flight). A general theory of conditionally normal stochastic variance and stochastic time change models is considered in [105], [106], [97], [75], [14].

Majority of papers discuss a one-dimensional case with applications to option pricing. However, multi-

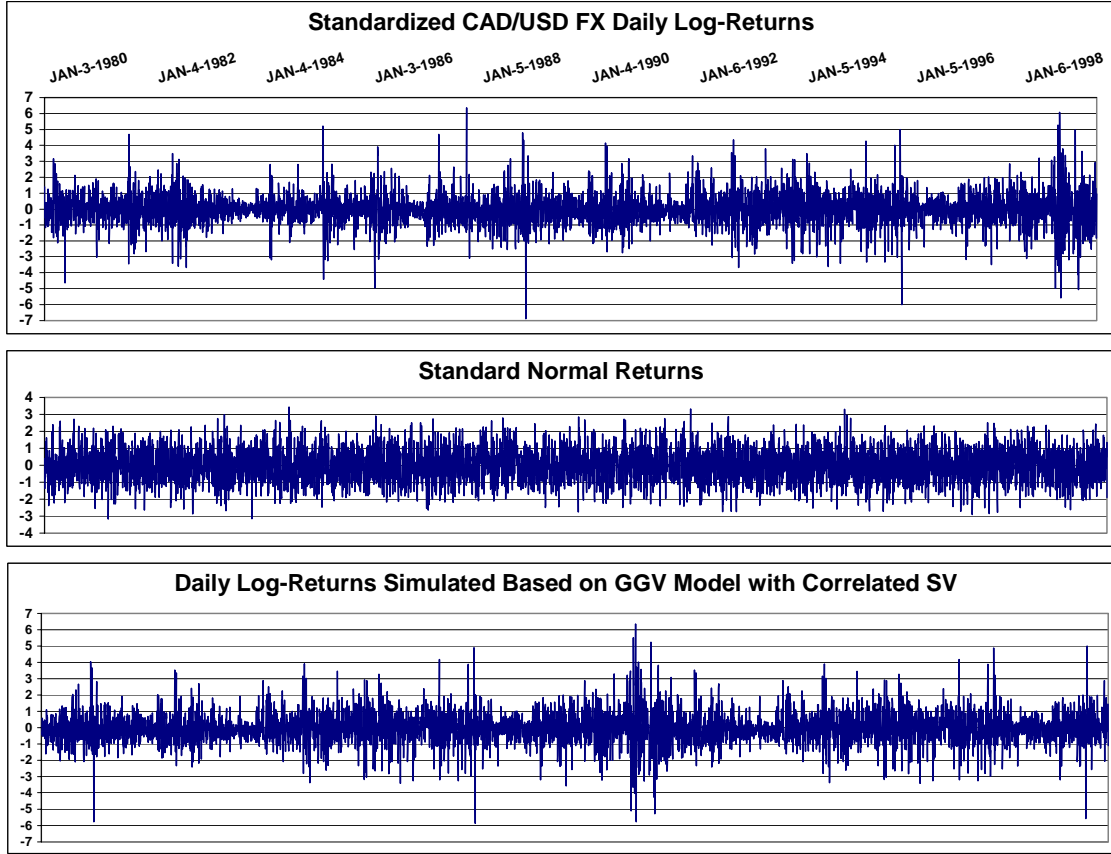


Figure 3: Volatility clustering and large deviations in CAD/USD FX rate daily returns

dimensional models with a large number of risk factors are of significance for Risk Management. This paper presents a new class of multivariate VaR models with the SV driven by Lévy processes.

2 Single-Factor Stochastic Variance Model

2.1 Maximum Entropy Approach and Lévy Processes

Let a risk factor X denote a t -day absolute return, relative return, or log-return of the underlying market variable. Value-at-Risk over a given holding period t with a specified confidence level q (usually, $q = 1\%$) is defined as a q -quantile of the distribution for the portfolio changes during the period t . For the standard model, a RF probability density function is normal with given constant mean and variance. We consider a class of conditional normal models where the variance V of the risk factor X is stochastic rather than constant. The stochastic variance of the underlying returns is not directly observable in the market. Generally, the most reliable information about the SV is its average value over some period of time. It can be estimated from the sampling variance of the underlying returns. Under conditions of uncertainty, it is reasonable to adopt a conservative approach, i.e. choose a probability distribution for the SV that provides the most uncertain outcomes given only information about the average value. A well-known measure of uncertainty associated with probability distributions is entropy [59]. Therefore, it is reasonable to determine the SV distribution from the Maximum Entropy principle.

A proposed single-factor SV model is based on the following assumptions [67]:

Assumption 1. The density function, $p_X(x, T)$, of the risk factor $X = X(T)$ for some holding period T is normal conditional upon the stochastic variance $V = V(T)$ that possesses a probability density function $p_V(v, T)$, $v \geq 0$, i.e.

$$p_X(x, T) = \int_0^{\infty} \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(x - \theta v - \mu T)^2}{2v}\right) p_V(v, T) dv \quad (1)$$

Parameter μT specifies a constant part of the mean for the conditional normal distribution, and parameter θ defines a shift in the mean proportional to the SV. As is shown later, θ determines the correlation between the RF and SV that results in a skewed RF distribution. The case $\theta = 0$ corresponds to a symmetric distribution. Linear dependence of the shift term θv from v in the mean of normal density is important for further construction of a Lévy process for the RF. The stochastic representation for X is as follows:

$$X(T) = \sqrt{V(T)}Z + \theta V(T) + \mu T, \quad Z \sim N(0, 1) \quad (2)$$

with Z being a standard normal random variable independent of $V(T)$.

Assumption 2. The average variance $E\{V(T)\}$ for the holding period T is known and equal to \bar{V} :

$$E\{V(T)\} = \int_0^{\infty} v p_V(v, T) dv = \bar{V} \quad (3)$$

Assumption 3. The probability density function $p_V(v, T)$ of the stochastic variance $V(T)$ is defined by the Maximum Entropy principle given the average variance(3):

$$H(p_V) = - \int_0^{\infty} p_V(v) \ln p_V(v) dv \longrightarrow \max_{p_V(v) \geq 0} \quad (4)$$

The optimization problem (4) for the SV density $p_V(v)$ subject to the constraint on the average variance (3) and standard normalization constraint $\int_0^{\infty} p_V(v) dv = 1$ has the exponential density

$$p_V(v) = \frac{1}{\bar{V}} \exp\left(-\frac{v}{\bar{V}}\right)$$

as a solution calculated by the Lagrange multiplier method [59]. According to the Law of Total Probability, the unconditional density (1) of the risk factor $X(T)$ has the following density:

$$p_X(x, T) = \frac{\lambda}{\bar{V}} \exp\left(-\frac{|x - \mu T|}{\lambda} + \theta(x - \mu T)\right), \quad \lambda = \sqrt{\frac{\bar{V}}{2 + \theta^2 \bar{V}}} \quad (5)$$

Distribution (5) is known as the skewed double exponential (Laplace) distribution [64]. This distribution has a sharp peak, exponential tails and non-zero skewness for $\theta \neq 0$. Kurtosis of a symmetric Laplace distribution is always equal to 6, in contrast to 3 for a normal distribution. Historical distributions of daily returns for many market variables, such as CAD/USD FX rate (Fig. 2), JPY/USD FX rate, S&P 500 Index, TSE 300 Index (Fig. 2), NYMEX Natural Gas futures prices, some LIBOR rates, etc., have a similar leptokurtic shape [67], [64].

In the case of a linear portfolio and symmetric Laplace distribution for the RF, the impact of non-normality on VaR can be estimated as

$$\frac{VaR_{Laplace}}{VaR_{Normal}} = \frac{\ln(2q)}{\sqrt{2}z_q}$$

where z_q is a standardized normal quantile for the confidence level q . For the case $q = 1\%$ ($z_q = 2.3263$), $VaR_{Laplace}$ for a linear portfolio is 19% higher than the standard VaR_{Normal} . The impact on VaR is

even more pronounced for non-linear instruments. For example, for a non-linear perfectly delta-hedged option portfolio, $\Pi(x)$, within Delta-Gamma approximation for the portfolio changes, $\delta\Pi(x) = 0.5\Gamma x^2$, the corresponding formulas for VaR are as follows:

$$VaR_{Laplace} = -\bar{V}\frac{\Gamma}{4}\ln^2(q), \quad VaR_{Normal} = -\bar{V}\frac{\Gamma}{2}(z_{\frac{q}{2}})^2$$

This results in 60% higher $VaR_{Laplace}$ number than VaR_{Normal} [67].

The exponential distribution for the SV was derived from the Maximum Entropy principle for some unspecified holding period T . To calculate VaR for different holding periods t , a stochastic process for the risk factor X is required. The standard normal model assumes that the risk factor X follows a Wiener process with independent stationary Gaussian increments. The simplest extension of this assumption is that the RF follows a Lévy process, i.e. a stochastic process with independent stationary (not necessarily Gaussian) increments. It can be shown [97] that within the class of conditionally normal models (2) this assumption is equivalent to the following assumption on the SV:

Assumption 4. *The total stochastic variance $V(t)$ in (2) follows a positive increasing Lévy process.*

The exponential distribution for the $V(T)$ is infinitely divisible. It uniquely determines a positive increasing pure jump Gamma process (see [102]) for the total stochastic variance $V(t)$, $t > 0$, with a Gamma probability density function

$$p_{V(t)}(v) = \frac{v^{\alpha t - 1}}{\Gamma(\alpha t)\beta^{\alpha t}} \exp\left(-\frac{v}{\beta}\right), \quad (6)$$

where $\alpha = 1/T$, $\beta = \bar{V}$. Assumptions 1 - 4 define the corresponding Lévy process for the risk factor $X(t)$ with the following probability density function:

$$p_X(x, t) = \sqrt{\frac{2}{\pi}} \frac{\lambda^{2\alpha t - 1} \exp(\theta\lambda y)}{\Gamma(\alpha t)\beta^{\alpha t}} |y|^{\alpha t - 1/2} K_{\alpha t - 1/2}(|y|), \quad y = \frac{x - \mu t}{\lambda} \quad (7)$$

Here λ is defined in (5), $\Gamma(\nu)$ is a gamma function, and $K_\nu(y)$ is the modified Bessel function of the third kind of the order ν , $K_{-\nu}(y) = K_\nu(y)$ [1]. Distribution (7) is known as a Bessel K-function distribution [55] or as a Generalized Laplace distribution [64]. Essentially, the SV model derived from the Maximum Entropy principle is equivalent to the Variance Gamma (VG) model (Gamma stochastic time change model, see [74], [73], [44]). The tail asymptotic behavior and behavior at the origin for the density (7) follows from known asymptotics for the modified Bessel function $K_\nu(y)$ [1]

$$K_\nu(y) \underset{y \rightarrow \infty}{\sim} \sqrt{\frac{\pi}{2y}} e^{-y}, \quad K_\nu(y) \underset{y \rightarrow 0}{\sim} \Gamma(\nu) 2^{\nu-1} y^{-\nu}, \quad \nu > 0, \quad K_0(y) \underset{y \rightarrow 0}{\sim} -\ln(y)$$

The RF density (7) has exponential tails for all t and a wide range of shapes at the origin, from almost normal "bell" shape (for large $\alpha \gg 1$) to a highly peaked ($0.5 < \alpha \leq 1$) and even unbounded shape ($0 < \alpha \leq 0.5$) (see Fig. 4). A skewed Laplace density (5) is a special case of (7) for $t = T$. The Bessel K-function family of distributions possesses finite moments of all orders. The characteristic function for the Gamma process has a simple form,

$$\phi_{X(t)}(\omega) = E\left\{e^{i\omega X(t)}\right\} = \frac{\exp(i\omega\mu t)}{(1 - i\omega\beta\theta + \omega^2\beta/2)^{\alpha t}} \quad (8)$$

The Lévy density from the Lévy-Khintchine representation of $\phi_{X(t)}(\omega)$ that characterizes the intensity of jumps of different sizes x has the following closed form (see [102]):

$$k(x) = \frac{\alpha}{|x|} \exp\left(-\sqrt{\frac{2}{\beta} + \theta^2}|x| + \theta x\right)$$

The RF distribution (7) tends to a normal distribution for $t \rightarrow +\infty$. This normalization effect is important for a proper VaR scaling from short holding periods to longer ones. The total variance $D_{X(t)}$

