Heavy-Tailed Innovations in the R Package stochvol

Gregor Kastner
WU Vienna University of Economics and Business

First version: May 15, 2015
This version: Aug 24, 2016

Abstract

We document how sampling from a conditional Student's t distribution is implemented in **stochvol**. Moreover, a simple example using EUR/CHF exchange rates illustrates how to use the augmented sampler. We conclude with results and implications.

Keywords: Student's t distribution, data augmentation, EUR/CHF exchange rates.

Preface

This note serves as a preliminary add-on to the more elaborate article "Dealing with Stochastic Volatility in Time Series using the R package **stochvol**" (Kastner 2016a). It discusses and relaxes the restriction to conditionally normal errors in the vanilla stochastic volatility (SV) model.

1. The SV model with Student's t errors

Several authors have suggested to use non-normal conditional innovation distributions for stochastic volatility modeling. Examples include the Student's t distribution (Harvey, Ruiz, and Shephard 1994), the extended Generalized Inverse Gaussian (Silva, Lopes, and Migon 2006), (semi-)parametric innovations (Jensen and Maheu 2010; Delatola and Griffin 2011), or the GH skew Student's t distribution (Nakajima and Omori 2012). In the following, we describe how the estimation of the SV model with Student's t errors is implemented in the R (R Core Team 2016) package **stochvol** (Kastner 2016b).

Let $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^\top$ be a vector of returns with mean zero. The SV model with Student's t errors (in short SV-t) is given through

$$\begin{aligned} y_t|h_t,\nu &\sim t_\nu(0, \exp h_t)\,, && (1)\\ h_t|h_{t-1},\mu,\phi,\sigma_\eta &\sim \mathcal{N}\big(\mu + \phi(h_{t-1}-\mu), \sigma_\eta^2\big)\,, && (2)\\ h_0|\mu,\phi,\sigma_\eta &\sim \mathcal{N}\big(\mu, \sigma_\eta^2/(1-\phi^2)\big)\,, && (3)\end{aligned}$$

i.e., conditionally on h_t, the data is assumed to follow a zero-mean non-standardized Student's t distribution with ν degrees of freedom and variance $(\nu \exp h_t)/(\nu-2)$ for $\nu > 2$. Following Chib, Nardari, and Shephard (2002), we assume that a priori the degrees of freedom parameter

$\nu \sim \mathcal{U}(a, b)$, i.e., follows a uniform distribution with support on the real interval (a, b). All other prior components are chosen as in Kastner (2016a).

2. Usage

Estimating a stochastic volatility model with conditional t errors via **stochvol** is very similar to estimating a model with standard Gaussian errors, differing only through specifying a non-`NA` argument `priornu`. This triggers the sampler specified in Section 3. To provide an example, we investigate the historical daily EUR/CHF exchange rates and display these in Figure 1.

```
R> library(stochvol)
R> data(exrates)
R> par(mfrow = c(2, 1), mar = c(1.7, 1.7, 1.7, 0.1), mgp = c(1.6, 0.6, 0))
R> plot(exrates$date, exrates$CHF, type = 'l', main = 'Price of 1 EUR in CHF')
R> dat <- logret(exrates$CHF, demean = TRUE)
R> plot(exrates$date[-1], dat, type = 'l', main = 'Demeaned log returns')
```

