Ordinal stochastic volatility and stochastic volatility models for price changes: An empirical comparison

Claudia Czado[1], Thi-Ngoc-Giau Nguyen[2], Gernot Müller[3]

Abstract

Ordinal stochastic volatility (OSV) models were recently developed and fitted by Müller and Czado (2008) to account for the discreteness of financial price changes, while allowing for stochastic volatility (SV). The model allows for exogenous factors both on the mean and volatility level. A Bayesian approach using Markov Chain Monte Carlo (MCMC) is followed to facilitate estimation in these parameter driven models. In this paper the applicability of the OSV model to financial stocks with different levels of trading activity is investigated and the influence of time between trades, volume, day time and the number of quotes between trades is determined. In a second focus we compare the performance of OSV models to SV models by developing model selection criteria. This analysis shows that the discreteness of price changes should not be ignored.

Keywords and phrases: Markov chain Monte Carlo; ordinal time series; stochastic volatility; ultra high frequency financial data

[1]Zentrum Mathematik / M4, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany, email: cczado@ma.tum.de

[2]Zentrum Mathematik / M4, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany, email: ntngiau2002@yahoo.com

[3]corresponding author; Zentrum Mathematik / M4, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany, email: mueller@ma.tum.de, phone: +49 89 289 17441, fax: +49 89 289 17435

1 Introduction

Modeling price changes in financial markets is a challenging task especially when models have to account for salient features such as fat tail distributions and volatility clustering. An additional difficulty is to allow for the discreteness of price changes. These are still present after the US market graduation to decimalization of possible tick sizes. Recently, Müller and Czado (2008) introduced the class of ordinal stochastic volatility (OSV) models, which utilizes the advantages of stochastic volatility (SV) models (see Ghysels, Harvey, and Renault (1996) and more lately Shephard (2006)) such as fat tails and persistence through autoregressive terms in the volatility process, while adjusting for the discreteness of the price changes.

OSV models are based on a threshold approach, where the hidden continuous process follows a SV model, thus providing a more realistic extension of the ordered probit model suggested by Hausman et al. (1992). In addition we allow for exogenous variables both on the mean and variance level of the hidden process. Parameter estimation in OSV models using maximum likelihood is not feasible, since first the hidden SV process has no closed form of the likelihood and second the threshold approach induces the need to evaluate multidimensional integrals with dimension equal to the length of the financial time series. Therefore Müller and Czado (2008) follow a Bayesian approach. Here Markov Chain Monte Carlo (MCMC) methods allow for sampling from the posterior distributions of model parameters and the hidden process variables.

While Müller and Czado (2008) provided the model specification, developed and implemented the necessary estimation techniques, this paper explores the applicability of the OSV model to financial stocks with different levels of trading activity. Especially, we investigate which exogenous factors such as volume, daytime, time elapsed between trades and the number of quotes between trades have influence on the mean and variance level of the hidden process and thus on the discrete price changes. A second focus of this paper is the comparison of OSV and SV model specifications to assess how large the gain of the OSV over the SV model is.

Alternative discrete price change models include rounding approaches and decomposition approaches. For the rounding approach Harris (1990) models discrete prices by assuming constant variances of the underlying efficient price, while Hasbrouck (2000) models efficient prices for bid and ask prices separately using GARCH dynamics for the volatility of the effcient price processes. Hasbrouck (2000) proposes to use non-Gaussian, non-linear state space estimation of Kitagawa (1987). Later works of Manrique and Shephard (1997), Hasbrouck (1999), Hasbrouck (2004a) and Hasbrouck (2004b) also use MCMC techniques for estimation.

A second class of models to deal with discrete price changes are decomposition models, where the price change is assumed to be a product of three random variables, namely price change indicator, its direction and the size of the price change. Rydberg and Shephard (2003) and Liesenfeld et al. (2006) follow this approach. Russell and Engle (2005) introduce a joint model of price changes and time elapsed between trades (duration) where price changes follow an autoregressive conditional multinomial (ACM) model and durations the autoregressive conditional duration (ACD) model of Engle and Russell (1998). A common feature of these models is that the time dependence is solely induced by lagged endogenous variables, while our OSV specification allows for parameter driven time dynamics.

The paper is organized as follows: Section 2 introduces OSV and SV specifications and its estimation using MCMC methods. It also considers the problem of model selection among OSV, among SV and between OSV and SV models. The data application to three NYSE stocks with different trading levels from the TAQ data base are given in Section 3. Special emphasis is given to model interpretation and model selection. The paper closes with a summary and outlines further research.

2 Ordinal stochastic volatility models

2.1 Model specification and interpretation

As introduced by Müller and Czado (2008) we consider the following stochastic volatility model for ordinal valued time series $\{Y_{t_i}, i = 1, \cdots, I\}$. In this model the response Y_{t_i} with K possible values is viewed as a censored observation from a hidden continuous variable $Y^*_{t_i}$ which follows a stochastic volatility model, i.e.

$$
\begin{aligned}
Y_{t_i} = k \Leftrightarrow \quad & Y^*_{t_i} \in [c_{k-1}, c_k)\,, \\
& Y^*_{t_i} = \mathbf{x}'_{t_i}\boldsymbol{\beta} + \exp(h^*_{t_i}/2)\epsilon^*_{t_i}, \\
& h^*_{t_i} = \mathbf{z}'_{t_i}\boldsymbol{\alpha} + \phi(h^*_{t_{i-1}} - \mathbf{z}'_{t_{i-1}}\boldsymbol{\alpha}) + \sigma\eta^*_{t_i},
\end{aligned} \tag{2.1}
$$

where $c_0 = -\infty < c_1 < \cdots < c_{K-1} < c_K = +\infty$ are unknown threshold parameters. Further $\mathbf{x}_{t_i}$ and $\mathbf{z}_{t_i}$ are p and q dimensional covariate vectors on the hidden mean and log volatility level, respectively. Associated with these covariate vectors are unknown regression parameters $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, respectively. For t_0 we assume $\mathbf{z}_0 := (0, ..., 0)'$ and h^*_0 follows a known distribution. The error terms $\epsilon^*_{t_i}$ and $\eta^*_{t_i}$ are assumed i.i.d. standard normal and independent of each other. The model specified by (2.1) is denoted by $OSV(X_1, \cdots, X_p; Z_1, \cdots, Z_q)$. For identifiability reasons we need to fix a threshold parameter, i.e. we set $c_1 = 0$. Further, ϕ is an unknown autocorrelation parameter and σ^2 an unknown variance parameter on the hidden log volatility scale.

To interpret such a model , denote the mean and variance of the hidden process at t_i by μ_{t_i} and $\sigma^2_{t_i}$, respectively. As μ_{t_i} is increased holding $\sigma^2_{t_i}$ fixed, we see that the probability of a large (small) category is increased (decreased). For fixed μ_{t_i}, we see that if $\sigma^2_{t_i}$ is increased the probability of extreme categories is increased. These two situations are illustrated in Figure 1.