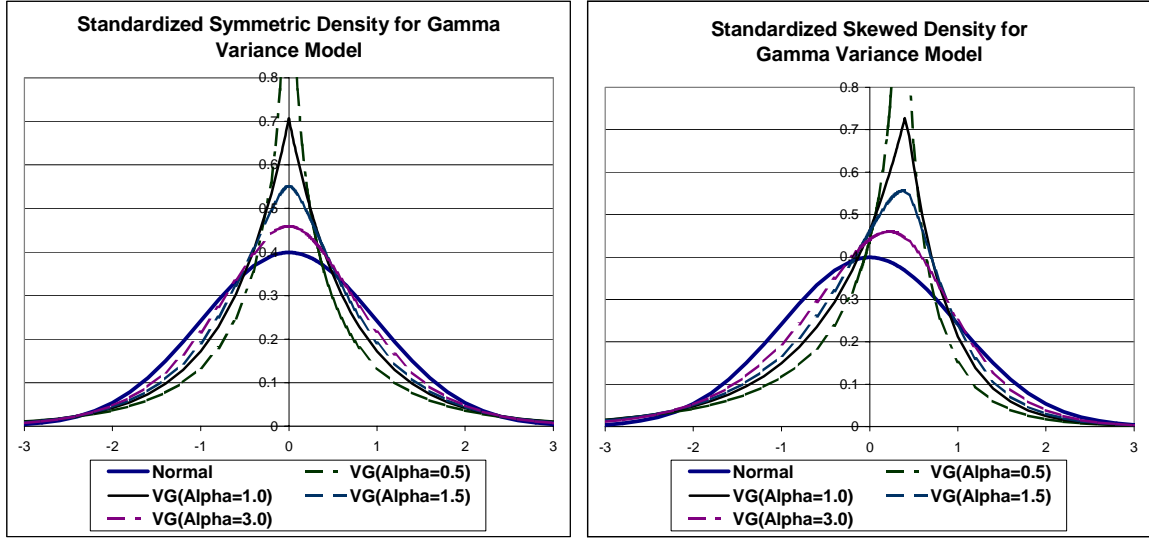


Figure 4: Probability densities for the Gamma SV model

is proportional to time, as it is for any Lévy process with finite variance [41] (a "square root of time" rule for the volatility is valid). However, contrary to the Gaussian case, the ratios of q -quantiles and standard deviation for the RF distributions (7) are not constant for different holding periods t . For example, the standardized 1%-quantile (VaR_{VG}) is higher for shorter holding period than the same 1%-quantile for longer holding period (Fig. 5).

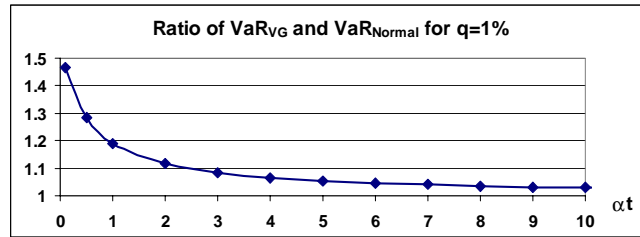


Figure 5: VG model 1%-VaR term structure w.r.t. 1% Normal VaR=2.33

The entropy for the SV distribution standardized by time t (the mean of a standardized SV is equal to 1 for all t) has the maximum at $t = T$ (Fig. 6) that corresponds to the exponential distribution. This property may be explained by transition of the standardized Gamma density from the delta-function at 0 to the delta-function at 1 as time t passes. Heuristically, this evolution of shape for the SV density corresponds to a transition from the state of maximum certainty at time 0 to the limiting state of maximum certainty at $t = \infty$ (with the limiting normal density for a standardized RF).

The following expressions provide a connection between the first four moments of the RF distribution and those of the SV distribution [69]:

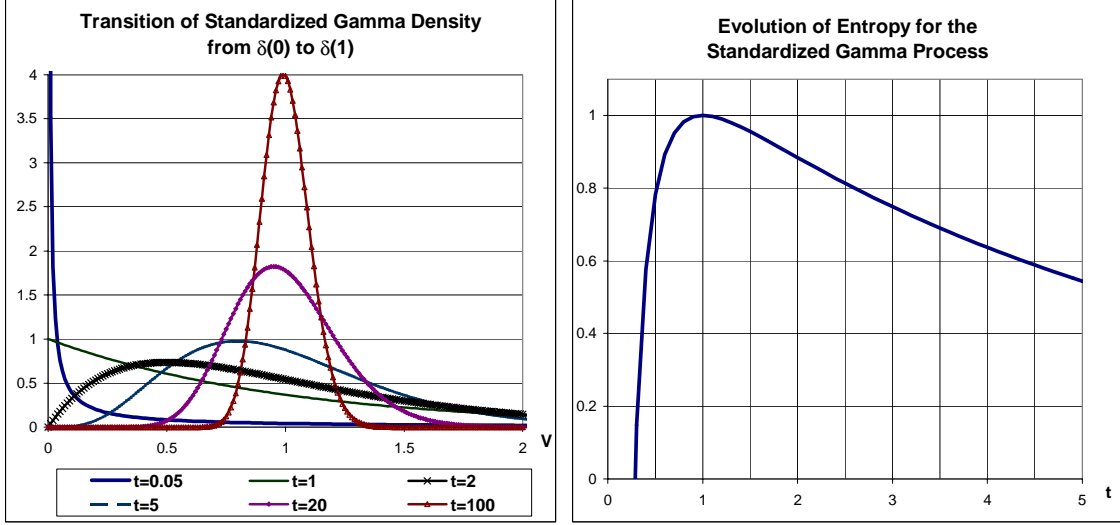


Figure 6: Evolution of standardized Gamma SV density and entropy

$$\begin{aligned}
 m_{X(t)} &= \mu t + \theta m_{V(t)} \\
 D_{X(t)} &= m_{V(t)} + \theta^2 D_{V(t)} \\
 m_{3,X(t)} &= \theta (3D_{V(t)} + \theta^2 m_{3,V(t)}) \\
 m_{4,X(t)} &= 3m_{V(t)}^2 + 3D_{V(t)} + 6\theta^2 m_{V(t)} D_{V(t)} + 6\theta^2 m_{3,V(t)} + \theta^4 m_{4,V(t)}
 \end{aligned} \tag{9}$$

The expressions (9) for the moments are valid for conditional normal models of the form (1) provided that the distribution $p_{V(t)}(v)$ for the stochastic variance $V(t)$ possesses moments up to the fourth order. Parameter θ controls skewness of the RF distribution and defines the correlation $\rho_{X,V}$ between the risk factor X and its stochastic variance V :

$$\rho_{X,V} = \theta \sqrt{\frac{D_V}{m_V + \theta^2 D_V}}$$

A parameter estimation procedure with respect to the four parameters, μ , θ , β , and α (model calibration), can be based either on the Maximum Likelihood approach or the method of moments. Given four sampling central moments for the T_1 -day underlying returns and analytical expressions for the moments of the Gamma stochastic variance [55]

$$m_{V(T_1)} = \alpha T_1 \beta, \quad D_{V(T_1)} = \alpha T_1 \beta^2, \quad m_{3,V(T_1)} = 2\alpha T_1 \beta^3, \quad m_{4,V(T_1)} = 3\alpha T_1 \beta^4 (\alpha T_1 + 2),$$

equations (9) can be used for the model calibration. Note that time T , corresponding to the maximum entropy for the SV density, can be recovered from the calibrated parameter α as $T = 1/\alpha$.

It follows from (6) and (9) that the term structure of the RF variance and kurtosis for the symmetric case of the Gamma-SV model ($\theta = 0$) is:

$$D_{X(t)} = \alpha \beta t, \quad Kurt_{X(t)} - 3 = \frac{3}{\alpha t} \tag{10}$$

2.2 Generalized Gamma Variance Model

Some market variables exhibit jumps as large as 5 to 10 daily standard deviations [39], [25], [80], [30]. Such events have significantly lower theoretical probability to occur for the corresponding period of observations

not only for the normal model, but even for the Gamma SV model with exponential tails. Extremely large jumps in the risk factors have often been described by distributions with polynomial tails, specifically by stable distributions [76], [77], [88], [96]. However, stable Paretian distributions do not have finite variance (volatility). This contradicts the majority of empirical observations (see [89]). Also, volatility is a main tool in financial risk management and pricing. Therefore, heavy tailed distributions with finite variance are of considerable interest for the finance applications. An example of such distribution widely discussed in the financial literature is Student t-distribution [93], [2], [96]. The family of distributions introduced below includes t-distribution as a special case. The symmetric Gamma SV model considered above has only one shape parameter, α , that controls both the tails and central part of the distribution. It seems that one shape parameter is insufficient to distinguish between sources of high kurtosis: whether it comes from heavy tails or high peak. It is possible to show that for a class of conditional normal models the tail asymptotics of the RF distribution depends upon the tail asymptotics of the corresponding SV distribution. Therefore, a more general SV model that allows for separate control for the tails and peak should more successfully describe large deviations of the risk factors.

Note, that the Gamma SV density (6) can be formally derived from the Maximum Entropy principle (4) without Assumption 4. Instead, one can use a constraint on the logarithmic moment $E\{\ln(V)\}$ in addition to the condition on the average variance $E\{V\}$ [59]. Essentially, this logarithmic constraint defines the power behavior of the SV density at the origin, while the constraint on $E\{V\}$ defines the exponential tail behavior. The condition on average variance can be replaced by a more flexible condition to accommodate information on a generalized moment of any power for the SV [69], [70]. For example, one can assume that the average volatility is known instead of average variance. This approximately corresponds to a constraint on the fractional moment $E\{\sqrt[\nu]{V}\}$ instead of $E\{V\}$. Hence, we can formally define the Entropy maximization problem (4) with two essential constraints

$$\int_0^{\infty} \ln(v) p_V(v) dv = c_0, \quad \int_0^{\infty} v^{1/\nu} p_V(v) dv = c_1 \quad (11)$$

and a standard normalization constraint for a probability density function. The use of the Maximum Entropy approach with a constraint on the generalized moment $E\{V^{1/\nu}\}$, $\nu \in \mathbb{R}^1$, allows for a desirable generalization of the Gamma SV model to a broad class of models with a wide range of heavy tails, from exponential and sub-exponential (stretched exponential) to polynomial [69], [70]. A solution of the maximization problem (4), (11) is the Generalized Gamma density for the stochastic variance V :

$$p_V(v) = \frac{v^{\frac{\alpha}{\nu}-1}}{|\nu| \Gamma(\alpha) \beta^\alpha} \exp\left(-\frac{v^{1/\nu}}{\beta}\right) \quad (12)$$

The corresponding stochastic representation for V is a ν -th power of the Gamma distributed random variable γ with the density (6) (see [55]):

$$V = \gamma^\nu \quad (13)$$

Stochastic representations (2) and (13) allow an effective Monte Carlo simulation procedure for the RF given well-known simulation procedures for normal and gamma random variables [43].

A Generalized Gamma distribution is a very flexible class of distributions with two shape parameters α and ν . This class includes Gamma ($\nu = 1$), Inverse Gamma ($\nu = -1$), and Weibull ($\alpha = 1$, $\nu > 0$) distributions as special cases. It is known that Generalized Gamma distribution is infinitely divisible for all these representatives (see [47], [52], [102]) and for positive $\nu \geq \max(\alpha, 1)$ [53]. Therefore, for all of these cases the Generalized Gamma distribution produces Lévy processes for the SV. We do not know if the Generalized Gamma distribution is infinitely divisible for arbitrary values of $\nu \in \mathbb{R}^1$, nor whether there is a closed form representation for the characteristic function. Hence, we apply the distribution (12) to describe the returns for the shortest holding period available, say one day, and then construct an additive SV process for a longer holding period, say 10 days, by summing up the independent Generalized Gamma distributed random variables. An analytical formula for moments of the Generalized Gamma distribution is readily

Constr. $E\{V^{1/\nu}\}$	SV Density & Stoch. Rep.	RF Density	RF Asymptotics $x \rightarrow \infty$
$E\{V\}, \nu = 1$	Gamma, $V = \gamma$	K-Bessel	$\sim x^{\alpha-1} e^{-cx}$
$E\{\sqrt{V}\}, \nu = 2$	Square of Gamma, $V = \gamma^2$	GGV(2, α)	$\sim x^{\frac{2\alpha}{3}-1} e^{-cx^{2/3}}$
$E\{1/V\}, \nu = -1$	Inverse Gamma, $V = 1/\gamma$	t-Distribution	$\sim x^{-(2\alpha+1)}$
$E\{V^{1/\nu}\}, \nu > 0$	Generalized Gamma, $V = \gamma^\nu$	GGV(ν, α)	$\sim x^{\frac{2\alpha}{(1+\nu)}-1} e^{-cx^{2/(1+\nu)}}$
$E\{V^{1/\nu}\}, \nu < 0$	Generalized (Inv.) Gamma, $V = \gamma^\nu$	GGV(ν, α)	$\sim x^{-\left(\frac{2\alpha}{ \nu }+1\right)}$
$\nu = 0$	SV degenerates to $V \equiv 1$	Normal	$\sim e^{-\frac{1}{2}x^2}$

Table 1: GGV Model Summary

available

$$E\{V^k\} = \frac{\beta^{k\nu} \Gamma(\alpha + k\nu)}{\Gamma(\alpha)}$$

(a condition for the k -moments to exist is $(\alpha + k\nu) > 0$).

