Estimating Correlated Jumps and Stochastic Volatilities

Jiří Witzany

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00   Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Estimating Correlated Jumps and Stochastic Volatilities

Jiří Witzany*

*University of Economics, Prague
E-mail: jiri.witzany@vse.cz

November 2011

Abstract:

We formulate a bivariate stochastic volatility jump-diffusion model with correlated jumps and volatilities. An MCMC Metropolis-Hastings sampling algorithm is proposed to estimate the model's parameters and latent state variables (jumps and stochastic volatilities) given observed returns. The methodology is successfully tested on several artificially generated bivariate time series and then on the two most important Czech domestic financial market time series of the FX (CZK/EUR) and stock (PX index) returns. Four bivariate models with and without jumps and/or stochastic volatility are compared using the deviance information criterion (DIC) confirming importance of incorporation of jumps and stochastic volatility into the model.

Keywords: jump-diffusion, stochastic volatility, MCMC, Value at Risk, Monte Carlo

JEL: C11, C15, G1

1. Introduction

A number of empirical studies confirmed that financial asset returns are not normal and exhibit fat tails (leptokurtic distribution). Many models going beyond the standard geometrical Brownian diffusion model have been proposed in order to accommodate the empirical facts. The most prominent are jump-diffusion models (see e.g. Cont, Tankov, 2004 for a review), models with stochastic volatility (see Shephard, 2004 for selected papers), or models combining both features, i.e. jump-diffusion models with stochastic volatility, or even models with jumps in volatility. Modeling of portfolio returns or valuation of various multi-asset derivatives requires generalizing of the models into multivariate setting.

We are going to consider a bivariate jump-diffusion model with stochastic volatilities incorporating possible correlation of jump occurrence, jump size, and of the stochastic volatilities. The main goal of the paper is to propose an MCMC estimation method that will be tested on artificial and real world data. We have chosen the two most important Czech financial markets time series, namely a series of FX (CZK/EUR) exchange rates and of the stock market (PX index) returns. Joint modeling of the two series might be important for and an asset manager exposed to the Czech stock market and the exchange rate, or in case of certain derivatives (e.g. quanto) modeling. In both cases, the goal is to model the distribution of future returns of a portfolio exposed to both factors.

For example, inspecting the development of the markets in 2004-2011 (Figure 4 - Figure 6) it seems that the returns of the time series exhibit many jumps and periods of low volatility that are followed by periods of high volatility or vice versa. Although the sample correlation of returns turns out to be almost zero, the volatilities appear to move in the same direction (Figure 6). The questions are: How do the two markets jump and how volatile are the stochastic volatilities? Is the non-Gaussian behavior explained by jumps or rather by stochastic volatility? Are the jumps in the two markets correlated in terms of occurrence and size? And moreover, are the stochastic volatilities correlated?

Our estimation methodology is based on the MCMC (Markov Chain Monte Carlo) approach following Jacquier et al. (2007) and Johannes, Polson (2009). The first break-through application of the Bayesian methods for the analysis of stochastic volatility models has been made in Jacquier et al. (1994). The authors applied the MCMC algorithm to estimate

parameters as well the latent states of the stochastic volatility model on the US stock return data. The estimation method is shown to outperform other known estimation approaches, as the Method of Moments or the Quasi-Maximum Likelihood Estimator. Since then extensive research on application of Bayesian methods on stochastic volatility models have appeared (see e.g. Shephard, 2004). Johannes, Kumar (1999) estimate state dependent jump models (on US stock data) in which arrival intensity and jump sizes depend on a given state variable including lagged jumps. Eraker et al. (2003) examine stochastic volatility models incorporating jumps in returns and volatility using US stock indices returns. Eraker (2004) utilizes in addition stock index option date and allows the diffusion and volatility processes being correlated.

The contribution of this paper is in specification and estimation of a bivariate model with correlated stochastic volatilities and jumps. Johannes, Polson (2009) consider a multivariate version of Merton's jump-diffusion model where jumps occur at the same times for all processes and the jump sizes have a multivariate distribution. Inspecting the times of probable jumps from the MCMC estimation of two univariate jump-diffusion models applied to the two considered return series we have noted that the jump times overlap only partially. So, in our specification we have two correlated Poisson processes and correlated jump sizes (if the jumps occur at the same time). Regarding the correlation of stochastic volatilities we may again firstly inspect the mean stochastic volatilities given by the MCMC algorithm applied to the two processes separately. Since the stochastic volatility residuals do not indicate any significant correlation but show a strong correlation in levels of the stochastic volatilities we propose a bivariate stochastic volatility jump-diffusion model with possible Granger causality between the stochastic volatilities as our full model specification (see Asai et al., 2006 or Yu, Meyer, 2006 for an overview of multivariate stochastic volatility models). The computationally difficult estimation of the model is based on our generalization of the method proposed firstly in Jacquier et al. (1994).

The proposed bivariate jump-diffusion with stochastic volatilities will be firstly tested on artificially generated data. The goal of the test is to demonstrate that the estimation procedure yields acceptable results with respect to generating parameters, in particular that it is able to identify existence of jumps, stochastic volatilities, and their correlations.

The methodology will be finally applied to estimate and compare four bivariate models on the FX and stock returns series, specifically: the ordinary diffusion model, the jump-diffusion model, the diffusion model with stochastic volatility, and the jump-diffusion model with stochastic volatility. Performance of the models will be compared using the deviance information criterion (DIC) generalizing according to Spiegelhalter et al (2002) the Akaike information criterion (AIC) and that is applicable to models with a large number of latent state variables. Importance of the choice of an appropriate model will be illustrated calculating VaR for various time horizons and confidence levels.

2. Methodology

In this section we are going to give a brief overview of the relevant stochastic models. We also outline the key elements of the MCMC methods and their implementation in case of the stochastic models under consideration. Finally we shortly describe the classical and Bayesian VaR estimation methodology.

2.1 Stochastic Asset Price Models

The most traditional (related to the Black-Scholes formula) continuous-time financial model is the geometric Brownian motion described by the stochastic differential equation (SDE)

$$(1) \qquad dS = \mu S dt + \sigma S dz$$

where $S(t)$ is an asset price, μ its drift, σ its volatility, and dz the Wiener process increment (see e.g. Shreve, 2004). The equation can be simplified applying the Ito's lemma into a generalized Wiener process equation

$$(2) \qquad d(\log S) = (\mu - \sigma^2 / 2) dt + \sigma dz .$$

The left hand side of (2) can be interpreted as the log return over a time interval of the length dt. In practice, a time (Euler) discretization is used in order estimate the parameters from an observed financial time series or in order to generate future returns. In case of the equation (2) the Euler discretization takes the simple form

$$(3) \qquad r_i = \tilde{\mu} + \tilde{\sigma}\breve{\eta}_i, \ \breve{\eta}_i \sim N(0,1) ,$$

where $r_i = \ln S_i / S_{i-1}$ is the log return over a regular time interval of length Δt, $\tilde{\mu} = (\mu - \sigma^2 / 2)\Delta t$, and $\tilde{\sigma} = \sigma\sqrt{\Delta t}$. According to (3), observed returns should have a normal distribution. However, many studies demonstrate that the returns have, typically, a leptokurtic distribution, in particular fat tails. Consequently, the geometric Brownian motion is proposed to be generalized in various directions, in particular, allowing for jumps (see Cont, Tankov, 2004) and stochastic volatility (see Shephard, 2004).

The jump-diffusion SDE can be written in the log-return form as

$$(4) \qquad d(\log S) = (\mu - \sigma^2 / 2 - \lambda\mu_J)dt + \sigma dz + dJ ,$$

where the jump term, $dJ = ZdN$, has a normally distributed $N(\mu_J, \sigma_J)$ jump-size component and a component given by the Poisson counting process N with intensity λ. Essentially, this component adds a mass to the tails of the returns distribution. We consider a time discretization where at most one jump can happen over a single time step:

$$(5) \qquad \begin{aligned} r_i &= \tilde{\mu} + \tilde{\sigma}\breve{\eta}_i + Z_i J_i \\ \breve{\eta}_i &\sim N(0,1), Z_i \sim N(\mu_J, \sigma_J), J_i \sim \text{Bern}(\lambda). \end{aligned}$$

According to Eraker et al (2003) this assumption does not introduce any bias in the parameter estimates.

Stochastic volatility models allow variance $V = \sigma^2$ or log-variance to evolve according to an SDE, i.e. the constant volatility σ in (2) or (4) is replaced by the stochastic volatility $\sqrt{V}$. For example the Heston's (1993) model sets

$$dV = \kappa(\theta - V)dt + \sigma_V \sqrt{V}\,dz_V .$$

Following Jacquier et al (1994) and Johannes, Polson (2009) will we rather consider the log-variance SDE:

(6) $\quad d\log V = \kappa(\theta - \log V)dt + \sigma_V dz_V .$

In the discrete setting, with $h_i = \log V_i$, the equation takes the form of AR(1) model:

$$h_i - h_{i-1} = \kappa(\theta - h_{i-1})\Delta t + \sigma_V \sqrt{\Delta t}\,\check{n}_i^V , \quad \check{n}_i^V \sim N(0,1) .$$

Rearranging the terms the equation can be written as

(7) $\quad h_i = \alpha + \beta h_{i-1} + \gamma \check{n}_i^V ,$

where $\alpha = \kappa\theta\Delta t$, $\beta = 1 - \kappa\Delta t$, and $\gamma = \sigma_V \sqrt{\Delta t}$.

If two or more financial assets is to be model than possible correlations needs to be considered. For example if S_1 and S_2 are two asset prices following (1) with appropriate indexes we have to admit possibility of correlated dz_1 and dz_2. If the prices follow the jump-diffusion process (4) then we have to admit a correlation between dJ_1 and dJ_2. Finally, in case of stochastic volatility (6) we might consider a correlation between dz_1^V and dz_2^V, but also a mutual correlation between dz_1 and dz_1^V, and the correlation between dz_2 and dz_2^V. Due to the mean-reverting form of (6) we should also consider a possible correlation between $\log V_1$ and $\log V_2$ that could be captured introducing $\ln V_2$ into the SDE for $\ln V_1$ and vice versa. A nonzero coefficient is then interpreted as Granger causality form one asset variance to another.

An analyst that needs to model the distribution of future returns of a portfolio or of a derivative payoff depending on two or even more assets stands in front of a difficult task: to choose an optimal model and at the same time to estimate in a feasible way its parameters using historical or currently observable data. Due to increasing complexity of the models we will focus on the MCMC Bayesian estimation and model comparison approach.

2.2 Markov Chain Monte Carlo (MCMC)

The Bayesian MCMC sampling algorithm has become a strong and frequently used tool to estimate complex models with multidimensional parameter vectors, including latent state variables. Examples are financial stochastic models with jumps, stochastic volatility processes, models with complex correlation structure, or switching-regime processes. For a more complete treatment of MCMC methods and applications we refer reader for example to Johannes, Polson (2009), Rachev et al. (2004), or Lynch (2010).

MCMC provides a method of sampling from multivariate densities that are not easy to sample from directly, by breaking these densities down into more manageable univariate or lower dimensional multivariate densities. To estimate a vector of unknown parameters $\Theta = (\theta_1,...,\theta_k)$ from a given dataset, where we are able to write down the Bayesian marginal densities $p(\theta_j \mid \theta_i, i \neq j, \text{data})$ but not the multivariate density $p(\Theta \mid \text{data})$, the MCMC *Gibbs sampler* works according to the following generic procedure:

0. Assign a vector of initial values to $\Theta^0 = (\theta_1^0,...,\theta_k^0)$ and set $j = 0$.

1. Set $j = j + 1$.

2. Sample $\theta_1^j \sim p(\theta_1 \mid \theta_2^{j-1},...,\theta_k^{j-1}, \text{data})$.

3. Sample $\theta_2^j \sim p(\theta_2 \mid \theta_1^j, \theta_3^{j-1},...,\theta_k^{j-1}, \text{data})$.

$\vdots$

k+1. Sample $\theta_k^j \sim p(\theta_k \mid \theta_1^j, \theta_2^j,...,\theta_{k-1}^j, \text{data})$ and return to step 1.