Further the OSV model allows to quantify the probability for observing a specific category at time t_i given by

$$
\begin{aligned}
p^1_{t_i} &:= P(Y_{t_i} = 1) = \Phi\left(\frac{c_1 - \mathbf{x}'_{t_i}\boldsymbol{\beta}}{\exp(h^*_{t_i})}\right) \\
p^k_{t_i} &:= P(Y_{t_i} = k) = \Phi\left(\frac{c_k - \mathbf{x}'_{t_i}\boldsymbol{\beta}}{\exp(h^*_{t_i})}\right) - \Phi\left(\frac{c_{k-1} - \mathbf{x}'_{t_i}\boldsymbol{\beta}}{\exp(h^*_{t_i})}\right), k = 2, \cdots, K-1 \\
p^K_{t_i} &:= P(Y_{t_i} = K) = 1 - \Phi\left(\frac{c_{K-1} - \mathbf{x}'_{t_i}\boldsymbol{\beta}}{\exp(h^*_{t_i})}\right),
\end{aligned}
$$

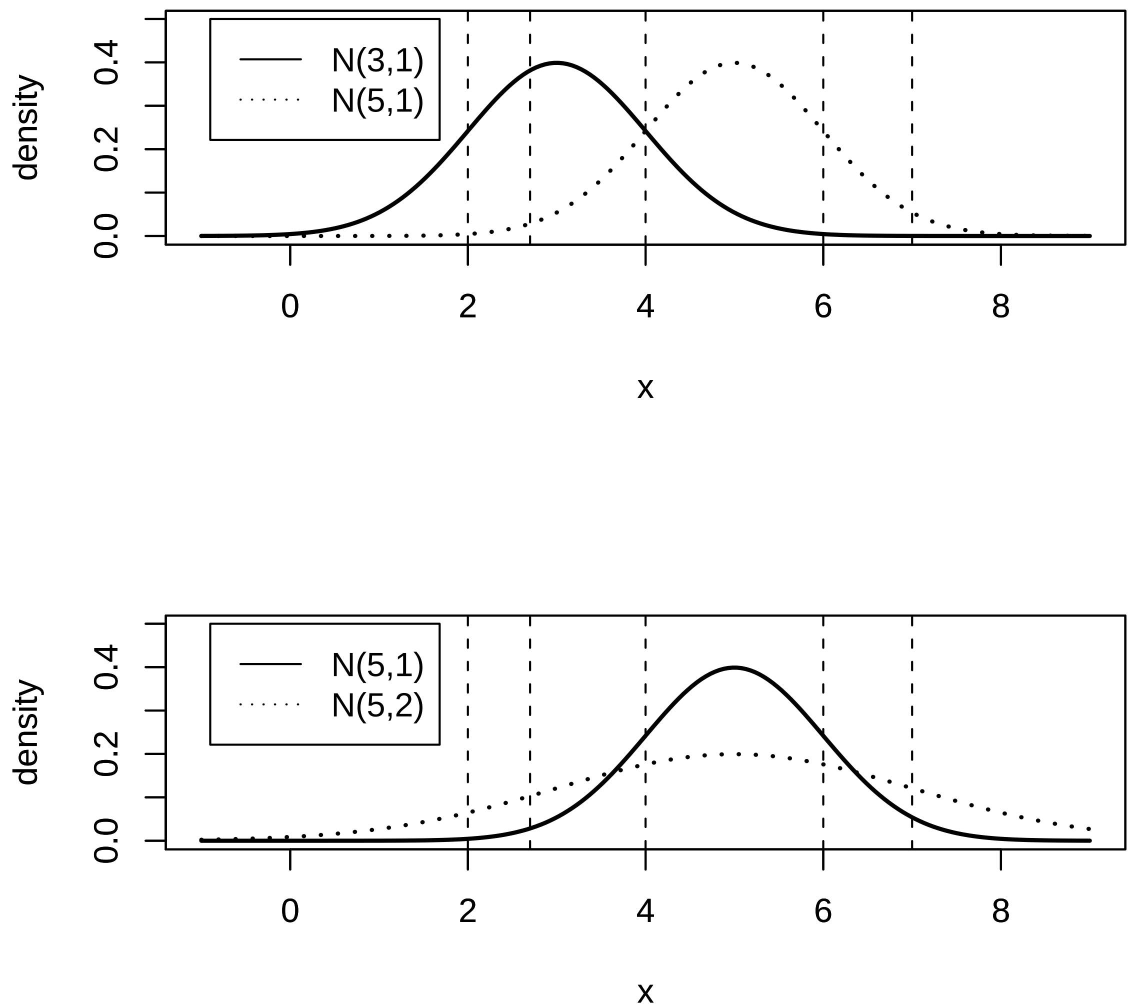

Figure 1: Category probabilities (visualized as area under the curve between adjacent threshold bounds) as mean and variance of a hidden $N(\mu, \sigma^2)$ process varies

where $\Phi()$ denotes the cumulative distribution function of a standard normal random variable. Therefore the model is able to identify time points where there is a large probability of extreme small or large category labels. Note that no symmetry assumptions about the occurrence of large/small categories are present in the model specification.

We conclude this subsection by presenting the ordinary stochastic volatility model. For a real valued time series $\{Y^c_{t_i}, i = 1, \cdots, I\}$, we assume the following model specification

$$
\begin{aligned}
Y^c_{t_i} &= \mathbf{x}'_{t_i}\boldsymbol{\beta} + \exp(h_{t_i}/2)\epsilon_{t_i} \\
h_{t_i} &= \mathbf{z}'_{t_i}\boldsymbol{\alpha} + \phi(h_{t_{i-1}} - \mathbf{z}'_{t_{i-1}}\boldsymbol{\alpha}) + \sigma\eta_{t_i},
\end{aligned}
\tag{2.2}
$$

where $\mathbf{x}_{t_i}, \boldsymbol{\beta}, \mathbf{z}_{t_i}, \boldsymbol{\alpha}, \phi$ and σ^2 are specified as in the OSV model. For the error terms ϵ_{t_i} and η_{t_i} we assume i.i.d. standard normal and independence between ϵ_{t_i} and η_{t_i}. The model specified by (2.2) we denote by $SV(X_1, \cdots, X_p; Z_1, \cdots, Z_q)$.

In our application we use $OSV(X_1, \cdots, X_p; Z_1, \cdots, Z_q)$ models for the category labels of the associated price change classes, while $SV(X_1, \cdots, X_p; Z_1, \cdots, Z_q)$ models are used for modeling directly the observed price changes.

2.2 Bayesian inference for OSV and SV models

Bayesian inference for the SV models was thoroughly investigated in Chib et al. (2002). They used an estimation procedure based on a state space approximation which we now just recall briefly. Obviously, in model (2.2) one can equivalently write

$$\log\left(Y^c_{t_i} - \mathbf{x}'_{t_i}\boldsymbol{\beta}\right)^2 = h_{t_i} + \log\epsilon^2_{t_i}.$$

Kim et al. (1998) have shown that the distribution of $\log\epsilon^2_{t_i}$ can be approximated very well by a seven-component mixture of normals. In particular, one can assume $\log\epsilon^2_{t_i} \approx \sum_{k=1}^{7} q_k u^{(k)}_{t_i}$ where $u^{(k)}_{t_i}$ is normally distributed with mean m_k and variance v^2_k independent of t_i. Moreover, the random variables $\{u^{(k)}_{t_i} \,|\, i = 1, \cdots, I, k = 1, \cdots, 7\}$ are independent. The quantity q_k denotes the probability that the mixture component k occurs. These probabilities are also independent of t and are given in Chib et al. (2002), Table 1, together with the corresponding means and variances. Let $s_{t_i} \in \{1, \cdots, 7\}$ denote the component of the mixture that occurs at time t_i and let $\pi(s_{t_i})$ denote the prior for s_{t_i}, where $\pi(s_{t_i}\!=\!k) = q_k$. Then, by setting $\tilde{Y}^c_{t_i} := \log\left(Y^c_{t_i} - \mathbf{x}'_{t_i}\boldsymbol{\beta}\right)^2$, one arrives at

$$\tilde{Y}^c_{t_i} = h_{t_i} + u^{(s_{t_i})}_{t_i}$$

which, together with the second equation of (2.2), gives the desired state space representation.