The corresponding RF density $p_X(x)$ is given by the integral (1) with SV density $p_V(v)$ being of the form (12). We call this density a Generalized Gamma Variance (GGV) density. Unfortunately, in the general case there is no closed analytical form for the density $p_X(x)$. However, we consider an effective numerical procedure for calculating the integral (1) to be as good as, for example, a ‘‘closed form’’ formula (7) involving special K-Bessel functions. Effective asymptotic expansion methods [91], [1] can be applied for the numerical calculations¹. In the case of a symmetric GGV density, there is an analytical formula for the moments of any fractional order k (finite for $(\alpha + \frac{k\nu}{2}) > 0$):

$$E\{|X|^k\} = \frac{2^{\frac{k}{2}} \beta^{\frac{k\nu}{2}} \Gamma(\frac{k+1}{2}) \Gamma(\alpha + \frac{k\nu}{2})}{\sqrt{\pi} \Gamma(\alpha)} \quad (14)$$

The moments cease to exist for some combinations of negative values of ν and $\alpha > 0$ because of polynomial tails for the GGV density.

Below, we provide some results for a symmetric density $p_X(x)$. There are some known special analytical cases for $p_X(x)$:

- i) $\nu = -1$ corresponds to the t -distribution with 2α degrees of freedom;
- ii) $\nu = 0$ corresponds to the Gaussian distribution;
- iii) $\nu = +1$ corresponds to the K-Bessel function distribution (7).

Table 1 presents a summary of results for the Generalized Gamma Variance model, including a constraint on the generalized moment in Maximum Entropy principle (column 1), SV stochastic representation (column 2), corresponding RF density (column 3), and asymptotics for the tails of the RF density (column 4).

Some market variables are better described by distributions with polynomial tails, while others are better described by distributions with semi-heavy tails (exponential and sub-exponential) (see [93], [96], [33]). The GGV model is capable of accommodating both types of behavior. A range of values $\nu < 0$ corresponds to a power law tails. GGV density is finite at zero for all $\nu < 0$. A range of values $\nu > 0$ corresponds to exponential and sub-exponential tails. In this case, tails are far lighter and the moments of all orders exist. The range $\nu > 1$ corresponds to a class of stretched exponential densities $p_X(x)$. The specific class of the stretched exponential distributions based on a modified Weibull density was considered in [103]. Fig. 7 shows the RF GGV density $p_X(x)$ for different values of parameters ν and α . Parameter ν brings an extra flexibility to the GGV density: it is seen that GGV model can accommodate a wide variety of shapes and tail behavior.

A statistical investigation of different SV models from a Generalized Hyperbolic family based on historical data for 15 stock market indices was presented in the paper [93]. The class of Generalized Hyperbolic distributions developed in [10], [12], [35][36] is also a two shape parameter family in symmetric case. All

¹Effective numerical procedure and software for the GGV density calculation was developed by Xiaofang Ma.

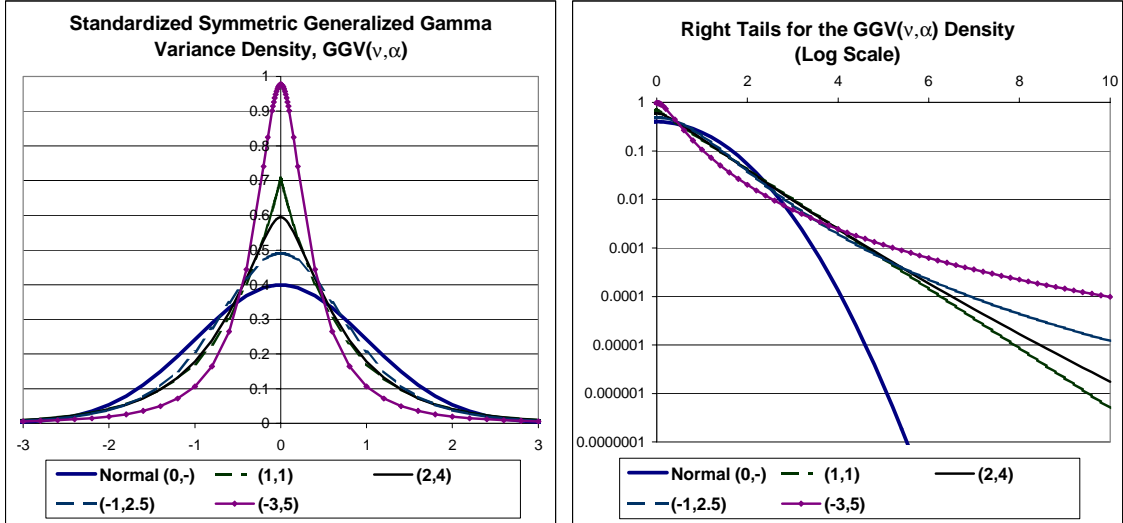


Figure 7: Generalized Gamma Variance densities

members of this family have exponential tails except the Student t-distribution, which has polynomial tails. For this specific case, the class of Generalized Hyperbolic distributions collapses to a one shape parameter (number of degrees of freedom) family. Four representatives from a Generalized Hyperbolic class (Student t-distribution, Normal Inverse Gaussian, Variance Gamma, and Hyperbolic distributions) were compared based on the Maximum Likelihood criteria. The last three of these distributions have exponential tails. Results presented in [93] show that all distributions having exponential tails fail to satisfy the Pearson χ^2 test. In contrast, the t-distribution has not been rejected on a 99% confidence level for ten of the fifteen indices. Two-parameter Paretian tail GGV distributions perform better than the t-distribution. As an example, Fig. 8 demonstrates a fit for the Canadian 3-month BA interest rate daily return density (1992 - 1998) by normal, Student-t, and GGV densities calibrated using Maximum Likelihood approach. It is seen that $GGV(\alpha, \nu)$ with optimal parameters $\nu = -5.5$ and $\alpha = 15$ outperforms t-distribution, and both GGV and t-distributions significantly outperform normal. The χ^2 value for the $GGV(15, -5.5)$ is about 80% less than χ^2 value for the calibrated t-distribution. It is interesting to note, that during the period 1992 - 1998, Canadian 3-month BA interest rate exhibited 14 large daily moves greater than four standard deviations (about 1% of all observations).

Another example (Fig. 9) shows a GGV model log-likelihood surface for S&P 500 Index as a function of parameters ν and α . A deep minimum for $\nu = 0$ corresponds to the normal distribution, while two wings correspond to the power law ($\nu < 0$) and stretched exponential ($\nu > 1$) tailed distributions. For this example, a stretched exponential sub-class produces almost the same maximum likelihood value as a power law sub-class.

2.3 Mean-Reverting Stochastic Variance Model

So far, we have considered a class of SV models driven by Lévy processes with independent, identically distributed, but not necessarily Gaussian increments. The model explains non-normality of the RF distributions. For any conditional normal SV model, expressions (9) provide an exact answer for the term structures of the risk factor variance and kurtosis

$$D_{X(t)} = m_{V(t)}, \quad Kurt_{X(t)} - 3 = 3 \frac{D_{V(t)}}{m_{V(t)}^2} \quad (15)$$

Here $V(t)$ is total variance. Since $m_{V(t)}$ and $D_{V(t)}$ are linear functions of time for any Lévy process for $V(t)$, the above expressions predict linear increase of the RF variance and hyperbolic decrease of its excess

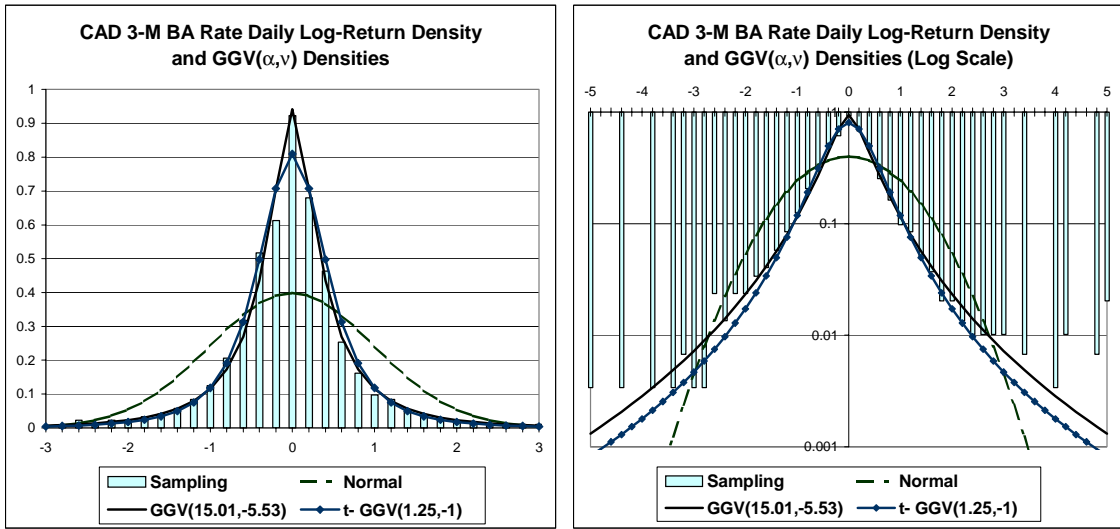


Figure 8: Historical and calibrated GGV densities for the CAD 3-month BA interest rate daily log-returns

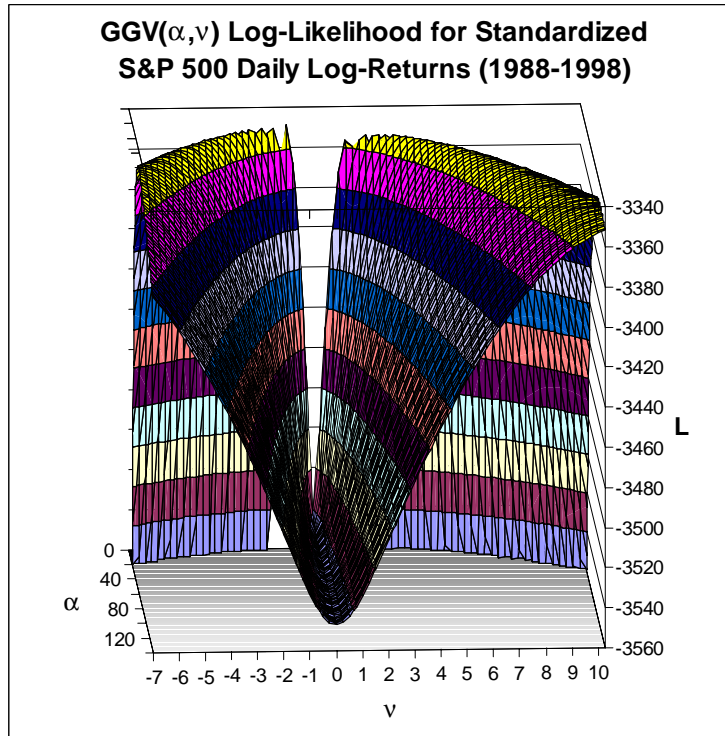


Figure 9: GGV model log-likelihood surface for S&P 500

kurtosis.

However, empirical investigations show that the underlying returns are almost uncorrelated, but not independent (see [25], [30], [89]). The easiest way to demonstrate this dependence is to consider the empirical correlations for the absolute values or squares of the returns. It is seen (Fig. 10) that autocorrelations of squares are statistically significant. This phenomenon is connected with a known volatility clustering effect (Fig. 3). Also, it is known that empirical term structure of kurtosis decreases slower than is predicted by a ‘‘Lévy term structure’’ model (15) (so called ‘‘anomalous decay’’, see [25], [30]). All this suggests that a better model for the instantaneous stochastic volatility is not a ‘‘white noise’’ kind of process, but rather a process with autocorrelation. One way to account for the autocorrelation structure of the SV is to consider regime-switching SV processes (see [62]). We will follow another approach to introduce the SV autocorrelation by considering Ornstein-Uhlenbeck (OU) type processes for the instantaneous SV [67], [69]. Such class of non-Gaussian OU type processes driven by positive Lévy noise was investigated in detail in [15], [16]. In this section we will only demonstrate that the empirically observed term structure of kurtosis can be consistently described by such models.