According to the Clifford-Hammersley theorem the conditional distributions $p(\theta_j \mid \theta_i, i \neq j, \text{data})$ fully characterize the joint distribution $p(\Theta \mid \text{data})$ and moreover under mild conditions the Gibbs sampler distribution converges to the target joint distribution (Johannes, Polson, 2003).

The conditional probabilities are typically obtained applying the Bayes theorem to the likelihood function and a prior density, e.g.

$$(8) \qquad p(\theta_1 \mid \theta_2^{j-1},...,\theta_k^{j-1}, \text{data}) \propto L(\text{data} \mid \theta_1, \theta_2^{j-1},...,\theta_k^{j-1}) \cdot \text{prior}(\theta_1 \mid \theta_2^{j-1},...,\theta_k^{j-1}).$$

We will often use uninformative priors, i.e. $\text{prior}(\theta_i) \propto 1$ and assume that the parameters are independent. In order to apply the Gibbs sampler the right hand side of the proportional relationship needs to be normalized, i.e. we need to be able to integrate the right hand side with respect to θ_1 conditional on $\theta_2^{j-1},...,\theta_k^{j-1}$.

Useful Gibbs sampling distributions are univariate or multivariate normal, Inverse Gamma[1] or Wishart, and the Beta distribution. If $\mathbf{y} = \langle y_1,...,y_T \rangle$ is an observed series assuming that iid $y_i \sim N(\mu, \sigma^2)$ with unknown parameters μ and σ then

---

[1] Inverse gamma probability distribution density function with the shape parameter α and scale parameter β is given by $IG(x; \alpha, \beta) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\beta / x)$ where Γ is the Gamma function. The mean of x is

$$p(\mu \mid \mathbf{y}, \sigma) \propto L(\mathbf{y} \mid \mu, \sigma) p(\mu) = \prod_{i=1}^{T} \varphi(y_i; \mu, \sigma) \propto \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^T \exp\left(-\frac{\sum(y_i - \mu)^2}{2\sigma^2}\right)$$

(9)

$$\propto \exp\left(-\frac{T\mu^2 - 2\mu\sum y_i}{2\sigma^2}\right) \propto \varphi\left(\mu; \frac{\sum y_i}{T}, \frac{\sigma}{\sqrt{T}}\right)$$

using the uninformative prior $p(\mu) \propto 1$ and

$$p(\sigma^2 \mid \mathbf{y}, \mu) \propto L(\mathbf{y} \mid \mu, \sigma^2) \cdot p(\sigma^2) = \frac{1}{\sigma^2} \prod_{i=1}^{T} \varphi(y_i; \mu, \sigma)$$

(10)

$$\propto \left(\sigma^2\right)^{-\frac{T}{2}-1} \exp\left(-\frac{\sum(y_i - \mu)^2}{2\sigma^2}\right) \propto IG\left(\sigma^2; \frac{T}{2}, \frac{\sum(y_i - \mu)^2}{2}\right)$$

using the prior $p(\sigma^2) \propto 1/\sigma^2$ equivalent to the uninformative prior $p(\log \sigma^2) \propto 1$. Hence the Bayesian distributions for μ and σ can be obtained by the Gibbs sampler iterating (9) and (10). The prior distributions are often specified in order to improve convergence but not to influence (significantly) the final results, typically a wide normal distribution conjugate prior distribution for μ and a flat inverse gamma distribution for σ^2.

If the series is multivariate normal then the distributions are generalized to multivariate normal and inverse Wishart (Lynch, 2007). A multivariate discrete-time diffusion process (3) is in fact equivalent to a multivariate normal return series model with iid $\mathbf{r}_i \sim N(\boldsymbol{\mu}, \Sigma)$, where $\mathbf{r}_i = (r_{i,1}, ..., r_{i,m})'$ is the vector of returns on m assets observed at time i, $\boldsymbol{\mu}$ is a vector of means, and Σ a covariance matrix. The marginal distributions are

(11) $\quad p(\boldsymbol{\mu} \mid \mathbf{r}, \Sigma) = \varphi\left(\boldsymbol{\mu}; \frac{1}{T}\sum_{i=1}^{T} \mathbf{r}_i, \frac{1}{T}\Sigma\right)$ and

(12) $\quad p(\Sigma \mid \mathbf{r}, \boldsymbol{\mu}) = IW(\Sigma; T, S) \propto |\Sigma|^{-(T+m+1)/2} \exp\left(-\frac{1}{2}\operatorname{tr}(S\Sigma^{-1})\right)$

Where $IW(\Sigma; T, S)$ denotes the inverse Wishart distribution, $S = \sum_{i=1}^{T} (\mathbf{r}_i - \boldsymbol{\mu})'(\mathbf{r}_i - \boldsymbol{\mu})$ is the scale matrix, and the improper prior $p(\Sigma) \propto |\Sigma|^{-\frac{m+1}{2}}$ analogous to the univariate case has been used.

$\mu = \dfrac{\beta}{\alpha - 1}$ for $\alpha > 1$ and the variance is $\sigma^2 = \dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$. Alternatively, given μ and σ^2 we

get $\alpha = \dfrac{\mu^2}{\sigma^2} + 2$ and $\beta = \mu\left(\dfrac{\mu^2}{\sigma^2} + 1\right)$.

If $\mathbf{b} = \langle b_1, \ldots, b_T \rangle$ is a binary series where $b_i \sim \mathrm{Bern}(\lambda)$ iid, then λ can be sampled using the beta[2] distribution:

$$(13) \qquad p(\lambda \mid \mathbf{b}) \propto L(\mathbf{b} \mid \lambda) \cdot p(\lambda) \propto \prod_{i=1}^{T} \lambda^{b_i} (1-\lambda)^{1-b_i} \propto \lambda^{n}(1-\lambda)^{T-n} \propto \mathrm{Beta}(\lambda; n+1, T-n+1)$$

with the uninformative prior $p(\lambda) \propto 1$. Generally, the beta distribution $\mathrm{Beta}(x; \alpha, \beta)$ would be a conjugate prior where α and β can be interpreted as prior "successes" and "failures."

If the integration on the right hand side of (8) is not analytically possible (which will be also our case) then the *Metropolis-Hastings algorithm* can be used. It is based on a rejection sampling algorithm. For example in step 2 the idea is firstly to sample a new proposal value of θ_1^j and then accept it or reject it (i.e. reset $\theta_1^j := \theta_1^{j-1}$) with appropriate probability so that, intuitively speaking, we rather move to the parameter estimates with higher corresponding likelihood values.

Specifically, step 1 is replaced with a two step procedure:

1. A. Draw θ_1^j from a proposal density $q(\theta_1 \mid \theta_1^{j-1}, \theta_2^{j-1}, \ldots, \theta_k^{j-1}, \mathrm{data})$,

   B. Accept θ_1^j with the probability $\alpha = \min(R, 1)$, where

$$(14) \qquad R = \frac{p\left(\theta_1^j \mid \theta_2^{j-1}, \ldots, \theta_k^{j-1}, \mathrm{data}\right) q\left(\theta_1^{j-1} \mid \theta_1^j, \theta_2^{j-1}, \ldots, \theta_k^{j-1}, \mathrm{data}\right)}{p\left(\theta_1^{j-1} \mid \theta_2^{j-1}, \ldots, \theta_k^{j-1}, \mathrm{data}\right) q\left(\theta_1^j \mid \theta_1^{j-1}, \theta_2^{j-1}, \ldots, \theta_k^{j-1}, \mathrm{data}\right)} .$$

In practice the step 1B is implemented by sampling a $u \sim U(0,1)$ from the uniform distribution and accepting θ_1^j if and only if $u < R$.

It is again shown (see Johannes, Polson, 2003) that under certain mild conditions the limiting distribution is the joint distribution $p(\Theta \mid \mathrm{data})$ of the parameter vector. Note that the limiting distribution does not depend on the proposal density, or on the starting parameter values. The proposal density and the initial estimates only make the algorithm more-or-less efficient.

A popular proposal density is the random walk, i.e. sampling by

$$(15) \qquad \theta_1^j \sim \theta_1^{j-1} + N(0, c) .$$

The algorithm is then called *Random Walk Metropolis-Hastings*. The proposal density is in this case symmetric, i.e. the probability of going from θ_1^{j-1} to θ_1^j is the same as the probability of going from θ_1^j to θ_1^{j-1} (fixing the other parameters), and so the second part of the fraction in

---

[2] The pdf of the beta distribution is $\mathrm{Beta}(x; \alpha, \beta) = \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ for $0 < x < 1$. The mean can be expressed as $\mu = \dfrac{\alpha}{\alpha+\beta}$ and the variance $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

the formula (14) for α in step 1B cancels out. Consequently, assuming non-informative prior, the acceptance or rejection is driven just by the likelihood ratio

$$R = \frac{L\left(\text{data} \mid \theta_1^j, \theta_2^{j-1}, ..., \theta_k^{j-1}\right)}{L\left(\text{data} \mid \theta_1^{j-1}, \theta_2^{j-1}, ..., \theta_k^{j-1}\right)}.$$

Another popular approach we shall use is the *Independence Sampling Metropolis-Hastings algorithm* where the proposal density $q\left(\theta_1^j\right)$ does not depend on θ_1^{j-1} (given the other parameters). The acceptance probability ratio (14) is slightly simplified but note that the proposal densities do not cancel out. In order to achieve efficiency the shape of the proposal density q should be close to the shape of the target density p which is known only up to a normalizing constant.

Typically, estimating complex stochastic models, we need to estimate the parameter vector with a few model parameters Θ, and a vector with a large number of state variables X (typically proportional to the number of observations). We know that $p(\Theta, X \mid \text{data}) \propto p(\text{data} \mid \Theta, X) \cdot p(X, \Theta)$ and so we may estimate iteratively the parameters and the state variables:

$$p(\Theta \mid X, \text{data}) \propto p(\text{data} \mid \Theta, X) \cdot p(X \mid \Theta) \cdot p(\Theta),$$
$$p(X \mid \Theta, \text{data}) \propto p(\text{data} \mid \Theta, X) \cdot p(\Theta \mid X) \cdot p(X).$$

The parameters and state variables are sampled step by step, or in blocks, often combining Gibbs and Metropolis-Hastings sampling.

2.3 Univariate jump-diffusion stochastic volatility model

The main goal of this section is to propose an MCMC estimation algorithm for a bivariate jump diffusion stochastic volatility model. However, it will be useful to outline firstly the univariate jump-diffusion model and the extended model with stochastic volatility. It is then easier to define the sampling steps of the bivariate model estimation algorithm. Moreover, given two financial series, it is useful to estimate firstly the univariate models separately on the two series. Since the latent variables, i.e. jumps and stochastic volatilities, are also estimated, the output can be used to make a preliminary analysis of correlations between the jumps and stochastic volatilities.

Let us firstly consider the *discrete-time jump diffusion model* (5) *with constant volatility.* Given the sequence of observed returns data=$\{r_i; i = 1, ..., T\}$, the parameters and latent state variables to be estimated are: $\mu, \sigma, \lambda, \mu_J, \sigma_J, \mathbf{Z}, \mathbf{J}$.

In this case we may use the pure Gibbs MCMC algorithm:

1. Sample reasonable initial values $\mu^{(0)}, \sigma^{(0)}, \lambda^{(0)}, \mu_J^{(0)}, \sigma_J^{(0)}, \mathbf{Z}^{(0)}, \mathbf{J}^{(0)}$

2. Sample $Z_i^{(g)} \propto \varphi(Z; \mu_J^{(g-1)}, \sigma_J^{(g-1)})$ if $J_i^{(g-1)} = 0$, and

$$Z_i^{(g)} \propto \varphi(r_i; \mu^{(g-1)} + Z, \sigma^{(g-1)})\varphi(Z; \mu_J^{(g-1)}, \sigma_J^{(g-1)}) \text{ if } J_i^{(g-1)} = 1.$$

3. Sample $J_i^{(g)} \in \{0,1\}$, $\Pr[J = 1] = p_1 / (p_0 + p_1)$, where

$$p_0 = \varphi(r_i; \mu^{(g-1)}, \sigma^{(g-1)}) \; (1 - \lambda^{(g-1)}),$$
$$p_1 = \varphi(r_i; \mu^{(g-1)} + Z, \sigma^{(g-1)})\lambda^{(g-1)}$$

4. Sample $\mu^{(g)}, \sigma^{(g)}$ based on the normally distributed time series $r_i - Z_i^{(g)} J_i^{(g)}$ according to (9) and (10).