The inference for the OSV models is even more complicated, since a straight-forward extension of algorithm by Chib et al. (2002) shows an unacceptable bad mixing of the chains. Therefore, Müller and Czado (2008) developed a grouped-move multigrid Monte Carlo (GM-MGMC) algorithm which exhibits fast convergence of the produced Markov chains. Since the SV model given by (2.2) is a submodel of the OSV model, we use the same sampling scheme also for the

SV model, of course reduced by the sampling of the cutpoints which do not appear in the SV model, and the variables $Y^*_{t_i}$, $i = 1, \cdots, I$, which are observed in the SV case.

Each iteration of the GM-MGMC sampler consists of three parts. In the first part, the parameter vector $\boldsymbol{\beta}$ is drawn in a block update from a $(p+1)$-variate normal distribution, the latent variables $Y^*_{t_i}$, $i = 1, \cdots, I$, from truncated univariate normals, and the cutpoints c_k, $k = 2, \cdots, K-1$, from uniform distributions. In the second part, the grouped move step is performed. Here one draws a transformation element γ^2 from a Gamma distribution and updates $\boldsymbol{\beta}$, $(Y^*_{t_1}, \cdots, Y^*_{t_I})$, and $\mathbf{c}$ by multiplication by the element $\gamma = \sqrt{\gamma^2}$. The third part starts with computation of the state space approximation, i.e. by computing $\tilde{Y}^*_{t_i} = \log(Y^*_{t_i} - \mathbf{x}'_{t_i}\boldsymbol{\beta})^2$ for $i = 1, \cdots, I$. Then s_{t_i}, $i = 1, \cdots, I$, are updated in single updates, and $(\boldsymbol{\alpha}, \phi, \sigma)$ by a Metropolis-Hastings step. Finally, the log volatilities $h^*_{t_1}, \cdots, h^*_{t_I}$ are drawn in one block using the simulation smoother of De Jong and Shephard (1995). For more details on the updates we refer to Müller and Czado (2008).

For the Bayesian approach one also has to specify the prior distributions for $\mathbf{c}$, $\boldsymbol{\beta}$, h^*_0, $\boldsymbol{\alpha}$, ϕ, and σ. Assuming prior independence the joint prior density can be written as

$$\pi(\mathbf{c}, \boldsymbol{\beta}, h^*_0, \boldsymbol{\alpha}, \phi, \sigma) = \pi(\mathbf{c})\pi(\boldsymbol{\beta})\pi(h^*_0)\pi(\alpha_1)\cdots\pi(\alpha_q)\pi(\phi)\pi(\sigma).$$

For $\boldsymbol{\beta}$ a multivariate normal prior distribution is chosen, for h^*_0 the Dirac measure at 0, and for the remaining parameters uniform priors. In particular,

$$\begin{aligned} \pi(\mathbf{c}) &= \mathbf{I}_{\{0<c_2<\cdots<c_{K-1}<C\}}, & \pi(\boldsymbol{\beta}) &= N_{p+1}(\boldsymbol{\beta} \mid \mathbf{b}_0, B_0), & \pi(h^*_0) &= \mathbf{I}_{\{h^*_0=0\}}, \\ \pi(\alpha_j) &= \mathbf{I}_{(-C_\alpha, C_\alpha)}(\alpha_j),\ j = 1, \cdots, q, & \pi(\phi) &= \mathbf{I}_{(-1,1)}(\phi), & \pi(\sigma) &= \mathbf{I}_{(0,C_\sigma)}(\sigma), \end{aligned}$$

where $C > 0$, $C_\alpha > 0$, and $C_\sigma > 0$ are (known) hyperparameters, as well as the mean vector $\mathbf{b}_0$ and the covariance matrix B_0.

2.3 Model selection between OSV models

Reasonable model specifications will be models where credible intervals do not contain the zero for all parameters. However model selection among such reasonable models is difficult since the likelihood cannot be evaluated simply for OSV models, thus the often used deviance information criteria (DIC) of Spiegelhalter et al. (2002) or score measures discussed in Gneiting and Raftery (2007) cannot be computed directly. Therefore we consider the following simple model selection criteria.

To choose among OSV models we first derive estimates of the ordinal categories for each t_i based on the MCMC iteration values. Note that the hidden observed volatilities are individually updated for each t_i, but it would be prohibitive to store all MCMC iterates for $h^*_{t_i}$, since the number of time points is too large in the applications considered. Therefore we keep only the average value of the hidden log volatilities at t_i over all MCMC iterations. These averages we denote by $\hat{h}^*_{t_i}$. These are used to derive fitted hidden process values. Let $\boldsymbol{\beta}^r, \boldsymbol{\alpha}^r, \sigma^r, \phi^r$ and $c^r_k, k = 2, \cdots, K-1$ the rth MCMC iterate of $\boldsymbol{\beta}, \boldsymbol{\alpha}, \sigma, \phi$ and $c_k, k = 2, \cdots, K-1$, respectively for $r = 1, \cdots, R$. The average hidden log volatilities $\hat{h}^*_{t_i}$ give now rise to fitted hidden process variables $y^{*r}_{t_i}$ defined by

$$y^{*r}_{t_i} := \mathbf{x}'_{t_i}\boldsymbol{\beta}^r + \exp(\hat{h}^*_{t_i}/2)\epsilon^{*r}_{t_i},$$

where $\epsilon_{t_i}^{*r}$ are i.i.d. standard normal observations. Finally find category k such that $y_{t_i}^{*r} \in [c_{k-1}^r, c_k^r)$ and set

$$y_{t_i}^r := k.$$

The ordinal category at time t_i is now fitted by the empirical median of $\{y_{t_i}^r, r = 1, \cdots, R\}$, which we denote as $\hat{y}_{t_i}$.

To construct interval estimates for the ordinal categories we define

$$\begin{aligned} y_{t_i,1-\alpha}^{*r} &:= \mathbf{x}_{t_i}'\boldsymbol{\beta}^r + \exp(\hat{h}_{t_i}^*/2) z_{1-\alpha} \\ y_{t_i,\alpha}^{*r} &:= \mathbf{x}_{t_i}'\boldsymbol{\beta}^r - \exp(\hat{h}_{t_i}^*/2) z_{\alpha}, \end{aligned}$$

where z_β is the β quantile of a standard normal distributed random variable. Then we find categories $k_{1-\alpha}$ such that $y_{t_i,1-\alpha}^{*r} \in [c_{k-1}^r, c_k^r)$ and k_α such that $y_{t_i,\alpha}^{*r} \in [c_{k-1}^r, c_k^r)$, respectively, and set

$$y_{t_i,1-\alpha}^r := k_{1-\alpha} \text{ and } y_{t_i,\alpha}^r := k_\alpha.$$

The interval estimate for a category at a time t_i is now defined as the interval $[\hat{y}_{t_i,\alpha}, \hat{y}_{t_i,1-\alpha}]$ where $\hat{y}_{t_i,\alpha}(\hat{y}_{t_i,1-\alpha})$ is the empirical median of $\{y_{t_i,\alpha}^r, r = 1, \cdots, R\}(\{y_{t_i,1-\alpha}^r, r = 1, \cdots, R\})$.

Alternatively we could consider a $100(1-\alpha)\%$ credible interval given by $[\hat{y}_{t_i,\alpha}^B, \hat{y}_{t_i,1-\alpha}]$, where $\hat{y}_{t_i,\alpha}^B(\hat{y}_{t_i,1-\alpha}^B)$ is the empirical α quantile ($(1-\alpha)$ quantile) of $\{y_{t_i}^r, r = 1, \cdots, R\}$. Since the fitted category $y_{t_i}^r$ of the rth MCMC iterate takes on only a few values, the empirical α and $(1-\alpha)$ quantiles are not well defined. Therefore we will not follow this approach.