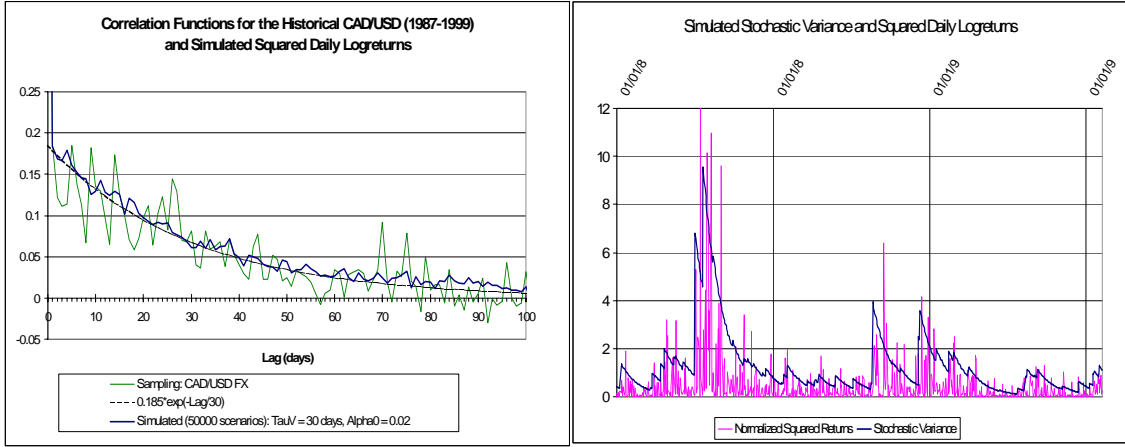


Figure 10: Autocorrelation in the squared CAD/USD FX daily returns and in the SV

Consider a stationary non-negative process $\xi(t)$ with autocorrelation function $R_\xi(\tau)$ that describes the instantaneous stochastic variance. For the total variance $V(t)$ being $V(t) = \int_0^t \xi(t') dt'$, it follows that

$$m_{V(t)} = m_{V(1)}t, \quad D_{V(t)} = 2 \int_0^t (t - \tau) R_\xi(\tau) d\tau$$

These expressions in conjunction with (15) can be used to calculate a term structure of the RF kurtosis. In particular, assume a mean-reverting process for the instantaneous stochastic variance $\xi(t)$ be a OU type process

$$d\xi(t) = -\lambda\xi(t)dt + \lambda dG(t) \quad (16)$$

where $G(t)$ is, for example, a Gamma process, and $\lambda > 0$ is a mean-reversion speed parameter. Expressions for $R_\xi(\tau)$ and variance $D_{V(t)}$ are as follows

$$R_\xi(\tau) = \frac{\lambda\alpha\beta^2}{2} e^{-\lambda|\tau|}, \quad D_{V(t)} = \frac{\alpha\beta^2 t}{2} \left(1 - \frac{1 - e^{-\lambda t}}{\lambda t} \right)$$

The autocorrelation function $R_\xi(\tau)$ is an exponential function for any OU model (16). It is seen that $D_{V(t)}$ is not a linear function of time contrary to the Lévy case (10). Previous formulas and formulas (15) result

in the following term structure of the RF kurtosis:

$$Kurt_{X(t)}^{OU} - 3 = \frac{3}{\alpha t} \left(1 - \frac{1 - e^{-\lambda t}}{\lambda t} \right)$$

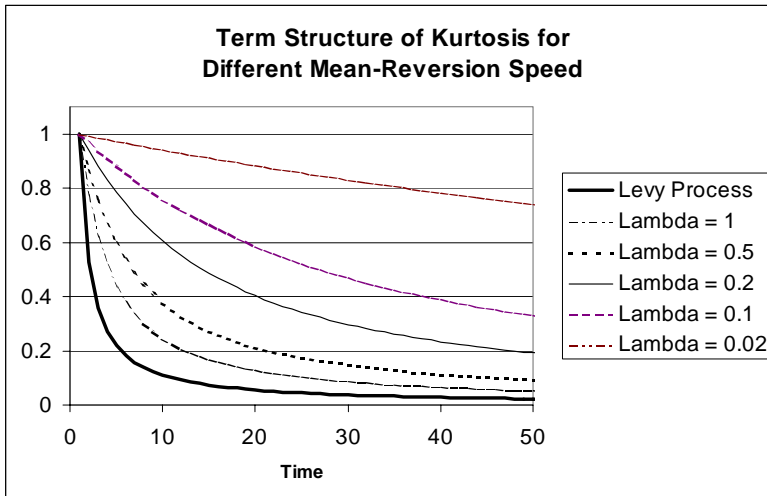


Figure 11: Term structure of the RF kurtosis for the model with autocorrelated SV

Fig. 11 shows a term structure of the RF kurtosis for different values of the mean-reversion speed parameter λ . As expected, the OU stochastic variance process provides slower decay of kurtosis vs. Lévy SV process. Reduction in decay can be significant depending upon the mean-reversion speed λ . This is equivalent to slower “normalization” effect. The right graph in Fig. 10 presents the time series for the simulated SV and empirical CAD/USD FX rate squared daily log-returns. The bottom graph in Fig. 3 presents the simulated RF time series. It is evident that the model produces large deviations for the FX rate and volatility clustering effect that is very similar to the one observed in the market (top graph in Fig. 3).

3 Multifactor Stochastic Variance Model

3.1 Requirements for Multifactor VaR Models

A realistic multifactor VaR model should consistently describe not only the correlation and volatility structure for the risk factors, but also different shapes of the marginal risk factor distributions and distributions in other “diagonal” directions. Also, a principal component analysis for daily returns in different markets (interest rate curves, commodity futures prices, implied volatility curves and surfaces), clearly indicates the presence of non-linear dependence between risk factors (principal components). For example, the squared daily changes of the principal components are significantly correlated, while daily changes themselves are uncorrelated. This non-linear dependence breaks conditions of the Central Limit Theorem and has an important impact on VaR calculation: even for well-diversified linear portfolios with a large number of instruments there is no full normalization of the portfolio return distributions [67], [68]. An example of such large diversified portfolio is the S&P500 Index. Its distribution is quite far from normal despite the portfolio averaging effect. Hence, a comprehensive model for multiple risk factors should additionally capture the following important features observed in the market:

- exact match of a given volatility and correlation structure of the risk factors;
- approximate match of shapes, kurtosis, and tails for different risk factors (marginal distributions);

- approximate match of shapes, kurtosis, and tails for different linear sub-portfolios (marginal distributions in diagonal directions).

The model should also allow for an effective Monte Carlo simulation procedure. To facilitate further multivariate analysis, in the sequel, we shall consider the case of symmetric joint probability distributions for the RF returns.

3.2 "Naïve" Multifactor Model

A very simple idea for constructing a multivariate conditionally Gaussian stochastic variance model is to define a distribution for the vector of risk factors $X(t) \in \mathbb{R}^N$ as a multivariate normal with some fixed correlation matrix R and independent stochastic variances $V_i(t)$, $i = 1, \dots, N$. A symmetric multivariate probability density function for the vector of risk factors is represented as:

$$p_{X(t)}(x) = \int \cdots \int_{V_1 \dots V_N} \frac{1}{\sqrt{(2\pi)^N \det(C)}} \exp\left(-\frac{x' C^{-1} x}{2}\right) p_V(V_1, \dots, V_N) dV_1 \dots dV_N, \quad (17)$$

$$C = \Sigma R \Sigma', \quad \Sigma = \text{diag}(\sqrt{V_1(t)}, \dots, \sqrt{V_N(t)}) \quad (18)$$

Here $p_V(V_1, \dots, V_N) = \prod_{i=1}^N p_{V_i}(V_i)$ is a probability density for independent stochastic variances $V_i(t)$, x' is transpose of x . The corresponding stochastic representation for the risk factors $X(t)$ is

$$X(t) = \text{diag}\left(\sqrt{V_1(t)}, \dots, \sqrt{V_N(t)}\right) AZ, \quad AA' = C, \quad Z \sim N(0, I), \quad (19)$$

where $Z \sim N(0, I)$ is independent of V standard normal vector with identity covariance matrix I . This representation allows for modelling marginal distributions with different leptokurtic shapes.

However, it can be shown that this "naïve" approach reduces the correlations between risk factors because of "randomization" for the covariance matrix [67]. Due to independence of the stochastic variances V_i , absolute values of the model correlations $\text{Corr}(X_i, X_j)$ are less than absolute values of the correlations R_{ij} used in (17):

$$\begin{aligned} \text{Cov}(X_i, X_j) &= \int \int x_i x_j p_X(x) dx_i dx_j = R_{ij} \int \int \sqrt{V_i} \sqrt{V_j} p_V(V_i, V_j) dV_i dV_j \\ &= R_{ij} \int \sqrt{V_i} p_{V_i}(V_i) dV_i \int \sqrt{V_j} p_{V_j}(V_j) dV_j = f_{ij} \sigma_{X_i} \sigma_{X_j} R_{ij}, \quad i \neq j \end{aligned} \quad (20)$$

Reduction factors f_{ij} , $i \neq j$, are less than one, because

$$\int \sqrt{V_i} \mathbf{1} p_{V_i}(V_i) dV_i < \sqrt{\int V_i p_{V_i}(V_i) dV_i} \sqrt{\int \mathbf{1} p_{V_i}(V_i) dV_i} = \sqrt{E\{V_i\}} = \sigma_{X_i}$$

It means that the sampling correlation matrix cannot be used as the matrix R in (17). For example, the reduction factors $f_{ij} < 1$, $i \neq j$, calculated explicitly for the case of the Gamma stochastic variances (6) are as follows:

$$\text{Corr}(X_i, X_j) = f_{ij} R_{ij}, \quad f_{ij} = f(\alpha_i, \alpha_j) = \frac{\Gamma(\alpha_i + 1/2) \Gamma(\alpha_j + 1/2)}{\Gamma(\alpha_i) \Gamma(\alpha_j) \sqrt{\alpha_i \alpha_j}}, \quad i \neq j$$

The underestimation of the correlations can be significant for some values of parameters α_i , α_j as it is shown in Table 2.

The randomization effect exists for any probability density functions $p_{V_i}(V_i)$ for independent stochastic variances. Usually, equations $\text{Corr}(\tilde{X}_i, \tilde{X}_j) = f_{ij} R_{ij}$ cannot be resolved with respect to correlations R_{ij} given sampling correlations $\text{Corr}(\tilde{X}_i, \tilde{X}_j)$ while preserving the necessary conditions $|R_{ij}| \leq 1$ or non-negative definite matrices R . Hence, this "naïve" model does not allow to preserve historical correlations between the risk factors.

$\alpha_i = \alpha_j$	0.5	1	2	5	10
$f(\alpha_i, \alpha_j)$	0.64	0.79	0.88	0.95	0.98

Table 2: Correlation reduction factors

Remark. Equation (20) and the inequality

$$\int \int \sqrt{V_i} \sqrt{V_j} p_V(V_i, V_j) dV_i dV_j < \sqrt{\text{Cov}(V_i, V_j) + E\{V_i\} E\{V_j\}}$$

show that the class of the SV models with the stochastic representation (18) for the covariance matrix preserves the RF correlation structure only if

$$\int \int \sqrt{V_i} \sqrt{V_j} p_V(V_i, V_j) dV_i dV_j = \sqrt{E\{V_i\}} \sqrt{E\{V_j\}},$$

which requires dependent stochastic variances with positive correlations. We do not investigate this direction in the article.