5. Sample $\lambda^{(0)}$ based on Bernoulli distributed $J_i^{(g)}$ binary time series according to (13).

6. Sample $\mu_J^{(g)}, \sigma_J^{(g)}$ based on the normally distributed time series $Z_i^{(g)}$ according to (9) and (10).

Secondly let us consider *a jump-diffusion model with stochastic volatility* following the equation (6). The discrete time specification is:

$$r_i = \mu + \sqrt{V_i}\,\check{\eta}_i + Z_i J_i$$

(16) $\log V_i = \alpha + \beta \log V_{i-1} + \gamma \check{\eta}_i^V$

$$\check{\eta}_i, \check{\eta}_i^V \sim N(0,1), \; Z_i \sim N(\mu_J, \sigma_J), \; J_i \sim \text{Bern}(\lambda), \text{iid}$$

In this case we need to estimate not only the latent state variables $\mathbf{Z}, \mathbf{J}$ but also the vector of latent stochastic variances $\mathbf{V}$. The MCMC estimation unfortunately requires application of the Metropolis-Hastings since the conditional distribution for the variance V_i (conditional on the other variances and parameters) is not a known one. It follows from (16) and the Bayes Theorem that:

(17) $p(V_i \mid \mathbf{V}_{(-i)}, \Theta, \mathbf{r}, \mathbf{Z}, \mathbf{J}) \propto p(r_i \mid V_i, \Theta, Z_i, J_i)p(V_i \mid V_{i-1}, \Theta)p(V_{i+1} \mid V_i, \Theta).$

Here the first part of the right hand side of (17) is inverse gamma in V_i:

(18) $p(r_i \mid V_i, \Theta, Z_i, J_i) = \varphi\left(r_i; \mu + Z_i J_i, V_i\right) \propto V_i^{-0.5} \exp\left(-0.5(r_i - \mu - Z_i J_i)^2 / V_i\right).$

But the remaining two factors are lognormal[3] in V_i:

$$\log V_i \;\square\; \varphi(\log V_i; \alpha + \beta \log V_{i-1}, \gamma), \text{ i.e.}$$

$$p(V_i \mid V_{i-1}, \Theta) \propto V_i^{-1} \exp\left(-(\log V_i - \alpha - \beta \log V_{i-1})^2 / (2\gamma^2)\right),$$

---

[3] Lognormal probability density fiction with parameters μ and σ is given by

$LN(x; \mu, \sigma) = \dfrac{1}{x\sigma\sqrt{2\pi}} \exp\left(-0.5(\log x - \mu)^2 / \sigma^2\right)$. It is useful to note that the mean of x is

$\exp(\mu + \sigma^2 / 2)$ and the variance is $(\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2)$.

and similarly

$$\log V_{i+1} \sim \varphi(\log V_{i+1}; \alpha + \beta \log V_i, \gamma), \text{ i.e.}$$

$$p(V_{i+1} | V_i, \Theta) \propto \exp\left(-(\log V_{i+1} - \alpha - \beta \log V_i)^2 / (2\gamma^2)\right) \text{ in terms of } V_i.$$

It is easy to verify that the product of the two lognormal distributions is proportional to the lognormal distribution with the corresponding normal distribution mean and standard deviation:

$$(19) \quad \begin{aligned} \mu_i &= \left(\alpha(1-\beta) + \beta(\log V_{i+1} + \log V_{i-1})\right) / \left(1 + \beta^2\right), \\ \sigma &= \gamma / \sqrt{1 + \beta^2}. \end{aligned}$$

In order to obtain a proposal distribution Jacquier et al (1994) suggest replacing the lognormal distribution with an inverse gamma distribution fitting the first two moments. It is confirmed empirically that the choice of a proposal distribution with a shape closely mimicking the original distribution is of key importance since high dimensionality of the variance state variable vector makes convergence of the MCMC algorithm difficult.

The product of two inverse gamma distribution density functions is an inverse gamma distribution density function, hence combining the inverse gamma distribution (18) and the fitted inverse gamma distribution we finally obtain the proposal density function:

$$(20) \quad \begin{aligned} q(V_i | \mathbf{V}_{(-i)}, \Theta, \mathbf{r}, \mathbf{Z}, \mathbf{J}) &= IG\left(V_i; \phi + 0, 5, (\phi - 1)\exp(\mu_i + 0.5\sigma^2) + 0.5(r_i - \mu - Z_i J_i)^2\right), \\ \text{where } \phi &= \frac{1 - 2\exp(\sigma^2)}{1 - \exp(\sigma^2)}. \end{aligned}$$

The proposal density is used in the Metropolis/Hasting algorithm within a new block, e.g. following the step 3 in the MCMC procedure for the jump-diffusion processes. This bloc updates all the variances $V_i, i = 1, ..., T$. For V_1 and V_T the formula (19) needs to modified slightly since V_0 and V_T are not known. The diffusion volatility σ is obviously replaced by the square root of the latest estimate of the variance V_i and we also need to add a new MCMC step for the AR(1) coefficients α, β and γ (for example following the step that updates $\mathbf{V}$). The coefficients α, β can be sampled with a bivariate normal distribution and γ with the inverse gamma distribution. The extended MCMC algorithm is in detail described as follows:

1. Sample reasonable initial values $\mu^{(0)}, \lambda^{(0)}, \mu_J^{(0)}, \sigma_J^{(0)}, \alpha^{(0)}, \beta^{(0)}, \gamma^{(0)}, \mathbf{V}^{(0)}, \mathbf{Z}^{(0)}, \mathbf{J}^{(0)}$

2. Sample $Z_i^{(g)} \propto \varphi(Z; \mu_J^{(g-1)}, \sigma_J^{(g-1)})$ if $J_i^{(g-1)} = 0$, and

   $$Z_i^{(g)} \propto \varphi(r_i; \mu^{(g-1)} + Z, \sqrt{V_i^{(g-1)}})\varphi(Z; \mu_J^{(g-1)}, \sigma_J^{(g-1)}) \text{ if } J_i^{(g-1)} = 1.$$

3. Sample $J_i^{(g)} \in \{0,1\}$, $\Pr[J = 1] = p_1 / (p_0 + p_1)$, where

$$
\begin{aligned}
(21) \quad & p_0 = \varphi(r_i; \mu^{(g-1)}, \sqrt{V_i^{(g-1)}}) \, (1-\lambda^{(g-1)}), \\
& p_1 = \varphi(r_i; \mu^{(g-1)} + Z, \sqrt{V_i^{(g-1)}}) \lambda^{(g-1)}
\end{aligned}
$$

4. Sample new stochastic variances $V_i^{(g)}$ for $i = 1,...,T$ using Metrolis-Hastings (14) with the proposal density given by (20)

5. Sample new stochastic volatility autoregression coefficients $\alpha^{(g)}, \beta^{(g)}, \gamma^{(g)}$ from $h_i = \log V_i^{(g)}$ for $i = 1,...,T$ using the Bayesian linear regression model (Lynch, 2007):

$$
\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{Xy}, \ \hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}, \text{ where } \mathbf{X} = \begin{pmatrix} 1....1 \\ h_1...h_{T-1} \end{pmatrix}', \mathbf{y} = (h_2...h_T)',
$$

$$
(22) \quad (\gamma^{(g)})^2 \propto IG\left(\frac{n-2}{2}, \frac{\hat{\mathbf{e}}'\hat{\mathbf{e}}}{2}\right),
$$

$$
\left(\alpha^{(g)}, \beta^{(g)}\right)' \propto \varphi\left((\alpha, \beta)'; \hat{\boldsymbol{\beta}}, (\gamma^{(g)})^2 (\mathbf{X'X})^{-1}\right).
$$

6. Sample $\mu^{(g)}$ based on the normally distributed time series $r_i - Z_i^{(g)} J_i^{(g)}$ with variances $V_i^{(g)}$, i.e.

$$
p\left(\mu^{(g)} \mid \mathbf{r}, \mathbf{Z}^{(g)}, \mathbf{J}^{(g)}, \mathbf{V}^{(g)}\right) \propto \varphi\left(\mu; \sum_{i=1}^{T} \frac{r_i - J_i^{(g)} Z_i^{(g)}}{V_i^{(g)}} \middle/ \sum_{i=1}^{T} \frac{1}{V_i^{(g)}}, 1 \middle/ \sum_{i=1}^{T} \frac{1}{V_i^{(g)}}\right).
$$

7. Sample $\lambda^{(0)}$ based on Bernoulli distributed $J_i^{(g)}$ binary time series according to (13).

8. Sample $\mu_J^{(g)}, \sigma_J^{(g)}$ based on the normally distributed time series $Z_i^{(g)}$ according to (9) and (10).

2.4 Bivariate jump-diffusion stochastic volatility model

Our ultimate goal is to study relationship between returns, jumps, and volatilities of two related financial series. First, we may estimate independently the parameters and latent variables $\mathbf{J}_1, \mathbf{Z}_1, \mathbf{V}_1$ and $\mathbf{J}_2, \mathbf{Z}_2, \mathbf{V}_2$ of (16) for two given series of returns $\mathbf{r}_1$ and $\mathbf{r}_2$. Since we also get estimates of the latent variables, i.e. jump times, jump sizes, and variances, we may inspect their relationship. For example we may analyze the overlap of probable jump times of the two processes, i.e. of the sets $\{i \mid P[J_{1,i} = 1] > 0.5\}$ and $\{i \mid P[J_{2,i} = 1] > 0.5\}$. Similarly the relationship between mean volatilities can be analyzed. Since $h_{1,i} = \log V_{1,i}$ is specified as a normal variable by the model, the mean volatility should be expressed by $\bar{\sigma}_{1,i} = \exp(\bar{h}_{1,i} / 2)$ where $\bar{h}_{1,i}$ is the MCMC estimation of the mean of $\log V_{1,i}$. The correlation between the series $\langle \bar{\sigma}_{1,i}; i = 1,...,T \rangle$ and $\langle \bar{\sigma}_{2,i}; i = 1,...,T \rangle$ (residuals or levels) indicates what the correlation between the stochastic volatilities might be.

However, to give a consistent answer we need to specify a full bivariate stochastic volatility jump diffusion model:

$$r_{1,i} = \mu_1 + \sqrt{V_{1,i}}\,\check{n}_{1,i} + Z_{1,i}J_{1,i}$$

$$\log V_{1,i} = \alpha_1 + \beta_{11}\log V_{1,i-1} + \beta_{12}\log V_{2,i-1} + \gamma_1 \check{n}_{1,i}^{V}$$

$$r_{2,i} = \mu_2 + \sqrt{V_{2,i}}\,\check{n}_{2,i} + Z_{2,i}J_{2,i}$$

(23) $\quad \log V_{2,i} = \alpha_2 + \beta_{21}\log V_{1,i-1} + \beta_{22}\log V_{2,i-1} + \gamma_2 \check{n}_{2,i}^{V}$

$$\begin{pmatrix} \check{n}_{1,i} \\ \check{n}_{2,i} \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \quad \begin{pmatrix} \check{n}_{1,i}^{V} \\ \check{n}_{2,i}^{V} \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & \rho_V \\ \rho_V & 1 \end{pmatrix}\right)$$

$$\begin{pmatrix} Z_{1,i} \\ Z_{2,i} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{J1} \\ \mu_{J2} \end{pmatrix}, \begin{pmatrix} \sigma_{J1}^2 & \rho_Z\sigma_{J1}\sigma_{J2} \\ \rho_Z\sigma_{J1}\sigma_{J2} & \sigma_{J2}^2 \end{pmatrix}\right)$$

$$J_{1,i} \sim \text{Bern}(\lambda_1), J_{2,i} \sim \text{Bern}(\lambda_2), \ \text{corr}\left(J_{1,i}, J_{2,i}\right) = \rho_J$$

The model does not take into account a possible correlation between the diffusion and log-variance residuals, but it incorporate possible Granger causality between the two log-variance processes. For example if $\beta_{12} > 0$ then high level of $\log V_2$ (Granger) causes $\log V_1$ to become larger. Our preliminary analysis of the data we intend to study indeed indicates that the volatility residuals $\check{n}_{1,i}^{V}$ and $\check{n}_{2,i}^{V}$ are not correlated, but $\log V_{1,i}$ and $\log V_{2,i}$ are correlated. We also compare the results of the stochastic volatility jump-diffusion model against the restricted bivariate jump-diffusion model with constant volatilities inspecting in particular the jump probabilities and the jump correlations estimated with and without stochastic volatilities.