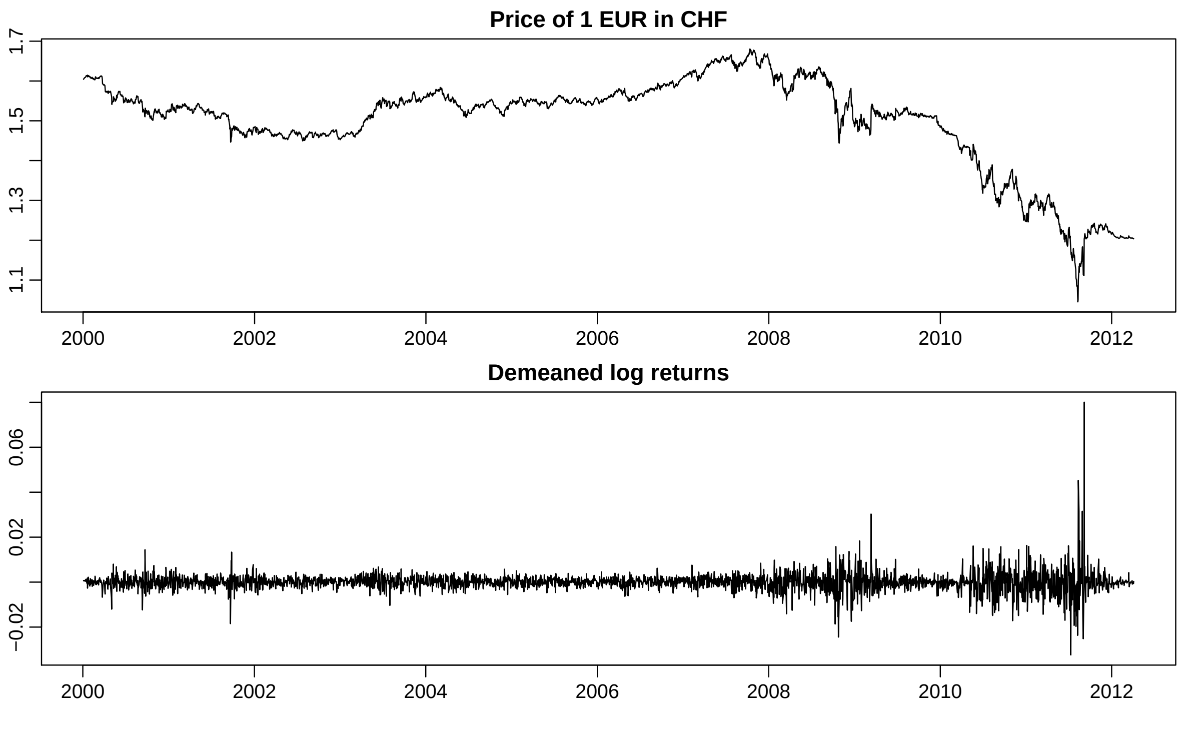

Figure 1: Levels and demeaned log returns of EUR in CHF.

By specifying the argument `priornu` (a two-element vector containing the lower and upper bounds of the uniform prior for ν), we can trigger the sampler to allow for heavy-tailed conditional innovations.

```
R> rest <- svsample(dat, priormu = c(-12, 1), priorphi = c(20, 1.1),
+    priorsigma = 0.1, priornu = c(2, 100), burnin = 2000)
R> plot(rest, showobs = FALSE)
```

Results are displayed in Figure 2, containing the output from the SV-t model. Row 1 depicts $\exp(h_t/2)$ for $t \in \{1, \ldots, n\}$; row 2 shows the time varying standard deviations given through $\sqrt{\nu/(\nu-2)}\exp(h_t/2)$ for $t \in \{1, \ldots, n\}$; row 3 portrays trace plots and row 4 outlines the corresponding smoothed kernel density estimates for the four parameters μ, ϕ, σ, and ν. It is worth noting that ν is estimated to lie between 6 and 18 with high posterior probability, indicating evidence for the presence of heavy tails even after catering for stochastic volatility. The extra flexibility of the SV-t sampler seems to allow for increased persistence ϕ and smaller variance of log-volatility σ^2, resulting in smoother time-varying volatility estimates.