3.3 Elliptical Stochastic Variance Model

The simplest extension of a single-factor SV model to the multifactor case is an Elliptical SV model. Elliptical models are widely used for representing non-normal multivariate distributions in finance (see [36], [64]). This class of models preserves the observed RF correlation structure. The model is similar by construction to the one-dimensional variance mixture of normals. An elliptical N -dimensional symmetric process $X^E(t)$ for N risk factors has a stochastic representation as a single variance mixture of multivariate normals with a given covariance matrix C :

$$X^E(t) = \sqrt{V(t)} Z^C, \quad Z^C \sim N(0, C) \quad (21)$$

Here $V(t)$ is a univariate SV process, Z^C is a multivariate normal N -dimensional vector independent of $V(t)$. The covariance matrix C is estimated from historical T_1 -day returns (e.g., daily returns), while the SV is normalized to satisfy a condition $m_V(T_1) = E\{V(T_1)\} = 1$. The unconditional density for the random vector of risk factors $X^E(t)$ is:

$$p_{X^E(t)}(x) = \int_0^\infty \frac{1}{\sqrt{(2\pi V)^N \det(C)}} \exp\left(-\frac{x^* C^{-1} x}{2V}\right) p_{V(t)}(V) dV$$

As an example, consider the case of Gamma stochastic variance $V(t)$. A closed analytical form for the unconditional elliptical Bessel K-function density for $X^E(T)$ is available in [64]. A characteristic function $\phi_{X^E(t)}(\omega)$ for the elliptical Lévy process $X^E(t)$ is represented as:

$$\phi_{X^E(t)}(\omega) = \left(1 + \frac{\beta}{2} \omega' C \omega\right)^{-\alpha t} \quad (22)$$

where ω is N -dimensional vector, ω' is a vector transposed to ω . Due to known properties of elliptical distributions (see [40]), all marginal one-dimensional distributions for the risk factors are univariate Bessel K-function distributions with the same shape parameter αt and the same kurtosis. They differ only by the standard deviations. The same property holds for all one-dimensional distributions of linear combinations $X_\Delta = \Delta' X^E(t)$ of the risk factors. These linear combinations correspond to the linear portfolios defined by Δ . The kurtosis of $X_\Delta(T_1)$ for any arbitrary Δ is equal to $k_\Delta = 3(1 + D_V(T_1)/m_V^2(T_1)) = 3(1 + D_V(T_1))$. Therefore, within a class of elliptical models there is no normalization effect at all for the distributions of large diversified portfolios. This is a result of violation of the conditions for the Central Limit Theorem: the risk factors are dependent through the common stochastic variance V . Such property is a drawback for all elliptical models. It is clear that the actual RF fluctuations are not driven by a single stochastic variance

("global market activity"). More realistic SV model should include a multidimensional processes for the SV to model different distributional shapes for the risk factors and linear sub-portfolios. Since sampling marginal RF distributions have different shapes, the calibration of elliptical model is restricted to fitting a distribution of some preselected portfolio. Hence, the calibration of elliptical models is portfolio dependent.

3.4 Independent Stochastic Variances for the Principal Components

One of the possible ways to model different shapes for the RF distributions while preserving a given correlation structure was considered in [67], [68]. An N -dimensional vector of the risk factors is represented as a linear combination of principal components (PC) with independent one-dimensional stochastic variances. The corresponding stochastic representation is as follows:

$$\begin{aligned} X^L(t) &= \tilde{A}Z^I(t), \quad Z_i^I(t) = \sqrt{V_i(t)}Z_i, \\ Z_i &\sim N(0, 1), \quad i = 1, \dots, M \end{aligned}$$

Here Z_i are independent standard normal variables, $V_i(t)$ are independent SV processes with a unit mean and some variances D_{V_i} for a specified time horizon T_1 . The columns of a constant matrix \tilde{A} are the principal components of a given covariance matrix C . The covariance matrix C is estimated from the historical T_1 -day returns. Matrix \tilde{A} is calculated based on eigenvalue decomposition of the covariance matrix C as [108]:

$$C = UDU', \quad U' = U^{-1}, \quad D = \text{diag}(d_1, \dots, d_N) \quad (23)$$

$$\tilde{A} = U_M D_M^{1/2}, \quad D_M = \text{diag}(d_1, \dots, d_M), \quad M \leq N, \quad C = \tilde{A}\tilde{A}' \quad (24)$$

Matrix U_M consists of the first M columns of the orthogonal matrix U , which correspond to the first M largest eigenvalues d_1, \dots, d_M of the matrix C . Number M may be chosen less than N if the matrix C is singular and has only M non-zero eigenvalues. Some numerical issues related to singularity of the matrix C were considered in [66]. It follows from the construction of the process $X^L(t)$ that $\text{Cov}(X^L(T_1)) = C$. This ensures an exact match of the sampling covariance matrix C . One can keep even a smaller number M of the principal components in (24) and recover the matrix C with the required accuracy.

A characteristic function for the model is a product of the characteristic functions of one-dimensional processes for the PCs. For example, a characteristic function for the Gamma SV model with independent SV has a form

$$\phi_{X^E(t)}(\omega) = \prod_{i=1}^M \left(1 + \frac{(\omega'(\tilde{A})_i)^2}{2} \right)^{-\alpha_i t}$$

where $(\tilde{A})_i$ is i -th column of the matrix \tilde{A} .

The matrix \tilde{A} can be defined up to an arbitrary orthogonal transformation H without change of the covariance matrix C since Z_i are independent standard normal variables, Z_i and $V_i(T_1)$ are independent and $E\{V_i(T_1)\} = 1$. Hence, $E\{Z^I(T_1)Z^I(T_1)'\} = I$ and

$$E\{X^L(T_1)X^L(T_1)'\} = \tilde{A}HE\{Z^I(T_1)Z^I(T_1)'\}H'\tilde{A}' = \tilde{A}HH^{-1}\tilde{A}' = C$$

for any orthogonal matrix H . However, the matrix H influences a matrix of the fourth moments of $X^L(t)$, $K_{ij} = E\{(X^L)_i^2(X^L)_j^2\}$. The orthogonal matrix H and shape parameters for the V_i can be determined to approximate a given sampling matrix $\{K_{ij}\}$ of the fourth moments for the RF distribution (all moments $E\{(X^L)_i^3(X^L)_j\}$ are equal to zero for symmetric distributions). An explicit calculation yields:

$$K_{ij} = E\{(X^L)_i^2(X^L)_j^2\} = 3 \sum_{k=1}^M a_{ik}^2 a_{jk}^2 D_{V_k} + C_{ii}C_{jj} + 2C_{ij}^2, \quad i, j = 1, \dots, N \quad (25)$$

where a_{ik} are the elements of the matrix $A = \tilde{A}H$. An effective method for calculating the matrix H and shape parameters is discussed in Section 3.6 below.

The model provides an exact match of the RF correlation and volatility structures and approximates different shapes and kurtosis of the marginal RF distributions contrary to the Elliptical model. However, there is a significant drawback for this model. Since the stochastic variances V_i are independent, there is a strong normalization effect in any "diagonal" direction. This means that some linear portfolios $X_\Delta(t) = \Delta' X^L(t)$ have almost normal distributions whenever the portfolio Delta, Δ , is not a marginal direction and the number of principal risk factors M is large enough. Described effect presents a real danger, because the non-normal marginal RF distributions may be well-approximated, while the modelled portfolio distributions (contrary to the actual sampling distributions) may be almost normalized and the VaR underestimated.

3.5 A model with Correlated Stochastic Variances

As it was pointed out above, a more general and realistic market model should incorporate the correlated stochastic variances that can correct the deficiencies of both Elliptical model and the model with independent SV for the principal components. The correlated SV structure should allow modeling of some general economic factors as well as idiosyncratic components that drive the SV processes for different risk factors and markets.

The model is defined via stochastic representation of the following form [69], [70]:

$$\begin{aligned} X^{CV}(t) &= AZ^I(t), \quad Z_i^I(t) = \sqrt{V_i(t)}Z_i \\ Z_i &\sim N(0, 1), \quad i = 1, \dots, M \end{aligned} \quad (26)$$

Here Z_i are independent standard normal variables, $V_i(t)$ are the correlated stochastic variance processes with a unit mean for a specified time horizon T_1 . The matrix $A \in \mathbb{R}^{N \times M}$ is defined as in the previous section through the eigenvalue decomposition for the covariance matrix C up to an arbitrary orthogonal transformation $H \in \mathbb{R}^{M \times M}$:

$$C = Cov(X^{CV}(T_1)) = AA' = \tilde{A}\tilde{A}', \quad A = \tilde{A}H, \quad H' = H^{-1}$$

Stochastic variances $V_i(t)$ are correlated to each other due to the following stochastic representation:

$$V_i(t) = \sum_{k=1}^L b_{ik}\xi_k(t), \quad \sum_{k=1}^L b_{ik} = 1, \quad b_{ik} \geq 0, \quad B \in R^{M \times L}, \quad (27)$$

where $\xi_k(t)$ are independent positive increasing Lévy processes with unit mean for the time horizon T_1 and different shape parameters, and B is a constant matrix with non-negative elements. The processes $\xi_k(t)$ are the drivers for the SV processes $V_i(t)$. For example, each driver ξ_k can be a Gamma process or Generalized Gamma process. Linear structure in (27) with $b_{ik} \geq 0$ ensures that $V_i(t)$ are positive increasing Lévy processes. The normalization conditions $E\{\xi_k(T_1)\} = 1$ and $\sum b_{ik} = 1$ ensure, as in Section 3.4, exact recovering of the sampling covariance matrix for the risk factors. It follows, that the vector of stochastic variances $V(T_1)$ has covariance matrix C_V equal to

$$C_V = Cov(V(T_1)) = BD_\xi B', \quad D_\xi = diag(D_{\xi 1}, \dots, D_{\xi L}), \quad D_{\xi k} = Var\{\xi_k(T_1)\} \quad (28)$$

The multivariate Generalized Stochastic Variance (GSV) model (26) - (27) has two levels of correlations. First level defines usual correlations across the risk factors described by the covariance matrix C . Second level defines the correlations across the stochastic variances described by the covariance matrix C_V . The second level of correlations provides a possibility to obtain an approximate, but consistent match of the higher order moments and shape of the RF multivariate distribution. The elliptical model and the model with independent stochastic variances are the special cases of the above GSV model. Elliptical model corresponds to the matrix B being equal to one column with all unit entries, $B = [1, \dots, 1]'$. The model with independent SV corresponds to the case when the matrix B is equal to the identity matrix, $B = I$.

There is no analytical form for the probability density function of the vector $X^{CV}(t)$ even for the Gamma drivers $\xi_k(t)$. However, a characteristic function $\phi_{X^{CV}(t)}(\omega)$ can be calculated as

$$\begin{aligned}\phi_{X^{CV}(t)}(\omega) &= \int_{\xi \in R_+^L} \exp\left(-\frac{1}{2}\omega' A \text{diag}(B\xi)A'\omega\right) p_{\xi(t)}(\xi) d\xi \\ &= \prod_{j=1}^L \int_0^{+\infty} \exp\left(-\frac{\xi_j}{2} \sum_{i=1}^M b_{ij} \left[\sum_{k=1}^N A_{ki}\omega_k\right]^2\right) p_{\xi_j(t)}(\xi_j) d\xi_j\end{aligned}$$

The expression for the characteristic function above is equivalent to

$$\phi_{X^{CV}(t)}(\omega) = \prod_{j=1}^L \int_0^{+\infty} \exp\left(-\frac{\xi_j}{2}\omega' \widehat{C}_j \omega\right) p_{\xi_j(t)}(\xi_j) d\xi_j$$

where \widehat{C}_j , $j = 1, \dots, L$, are certain positive semi-definite matrices. The latter expression for the characteristic function allows for a different interpretation of the GSV model. It shows that the process for the risk factors $X^{CV}(t)$ can be presented as a sum of L independent elliptical Lévy processes. In turn, each of these elliptical processes has a multivariate conditional normal distribution with a covariance matrix proportional to \widehat{C}_j and the corresponding stochastic variance $\xi_j(t)$.