Correlated jumps

In order to incorporate correlated jumps and jump sizes into the MCMC algorithm given in Section 2.3 we need to modify the steps 2 and 3 given as follows:

Sample the new jump sizes (omitting the upper indices g and $g-1$)

$$p\left(Z_i \mid r_i, J_i, \Theta\right) \propto p\left(r_i \mid Z_i, J_i, \Theta\right)p\left(Z_i \mid \Theta\right)$$

$$\propto \exp\left(-0.5\left(y_i{'}\Sigma_i y_i + (Z_i - \mu_Z)'\Sigma_J(Z_i - \mu_Z)\right)\right)$$

$$\propto \exp\left(-0.5(Z_i - m_i)'V_i^{-1}(Z_i - m_i)\right)$$

where

$$r_i = \begin{pmatrix} r_{1,i} \\ r_{2,i} \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, Z_i = \begin{pmatrix} Z_{1,i} \\ Z_{2,i} \end{pmatrix}, J_i = \begin{pmatrix} J_{1,i} & 0 \\ 0 & J_{2,i} \end{pmatrix}, y_i = r_i - \mu - J_i Z_i,$$

(24) $\quad \Sigma_i = \begin{pmatrix} V_{1,i} & \rho\sqrt{V_{1,i}V_{2,i}} \\ \rho\sqrt{V_{1,i}V_{2,i}} & V_{2,i} \end{pmatrix}, \Sigma_J = \begin{pmatrix} \sigma_{J1}^2 & \rho_Z\sigma_{J1}\sigma_{J2} \\ \rho_Z\sigma_{J1}\sigma_{J2} & \sigma_{J2}^2 \end{pmatrix}$, and

$$V_i = \left(J_i\Sigma_i^{-1}J_i + \Sigma_J^{-1}\right)^{-1}, m_i = V_i\left(J_i\Sigma_i^{-1}(r_i - \mu_i) + \Sigma_J^{-1}\right).$$

12

Next we need to resample the correlated jump times. Since by definition $\rho_J = \dfrac{p_{11} - \lambda_1 \lambda_2}{\sqrt{\lambda_1 \lambda_2 (1-\lambda_1)(1-\lambda_2)}}$ where $p_{11} = \Pr[J_{1,i} = 1 \,\&\, J_{2,i} = 1]$ we have

$$\Pr[J_{1,i} = 1 \mid J_{2,i} = 1] = \lambda_1 + \rho_J \frac{\sqrt{\lambda_1 (1-\lambda_1)(1-\lambda_2)}}{\sqrt{\lambda_2}} \text{ and}$$

$$\Pr[J_{1,i} = 1 \mid J_{2,i} = 0] = \lambda_1 - \rho_J \frac{\sqrt{\lambda_1 \lambda_2 (1-\lambda_1)}}{\sqrt{1-\lambda_2}}.$$

Similarly to (21) (again omitting the upper indices g and $g-1$) we set

$$p_0 = \varphi\left(r_i ; \left\langle \mu_1, \mu_2 - Z_{2,i} J_{2,i} \right\rangle', \Sigma_i \right) \Pr[J_{1,i} = 0 \mid J_{2,i}],$$

$$p_1 = \varphi\left(r_i ; \left\langle \mu_1 - Z_{1,i}, \mu_2 - Z_{2,i} J_{2,i} \right\rangle', \Sigma_i \right) \Pr[J_{1,i} = 1 \mid J_{2,i}],$$

and sample $J_{1,i} \in \{0,1\}$, $\Pr[J_{1,i} = 1] = p_1 / (p_0 + p_1)$.

Finally we need to add two additional steps re-sampling the correlations ρ_J and ρ_Z. The jump size correlation can be in fact sampled in one step with jump size volatilities Σ_J using the inverse Wishart distribution (12) as the series Z_i is assumed to be bivariate normal. Regarding ρ_J we use the random walk Metropolis-Hastings step and the relationship

$$p(\rho_J \mid \mathbf{J}_1, \mathbf{J}_2, \lambda_1, \lambda_2) \propto L(\mathbf{J}_1, \mathbf{J}_2 \mid \rho_J, \lambda_1, \lambda_2) = \prod_{i=1}^{T} p(J_{1,i}, J_{2,i}), \text{ where}$$

$$p(1,1) = \lambda_1 \lambda_2 + \rho_J \sqrt{\lambda_1 \lambda_2 (1-\lambda_1)(1-\lambda_2)},\ p(1,0) = \lambda_1 - p(1,1),$$
$$p(0,1) = \lambda_2 - p(1,1),\ p(0,0) = 1 - \lambda_1 - p(0,1).$$

Regarding the diffusion residuals correlation coefficient ρ it can be easily re-sampled using the random walk Metropolis-Hastings:

$$p\left(\rho \mid \mathbf{r}_1, \mathbf{J}_1, \mathbf{Z}_1, \mathbf{V}_1, \mathbf{r}_2, \mathbf{J}_2, \mathbf{Z}_2, \mathbf{V}_2, \Theta \right) \propto \prod_{i=1}^{T} \varphi\left(r_i ; \mu, \Sigma_i \right)$$

with the bivariate normal density function and the notation (24).

Correlated stochastic volatilities

Correlation between stochastic volatilities can be captured by correlated residuals and by correlated levels of the log volatilities expressed in the VAR(1) model (23).

In order to simplify the notation set $h_{k,i} = \log V_{k,i}$, $(k = 1, 2;\ i = 1,...,T)$. Re-sampling the variance $V_{1,i}$ conditional on the other latent variables and parameters MCMC needs to take into

account its relationship to $r_{1,i}$ and to $V_{1,i-1}, V_{2,i-1}$ given by the VAR equation for $h_{1,i}$, but also to $V_{1,i+1}$ and $V_{2,i+1}$ given by the equations for $h_{1,i+1}$ and $h_{2,i+2}$. Applying the Bayes theorem:

$$p(V_{1,i} \mid \mathbf{V}_{1,(-i)}, \mathbf{V}_2, \Theta, \mathbf{r}_1, \mathbf{Z}_1, \mathbf{J}_1)$$
$$\propto p(r_{1,i} \mid V_{1,i}, \Theta, Z_{1,i}, J_{1,i}) p(V_{1,i} \mid V_{1,i-1}, V_{2,i-1}, \Theta) p(V_{1,i+1} \mid V_{1,i}, V_{2,i}, \Theta) p(V_{2,i+1} \mid V_{1,i}, V_{2,i}, \Theta).$$

The first probability distribution on the right hand side is again inverse gamma in $V_{1,i}$, while the remaining ones are lognormal. The product of the three lognormal densities is lognormal with the corresponding normal distribution mean and variance:

(25)
$$\mu_i = \frac{\gamma_2^2(\alpha_1 + \beta_{11}h_{1,i-1} + \beta_{12}h_{2,i-1}) + \gamma_2^2\beta_{11}(h_{1,i+1} - \alpha_1 - \beta_{12}h_{2,i}) + \gamma_1^2\beta_{21}(h_{2,i+1} - \alpha_2 - \beta_{22}h_{2,i})}{\gamma_2^2(1+\beta_{11}^2) + \gamma_1^2\beta_{21}^2},$$

$$\sigma^2 = \frac{\gamma_1^2\gamma_2^2}{\gamma_2^2(1+\beta_{11}^2) + \gamma_1^2\beta_{21}^2}.$$

The Metropolis-Hastings proposal density for updated $V_{1,i}$ is then given by (20). The second process variance $V_{2,i}$ is resampled analogously.

Finally, given $\mathbf{V}_1$ and $\mathbf{V}_2$ (omitting the upper index g) we need to update the coefficients $\alpha_1, \beta_{11}, \beta_{12}, \gamma_1$ and $\alpha_2, \beta_{21}, \beta_{22}, \gamma_2$. We shall use the Bayesian linear regression model as in the univariate case:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{Xy}, \ \hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}, \text{ where } \mathbf{X} = \begin{pmatrix} 1....1 \\ h_{1,1}...h_{1,T-1} \\ h_{2,1}...h_{2,T-1} \end{pmatrix}', \ \mathbf{y} = \left(h_{1,2}...h_{1,T}\right)',$$

$$\gamma_1^2 \propto IG\left(\frac{n-3}{2}, \frac{\hat{\mathbf{e}}'\hat{\mathbf{e}}}{2}\right),$$

$$\left(\alpha_1, \beta_{11}, \beta_{12}\right)' \propto \varphi\left(\left(\alpha_1, \beta_{11}, \beta_{12}\right)'; \hat{\boldsymbol{\beta}}, \gamma_1^2(\mathbf{X'X})^{-1}\right).$$

Similarly we proceed for $\alpha_2, \beta_{21}, \beta_{22}, \gamma_2$.

2.5 Model Comparison and Value at Risk Estimations

In order to compare the different models we will use the deviance information criterion (DIC) of Spielgelhalter et al (2002). It is shown to generalize the Akaike information criterion (AIC) that is not appropriate to compare stochastic volatility or jump-diffusion models (Yu, Meyer, 2006). In order to calculate AIC one needs to specify the number of free parameters. In case of the stochastic volatility and jump-diffusion models the number of additional latent variables estimated has an order of T. On the other hand, since the variables are not independent it is not clear what should be the number of free variables. Likewise AIC, DIC has two components, a term that measures goodness-of-fit and a penalty term for increasing model complexity:

$$DIC = \bar{D} + p_D.$$

The first term is defined as the posterior mean of the deviance $D(\Theta) = -2 \log L(\text{data} \mid \Theta)$, i.e.

$$\bar{D} = E_{\Theta \mid \text{data}}[D(\Theta)].$$

Here Θ includes all the model parameters and estimated state variables. Since the likelihood function $L(\text{data} \mid \Theta)$ is known the measure is estimated just averaging the likelihood over MCMC sample estimates of $\Theta^{(g)}$.

The second term measuring the effective number of parameters is defined as the difference between the posterior mean of the deviance and the deviance evaluated at the estimated parameters $\hat{\Theta}$, i.e.

$$p_D = \bar{D} - D(\hat{\Theta}).$$

The parameters are, in our case, estimated as means of the sampled values discarding an initial period of the MCMC sampling. The stochastic volatilities are estimated as $\hat{\sigma}_{k,i} = \exp(\bar{h}_{k,i} / 2)$ where $\bar{h}_{k,i}$ is the mean of sampled values. In case of jump occurrence we set $\hat{J}_{k,i} = 1$ if and only if the mean of sampled jump indicators is larger than 0.5, i.e. the probability of jump occurrence is estimated to be larger than 50%. The jump size is again estimated as the sample mean of sampled jumps.

Other classical classical Bayesian goodness-of-fit measures to be mentioned are the Bayes factors or marginal likelihood. We follow Yu and Meyer (2006) who recommend the DIC measure that is shown to have a consistent performance with respect to the two standard measures and is relatively easy to compute.