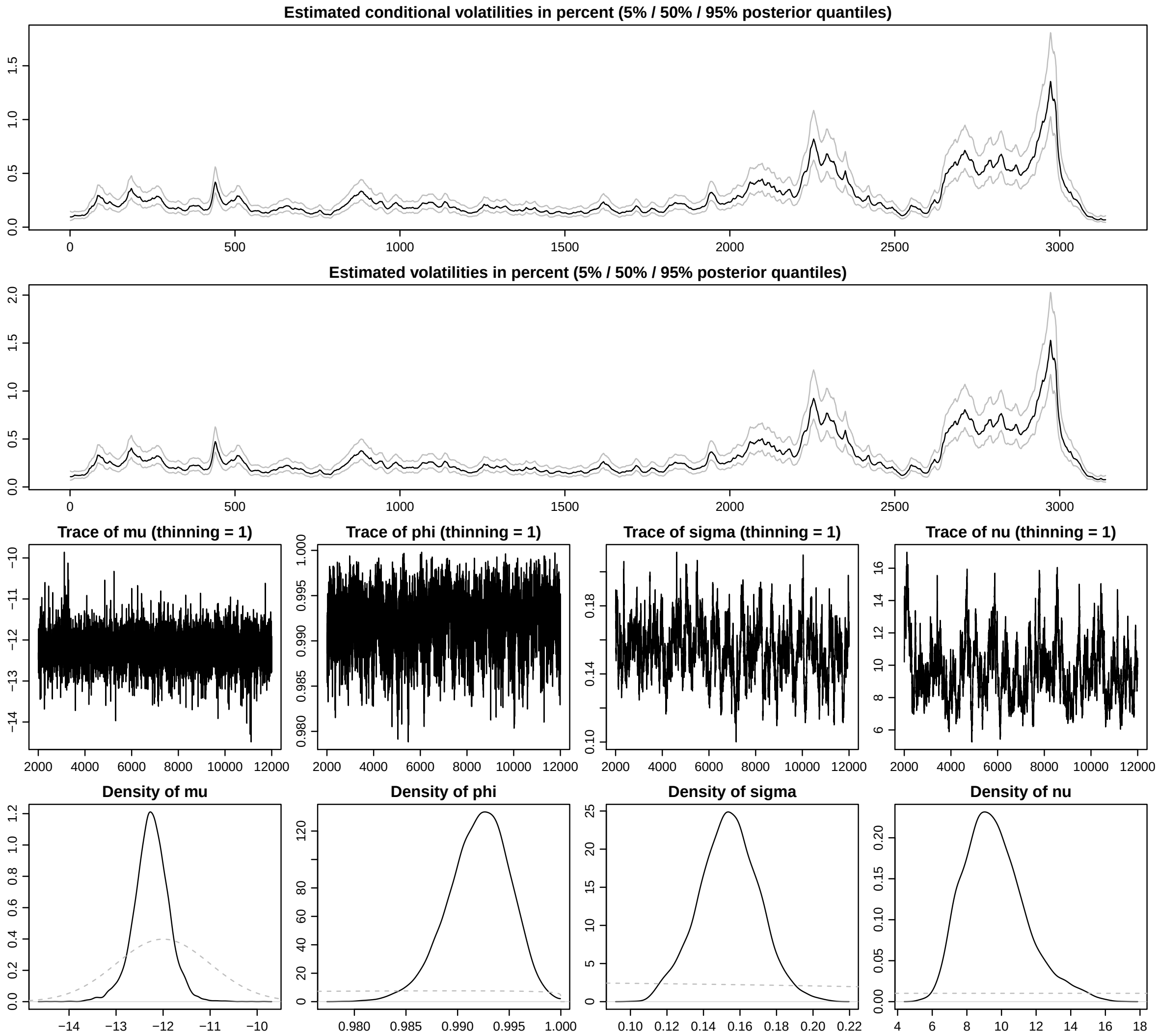

Figure 2: Standard output of the `plot` method when applied to an `svdraws` object containing posterior draws from an SV model with Student's t errors.

We investigate one-day-ahead out-of-sample predictive performance of an AR(1) model with (a) homoskedastic, (b) SV, (c) SV-t, (d) GARCH(1,1) errors, applied to the raw exchange rate data. Details about this procedure are provided in Chapter 5 of Kastner (2016a). The results, summarized in Figure 3, speak in favor of the SV-t model for this dataset.