To choose among several OSV specifications we now count the times the observed category coincides with the fitted category as well as how many times the interval estimate covers the observed category. We choose the model with the highest correctly fitted and covered categories as the best model. Note that the observed coverage percentage is not identical with $100(1-\alpha)$ for the α value used in the construction of the interval estimates, since category values for different time points are dependent.

2.4 Model selection between SV models

A similar approach as for the OSV models is followed for the SV models. First let $\hat{h}_{t_i}^c$ denote the average value of all log volatilities over all MCMC iterations. Again let $\boldsymbol{\beta}^r, \boldsymbol{\alpha}^r, \sigma^r, \phi^r$ and $c_k^r, k = 2, \cdots K-1$ the rth MCMC iterate of $\boldsymbol{\beta}, \boldsymbol{\alpha}, \sigma, \phi$ and $c_k, k = 2, \cdots, K-1$ for $r = 1, \cdots, R$ for the SV model, respectively. Define

$$\begin{aligned} y_{t_i}^{c,r} &:= \mathbf{x}_{t_i}'\boldsymbol{\beta}^r + \exp(\hat{h}_{t_i}^c/2)\epsilon_{t_i}^r \\ y_{t_i,1-\alpha}^{c,r} &:= \mathbf{x}_{t_i}'\boldsymbol{\beta}^r + \exp(\hat{h}_{t_i}^c/2) z_{1-\alpha} \\ y_{t_i,\alpha}^{c,r} &:= \mathbf{x}_{t_i}'\boldsymbol{\beta}^r - \exp(\hat{h}_{t_i}^c/2) z_{\alpha}, \end{aligned}$$

where $\epsilon_{t_i}^r$ are i.i.d. standard normal. Now determine the median of $\{y_{t_i}^{c,r}, r = 1, \cdots, R\}$, $\{y_{t_i,\alpha}^{c,r}, r = 1, \cdots, R\}$ and $\{y_{t_i,\alpha}^{c,r}, r = 1, \cdots, R\}$, and denote them by $\hat{y}_{t_i}^c, \hat{y}_{t_i,\alpha}^c$ and $\hat{y}_{t_i,1-\alpha}^c$, respectively. Since $\hat{y}_{t_i}^c$ is real-valued, it is not informative to count the times the observed value is equal the fitted value $\hat{y}_{t_i}^c$ for all t_i, we only count the number of times the observed value is covered by the interval $[\hat{y}_{t_i,\alpha}^c, \hat{y}_{t_i,1-\alpha}^c]$ for all t_i.

2.5 Model selection between OSV and SV models

The coverage percentage by the interval estimate for the OSV and SV, respectively, is used as a measure how good the model explains the observed values. A larger percentage gives a better fit.

		Minimum	Median	Maximum
FMT	number of trades per day	28.00	52.50	133.00
	price	2.44	4.31	5.31
	volume per trade	100.00	1000.00	122400.00
	price difference between t_{i-1} and t_i	-0.25	0.00	0.12
	time difference between t_{i-1} and t_i	0.00	192.00	4001.00
	number of quotes between t_{i-1} and t_i	0.00	1.00	24.00
Agilent	number of trades per day	796.00	1348.00	2056.00
	unit price	38.06	46.19	53.94
	volume per trade	100.00	500.00	247000.00
	price difference between t_{i-1} and t_i	-0.69	0.00	0.50
	time difference between t_{i-1} and t_i	0.00	11.00	276.00
	number of quotes between t_{i-1} und t_i	0.00	1.00	14.00
IBM	number of trades per day	1937.00	2478.00	3091.00
	unit price	91.62	99.44	104.30
	volume per trade	100.00	1000.00	225000.00
	price difference between t_{i-1} and t_i	-0.81	0.00	0.88
	time difference between t_{i-1} and t_i	0.00	7.00	150.00
	number of quotes between t_{i-1} and t_i	0.00	1.00	24.00

Table 1: Observed characteristics of the FMT, Agilent and IBM stocks between Nov. 1 - 30, 2000

3 Application

3.1 Data

To investigate the gain of the OSV model over a corresponding SV model for the price changes we selected 3 stocks traded at the NYSE, reflecting stocks which are traded at a low, medium and high level. In particular we chose the Fremont General Corporation (FMT), the Agilent Technologies (Agilent) and the International Business Machine Cooperation (IBM) from the TAQ data base for a low, medium and high level of trading, respectively. The data was collected between November 1-30, 2000 excluding November 23, 24 (thanksgiving) and between 9:30am until 4 pm to avoid opening and closing effects.

Table 1 gives trading characteristics for the three stocks during the investigated time period. It illustrates the different levels of trading activity. Further the absolute size of extremal price changes is increased as trading activity is increased, indicating a higher volatility for higher traded stocks. As expected the median time between trades decreases as the level of trading increases. The same is true for the maximum time between trades. For the number of quotes between trades we see a different behavior; while the medium number of quotes remains constant, the maximal number of quotes is the same for low and high trading stocks, while it is lower for medium traded stocks. A similar behavior is observed for volume.

To illustrate the severe discreteness of the observed price changes we recorded the number of occurrences of tick changes of size $\leq -\frac{3}{16}, -\frac{2}{16}, -\frac{1}{16}, 0, \frac{1}{16}, \frac{2}{16}, \geq \frac{3}{16}$ together with their percentages. For each of the tick change size we associate a category label necessary for the OSV formulation also given in Table 2. We observe that the observed price changes are quite symmetric around 0 during the investigated time period and that a zero price change is observed most often.

	price difference	<= −3/16	−2/16	−1/16	0	1/16	2/16	>=3/16
FMT	category	1	2	3	4	5	6	7
	frequency	3	25	229	755	227	28	0
	percent	0.002	0.019	0.181	0.596	0.179	0.023	0
Agilent	category	1	2	3	4	5	6	7
	frequency	196	939	4662	16599	4747	863	216
	percent	0.007	0.033	0.165	0.588	0.168	0.031	0.008
IBM	category	1	2	3	4	5	6	7
	frequency	585	3090	10251	22286	11161	2546	613
	percent	0.012	0.061	0.203	0.441	0.221	0.05	0.012

Table 2: Observed price changes together with category label, frequency and percent for the FMT, Agilent and IBM stocks from Nov. 1-30, 2000

The considered OSV and SV models allow for covariates on the mean and volatility level. To get an idea of possible day time effects we record the observed median values of the number of trades, price, volume, price change and time between trades (see Table 3), respectively. All stocks show larger (smaller) time intervals between trades during midday (opening and closing times), however the median price change is constant over the day time indicating no effect on the mean level of the hidden process. With regard to the volatility we also recorded the minimal and maximal price changes during trading hours in Table 4. Here we see less volatility changes for different trading hours for FMT and Agilent stocks compared the the IBM stock. This indicates a possible day time effect on the volatility level for IBM stocks, which is detected by a corresponding OSV model specification.