Value at Risk (VaR)

Finally we intend to illustrate difference between the estimated bivariate models calculating VaR for different time horizons and confidence levels. If the true data were generated by a simple process, e.g. the pure diffusion one, then the estimations yielded by more general models (including jumps and/or stochastic volatilities) calibrated on the same dataset should be similar. However if the data contain "true" jumps and/or stochastic volatility, then the VaR estimations will probably significantly differ and the choice of model becomes important.

Formally, Value at Risk, $\text{VaR}(N, 1-\alpha)$ is defined as $X - q_\alpha[X]$ where $q_\alpha[X]$ denotes the α-quantile of the random variable X modeling the market value of a portfolio, N (business) days to the future. The variable X can be also defined as the future return relative to the initial investment.

We will consider a simple foreign stock investment where the domestic currency value of the portfolio is $V^{DC} = S \cdot V^{FC}$, S is the exchange rate measuring foreign currency FC in terms of domestic currency DC, and V^{FC} is the foreign currency value. The domestic currency log-return over a time horizon can be expressed as the sum $r^{DC} = r^S + r^{FC}$ of the exchange rate

log-return and the foreign currency stock return. Thus, in order to get the distribution and quantiles of r^{DC} we need to model the joint distribution of r^S and r^{FC}.

The simplest and widely used (parametric normal VaR) method calculates the variances and the covariance of r^S and r^{FC}, combines them to get the standard deviation of $r^{DC} = r^S + r^{FC}$, and multiplies it with a standard normal distribution $1 - \alpha$ quantile to get an estimate of $\text{VaR}(1, 1 - \alpha)$. To get N day estimate the value is scaled by $\sqrt{N}$. We will use the estimation with an EWMA (exponentially moving weighted average) covariance matrix as a benchmark value.

Given a specific stochastic model and its parameters we may simulate the returns $r_{T+1}, ..., r_{T+N}$ one to N days ahead, and so the compounded N-days log-return $r = \sum_{i=1}^{N} r_{T+i}$. Note that in case of the stochastic volatility model (16) the initial values must also include an estimate of the last variance V_T. The bivariate model jump-diffusion model with stochastic volatility model (23) allows us to simulate the joint distribution of r^S and r^{FC} over an N days horizon.

The Bayesian estimations in fact yields an empirical distribution of the parameters approximated by a sequence of MCMC simulated values $\Theta_1, ..., \Theta_K$ (discarding an initial sampling period). The Value at Risk estimated as above is then conditional on the estimate $\hat{\Theta}$, i.e. in fact we get $\overset{\triangledown}{V}\text{aR}\left(N, 1 - \alpha \mid \hat{\Theta}\right)$. Since the parameters are uncertain and modified parameters might influence significantly the VaR measure we should rather calculate a VaR estimate incorporating the parameter uncertainty. This can be easily done in the Bayesian framework sampling firstly Θ_i from $\Theta_1, ..., \Theta_K$ and then simulating the returns given the parameter vector Θ_i. The distribution of simulated returns can be used to obtain the Bayesian Value at Risk denoted as $\overset{\triangledown}{V}\text{aR}\left(N, 1 - \alpha \mid \text{data}\right)$, i.e. VaR conditional on the data used to estimate the parameters rather than on a specific vector of point estimates.

3. Empirical Study

The main goal of this section is to apply the binomial jump-diffusion model stochastic volatility model, its submodels (binomial diffusion, jump-diffusion, and stochastic volatility), and the proposed estimation method to the FX and stock market time series returns. However, before doing so we are going to test the estimation procedures on artificially generated data in order to demonstrate that the model is able to identify or reject jumps and stochastic volatility parameters. The MCMC algorithm and all calculations have been implemented in Matlab.

3.1 An Empirical Test

We are going to sample $T = 2000$ returns according to the bivariate jump-diffusion stochastic volatility model (23) with several sets of the model parameters (and starting with initial log-

variances $\log V_{1,0} = \log V_{2,0} = -7$). The proposed MCMC estimation method is then applied to the generated data and estimated parameters are compared to the original parameters that were used to generate the data. The "true" parameters are expected to lie with the Bayesian confidence intervals.

The upper part of Table 1 shows parameters of the first artificially generated bivariate process. There are no jumps, there is no correlation between the diffusion residuals, and no correlation between the variances. The lower part of the table shows the MCMC estimates with standard deviations in parenthesis. The MCMC procedure has been run 3000 times and first 1000 estimations have been discarded. Figure 1 shows, for example, relatively fast convergence of the first stochastic volatility equation coefficients β_{11} and β_{12}.

Generated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.001	0	-	-	-0.14	0.98	0	0.13
Generated jump-diffusion process 2 with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.003	0	-	-	-0.28	0	0.96	0.16
Correlations							
ρ			ρ_J			ρ_Z	
-			-			-	

| Estimated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.0022 (4.477e-004)	0.0026 (0.0050)	-0.0041 (0.0421)	0.0893 (0.0108)	-0.1337 (0.0595)	0.9794 (0.0069)	0.0020 (0.0066)	0.1233 (0.0161)
Estimated jump-diffusion process 2 with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.0023 (4.606e-004)	0.0045 (0.0025)	0.0304 (0.0224)	0.1021 (0.0242)	-0.1825 (0.0603)	0.0011 (0.0059)	0.9730 (0.0074)	0.1332 (0.0138)
Estimated correlations							
ρ			ρ_J			ρ_Z	
-0.0056 (0.0243)			0.4275 (0.1920)			0.0764 (0.5542)	

Table 1. Generated bivariate stochastic volatility process (upper table) and the MCMC estimated parameters.

The estimated stochastic volatility parameters are consistent with the "true" parameters, indeed the estimated coefficients β_{12} and β_{21} are not significantly different from zero and so Granger causality would not be detected. The estimated means μ_1 and μ_2 seem to differ from the "true" values significantly, but we have to take into account the error caused by the data generation process itself possibly correlated to a relatively large volatility. The mean of the first generated return series turns out to be 0.0023 and of the second 0.0019 in line with the estimations. The zero jump intensities fall within the 95% confidence intervals around the

estimated values. In fact no jumps with probability larger than 0.5% are identified by the sampled jump probabilities (averaging the jump indicator $J_{k,i}$) – see Figure 2.
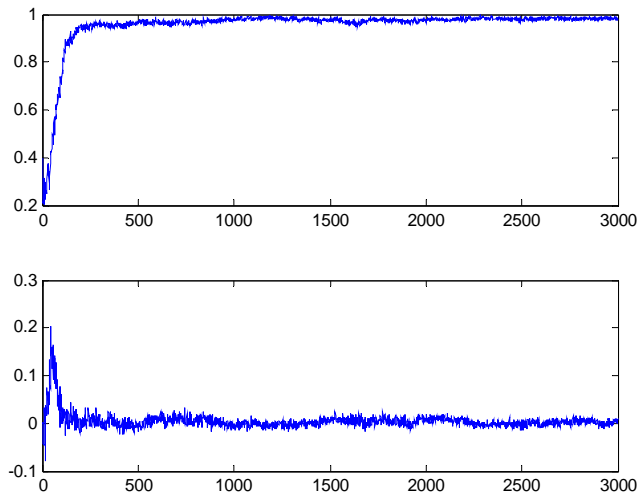
Figure 1. Convergence of the coefficients β_{11} and β_{12}
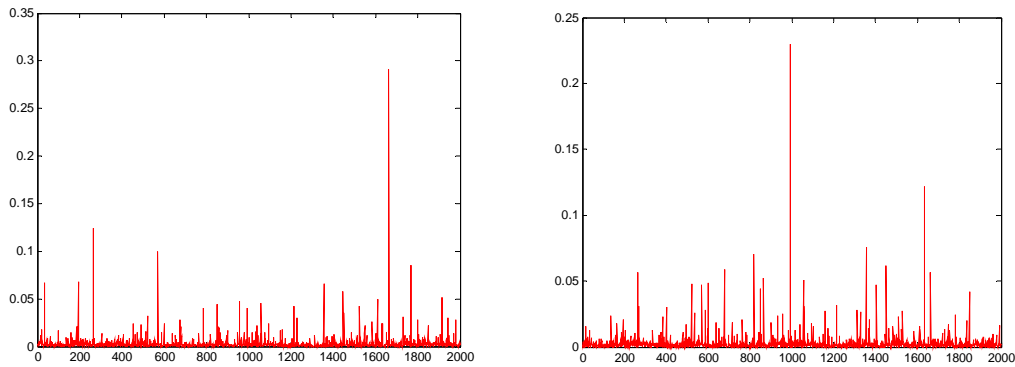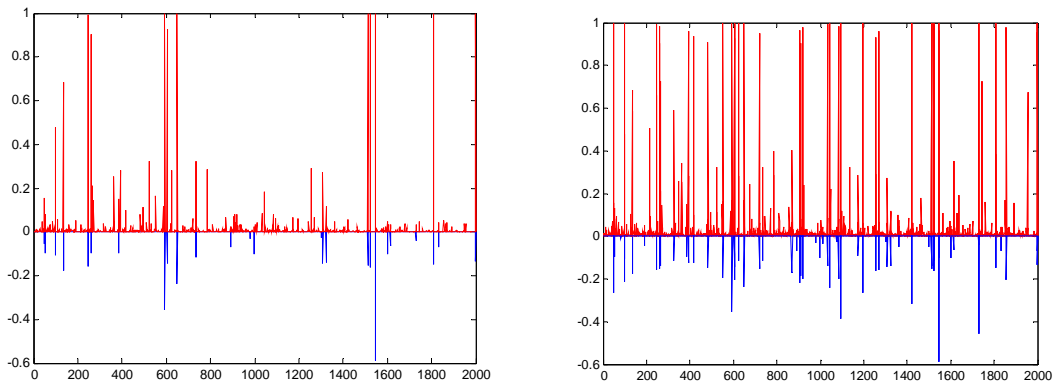

Figure 2. Sampled jump probabilities

Table 1 shows the generating and estimated parameters of a process with nonzero jump intensities, with a positive jump occurrence correlation, but still with zero variance correlation. The estimates are again more-or-less consistent with the "true" parameters. Figure 3 tries to compare the simulated jumps and their absolute sizes with the estimated jump probabilities. The point is that if a simulated jump size is too small than the jump can be ex post hardly identified and the estimated jump probability is low. This may explain the difference between the true jump probability $\lambda_1 = 0.013$ and the estimated value 0.0037 (0.0127). The precision of the jump intensity estimates should improve with a larger number of observations.

Generated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.003	0.0130	0.03	0.15	-0.14	0.98	0	0.13
Generated jump-diffusion process 2 with stochastic volatility							

μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.001	0.0205	-0.02	0.17	-0.21	0	0.97	0.16
Correlations							
ρ		ρ_J			ρ_Z		
0		0.2326			0		

Estimated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.0026 (4.876e-004)	0.0037 (0.0127)	0.0344 (0.0392)	0.1737 (0.0168)	-0.0941 (0.0505)	0.9836 (0.0053)	0.0029 (0.0053)	0.1206 (0.0161)
Estimated jump-diffusion process 2 with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
4.7705e-004 (4.574e-004)	0.0290 (0.0052)	-0.0361 (0.0223)	0.1463 (0.0144)	-0.3125 (0.0859)	-0.0022 (0.0080)	0.9588 (0.0101)	0.1794 (0.0197)
Estimated correlations							
ρ		ρ_J			ρ_Z		
-5.1738e-004 (0.0243)		0.4026 (0.1037)			0.0249 (0.2574)		

Table 2. Generated bivariate jump-diffusion stochastic volatility process (upper table) and the MCMC estimated parameters.

Figure 3. Blue bars below the x-axis show the simulated jumps and their absolute size, red bars above the x-axis show the MCMC estimated jump probabilities (left – process 1, right – process 2)

Finally Table 3 shows parameters and estimates of a bivariate jump-diffusion process with correlated stochastic volatilities. We have still set some parameters to zero, e.g β_{12} or ρ, in order to check that he estimation is able to confirm or reject significance of the individual parameters.