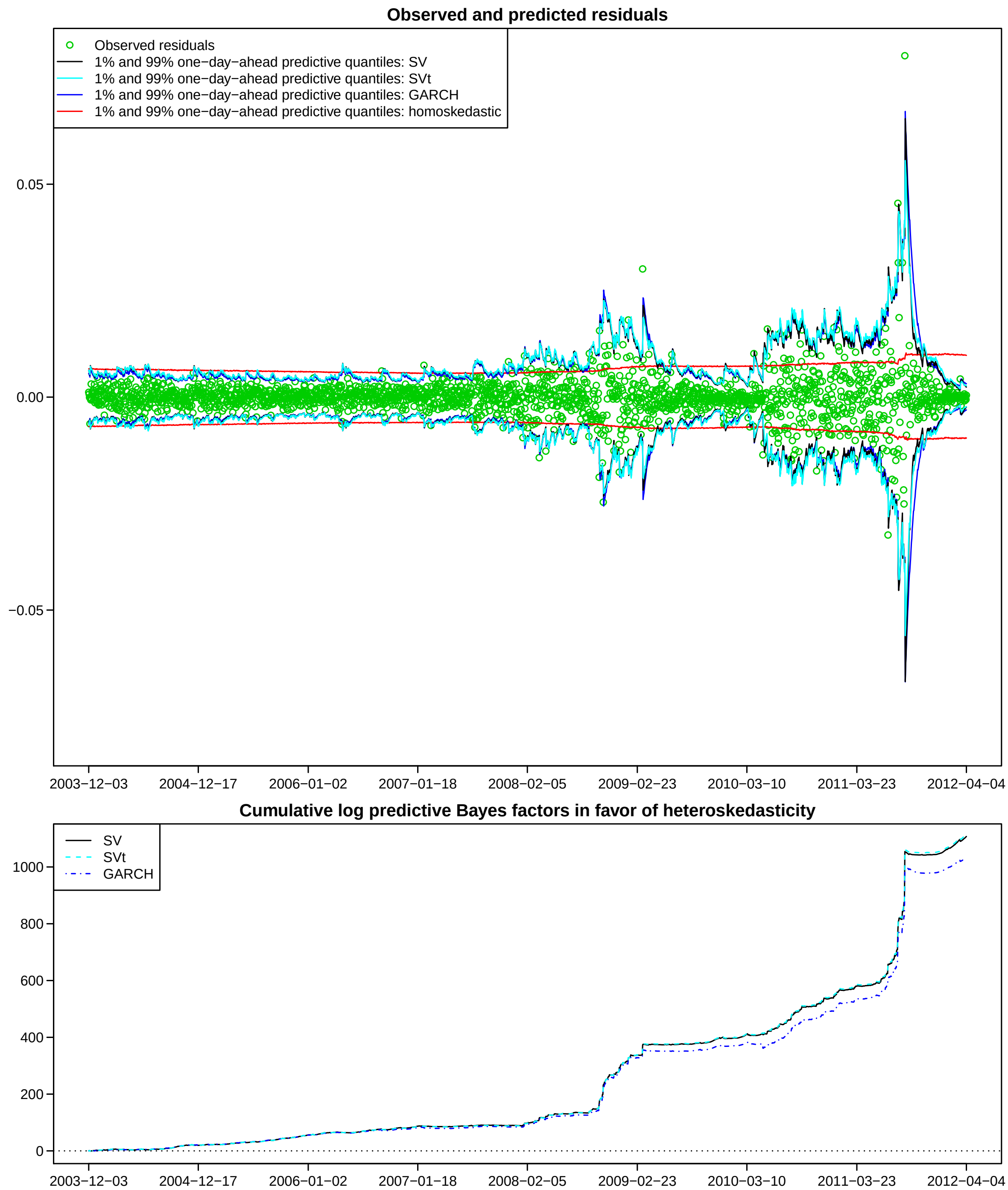

Figure 3: Log predictive one-day-ahead Bayes factors in favor of SV, SV-t, and GARCH errors over the homoskedastic model. The final log predictive Bayes factors aggregate to 1107.44 (SV), 1115.36 (SV-t), and 1033.53 (GARCH), respectively, thus providing strong evidence for the SV-t model.

3. Necessary modifications in the sampling scheme

The Student's t distribution appearing in Equation 1 can be conveniently expressed as a scale mixture of normal distributions,

$$\begin{aligned} y_t|h_t,\tau_t &\sim \mathcal{N}(0, \tau_t \exp h_t)\,, \\ \tau_t|\nu &\sim \mathcal{G}^{-1}(\nu/2, \nu/2)\,, \end{aligned}$$

where $\mathcal{N}\big(\mu, \sigma_\eta^2\big)$ denotes the normal distribution with mean μ and variance σ_η^2, and $\mathcal{G}^{-1}(a,b)$ denotes the inverse gamma distribution with shape and scale parameters a and b, respectively. Treating $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)^\top$ as latent data and letting $\tilde{y}_t = y_t/\sqrt{\tau_t}$ for $t \in \{1, \ldots, n\}$, we have

$$\tilde{y}_t|h_t \sim \mathcal{N}(0, \exp h_t)\,,$$

and the AWOL sampler described in Kastner and Frühwirth-Schnatter (2014) can directly be applied to the transformed data. To obtain draws from the newly introduced variables $\boldsymbol{\tau}$ and ν, two additional steps are required.