Comparing Table 3 with Table 4 we might identify covariates on the volatility level. For example the median volume value exhibits a similar pattern as the pattern of volatility changes for the FMT and IBM stocks, indicating that volume has some explanatory power for the volatility of the price changes. For Agilent stocks the patterns of volume and volatility of the price changes do not match as well. For the other covariates the identification is less pronounced, so we consider them all as potentially useful covariates and let the statistical models used later identify them.

	day time	9:30-10	10-11	11-12	12-1	1-2	2-3	3-4
FMT	no. of trades	120.00	412.00	366.00	322.00	344.00	394.00	530.00
	price	4.50	4.19	4.25	4.25	4.50	4.62	4.19
	volume	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00
	price change	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	time diff.	89.00	176.00	210.00	256.00	182.50	208.00	165.00
	no. of quotes	1.00	1.00	1.00	2.00	1.00	1.00	1.00
Agilent	no. of trades	2543.00	4860.00	4260.00	3474.00	3461.00	3974.00	4898.00
	price	46.50	46.56	46.12	46.25	45.88	46.12	46.25
	volume	600.00	600.00	500.00	500.00	500.00	500.00	500.00
	price change	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	time diff.	7.00	10.00	11.00	12.00	12.00	11.00	10.00
	no. of quotes	1.00	1.00	1.00	1.00	1.00	1.00	1.00
IBM	no. of trades	4935.00	9165.00	7450.00	5643.00	5731.00	7479.00	9143.00
	price	99.19	99.56	99.50	99.69	99.69	99.44	99.25
	volume	1300.00	1000.00	800.00	600.00	700.00	800.00	1000.00
	price change	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	time diff.	6.00	6.00	7.00	9.00	9.00	7.00	6.00
	no. of quotes	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3: Observed median number of trades, price, volume, price change, time between trades and number of quotes between trades for different trading hours of the FMT, Agilent and IBM stock between Nov. 1 -30, 2000

	day time	9:30-10	10-11	11-12	12-1	1-2	2-3	3-4
FMT Min.	price change	-0.12	-0.12	-0.12	-0.25	-0.19	-0.12	-0.12
Max.	price change	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Agilent Min.	price change	-0.38	-0.31	-0.25	-0.69	-0.31	-0.25	-0.50
Max.	price change	0.50	0.38	0.31	0.31	0.25	0.44	0.44
IBM Min.	price change	-0.81	-0.50	-0.25	-0.56	-0.25	-0.31	-0.50
Max.	price change	0.88	0.62	0.31	0.62	0.25	0.38	0.56

Table 4: Minimal and maximal price changes for different trading hours of the FMT, Agilent and IBM stock between Nov. 1-30, 2000

3.2 OSV models

As response we choose the category corresponding to the price change at trading time t_i denoted by y_{t_i}. Since the current price is highly dependent on the previous price we model this dependency by letting the current price change category depend also on its previous one. This is the only significant covariate on the mean level and we denote it by LAG1. We allow for an intercept parameter on the mean level. For possible covariates on the volatility level we use volume (V), daytime (D), time elapsed between trades (T) and the number of quotes between trades (Q). For numerical stability we use centered and standardized versions of these variables in our analysis. Further no intercept term is included for the linear predictor $\mathbf{z}'_{t_i}\boldsymbol{\alpha}$ to avoid non identifiability.

For all three stocks we run a variety of models involving V,D,T and Q as well as quadratic functions of these. In the following we only present models where all covariates are significant, i.e. their individual 80% credible interval does not contain zero. For all models we run 20000 MCMC iterations of the GM-MGMC algorithm. Appropriate burnin values were determined using trace plots. Further the estimated autocorrelations among the MCMC iterations suggested a subsampling of every 20th iteration.

Fremont General Cooperation

The left panel of Table 5 presents the estimated posterior means and medians of each parameter together with a 80% credible interval for the subsampled MCMC iterations after burnin. Estimated posterior densites for all $OSV(1, LAG1; V, T)$ parameters given in Figure 2. We see symmetric behavior of the posteriors for the cutpoint parameters and regression parameters and skewed distributions for σ and ϕ. The posterior density estimates for the remaining two OSV specification show a similar behavior and are therefore obmitted.

Interpreting the results for the OSV specifications, we see that a higher previous price change category decreases the probability of a lower current price change category compared to the case where a lower previous price change category is observed. A higher volume, a larger time interval between trades and a larger number of quotes increase the hidden log volatility, thus the probability of observing an extreme positive or negative price change is increased.

It remains to choose among these OSV specifications. Since $OSV(1, LAG1; V, T)$ ($OSV(1, LAG1; T, Q)$) is nested within $OSV(1, LAG1; V, T, Q)$, the significance of the parameter estimates established by the credible intervals indicate that $OSV(1, LAG1; V, T, Q)$ is the preferred model specification. This is also confirmed, when we consider first the fitted price change categories (see Section 2.3) and compare them to the observed price categories. Secondly we determine fitted interval bounds for the price change category and check how many times they are covering the observed price change category. The percentage of correctly fitted categories is $59.67\%, 59.43\%$ and 59.75% for the $OSV(1, LAG1; V, T), OSV(1, LAG1; T, Q)$ and $OSV(1, LAG1; V, T, Q)$ model, respectively. The corresponding values for the percentage of correctly covered categories are $96.45\%, 96.37\%$ and 96.53%, respectively. This shows a slight preference for the large model.

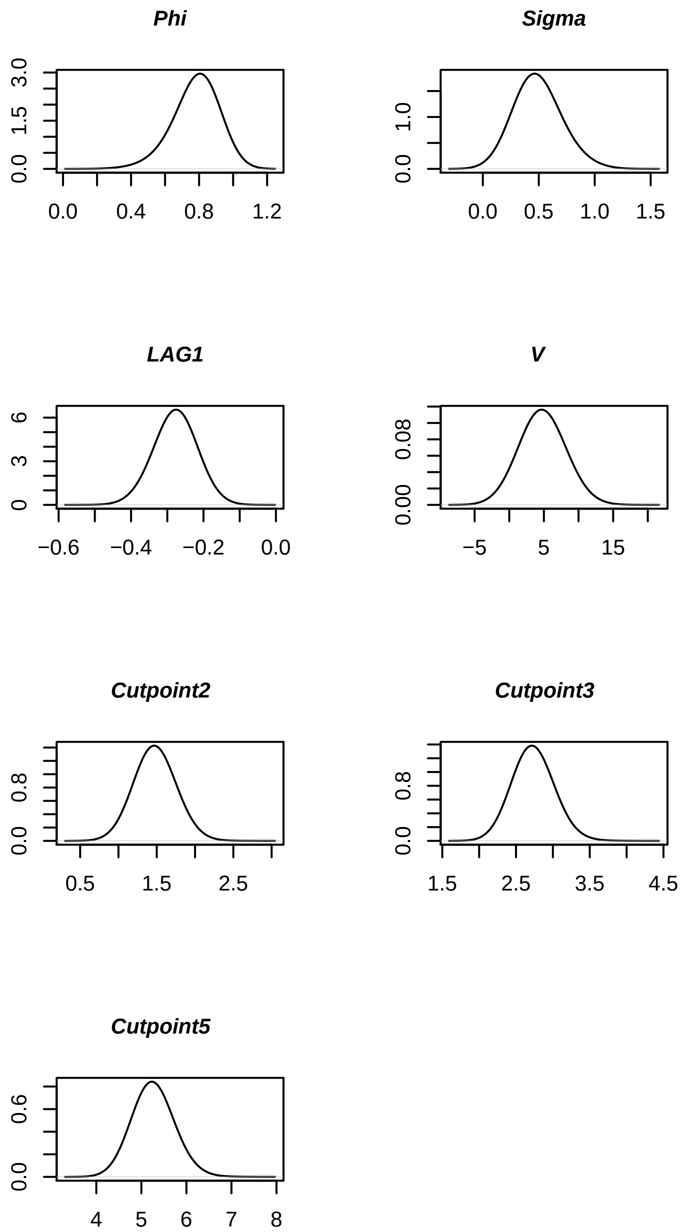

Figure 2: Estimated posterior density for $OSV(1, LAG1; V, T)$ parameters for FMT stocks