The kurtosis k_Δ of a linear combination of the risk factors $X_\Delta(T_1) = \Delta' X^{CV}(T_1)$ for any given direction Δ can be calculated analytically:

$$\begin{aligned}k_\Delta - 3 &= \frac{E\{X_\Delta^4(T_1)\}}{E\{X_\Delta^2(T_1)\}^2} - 3 = 3\eta' C_V \eta = 3\eta' B D_\xi B' \eta \\ \eta &\in R^M, \quad \eta_k = \frac{(\Delta' A)_k^2}{\|\Delta' A\|^2}, \quad k = 1, \dots, M\end{aligned}\tag{29}$$

The above expression provides a link between the covariance matrix C_V and the kurtosis k_Δ , that characterizes the shape of the RF multivariate distribution for the linear portfolio with Delta equal to Δ . Another useful quantity that clarifies the role of the correlated variances V_i is a standardized matrix of the fourth moments. This matrix, $\{k_{ij}\}$, is a multidimensional analog for kurtosis

$$k_{ij} = \frac{E\{(X^L)_i^2 (X^L)_j^2\}}{E\{(X^L)_i^2\} E\{(X^L)_j^2\}}\tag{30}$$

The matrix $\{k_{ij}\}$ incorporates kurtosis in all marginal and all pair-wise diagonal directions in the original risk factor space. It is expressed as

$$\begin{aligned}k_{ij} - (1 + 2\rho_{ij}^2) &= \sum_{k=1}^M \sum_{l=1}^M \lambda_{ik}^2 \lambda_{jl}^2 \text{Cov}(V_k V_l) + 2 \sum_{k=1}^M \sum_{l=1}^M \lambda_{ik} \lambda_{jk} \lambda_{il} \lambda_{jl} \text{Cov}(V_k V_l) \\ \lambda_{ik} &= \frac{a_{ik}}{\|a_i\|}, \quad \|a_i\|^2 = \sum_{k=1}^M a_{ik}^2\end{aligned}\tag{31}$$

where ρ_{ij} is a correlation between i -th and j -th risk factors. Formulas (29) and (31) clearly indicate that the correlation structure of the SV is embedded into the correlation structure of the fourth moments of the RF distribution. This connection will be used as the base for the GSV model calibration. A number L of the SV drivers can be chosen significantly smaller than a number of stochastic variances M and risk factors N . These SV drivers may be thought as “stochastic activities” for different countries, industries, sectors, etc.

The GSV model with the correlated stochastic variances is, in fact, a general framework. It can incorporate any reasonable processes to represent the SV drivers $\xi_k(t)$, $k = 1, \dots, L$. Some examples of suitable one-dimensional SV driver distributions are: Inverse Gaussian distribution [11], Gamma distribution [74], [67],

Lognormal distribution [29], or considered above class of Generalized Gamma distributions. The GSV model is practical in terms of effective Monte Carlo simulation: it is based on the simulation of one-dimensional SV processes and standard multivariate normal variables.

3.5.1 Example 1. Joint distribution for DEM/USD and JPY/USD FX rates

First example presents a bivariate GSV model applied to the foreign exchange market data. Four bivariate models were examined for DEM/USD and JPY/USD FX rate daily returns: Standard Gaussian model, Elliptical Gamma Variance model, model with independent stochastic variances for PCs, and the model with correlated SV. Fig. 12 and 13 show a 3-D plot and a contour plot of the joint probability density for the historical data and four types of the models considered. All three SV models provide a far better fit than the Gaussian distribution. However, the most convincing fit is provided by the GSV model with the correlated stochastic variances. Marginal distributions for DEM/USD and JPY/USD FX rates have kurtosis 5.2 and 6.9 respectively. Fig. 14 and 15 show that the latter model is able to capture kurtosis and shape of marginal distributions in different directions.

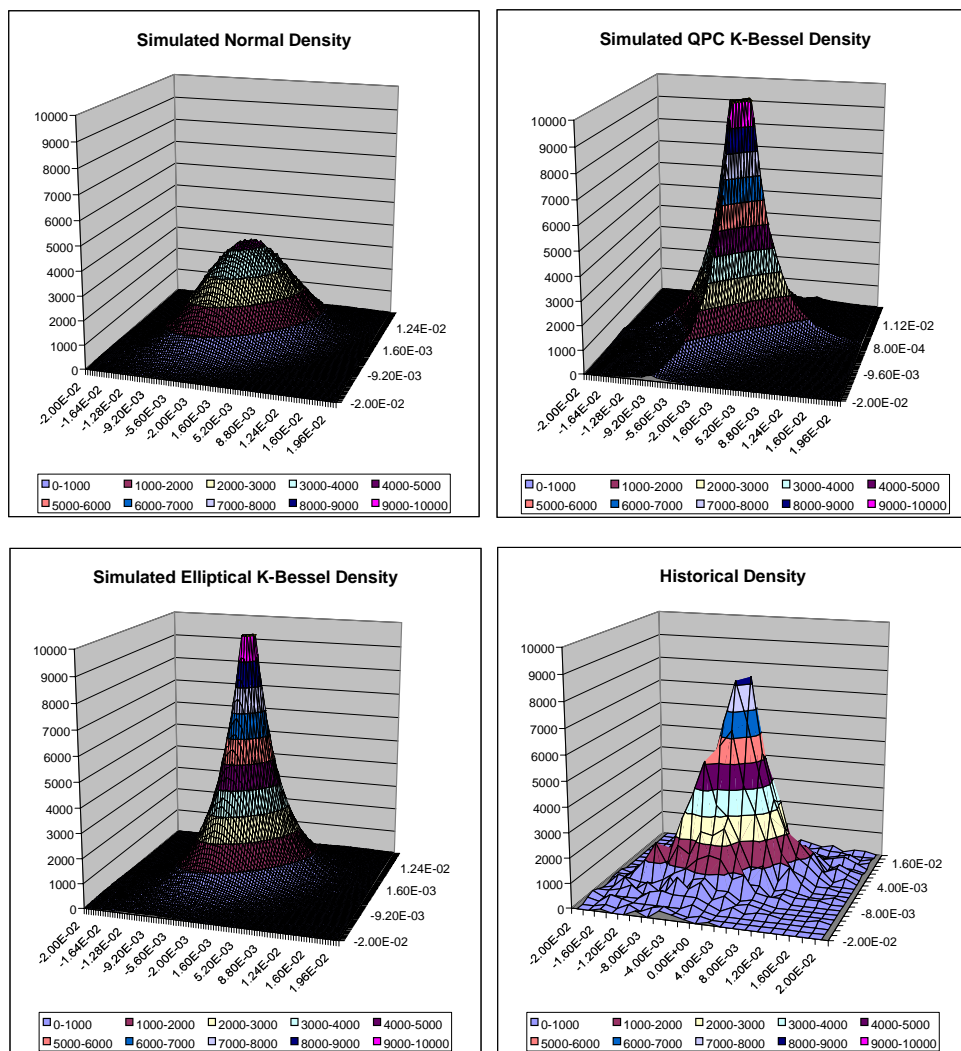


Figure 12: Joint density for DEM/USD and JPY/USD FX rate

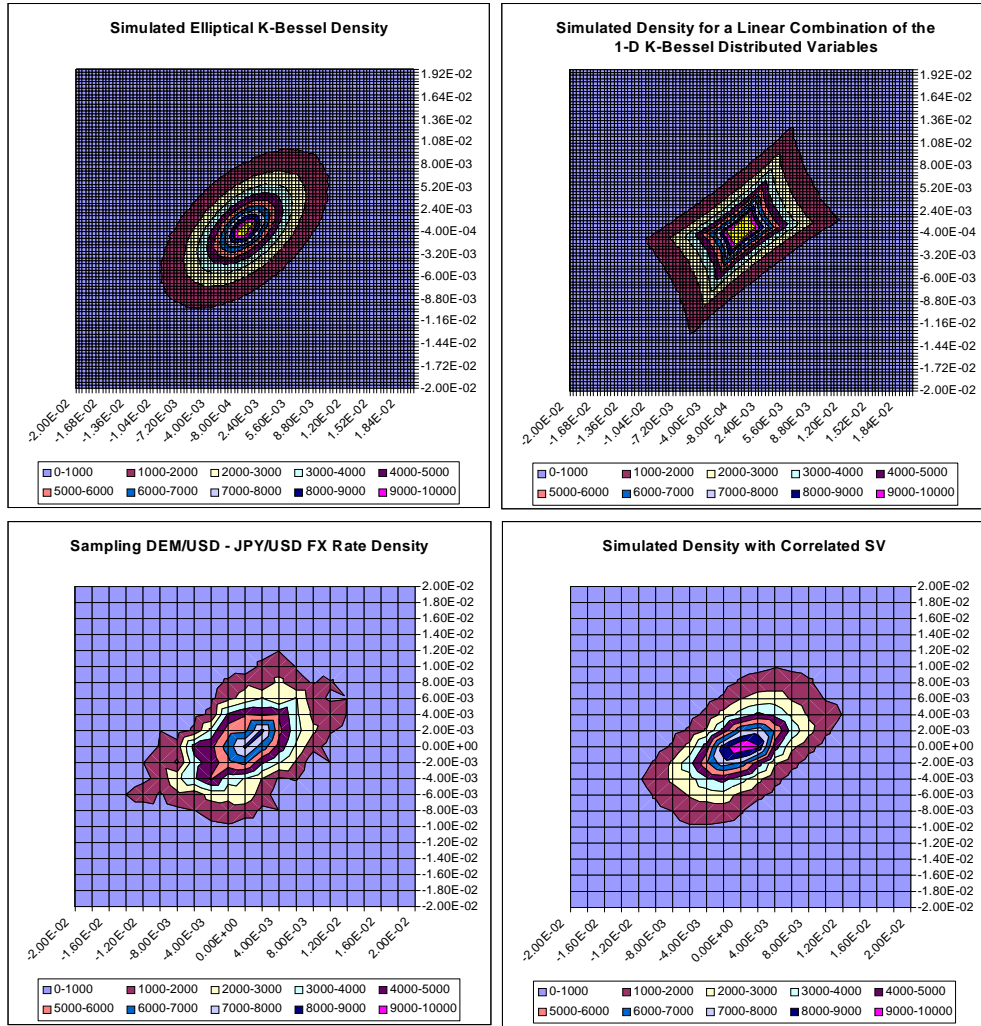


Figure 13: Contour plot for the DEM/USD - JPY/USD FX rate joint density

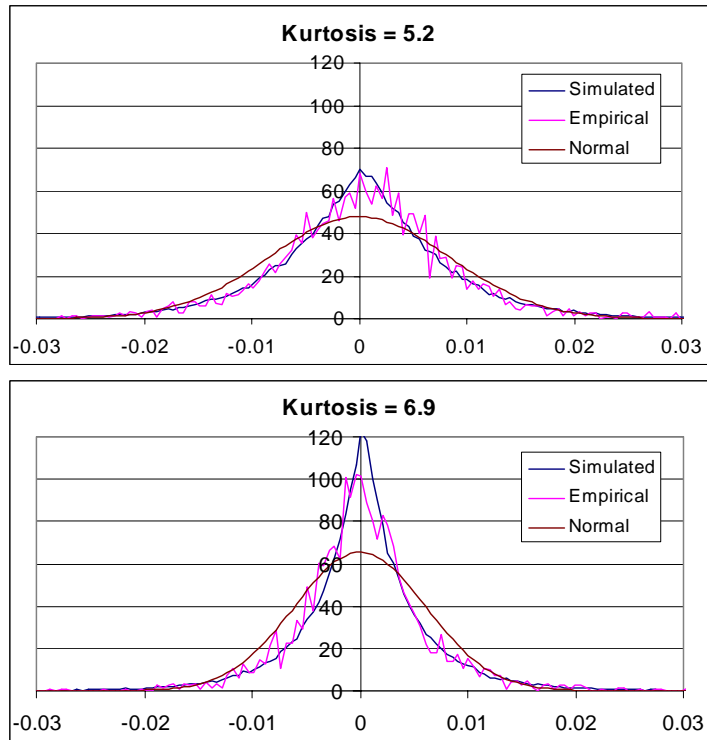


Figure 14: DEM/USD and JPY/USD FX marginal distributions

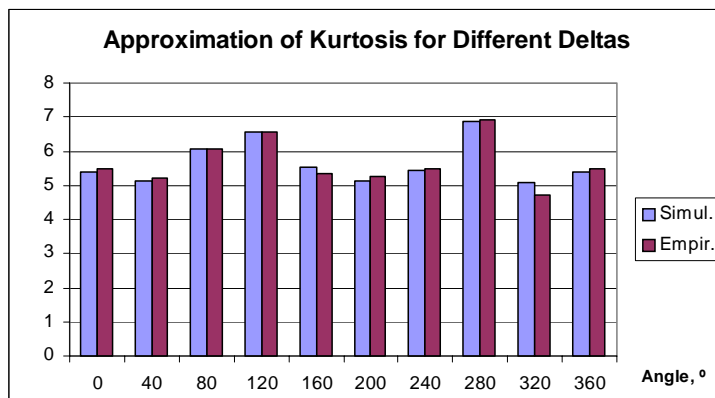


Figure 15: Fit of the kurtosis for different sub-portfolios

3.5.2 Example 2. Twenty risk factors

Second example examines a 20-dimensional GSV model with correlated SV applied to the data from the interest rate, FX rate, and equity markets. The USD and CAD zero interest rate curves each consisting of nine interest rates, CAD/USD FX rate, and S&P500 Index were chosen as a representative set of the risk factors. There were 5 years (1994 - 1999) of daily historical data used for the model calibration (about 1,250 data points). Fig. 16 presents statistical results for principal component analysis and the correlation matrix for squares of the first three PCs. These results indicate that uncorrelated PCs neither are normal nor independent. The first three "largest" PCs per zero curve were used for the GSV model calibration and simulation. Three Gamma distributed drivers ξ_k , $k = 1, 2, 3$, with different shape parameters were utilized to represent each stochastic variance V_i , $i = 1, 2, 3$, for PCs. Therefore, the following values for parameters were assigned: number of risk factors $N = 9 \times 2 + 1 + 1 = 20$, number of principal risk factors $M = 3 \times 2 + 1 + 1 = 8$, number of SV drivers $L = 3$.