3.1. Sampling the auxiliary variables

It is easy to see that the marginal posterior is given through

$$\tau_t|y_t, h_t, \nu \sim \mathcal{G}^{-1}\left(\frac{\nu+1}{2}, \frac{\nu + y_t^2 \exp(-h_t)}{2}\right),$$

independently for each $t \in \{1, \ldots, n\}$. Obtaining draws from this distribution is straightforward.

3.2. Sampling the degrees of freedom parameter

The full conditional posterior of the degrees of freedom parameter, $\nu|\cdot$, only depends on $\boldsymbol{\tau}$. Its density is given through the product of n univariate truncated inverse gamma densities which may be written as

$$p(\nu|\cdot) = p(\nu|\boldsymbol{\tau}) \propto \left(\frac{\nu}{2}\right)^{n\nu/2} \Gamma\left(\frac{\nu}{2}\right)^{-n} \left(\prod_{t=1}^n \tau_t\right)^{-\nu/2} \exp\left\{-\frac{\nu}{2}\sum_{t=1}^n \frac{1}{\tau_t}\right\} \tag{4}$$

for $\nu \in (a,b)$ and zero elsewhere.

For obtaining draws from this distribution, we use an independence Metropolis-Hastings update. We follow Chib and Greenberg (1994), who introduced the idea of specifying an independence proposal through numerical maximization of the log-density. For the problem at hand, we consequently aim for optimizing

$$\log p(\nu|\boldsymbol{\tau}) = \frac{n\nu}{2}\log(\nu/2) - n\log\Gamma(\nu/2) - \frac{\nu}{2}\sum_{t=1}^n\left(\log\tau_t + \frac{1}{\tau_t}\right) + C, \tag{5}$$

with first and second derivatives given through

$$\begin{aligned} \frac{\partial \log p(\nu|\boldsymbol{\tau})}{\partial \nu} &= \frac{n}{2}\Big(1 + \log(\nu/2) - \psi^{(0)}(\nu/2)\Big) - \frac{1}{2}\sum_{t=1}^n\left(\log\tau_t + \frac{1}{\tau_t}\right), &(6) \\ \frac{\partial^2 \log p(\nu|\boldsymbol{\tau})}{\partial \nu^2} &= \frac{n}{2\nu} - \frac{n}{4}\psi^{(1)}(\nu/2), &(7) \end{aligned}$$

where $\psi^{(m)}$ denotes the polygamma function of order m. Using the above, it is easy to numerically find

$$\begin{aligned}\hat{\nu} &= \arg\max_{\nu} \log p(\nu|\boldsymbol{\tau}),\\ B_{\hat{\nu}} &= -1 \Bigg/ \frac{\partial^2 \log p(\nu|\boldsymbol{\tau})}{\partial \nu^2}\Bigg|_{\nu=\hat{\nu}},\end{aligned}$$

and a proposal candidate ν_{prop} may be drawn from a normal distribution with mean $\hat{\nu}$ and variance $B_{\hat{\nu}}$ (the Laplace approximation). Letting $\phi(x|\hat{\nu}, B_{\hat{\nu}})$ denote the corresponding density function, the acceptance probability is equal to $\min\{1, R\}$ with

$$R = \frac{p(\nu_{\text{prop}}|\boldsymbol{\tau})}{p(\nu_{\text{old}}|\boldsymbol{\tau})} \times \frac{\phi(\nu_{\text{old}}|\hat{\nu}, B_{\hat{\nu}})}{\phi(\nu_{\text{prop}}|\hat{\nu}, B_{\hat{\nu}})}.$$

4. Conclusion

We have shown how a simply data augmentation trick can be utilized to generalize the core sampler in **stochvol** in order to cater for potentially heavier-tailed innovation distributions. However, several caveats are called for:

- Even though the uniform prior for ν has been used widely, more robust alternatives are probably preferred, cf. Frühwirth-Schnatter and Pyne (2010) and the references therein.
- Sampling the degrees of freedom parameter ν conditionally on $\boldsymbol{\tau}$ can be very inefficient if ν becomes large (and thus the plain vanilla SV model suffices). We recommend to resort to the original SV sampler in this case.
- Leaving aside the additional computational burden, it is trivial to incorporate this extension into samplers employing **stochvol** as part of a larger MCMC scheme (e.g. Huber 2014; Kastner, Frühwirth-Schnatter, and Lopes 2014; Dovern, Feldkircher, and Huber 2015). Nevertheless, at the current stage of development, this should be conducted with caution by carefully investigating the convergence of the posterior draws.