Generated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.003	0.0130	0.03	0.15	-0.07	0.99	0	0.13
Generated jump-diffusion process 2 with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.001	0.0260	-0.02	0.17	-0.21	0.02	0.95	0.16
Correlations							
ρ		ρ_J			ρ_Z		
0		0.3418			0		

Estimated jump-diffusion process 1 with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
0.0018 (5.261e-004)	0.0039 (0.0133)	0.0256 (0.0451)	0.1589 (0.0209)	-0.0931 (0.0613)	0.9703 (0.0090)	0.0157 (0.0125)	0.1527 (0.0174)
Estimated jump-diffusion process 2 with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.0013 (5.543e-004)	0.0288 (0.0051)	-0.0280 (0.0275)	0.1800 (0.0199)	-0.2760 (0.0931)	0.0280 (0.0099)	0.9326 (0.0186)	0.1676 (0.0254)
Estimated correlations							
ρ		ρ_J			ρ_Z		
0.0079 (0.0241)		0.3621 (0.1264)			-0.7194 (0.2454)		

Table 3. Generated bivariate jump-diffusion process (correlated) stochastic volatility model (upper table) and the MCMC estimated parameters.

During implementation of the models many empirical tests have been performed. We have shown only a few to illustrate a good performance of the proposed estimation algorithm.

3.2 FX and Stock Market Data and Empirical Results

Our data set consists of CZK/EUR exchange rates and the Czech PX stock index values from Sep 1, 2004 to Feb 2, 2011 (Figure 4). Note that the CZK/EUR is quoted as the inverse of the normal EUR/CZK exchange rate, i.e. as a direct quotation of CZK in terms of EUR from the perspective of an EUR investor. The period of strong decline of the stock index and depreciation of CZK/EUR corresponds obviously to the financial crisis in 2007-2009.
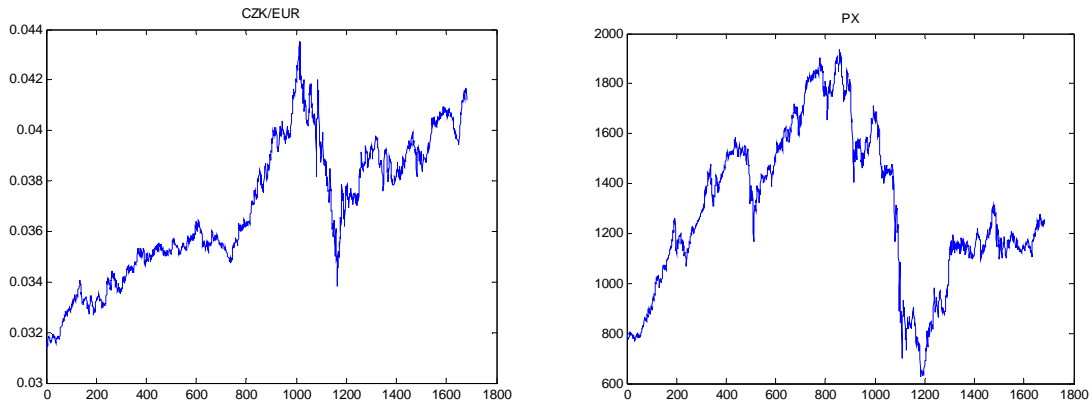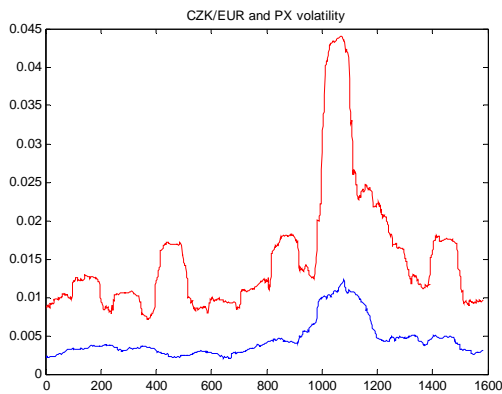
Figure 4. CZK/EUR exchange rate and PX stock index value Sep 1, 2004 to Feb 2, 2011.

Time series of daily returns (Figure 5) and 100 days moving window volatility (Figure 6) visually shows many jumps and overlapping periods of relatively high volatility.
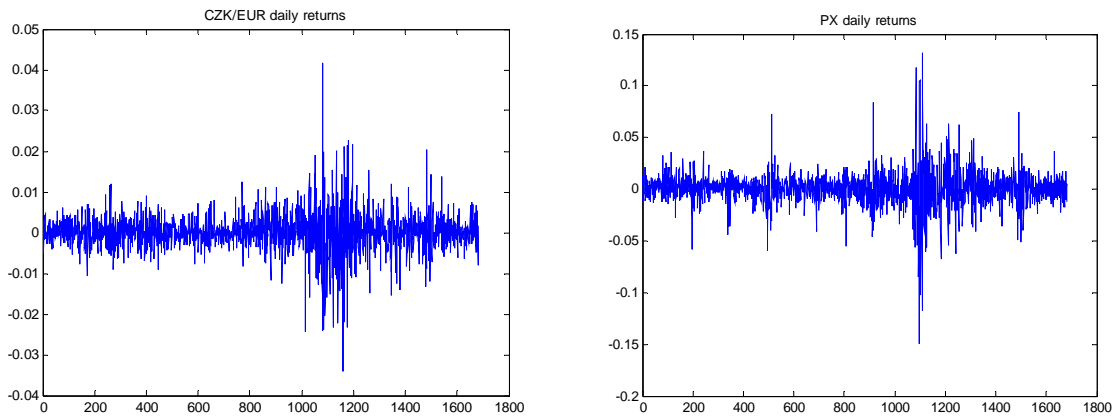
Figure 5. CZK/EUR exchange rate and PX stock index daily returns

Figure 6. CZK/EUR (blue line) and PX (red line) 100 days moving window volatility

Standard tests, e.g. Jarque-Bera test, reject normality of both series at the 0.1% confidence level (Table 4). The Person's correlation of the two return series is slightly negative -1.21% with p-value 0.62, i.e. the correlation does not significantly differ from zero.

21

	CZK/EUR daily returns	PX daily returns
Mean	1.5949e-004	4.0961e-004
Standard deviation	0.0049	0.0170
Skewness	0.0292	-0.1639
Kurtosis	11.2200	15.8395
Jarque-Bera statistic	4.7357e+003	1.1561e+004
p-value	<0.001	<0.001

Table 4. CZK/EUR exchange rate and PX stock index daily returns descriptive statistics

It seems obvious that the normal VaR based on the assumption of normality and no correlation would underestimate the true risk of a combined CZK/EUR and PX investment. In order to improve our ability to predict the risk we are going to estimate and compare the four bivariate models described in Section 2.

Model 1: Bivariate pure diffusion model (2Diff)

The simplest model we consider is the bivariate diffusion model. Table 5 shows MCMC estimated coefficients based on 3000 iterations when the first 1000 have been discarded. The results are in line with the descriptive statistics given in Table 4. The correlation between the two series does not significantly differ from zero in the context of this model.

|	μ_k	σ_k
FX returns (*k=1*)	1.5345e-004 (1.2927e-004)	0.0049 (8.6122e-005)
PX returns (*k=2*)	4.1621e-004 (4.5433e-004)	0.0172 (2.6047e-004)
ρ	-0.0109 (0.0265)	

Table 5. Estimated parameters for the CZK/EUR and PX pure diffusion bivariate model

Model 2: Bivariate jump-diffusion model (2JD)

Table 6 shows MCMC estimates of the bivariate jump-diffusion model (23) with correlated jumps, but with constant volatilities. We have again used 3000 MCMC simulations, non-informative priors, and discarded the first 1000 ones. The initial means and correlations were set to 0, initial diffusion volatilities σ to 0.01 and jump standard deviations σ_J to 0.1. Figure 7 shows relatively fast convergence of the jump correlation ρ_J to a surprisingly high level over 50% and the simulated kernel smoothed density. Note that the diffusion correlation ρ and the jump-size correlation ρ_Z do not significantly differ from 0.

CZK/EUR jump-diffusion process				
μ	σ	λ	μ_J	σ_J
2.2486e-004 (9.8678e-005)	0.0030 (1.0420e-004)	0.1737 (0.0281)	-2.9636e-004 (7.1327e-004)	0.0095 (8.3013e-004)
PX jump-diffusion process				
μ	σ	λ	μ_J	σ_J
9.9833e-004 (3.1909e-004)	0.0101 (3.3542e-004)	0.1549 (0.0244)	-0.0038 (0.0025)	0.0353 (0.0027)
Correlations				
ρ		ρ_J		ρ_Z
-0.0412 (0.0337)		0.5401 (0.0675)		0.0086 (0.0896)

Table 6. Estimated parameters for the CZK/EUR and PX jump-diffusion models with constant volatility
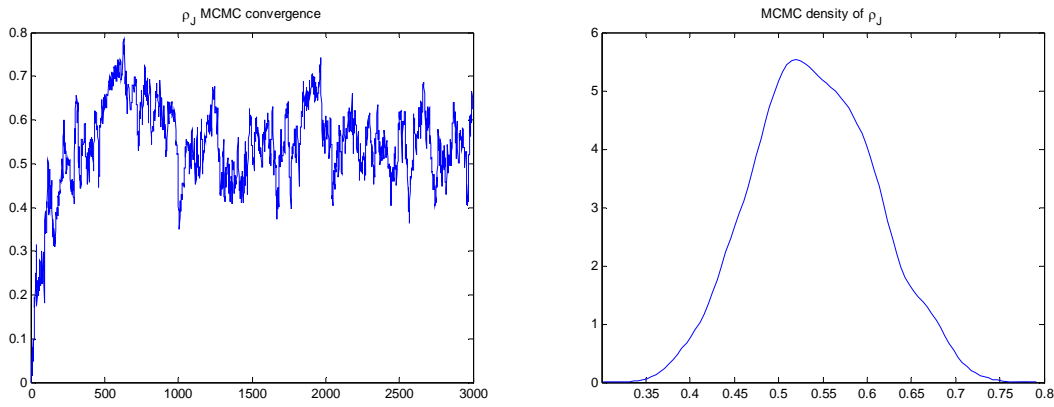
Figure 7. Convergence and the Bayesian density of the jump correlation ρ_J

Probably the most interesting finding is that the probabilities of jumps λ are relatively high, over 15% for both return series. However, the outcome should not be so surprising looking at the very high kurtosis (Table 4) of both series. The jump-diffusion model in fact decomposes the distribution into a mix of two normal distributions, one with the lower standard deviation σ and the other with a higher (more than three times) standard deviation σ_J. While CZK/EUR mean jump size does not significantly differ from zero, the mean PX jump size is negative showing that the stock index tends to jump down rather than up, as one would expect.