	$OSV(1, LAG1; V, T)$				$SV(1; V, T)$			
Parameter	10%	90%	Median	Mean	10%	90%	Median	Mean
ϕ	0.64	0.89	0.80	0.78	0.75	0.79	0.77	0.77
σ	0.32	0.70	0.47	0.49	9.17	11.86	9.90	10.21
c_2	1.25	1.72	1.47	1.48				
c_3	2.50	2.98	2.72	2.73				
c_4	3.73	4.29	3.99	4.00				
c_5	4.88	5.68	5.24	5.26				
$\mathbf{1}$	4.11	4.86	4.46	4.47	1.1^{-6}	7.1^{-6}	4.1^{-6}	4.1^{-6}
LAG1	−0.33	−0.23	−0.28	−0.28				
V	1.86	7.83	4.66	4.79	14.14	24.96	19.48	19.51
T	3.45	9.46	6.32	6.39	25.09	36.75	31.09	31.04
	$OSV(1, LAG1; V, Q)$				$SV(1; V, Q)$			
ϕ	0.44	0.85	0.74	0.69	0.75	0.79	0.77	0.77
σ	0.39	0.93	0.58	0.63	9.18	11.82	9.94	10.19
c_2	1.47	2.00	1.72	1.73				
c_3	2.78	3.42	3.07	3.08				
c_4	4.01	4.73	4.34	4.36				
c_5	5.21	6.20	5.65	5.68				
$\mathbf{1}$	4.45	5.38	4.87	4.89	1.1^{-6}	7.1^{-6}	4.1^{-6}	4.1^{-6}
LAG1	−0.34	−0.24	−0.30	−0.29				
V	1.49	7.57	4.76	4.64	13.67	25.19	19.28	19.44
Q	1.72	6.84	4.24	4.27	15.44	26.29	21.36	21.10
	$OSV(1, LAG1; V, T, Q)$				$SV(1; V, T, Q)$			
ϕ	0.72	0.91	0.83	0.82	0.76	0.79	0.77	0.77
σ	0.26	0.58	0.40	0.42	9.12	11.71	9.88	10.11
c_2	1.01	1.58	1.24	1.27				
c_3	2.14	2.84	2.42	2.46				
c_4	3.38	4.13	3.71	3.74				
c_5	4.52	5.45	4.94	4.96				
1	3.77	4.65	4.18	4.19	1.1^{-6}	7.1^{-6}	4.1^{-6}	4.1^{-6}
LAG1	−0.33	−0.22	−0.27	−0.27				
V	1.87	7.54	4.58	4.59	13.82	24.53	19.71	19.35
T	1.55	6.42	4.02	4.00	15.06	26.38	20.73	20.74
Q	3.63	8.93	6.26	6.30	25.13	36.67	31.00	30.98

Table 5: Estimated posterior means, medians and quantiles of three OSV (left panel) and three SV (right panel) model specifications with significant parameters fitted for FMT stocks based on the subsampled MCMC iterations after burnin

Agilent Technologies

For the Agilent stock we found only a single OSV specification with significant parameter estimates, whose summary statistics are given in the left panel of Table 6. It is a different specification as for FMT stocks. While the effect of the previous price change category for the Agilent stocks are similar to the one from the FMT stocks, the size of σ is larger and the hidden log volatilities are stronger autocorrelated. A notable difference is the effect of the number of quotes between trades on the price change categories. Here the parameter estimate has a negative sign thus the probability of extreme price change categories is decreased when the number of quotes is increased.

	$OSV(1, LAG1; T, Q)$				$SV(1; T)$			
Parameter	10%	90%	Median	Mean	10%	90%	Median	Mean
ϕ	0.80	0.84	0.82	0.82	.74	.80	.77	.77
σ	0.46	0.52	0.49	0.49	9.12	10.37	9.67	9.72
c_2	1.28	1.33	1.30	1.30				
c_3	2.24	2.31	2.28	2.28				
c_4	3.42	3.52	3.47	3.47				
c_5	4.39	4.52	4.46	4.46				
c_6	5.36	5.56	5.47	5.47				
$\mathbf{1}$	3.68	3.81	3.74	3.74	1.1^{-6}	7.1^{-6}	4.1^{-4}	3.6^{-6}
LAG1	-0.23	-0.21	-0.22	-0.22				
T	22.12	27.43	24.78	24.80	62.57	181.79	121.86	121.89
Q	-10.11	-5.25	-7.64	-7.64				

Table 6: Estimated posterior means, medians and quantiles of the $OSV(1, LAG1; T, Q)$ (left panel) and $SV(1; T)$ fitted for Agilent stocks based on subsampled MCMC iterations after burnin

.

International Business Machines Cooperation

For the highly traded IBM stocks we have two OSV model specification where all parameter estimates are significant (see the left panel of Table 7). The effect of the number of quotes is similar to the medium traded Agilent stock. The full specification also includes significant negative daytime parameter, indicating a lower probability of extreme price change categories for later in the day than in the morning. However the percentage of correctly fitted response categories is 41.54 % for the $OSV(1, LAG1; V, Q)$ model compared to 41.48 % for the $OSV(1, LAG1; V, T, Q, D)$ model. Also the percentage of correctly covered response categories is 92.94 % for the for the $OSV(1, LAG1; V, Q)$ model compared to 92.93 % for the $OSV(1, LAG1; V, T, Q, D)$ model. Both model fit measures show a slight preference for the simpler OSV model specification.

In summary, we see different OSV model specification occur for the different stocks. While there is a negative parameter estimate for the number of quotes between for medium and highly traded stocks, the opposite is true for the less traded FMT stock. Therefore the probability of extreme price changes is decreased for high and medium level traded stocks when the number of quotes increases between trades, while increased for low level traded stocks. In addition, the level of trading influences the magnitude of autocorrelation present in the hidden log volatilities. It increases as the level of trading increases. Daytime effects on the hidden volatility for the price change categories are neglible in the stocks we have investigated. The effect of time elapsed between trades also depends on the level of trading. A higher trading level decreases the size of the corresponding regression parameter for the hidden log volatility, which induces that the probabilities of extreme price changes is decreased for larger time intervals. The effect of the trading level on the influence of volume is less pronounced. The positive regression coefficient for volume in the hidden log volatilities induces a larger hidden volatility for larger volume trades, which results in higher probabilities for the occurrence of extreme price change categories. This effect is somewhat increased when the trading level is increased.

3.3 SV models

For the SV setup we use the observed price changes as response and ignore their discrete nature. For each of the three stocks we investigated different SV specification. A first difference to the OSV specifications are that none of the covariates LAG1, V, T, Q, and D for the mean level are significant. Therefore all SV models include only an intercept parameter, which is significant but very close to zero. For the log volatilities we find significant covariates, which we present in the following. Again we run 20000 MCMC iterations and determine appropriate burnin values and subsampling rates.

Fremont General Corporation

Three significant SV specifications were found for the FMT stocks and the results are summarized in the right panel of Table 5. The highest coverage percentage is achieved using the $SV(1; V, T, Q)$, which we select as best model among the SV models for the FMT stocks.

Agilent Technologies

For the Agilent stocks only a single SV specification produces significant parameter estimates and the results are presented in the right panel of Table 6. From this we see that only the time elapsed between trades has a significant effect on the price changes. A larger time interval between trades produces a larger volatility, i.e. extreme price changes become more likely.

International Business Machines Cooperation

For the highly traded only the $SV(1; V, T)$ produces significant posterior parameter estimates. The results presented in right panel of Table 7 show that both volume and time elapsed between trades increase the volatility, thus making more extreme price changes more likely.