Acknowledgments

The author would like to thank the attendants of the Spring 2015 Brown Bag Seminar of the Institute for Statistics and Mathematics, WU Vienna University of Economics and Business, in particular Sylvia Frühwirth-Schnatter, Mark Jensen, and Kurt Hornik, for stimulating comments and suggestions.

References

Chib S, Greenberg E (1994). "Bayes Inference in Regression Models with ARMA(p, q) Errors." *Journal of Econometrics*, **64**, 183–206. doi:10.1016/0304-4076(94)90063-9.

Chib S, Nardari F, Shephard N (2002). "Markov Chain Monte Carlo Methods for Stochastic Volatility Models." *Journal of Econometrics*, **108**, 281–316. doi:10.1016/S0304-4076(01)00137-3.

Delatola EI, Griffin JE (2011). "Bayesian Nonparametric Modelling of the Return Distribution with Stochastic Volatility." **6**, 901–926. doi:10.1214/11-BA632.

Dovern J, Feldkircher M, Huber F (2015). "Does Joint Modeling of the World Economy Pay Off? Evaluating GVAR Forecasts from a Multivariate Perspective." *Discussion Paper Series 590*, University of Heidelberg, Department of Economics. URL http://www.ub.uni-heidelberg.de/archiv/18586.

Frühwirth-Schnatter S, Pyne S (2010). "Bayesian inference for finite mixtures of univariate and multivariate skew-normal and skew-t distributions." *Biostatistics*, **11**(2), 317–336. doi:10.1093/biostatistics/kxp062.

Harvey AC, Ruiz E, Shephard N (1994). "Multivariate Stochastic Variance Models." *The Review of Economic Studies*, **61**(2), 247–264. doi:10.2307/2297980.

Huber F (2014). "Density Forecasting using Bayesian Global Vector Autoregressions with Common Stochastic Volatility." *Department of Economics Working Paper Series 179*, WU Vienna University of Economics and Business. URL http://epub.wu.ac.at/id/eprint/4280.

Jensen MJ, Maheu JM (2010). "Bayesian semiparametric stochastic volatility modeling." *Journal of Econometrics*, **157**(2), 306–316. doi:10.1016/j.jeconom.2010.01.014.

Kastner G (2016a). "Dealing with Stochastic Volatility in Time Series Using the R Package **stochvol**." *Journal of Statistical Software*, **69**(5), 1–30. doi:10.18637/jss.v069.i05.

Kastner G (2016b). ***stochvol**: Efficient Bayesian Inference for Stochastic Volatility (SV) Models*. R package version 1.2.3, URL http://CRAN.R-project.org/package=stochvol.

Kastner G, Frühwirth-Schnatter S (2014). "Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models." *Computational Statistics & Data Analysis*, **76**, 408–423. doi:10.1016/j.csda.2013.01.002.

Kastner G, Frühwirth-Schnatter S, Lopes HF (2014). "Analysis of Exchange Rates via Multivariate Bayesian Factor Stochastic Volatility Models." In E Lanzarone, F Ieva (eds.), *The Contribution of Young Researchers to Bayesian Statistics – Proceedings of BAYSM2013*, volume 63 of *Springer Proceedings in Mathematics & Statistics*, pp. 181–185. Springer-Verlag. doi:10.1007/978-3-319-02084-6_35.

Nakajima J, Omori Y (2012). "Stochastic Volatility Model with Leverage and Asymmetrically Heavy-Tailed Error Using GH Skew Student's t-Distribution." **56**(11), 3690–3704. doi:10.1016/j.csda.2010.07.012.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Silva RS, Lopes HF, Migon HS (2006). "The Extended Generalized Inverse Gaussian Distribution for Log-Linear and Stochastic Volatility Models." *Brazilian Journal of Probability and Statistics*, **20**(1), 67–91.

Affiliation:

Gregor Kastner
Institute for Statistics and Mathematics
Department of Finance, Accounting and Statistics
WU Vienna University of Economics and Business
Welthandelsplatz 1, Building D4, Level 4
1020 Vienna, Austria
Telephone: +43/1/31336-5593
Fax: +43/1/31336-90-5593
E-mail: `gregor.kastner@wu.ac.at`
URL: `http://statmath.wu.ac.at/~kastner/`