The high probabilities of jumps contradict to our intuition of jumps being rare events. The MCMC estimations give us also simulated distributions of jump occurrences and jumps sizes that allow us to analyze the jumps in more detail. Each run of the MCMC simulation samples specific jumps times and jump sizes. Hence for each day we may calculate the empirical probability of jump and identify days where "a jump probably happened", i.e. where the probability of jump is larger than 50%. For those days it makes sense to inspect the mean value of the simulated jump sizes (conditional on the jump occurrence). The results are shown in Figure 8. It is obvious that the jumps are not distributed evenly being clustered especially in

the financial crises period. We will see that the jump clustering is essentially filtered out in the model with stochastic volatility.
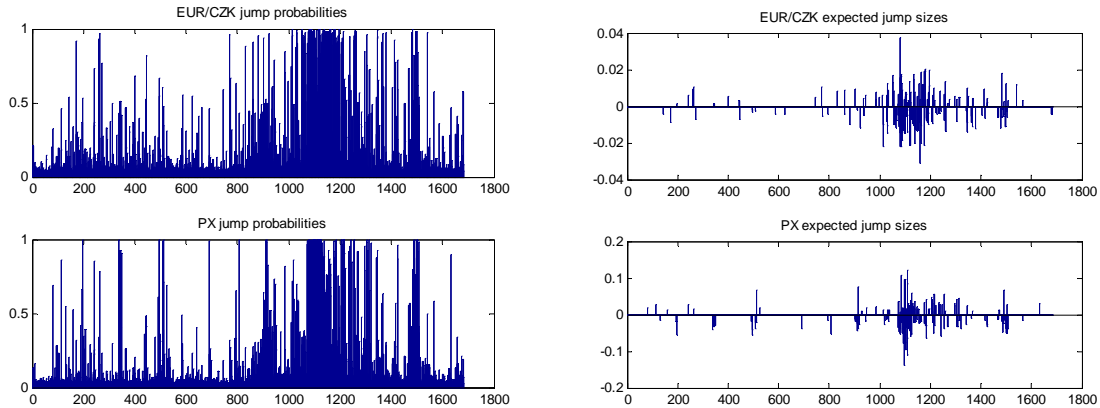

Figure 8. CZK/EUR and PX returns jump probabilities and mean jump sizes

Model 3: Bivariate jump-diffusion with stochastic volatility (2JD SV)

Before estimating the bivariate stochastic volatility model we firstly implement the univariate model (16) for both return series in order to inspect the relationship between the two latent stochastic volatilities time series. The estimated parameters and their standard deviations based on 3 000 MCMC simulations are shown in Table 7. Figure 9 shows for the sake of illustration relatively fast convergence of the coefficient β in case of the PX return process. The parameter means and standard deviations in Table 7 are based on the last 2 500 simulations discarding the first 500. Note that by introducing the stochastic volatility into the model the probabilities of jumps have been significantly reduced to less than 3% and the jump size standard deviation went up. The high value of the stochastic volatility (log-variance) autocorrelation coefficient β, almost 99% for CZK/EUR and almost 98% for PX, shows a high persistence of stochastic volatilities that is in line with other empirical studies on US stock market data (e.g. Jacquier et al, 1994, Eraker et al, 2003). The volatility of the stochastic volatility, i.e. the coefficient γ, around 13% for CZK/EUR and over 21% for PX, is also in the range estimated on US data.

CZK/EUR univariate jump-diffusion process with stochastic volatility						
μ	λ	μ_J	σ_J	α	β	γ
1.8506e-004	0.0284	-2.2616e-004	0.0117	-0.1205	0.9893	0.1313
(6.374e-005)	(0.0083)	(0.0024)	(0.0018)	(0.0545)	(0.0048)	(0.0193)
PX univariate jump-diffusion process with stochastic volatility						
μ	λ	μ_J	σ_J	α	β	γ
0.0012	0.0237	0.0011	0.0427	-0.1957	0.9781	0.2119
(1.921e-004)	(0.0068)	(0.0079)	(0.0066)	(0.0613)	(0.0069)	(0.0247)

Table 7. Estimated parameters for the CZK/EUR and PX univariate jump diffusion models with stochastic volatility

24
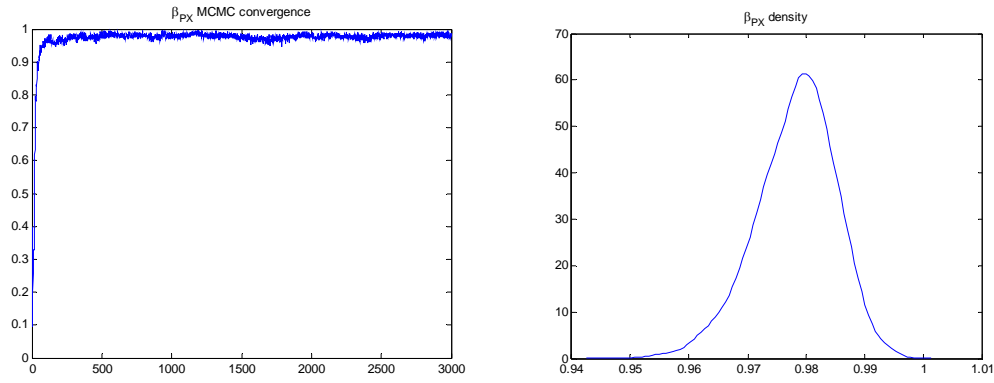
Figure 9. Convergence and MCMC simulated density of the parameter β for PX returns

The latent stochastic volatilities are sampled at each MCMC simulation run and we get a distribution for each particular day. In order to investigate the relationship between the CZK/EUR and PX volatilities we use the mean estimates, specifically given by the equation $\hat{\sigma}_i = \exp(\overline{h}_i / 2)$, where $\overline{h}_i$ is the MCMC mean of normally distributed $\log V_i$ sampled values. Figure 10 shows that the mean stochastic volatility for both series copy well the pattern of the observed returns. The figures also explain why many jumps identified in the constant volatility model have been filtered out in the stochastic volatility model.
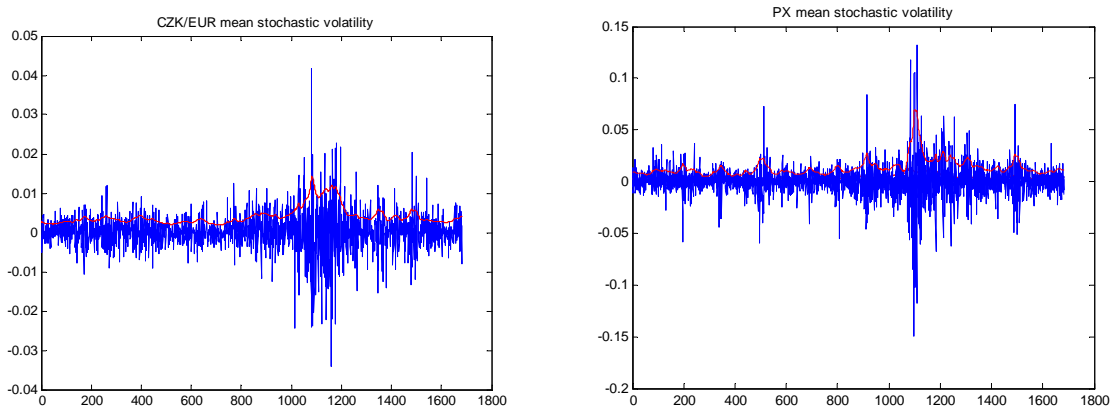
Figure 10. CZK/EUR and PX returns (blue bars) and mean stochastic volatilities (red lines)

The mean stochastic volatility series and mean estimated coefficients $\alpha, \beta,$ and γ may be used to obtain the residuals of the two series $e^V_{FX,i}$ and $e^V_{PX,i}$. The Pearson's correlation of the residuals is relatively low 5.7% and not significant at 1% confidence level, however the correlation of $\overline{h}_{FX,i}$ and $\overline{h}_{PX,i}$ comes out highly significant 61.66% as indicated visually by Figure 10. Consequently, having parsimony in mind, it is reasonable to specify the bivariate jump-diffusion model with correlated stochastic volatilities just through the mutual Granger causality and without correlated volatility residuals, i.e. according to the model (23).

The model is estimated by the methodology has been outlined in Section 2.4 and the resulting estimates are shown in Table 8. We have run again 3000 simulations, used the last 2000, and discarded the first 1000. Figure 11 shows relatively fast convergence of the interesting

coefficient β_{21}. The significant positive value 0.035 proves there is a Granger dependence of the PX stochastic volatility on the CZK/EUR stochastic volatility. On the other hand the coefficient β_{12} reflecting the causality in the opposite direction turns out not to be significant. This corresponds to our intuition: the FX market is more liquid and closely linked to the global markets while the Czech stock market liquidity is relatively low and in a sense behind the global FX market. It can be verified that the VAR(1) process estimated coefficients imply a high Pearson's correlation around 58% close to our finding based on the univariate models.

| CZK/EUR jump-diffusion process with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
1.5188e-004	0.0129	-9.18e-004	0.0127	-0.2694	0.9692	0.0086	0.2112
(5.994e-005)	(0.0057)	(0.0057)	(0.0016)	(0.0972)	(0.0125)	(0.0087)	(0.0316)
PX jump-diffusion process with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.0013	0.0117	-0.0199	0.0447	0.0220	0.0350	0.9582	0.2268
(1.809e-004)	(0.0048)	(0.0143)	(0.0060)	(0.0827)	(0.0092)	(0.0078)	(0.0174)
Correlations							
ρ			ρ_J			ρ_Z	
-0.0394 (0.0267)			0.2638 (0.1927)			0.4713 (0.3593)	

Table 8. Estimated parameters for the bivariate jump diffusion models with Granger related stochastic volatilities
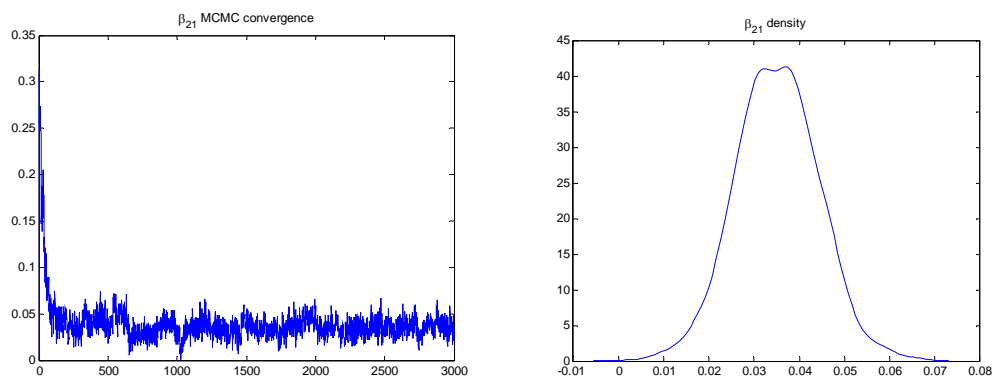
Figure 11. MCMC convergence and density of the coefficient β_{21} reflecting the Granger dependence of the PX stochastic volatility on the CZK/EUR stochastic volatility

Finally we may inspect the behavior of jumps. It is interesting to note that the jump probability fell to 1.3% in case of CZK/EUR and 1.2% in case of PX returns. Figure 12 shows the MCMC empirical daily jump probabilities. For both series there is only one day with jump probability higher than 50% and a few days with jump probabilities over 20%. The stochastic volatility component has clearly filtered out the clustering and the jumps seem to be distributed evenly. There is still a positive jump occurrence correlation 0.26 and it is interesting to note that in this case the jump size correlation is positive 0.47 (though not highly significant) meaning that if there is a coincidence of jumps on both markets than the jumps probably go in the same direction.