The model was calibrated to match kurtosis (in the least squares sense) for all 20 risk factors and kurtosis for 15 additional linear sub-portfolios. Sampling kurtosis varies within a wide range from 5 to 25. Typically, kurtosis for short-term interest rates is much higher than kurtosis for long-term rates. It is seen (Fig. 17) that the GSV model reproduces this typical decreasing kurtosis term structure quite well. It is also seen that the model adequately matches kurtosis of the FX rate and S&P500 Index, as well as kurtosis of different linear sub-portfolios. To compare, the standard multi-dimensional Gaussian model produces a flat kurtosis term structure identically equal to three.

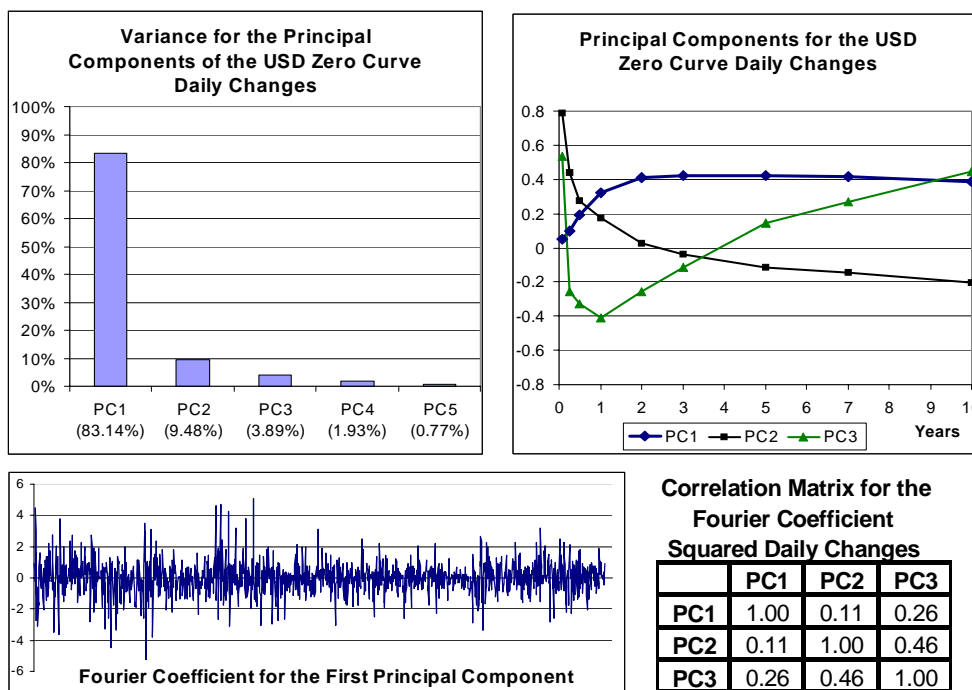


Figure 16: PCA for USD zero curve

3.6 Calibration for the GSV Model

The GSV model is a two-level model that incorporates a traditional variance-covariance structure of the risk factors and novel variance-covariance structure of the RF stochastic variances. The GSV model with correlated SV automatically preserves the RF covariance matrix C . At the second level, it is necessary to calibrate the SV covariance matrix C_V to approximate the fourth moments of the multivariate RF distribution.

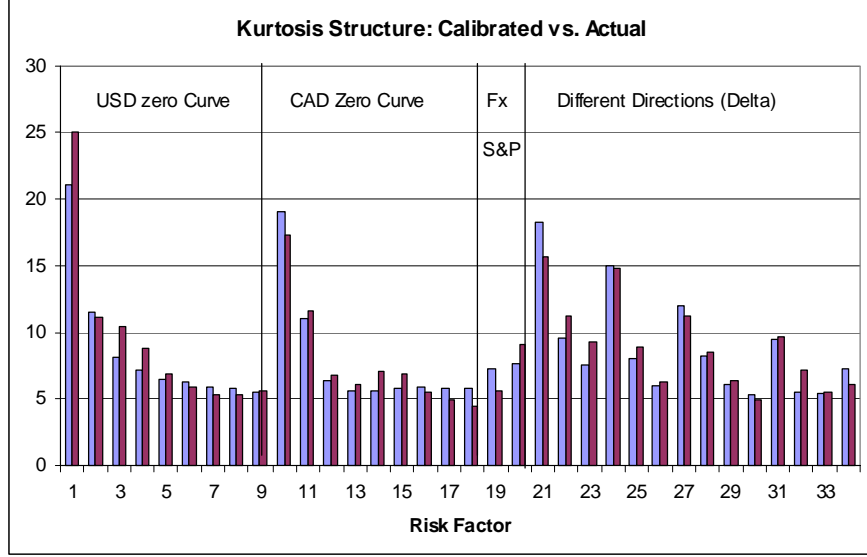


Figure 17: Fit of the kurtosis

The main steps of the model calibration procedure are as follows:

1. Calculate a sampling covariance matrix $C \in \mathbb{R}^{N \times N}$ for a given set of risk factors X . The time window usually used for calibration of the covariance matrix C is about 1 - 2 years. Exponentially weighted averaging or uniform sliding window are the usual methods for the covariance matrix calculation [58].
2. Decompose the sampling covariance matrix C using a standard eigenvalue decomposition procedure and form a matrix $\tilde{A} \in \mathbb{R}^{N \times M}$ from a set of M eigenvectors corresponding to M largest eigenvalues. Number M has to be chosen to recover the matrix C with a required accuracy.
3. Calculate sampling fourth order moments for the risk factors X (the matrix k_{ij} in (30)) and kurtosis k_{Δ} for any preselected set of directions (linear sub-portfolios) $\{\Delta\}$. The time window typically required for calculation of the fourth moments is of the order 5-10 years. This period of observations has to be much longer than the one for the second order moments. This is necessary to incorporate relatively rare extreme events into the calibration. Longer time horizon allows for an adequate approximation of the tails and general shape of the multivariate RF distribution.
4. Calculate matrices H , B , and D_{ξ} using the least squares approach:

$$\sum_i w_i \left(\widehat{k}_{\Delta_i} - k_{\Delta_i}(H, B, D_{\xi}) \right)^2 + \sum_i \sum_j w_{ij} \left(\widehat{k}_{ij}^e - k_{ij}^e(H, B, D_{\xi}) \right)^2 \rightarrow \min_{H, B, D_{\xi}} \quad (32)$$

where w_i , and w_{ij} are some predefined weights (these weights may be chosen depending on the importance of particular risk factors and sub-portfolios), \widehat{k}_{Δ_i} is the sampling kurtosis for the direction Δ_i , $k_{\Delta_i}(H, B, D_{\xi})$ is the analytical estimate (29), \widehat{k}_{ij}^e is the sampling matrix of the fourth moments, and $k_{ij}^e(H, B, D_{\xi})$ is its analytical estimate (31).

The minimization problem above is a subject to constraints imposed on the matrices H , B , D_{ξ} . The most difficult condition to satisfy is orthogonality of the matrix H . It follows from the analysis of expressions (29) and (31) that $M \times M$ orthogonal matrix H can be constructed as a product of $M \times (M - 1)/2$ elementary rotation matrices [108]. It can be shown that for the problem (29), a representation for the orthogonal matrix H does not require reflections. The diagonal matrix D_{ξ} is subject to simple non-negativity constraints. The matrix B is subject to constraints (27). Hence, the non-linear optimization problem (32) can be reformulated with respect to $M \times (M - 1)/2$ angles φ_m for the elementary rotation matrices with simple constraints $-\pi \leq \varphi_m \leq \pi$ and elements of the matrices B and D_{ξ} with mentioned above simple constraints.

5. If the Gamma Variance model for the SV drivers ξ_k is adopted, the diagonal matrix D_ξ and conditions $E\{\xi_k(T_1)\} = 1$ determine the shape and scale parameters α_k and β_k in (6). For the GGV model, the powers $\nu_k \in \mathbb{R}^1$ have to be additionally specified. As a practical approach, the following methodology has been adopted: a set of parameters $\{\nu_k\}$ is fixed in such a way that it covers a reasonably wide range of values ν_k . For example, the set of ν_k can be chosen as

$$\{\nu_k\} = \{-2, -1, +1, +2\}$$

This choice is justified by the fact that the SV drivers ξ_k with negative values of ν_k will produce the RF probability density function with heavy polynomial tails. On the other hand, positive values of ν_k can produce the RF distributions with semi-heavy exponential and sub-exponential tails, but with unbounded peaks at the origin. However, it is quite possible that a more flexible and adjustable structure for the set of parameters $\{\nu_k\}$ is more beneficial for the model calibration.

Acknowledgement. *Authors thank C. Albanese, O. Barndorff-Nielsen, D. Duffie, P. Embrechts, H. Geman, D. Madan, J. Nolan, and, especially, M. Taqqu for interesting discussions and useful comments related to presented models.*

References

- [1] Abramowitz, M. and Stegun, I. A. (eds). (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, DC.
- [2] Albanese, C., Levin, A. and Ching-Ming Chao, J. (1997). Bayesian Value at Risk, back-testing and calibration. Working paper, Bank of Montreal - University of Toronto, Toronto.
- [3] Albanese, C., Jaimungal, S. and Rubisov, D. H. (2001). A jump model with binomial volatility. Working Paper, University of Toronto - Math-Point, Toronto.
- [4] Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). Wiley, New York.
- [5] Ané, T. and Geman, H. (1999). Stochastic volatility and transaction time: an activity-based volatility estimator. *The Journal of Risk*, **2**, 57-69.
- [6] Ané, T. and Geman, H. (2000). Order flow, transaction clock and normality of asset returns. *The Journal of Finance*, **55**, 2259-2284.
- [7] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, No. 3, 203-228.
- [8] Bachelier, L. (1900). Theory of Speculation. Ph.D. Thesis, English Transl. In: Cootner, P. H. (ed.). *The Random Character of Stock Market Prices*. (1964). MIT Press, Cambridge, MA.
- [9] Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proc. of the Royal Society London, A*, **353**, 401-419.
- [10] Barndorff-Nielsen, O. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, **5**, 151-157.
- [11] Barndorff-Nielsen, O. (1997). Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, **24**, 1-14.
- [12] Barndorff-Nielsen, O. (1998). Processes of Normal Inverse Gaussian Type. *Finance and Stochastics*, **2**, 41-68.