	$OSV(1, LAG1; V, Q)$							
Parameter	10%	90%	Median	Mean	10%	90%	Median	Mean
ϕ	0.93	0.94	0.94	0.94				
σ	0.20	0.23	0.21	0.21				
c_2	0.93	0.95	0.94	0.94				
c_3	1.65	1.69	1.67	1.67				
c_4	2.52	2.57	2.54	2.54				
c_5	3.33	3.40	3.37	3.37				
c_6	4.10	4.21	4.16	4.16				
1	3.04	3.12	3.08	3.08				
LAG1	−0.25	−0.24	−0.24	−0.24				
V	5.54	10.25	7.98	7.98				
Q	−9.17	−5.07	−7.12	−7.12				
	$OSV(1, LAG1; V, T, Q, D)$				$SV(1; V, T)$			
ϕ	0.92	0.94	0.93	0.93	.67	.69	.68	.68
σ	0.21	0.24	0.22	0.23	.07	.14	.09	.10
c_2	−8.37	−3.90	0.92	0.91				
c_3	−35.08	−22.16	1.68	1.68				
c_4	0.89	0.93	2.53	2.53				
c_5	1.66	1.70	3.38	3.38				
c_6	2.50	2.56	4.20	3.38				
1	3.04	3.12	3.08	3.08	1.9^{-5}	7.5^{-4}	3.9^{-4}	3.9^{-4}
LAG1	−0.25	−0.24	−0.24	−0.24				
V	5.22	9.76	7.43	7.46	.54	9.6	4.91	4.99
T	33.36	38.85	36.08	36.02	23.15	37.59	29.28	29.99
Q	−8.37	−3.90	−6.06	−6.07				
D	−35.08	−22.16	−28.12	−28.26				

Table 7: Estimated posterior means, medians and quantiles of two OSV (left panel) and one SV (right panel) model specifications with significant parameters fitted for IBM stocks based on recorded MCMC iterations

3.4 Comparison between OSV and SV models

We now compare by using the coverage percentages for all selected OSV and SV models given in Table 8. From the table we see that there is a clear preference for the OSV specifications for Agilent and IBM stocks, while for the FMT stock a slight preference for the SV specification is visible. A graphical illustration of this is given in Figure 3 where the interval estimates are plotted for the last 100 observation together with the observed values.

	OSV	SV
FMT	$OSV(1, LAG1; V, T, Q)$	$SV(1, LAG1; V, T, Q)$
	$1223/1267 = 96.53$ %	$1267/1267 = 100.00$%
Agilent	$OSV(1, LAG1; T, Q)$	$SV(1; T)$
	$26738/28222 = 94.74$ %	$20980/28222 = 74.34$ %
IBM	$OSV(1, LAG1; V, Q)$	$SV(1; V, T)$
	$46965/50532 = 92.94$ %	$42811/50532 = 84.72$ %

Table 8: Percentage of correctly covered observations of different OSV and SV specifications for FMT, Agilent and IBM stocks

As a final comparison we estimate posterior densities of the volatilities for each price change category using the competing OSV and SV specifications for all three stocks. As expected the OSV specification nicely identify different volatility patterns. In particular extreme price categories correspond to larger volatilities. In contrast the competing SV specifications are not able to identfy these patterns for FMT and Agilent stocks.

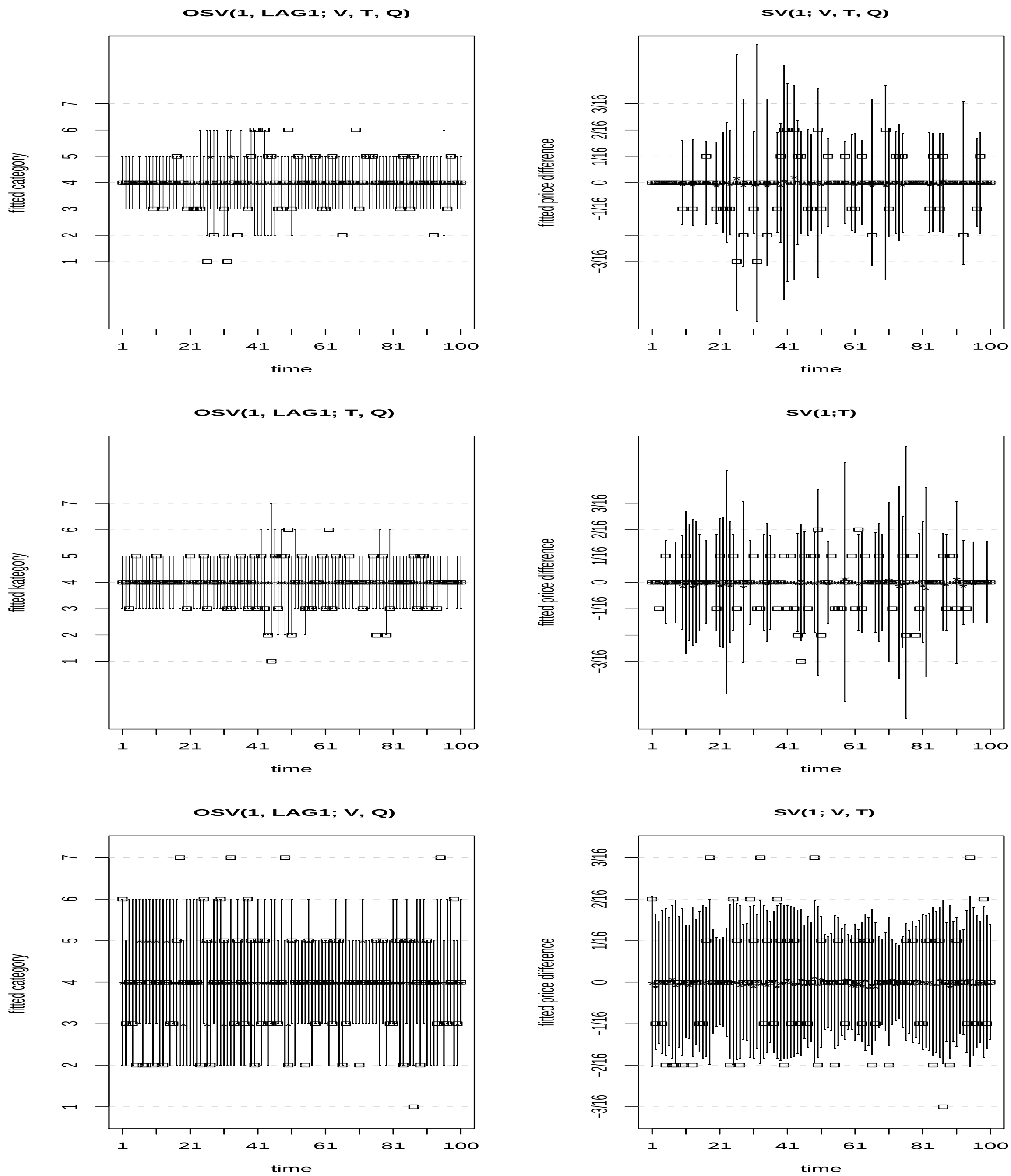

Figure 3: Fitted categories and fitted price differences of OSV and SV model of the last 100 observations together with interval estimates for FMT (top row), Agilent (middle row) and IBM (bottom row) stocks, respectively