Looking at the jump analysis in the jump-diffusion models with and without stochastic volatility it is obvious that jump identification strongly depends on the stochastic model chosen.
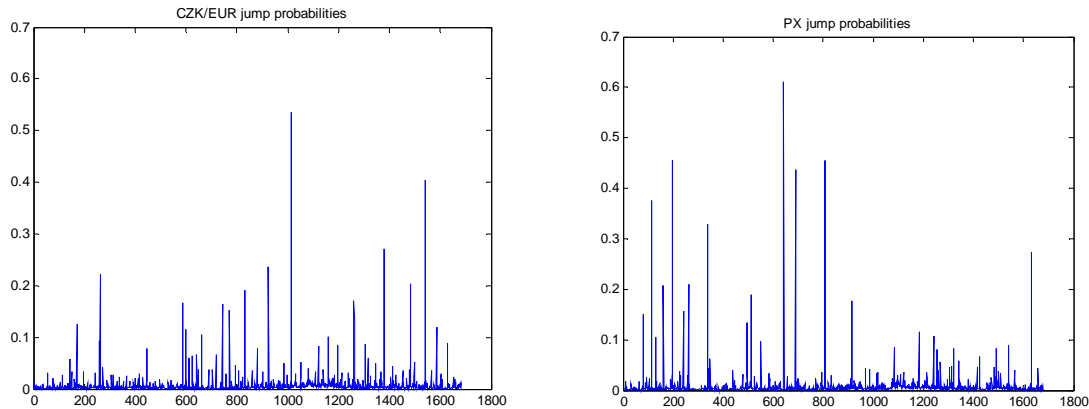
Figure 12. CZK/EUR and PX returns jump probabilities in the context of the bivariate stochastic volatility model

Model 4: Bivariate diffusion model with stochastic volatility (2SV)

Since the estimated jump probabilities in the previous models have been low and not highly significant Table 9 gives, for the sake of completes, estimates of the bivariate stochastic volatility model without jumps. It is interesting to note that the estimated coefficients do not differ significantly from the results in Table 8 (with jump parameters missing).

diffusion process FX with stochastic volatility							
μ_1	λ_1	$\mu_{1,J}$	$\sigma_{1,J}$	α_1	β_{11}	β_{12}	γ_1
1.7474e-004 (5.2611e-004)	-	-	-	-0.1981 (0.0951)	0.9791 (0.0121)	0.0041 (0.0077)	0.1750 (0.0393)
diffusion process PX with stochastic volatility							
μ_2	λ_2	$\mu_{2,J}$	$\sigma_{2,J}$	α_2	β_{21}	β_{22}	γ_2
0.0013 (0.0013)				-0.0111 (0.0867)	0.0387 (0.0100)	0.9501 (0.0105)	0.2564 (0.0250)
Correlations							
ρ		ρ_J			ρ_Z		
-0.0371 (0.0257)		-			-		

Table 9. Estimated parameters for the bivariate stochastic volatility model without jumps

3.3 Comparison of the Models

The four estimated models can be compared using the deviance information criterion (DIC) measure shown in Table 10. The last two columns give the goodness-of-fit and the model complexity measures. For example, in case of the bivariate diffusion model p_D equals to 5 (after rounding to units), i.e. to the number of the parameters estimated, as expected. It is more difficult to interpret the p_D values in case of the other models. The high complexity of the Model 2 (2JD) could be explained by existence of many jumps where the size needs to be

in addition estimated. The high value of p_D for the Model 4 (2SV) compared to the Model 3 (2JD SV) is, however, slightly puzzling.

An absolute difference in DIC over 10 is already considered important (Spiegelhalter et al, 2002, Asai et al, 2006), and so the models with jumps and/or stochastic volatility strongly outperform, in terms of DIC, the pure diffusion model. The best ranked model is the Model 2 (2JD) followed by the Model 3 (2JD SV). The difference between the DIC of 2JD and 2JDSV might be disappointing from the perspective of the stochastic volatility modeling and the estimation effort. However the 2JD model estimates many jump sizes significantly improving the goodness-of-fit with respect to the estimation dataset, but the jump sizes estimates do not have any added value for predictions of future returns (distributions). This is not the case the 2JDSV where the last estimated stochastic volatility is used to predict future volatilities and return distributions.

The ability of the different models to predict future distributions (in particular VaR) could be compared using a back-testing procedure. This is unfortunately unfeasible at the moment since one MCMC estimation takes more than one or two hours on a relatively powerful desk-top computer and back-testing would require repeating the procedure hundreds or thousands of times.

	DIC	$\bar{D}$	p_D
Model 1 (2Diff)	-22 019	-22 024	5
Model 2 (2JD)	-25 294	-25 479	185
Model 3 (2JD SV)	-24 369	-24 430	61
Model 4 (2SV)	-24 068	-24 228	160

Table 10. Deviance information criterion of the four estimated bivariate models

VaR Estimations

We are going to show VaR estimations given by the four tested models and by a benchmark model to illustrate importance of the model choice. Table 11 provides a comparison of the estimates VaR measures for the sum of lognormal returns $r_{CZK/EUR} + r_{PX}$ calculated using the different models and by the methodology described in Section 2.5. . The returns could be interpreted for example as returns on a PX stock index investment from the perspective of a EUR based investor as of February 2, 2011. The VaR measure has been calculated in 1, 10, and 30 days horizon, on the 95% and 99% probability level, and conditional on the point parameter estimates (Table 5 -Table 9 and the least mean variance for the SV models) or on the corresponding MCMC parameter distributions. The last line gives EWMA VaR estimates used frequently in practice, for the sake of comparison with our models. It is based on exponentially weighted moving average (EWMA) covariance matrix with the weight set at 0.97. The number of Monte Carlo simulations has been set to 40 000.

Since the jump-diffusion (2JD) model is using constant volatilities, it is comparable rather to the diffusion model (2Diff), corresponding to a standard normal VaR estimate. On the other hand the jump-diffusion model with stochastic volatility (2JDSV) starts with the last

estimated volatility (as of Feb 2, 2011) and so it can be meaningfully compared rather against the benchmark EWMA model that works a recent volatility trend.

The 2JD VaR does not differ significantly from the 2Diff VaR on the 95% probability level, but, not surprisingly, the difference is more significant on the 99% probability level and in the short 1- day horizon. Over longer time horizons the jumps mix better with the normal returns and the two approaches become closer. In case of the jump-diffusion model with stochastic volatility we observe the opposite. The 2JDSV and 2SV VaR is similar to EWMA VaR in the 1-day horizon, but the difference becomes more pronounced in longer time horizon due to the stochastic volatility effect. The 30 days JDSV and 2SV VaR on the 99% probability level exceeds the EWMA VaR by almost 40%.

It should be pointed that no definite conclusion can be drawn from the differences in VaR estimates. Nevertheless, if the true process was a pure diffusion one then the choice of the model should not matter. Significant parameters of the models with jumps and stochastic volatilities, the goodness-of-fit comparison, and the different VaR values demonstrate that the choice model is important.

VaR / Bayes VaR	(1,95%)	(1,99%)	(10,95%)	(10,99%)	(30,95%)	(30,99%)
Model 1 (2Diff)	0.0299 / 0.0297	0.0421 / 0.0417	0.0985 / 0.0986	0.1361 / 0.1361	0.1780 / 0.1787	0.2451 / 0.2433
Model 2 (2JD)	0.0240 / 0.0242	0.0520 / 0.0532	0.0969 / 0.0962	0.1414 / 0.1405	0.1743 / 0.1781	0.2437 / 0.2480
Model 3 (2JDSV)	0.0177 0.0179	0.0266 0.0284	0.0654 0.0678	0.1041 0.1105	0.1272 0.1280	0.1957 0.1988
Model 4 (2SV)	0.0213 / 0.0214	0.0293 / 0.0300	0.0763 / 0.0769	0.1040 / 0.1086	0.1499 / 0.1505	0.2048 / 0.2092
EWMA	0.0185	0.0262	0.0586	0.0828	0.1014	0.1435

Table 11. Comparison of Value at Risk in the 1, 10, or 30 day horizon, based on point parameter estimates / Bayes simulation, and on the 95%, 99% probability level estimated using the different models

3 Conclusion

We have proposed and tested a bivariate jump-diffusion model with stochastic volatilities where the jump occurrence, jump-sizes, and the volatilities are, in general, correlated. The stochastic volatility correlation is captured through the Granger causality in both directions. Our proposed MCMC estimation technique extends the one of Jacquier et al (2007). The tests have demonstrated ability of the algorithm to recover consistently the true parameters used to generated testing data. The model has been, for the sake of empirical illustration, applied on the CZK/EUR and the Czech stock index PX data (2004-2011). The results show that there is Granger causality in stochastic volatility going from the CZK/EUR market to the Czech stock market but not vice versa. The estimated probabilities of jumps are around 1.2%. The jumps and jump sizes appear to be strongly correlated. We have compared the model with the pure diffusion model, with the stochastic volatility model without jumps, and with the jump-diffusion model with constant volatility. It is interesting to note that the probabilities of jumps

in the pure jump diffusion model differ dramatically being more than 15% for both processes, if the volatility is constant. The four tested models have been compared applying the deviance information criterion (DIC) with the pure jump-diffusion and the stochastic volatility jump-diffusion models performing the best. However, it can be argued that the proposed stochastic volatility jump-diffusion model giving forward looking volatility estimates is the most appropriate for VaR calculations and generally for future returns distribution modeling.

The research can be extended in several directions: the model and the estimation technique can be applied to other time series, to high frequency data, and generalized to the general multivariate case requiring algorithm efficiency improvements.

Literature

Asai M., McAleer M., Jun Yu J. (2006): Multivariate Stochastic Volatility: A Review, Econometric Reviews, 25(2–3), p.145–175.

Cont R., Tankov P. (2004): Financial Modelling With Jump Processes, Chapman & Hall/CRC Financial Mathematics Series, pp. 535.

Eraker B. (2004): Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option prices, the Journal of Finance, Vol. LIX, No. 3, June, p. 1367-1403.

Eraker B., Johannes M., Polson N. (2003): The impact of Jumps in Volatility and Returns, The Journal of Finance, Vol. LVIII, No. 3, June, p. 1269-1300.

Heston, Steven (1993): Closed-form solution of options with stochastic volatility with application to bond and currency options, Review of Financial Studies 6, 327-343.

Jacquier E, Johannes M., Polson N. (2007): MCMC maximum likelihood for latent state models, Journal of Econometrics 137, p. 615–640.

Jacquier E., Polson N., Rossi P. (1994): Bayesian analysis of Stochastic Volatility Models, Journal of Business & Economic Statistics, Vol. 12, No. 4, October, p. 69-87.

Johannes M., Kumar R. (1999): State Dependent Jump Models: How do US Equity indices Jump? , Working paper, Columbia University, available at http://faculty.chicagobooth.edu/nicholas.polson/research/papers/statedep.pdf, pp. 13.

Johannes M., Polson N. (2009): MCMC Methods for Financial Econometrics , Handbook of Financial Econometrics (eds. Ait-Sahalia and L.P. Hansen), p. 1-72.

Lynch, S. M. (2007): Introduction to Applied Bayesian Statistics and Estimation for Social Scientists, Springer, pp. 359.

Rachev S.T., Hsu J.S, Bagasheva B.S., Fabozzi F.J. (2008): Bayesian Methods in Finance, The Frank J. Fabozzi Series, Wiley, pp. 329.

Shephard, N. (2004): Stochastic Volatility: Selected Readings. Oxford University Press: Oxford, pp. 534.

Shreve, S.E. (2004): Stochastic calculus for Finance II / Continuous Time Models, Springer, pp. 550.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A. (2002): Bayesian measures of model complexity and fit (with discussion). Journal of the Royal Statistical Society, Series B., 64,583–639.

Yu, J. and Meyer, R. (2006): Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison, Econometric Reviews, 25(2–3), p. 361–384.

IES Working Paper Series

2011

21. Josef Brechler, Adam Geršl : *Political Legislation Cycle in the Czech Republic*
22. Jozef Baruník, Lukáš Vácha, Ladislav Krištoufek *: Comovement of Central European stock markets using wavelet coherence: Evidence from high-frequency data*
23. Michal Skořepa : *A convergence-sensitive optimum-currency-area index*
24. Marek Rusnák, Tomáš Havránek, Roman Horváth : *How to Solve the Price Puzzle? A Meta-Analysis*
25. Marián Dinga, Vilma Dingová : *Currency Union and Investment Flows: Estimating the Euro Effect on FDI*
26. Krenar Avdulaj : *The Extreme Value Theory as a Tool to Measure Market Risk*
27. Ivo Jánský, Milan Rippel : *Value at Risk forecasting with the ARMA-GARCH family of models in times of increased volatility*
28. Pavel Ryska, Jan Průša : *Efficiency Wages in Heterogenous Labour Markets*
29. Peter Kukuk, Adam Geršl : *Political Pressure on the National Bank of Slovakia*
30. Jiří Schwarz : *Impact of Institutions on Cross-Border Price Dispersion*
31. Pavel Doležel : *Volební systémy pro volby do Poslanecké sněmovny Parlamentu ČR založené na matematickém programování*
32. Martin Gregor: *Corporate lobbying: A review of the recent literature*
33. Karel Janda, Ladislav Kristoufek, David Zilberman : *Modeling the Environmental and Socio-Economic Impacts of Biofuels*
34. Tomáš Havránek, Zuzana Iršová : *How to Stir Up FDI Spillovers: Evidence from a Large Meta-Analysis*
35. Jiří Witzany : *Estimating Correlated Jumps and Stochastic Volatilities*

All papers can be downloaded at: http://ies.fsv.cuni.cz