- [13] Barndorff-Nielsen, O. and Levendorskii, S. (2001). Feller processes of Normal Inverse Gaussian type. *Quantitative Finance*, **1**, 318-331.
- [14] Barndorff-Nielsen, O. and Pérez-Abreu, V. (2000). Multivariate type \mathbf{G} distributions. Working Paper, Centre for Mathematical Physics and Stochastics, University of Aarhus, Aarhus.
- [15] Barndorff-Nielsen, O. and Shephard, N. (2000). Non-Gaussian OU based models and some of their use in Financial Economics. Working Paper. Centre for Mathematical Physics and Stochastics, University of Aarhus, Aarhus.
- [16] Barndorff-Nielsen, O. and Shephard, N. (2000). Modelling by Lévy processes for Financial Econometrics. Working Paper. Centre for Mathematical Physics and Stochastics, University of Aarhus, Aarhus.
- [17] Basle Committee on Banking Supervision (1997). *International Convergence of Capital Measurements and Capital Standards*. July 1988, amended in April 1997, (<http://www.bis.org>).
- [18] Basle Committee on Banking Supervision (1996). *Supervisory Framework for the Use of Backtesting in Conjunction with the Internal Models Approach to Market Risk Capital Requirements*. January, (<http://www.bis.org>).
- [19] Bates, D. (1991). The crash of '87: was it expected? The evidence from option markets. *Journal of Finance*, **46**, 1009-1044.
- [20] Bates, D. (1996). Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, **9**, 69-107.
- [21] Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press, Cambridge.
- [22] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, 637-654.
- [23] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307-327.
- [24] Bollerslev, T., Engle, R. and Nelson, D. (1994). ARCH Models. In: Engle, R. and McFadden, D. (eds). *Handbook of Econometrics*, 4. Elsevier Sci. B.V., Amsterdam.
- [25] Bouchaud, J-P. and Potters, M. (2000). *Theory of Financial Risks*. Cambridge University Press, Cambridge.
- [26] Boyarchenko, S. and Levendorskii, S. (2000). Option pricing for truncated Lévy processes. *International Journal of Theoretical and Applied Finance*, **3**, 549-552.
- [27] Buchen, P. and Kelly, M. (1996). The Maximum Entropy distribution of an asset inferred from option prices. *Journal of Financial and Quantitative Analysis*, **31**, 143-159.
- [28] Carr, P., Geman, H., Madan, D. and Yor, M. (2000). The fine structure of asset returns: an empirical investigation. Working paper, University of Maryland, College Park.
- [29] Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, **41**, 135-155.
- [30] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1**, 223-236.
- [31] Cont, R., Potters M. and Bouchaud J-P. (1997). Scaling in stock market data: stable laws and beyond. In: Dubrulle, B., Graner, F. and Sornette, D. (eds). *Scale Invariance and Beyond*, Proceedings of the CNRS Workshop on Scale Invariance, Springer.
- [32] Crouhy, M., Galai, D. and Mark, R. (2001). *Risk Management*. McGraw-Hill, New York.

- [33] Duffie, D. and Pan, J. (1997). An overview of Value at Risk. *The Journal of Derivatives*, **4**, 7-49.
- [34] Duffie, D. and Pan, J. (2001). Analytical Value-at-Risk with jumps and credit risk. *Finance and Stochastics*, **5**, 155-180.
- [35] Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in Finance. *Bernoulli*, **1**, 281-299.
- [36] Eberlein, E., Keller, U. and Prause, K. (1998). New insights into smile, mispricing and Value at Risk: the Hyperbolic model. *Journal of Business*, **71**, 371-405.
- [37] Eberlein, E. and Raible, S. (1999). Term structure models driven by general Lévy processes. *Mathematical Finance*, **9**, No. 1, 31-53.
- [38] Embrechts, P., McNeil, A. and Straumann, D. (1999). Correlation and dependency in Risk Management: properties and pitfalls. ETH Preprint, Zürich.
- [39] Fama, E. (1965). The behavior of stock market prices. *Journal of Business*, **38**, 34-105.
- [40] Fang, K., Kotz, S. and Ng, K. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.
- [41] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. Vol. 2. John Wiley & Sons, New York.
- [42] Feuerverger, A. and McDunnough, P. (1981). On efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society, B*, **43**, No. 1, 20-27.
- [43] Fishman, G. S. (1996). *Monte Carlo: Concepts, Algorithms and Applications*. Springer-Verlag, New York.
- [44] Geman, H. and Ané, T. (1996). Stochastic subordination. *Risk*, **9**, September, 12-16.
- [45] Geman, H., Madan, D. and Yor, M. (1998). Asset prices are Brownian motion: only in business time. Working paper, University of Maryland, College Park.
- [46] Geman, H., Madan, D. and Yor, M. (1999). Time changes for Lévy processes. Working paper, University of Maryland, College Park.
- [47] Grosswald, E. (1976). The Student t-distribution of any degree of freedom is infinitely divisible. *Zeit. Wahrsch. Verw. Gebiete*, **36**, 103-109.
- [48] Heston, S. (1993). A closed-form solution for options with stochastic volatility, with applications to bond and currency options. *Review of Financial Studies*, **6**, 327-344.
- [49] Hull, J. C. (1999). *Options, Futures, and Other Derivatives* (4th edition). Prentice Hall, Upper Saddle River, NJ.
- [50] Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, **42**, 281-300.
- [51] Hull, J. and White, A. (1998). Value at Risk when daily changes in market variables are not normally distributed. *The Journal of Derivatives*, **5**, 9-19.
- [52] Ismail, M. (1977). Bessel functions and the infinite divisibility of the Student t-distribution. *Ann. Probab.*, **5**, 582-585.
- [53] Ismail, M. and Kelker, D. (1979). Special functions, Stieltjes transforms and infinite divisibility. *SIAM Journal Math. Analysis*, **10**, 884-901.
- [54] Janicki, A. and Weron, A. (1994). *Simulation and Chaotic Behavior of α -Stable Stochastic Processes*. Marcel Dekker, New York.

- [55] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Vol. 1 (2nd edition). John Wiley & Sons, New York.
- [56] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Vol. 2 (2nd edition). John Wiley & Sons, New York.
- [57] Jorion, P. (2001). *Value-at-Risk: The New Benchmark for Managing Financial Risk* (3rd edition). McGraw-Hill, New York.
- [58] JP Morgan (1996). *RiskMetricsTM Technical Document* (4th edition). JP Morgan, New York.
- [59] Kagan, A. M., Linnik, Yu. V. and Rao, C.R. (1973). *Characterization Problems in Mathematical Statistics*. John Wiley & Sons, New York.
- [60] Kelker, D. (1971). Infinite divisibility and variance mixtures of the normal distributions. *The Annals of Mathematical Statistics*, **42**, 802-808.
- [61] Khindanova, I., Rachev, S. and Schwartz, E. (2000). Stable modelling of Value at Risk. Working Paper, University of California, Santa Barbara.
- [62] Konikov, M. and Madan, D. (2000). Pricing options of all strikes and maturities using a generalization of the VG model. Working paper, University of Maryland, College Park.
- [63] Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate distributions*. Vol. 1 (2nd edition). John Wiley & Sons, New York.
- [64] Kotz, S., Kozubowski, T. and Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser, Boston.
- [65] Kou, S. G. (2000). A jump diffusion model for option pricing with three properties: leptokurtic feature, volatility smile, and analytical tractability. Working Paper, Columbia University, New York.
- [66] Kreinin, A. and Levin, A. (2000). Robust Monte Carlo simulation for approximate covariance matrices and VaR analysis. In: Uryasev, S. (ed.). *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Boston, 166-178.
- [67] Levin, A. and Tchernitser, A. (1999). Multifactor stochastic variance Value-at-Risk model. Presentation at the Workshop "Probability in Finance. Toronto, January 26-30, 1999". The Fields Institute, Toronto.
- [68] Levin, A. and Tchernitser, A. (1999). Multifactor Gamma stochastic variance Value-at-Risk model. Presentation at the Conference "Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics. Washington, DC, June 3-5, 1999". American University, Washington, DC.
- [69] Levin, A. and Tchernitser, A. (2000). A class of multifactor stochastic variance VaR models: Maximum Entropy approach and jump processes. Presentation at the First World Congress of the Bachelier Finance Society, Paris, June 28 - July 1, 2000.
- [70] Levin, A. and Tchernitser, A. (2000). Stochastic volatility and jump Lévy processes in Value-at-Risk modelling. Presentation at the Risk Conference "Math Week'2000. London, November 27 - December 1, 2000".
- [71] Lukacs, E. (1970). *Characteristic Functions*. Griffin, London.
- [72] Madan, D. (1999). Purely discontinuous asset price processes. Working paper, University of Maryland, College Park.
- [73] Madan, D. and Milne, F. (1991). Option pricing with VG martingale components. *Mathematical Finance*, **1**, No. 4, 39-55.

- [74] Madan, D. and Seneta, E. (1990). The Variance Gamma (VG) model for share market returns. *Journal of Business*, **63**, 511-524.
- [75] Maejima, M. and Rosinski, J. (2000). Type G distributions on R^d . Working Paper, Keio University.
- [76] Mandelbrot, B. (1960). The Pareto-Lévy law and the distribution of income. *Internat. Econ. Rev.*, **1**, 79-106.
- [77] Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, **36**, 394-419.
- [78] Mandelbrot, B. (1997). *Fractal and Scaling in Finance: Discontinuity, Concentration, Risk*. Springer, New York.
- [79] Mantegna, R. N. and Stanley, H. E. (1994). Stochastic process with ultraslow convergence to a Gaussian: the Truncated Lévy Flight. *Physical Reviews Letters*, **73**, 2946-2949.
- [80] Mantegna, R. N. and Stanley, H. E. (2000). *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge.
- [81] Marinelli, C., Rachev, S. and Roll, R. (1999). Subordinated exchange rate models: evidence for heavy tailed distributions and long-range dependence. Working Paper, University of Padova.
- [82] Mausser, H. and Rosen, D. (2000). Managing risk with Expected Shortfall. In: Uryasev, S. (ed.). *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Boston, 204-225.
- [83] McCulloch, J. (1978). Continuous time processes with stable increments. *Journal of Business*, **36**, 394-419.
- [84] McCulloch, J. (1996). Financial applications of stable distributions. In: Maddala, G. and Rao, C. (eds). *Statistical Methods in Finance*, ElsevierScience, Amsterdam, 393-425.
- [85] Melino, A. and Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, **45**, 239-265.
- [86] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, **3**, 125-144.
- [87] Merton, R. (1990). *Continuous-Time Finance*. Blackwell, Cambridge.
- [88] Mittnik, S. and Rachev, S. (1989). Stable distributions for asset returns. *Appl. Math. Lett.*, **213**, 301-304.
- [89] Müller, U., Dacorogna, M. and Pictet, O. (1998). Heavy tails in high-frequency financial data. In: Adler, R., Feldman, R. and Taqqu, M. (eds). *A Practical Guide to Heavy Tails*. Birkhäuser, Boston.
- [90] Nolan, J. (1998). Multivariate stable distributions: approximation, estimation, simulation and identification. In: Adler, R., Feldman, R. and Taqqu, M. (eds). *A Practical Guide to Heavy Tails*. Birkhäuser, Boston, 509-525.
- [91] Olver, F. (1974). *Introduction to Asymptotics and Special Functions*. Academic Press, New York
- [92] Osborne, M. (1959). Brownian motion in the stock market. *Operations Research*, **7**, 145-173.
- [93] Platen, E. (1999). A Minimal Share Market model with stochastic volatility. Research Paper 21, University of Technology, Sydney.
- [94] Prudnikov, A., Brychkov, Yu. and Marichev, O. (1986). *Integrals and Series*. Vol. 1. Gordon & Breach Sci. Publ., New York.

- [95] Rachev, S. and SenGupta, A. (1993). Laplace-Weibull mixtures for modelling price changes. *Management Science*, **39**, 1029-1038.
- [96] Rachev, S. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. John Wiley & Sons, Chichester, New York.
- [97] Rosinski, J. (1991). On a class of infinitely divisible processes represented as mixtures of Gaussian processes. In: Cambanis, S., Samorodnitsky, G. and Taqqu, M. (eds). *Stable Processes and Related Topics*, Birkhäuser, Boston, 27-41.
- [98] Rosinski, J. (2001). Series representations of Lévy processes from the perspective of point processes In: Barndorff-Nielsen, O., Mikosch, T. and Resnick, S. (eds). *Lévy Processes - Theory and Applications*, Birkhäuser, Boston, 27-41.
- [99] Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, New York.
- [100] Samuelson, P. (1965). Rational theory of warrant pricing. *Industrial Management Review*, **6**, 13-32.
- [101] Schoenberg, I. (1938). Metric spaces and completely monotonic functions. *Annals of Math.*, **39**, 811-841.
- [102] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge.
- [103] Sornette, D., Simonetti, P. and Andersen, J. V. (2000). ϕ^q -field theory for portfolio optimization: "fat tails" and non-linear correlations. *Physics Repts*, **335**, No. 2, 19-92.
- [104] Stein, E. and Stein, J. (1991). Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, **4**, 727-752.
- [105] Steutel, F. (1970). Preservation of infinite divisibility under mixing and related topics. *Math. Centre Tracts*, No. **33**, Math. Centrum, Amsterdam.
- [106] Steutel, F. (1973). Some recent results in infinite divisibility. *Stochastic Processes and Their Applications*, **1**, 125-143.
- [107] Wiggins, J. (1987). Option values under stochastic volatility: theory and empirical estimates. *Journal of Financial Economics*, **19**, 351-372.
- [108] Wilkinson, J. H. and Reinsch, C. (1971). *Linear Algebra*. Vol. 2 of Handbook for Automatic Computation. Springer-Verlag, New York.
- [109] Willinger, W., Taqqu, M. and Teverovsky, V. (1999). Stock market prices and long-range dependence. *Finance and Stochastics*, **3**, 1-13.