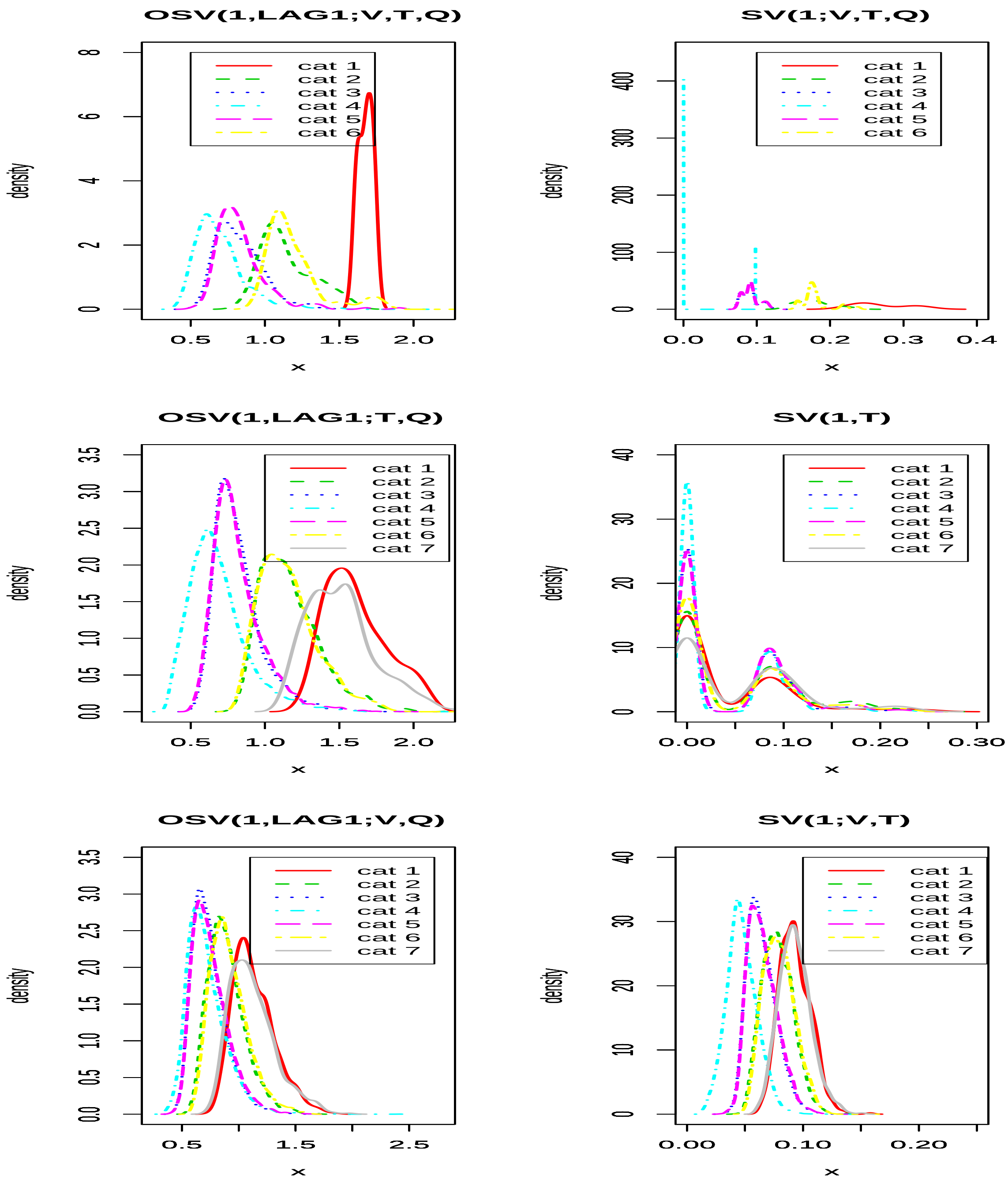

Figure 4: Estimated posterior densities of the (hidden) volatilities for each category of OSV and SV model for FMT (top row), Agilent (middle row) and IBM (bottom row) stocks, respectively

4 Summary and discussion

In this paper we presented the results of a Bayesian analysis of two model class specifications for financial price changes. Estimation is facilitated using MCMC methods. The OSV specification explicitly accounts for the discrete values of the price changes, while the SV specification ignores it. The OSV model identifies the previous price change, while the SV fails to identify this influence. In addition we see that volume, time between trades and the number of quotes between trades are important factors determining the volatility. Useful model specifications depend on the trading activity of the stock. In particular a higher number of quotes between trades increases the volatility for low traded stocks, while the opposite pattern is observed for stocks which are mediumly and highly traded. As expected a larger duration between trades increases the volatility. The coefficient of volume in the log volatility regression is lower than the one corresponding of time between trades, indicating that time between trades influences the volatility more than volume. For this comparison we need the fact that we use centered and standardized covariates. A quadratic day time effect was not significant indicating that a corresponding volatility smile was not present in the data.

When comparing the OSV and SV models we see that the OSV models perform better than the the SV models with regard to the coverage proportion of interval estimates. More precise model comparison criteria for comparing non nested models with numerical intractable likelihoods in a Bayesian setup are needed and subject to current research.

Finally, the OSV model specifications can clearly identify volatility differences between the different price change classes, while the SV specifications might fail to do so.

Overall we conclude that the discreteness of the price changes does matter and any useful model has to account for this discreteness. In the future we want to investigate the predictive capability of the OSV model.

Acknowledgement

Claudia Czado is supported by the *Deutsche Forschungsgemeinschaft*.

References

Chib, S., F. Nardari, and N. Shephard (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics 108*, 281–316.

De Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika 82*, 339–350.

Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration; a new model for irregularly spaced transaction data. *Econometrica 66*, 1127–1162.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic volatility, in *Statistical Methods in Finance* (eds. C.R.Rao and G.S.Maddala), 119-191. Amsterdam: North-Holland.

Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102*, 359–378.

Harris, L. (1990). Estimation of stock variances and serial covariances from discrete observations. *Journal of Financial and Quantitative Analysis 25*, 291–306.

Hasbrouck, J. (1999). Security bid/ask dynamics with discreteness and clustering. *Journal of Financial Markets 2*, 1–28.

Hasbrouck, J. (2000). The dynamics of discrete bid and ask quotes. *Journal of Finance 56*, 2109–2142.

Hasbrouck, J. (2004a). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of Financial and Quantitative Analysis 39*, 305–326.

Hasbrouck, J. (2004b). Markov Chain Monte Carlo methods for Bayesian estimation of microstructure models, computational appendix to "Liquidity in the futures pits: Inferring market dynamics from incomplete data". *Journal of Financial and Quantitative Analysis 39*, 305–326.

Hausman, J. A., A. W. Lo, and A. C. MacKinlay (1992). An ordered probit analysis of transaction stock prices. *Journal of Financial Economics 31*, 319–379.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies 65*, 361–393.

Kitagawa, G. (1987). Non-gaussian state space modeling of nonstationary time series. *Journal of the American Statistical Association 82*, 1032–1041.

Liesenfeld, R., I. Nolte, and W. Pohlmeier (2006). Modelling financial transaction price movements: a dynamic integer count data model. *Empirical Economics 30*, 795–825.

Manrique, A. and N. Shephard (1997). Likelihood analysis of a discrete bid/ask price model for a common stock quoted on the nyse. *Working Paper, Nuffield College, Oxford University*.

Müller, G. and C. Czado (2008). Stochastic volatility models for ordinal valued time series with application to finance. *Statistical Modelling*, to appear.

Russell, J. R. and R. F. Engle (2005). A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model. *Journal of Business and Economic Statistics 23*(2), 166–180.

Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: decomposition and models. *Journal of Financial Econometrics 1*, 2–25.

Shephard, N. (2006). Stochastic Volatility. In *New Palgrave Dictionary of Economics, 2nd ed.* (eds. S. Durlauf and L. Blume). Palgrave Macmillan.

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