



TI 2008-011 /4

Tinbergen Institute Discussion Paper

Model-based Estimation of High Frequency Jump Diffusions with Microstructure Noise and Stochastic Volatility

Charles S. Bos

VU University Amsterdam, and Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

Model-based Estimation of High Frequency Jump Diffusions with Microstructure Noise and Stochastic Volatility

Charles S. Bos*
Tinbergen Institute and
VU University Amsterdam

January 21, 2008

Abstract

When analysing the volatility related to high frequency financial data, mostly non-parametric approaches based on realised or bipower variation are applied. This article instead starts from a continuous time diffusion model and derives a parametric analog at high frequency for it, allowing simultaneously for microstructure effects, jumps, missing observations and stochastic volatility.

Estimation of the model delivers measures of daily variation outperforming their non-parametric counterparts. Both with simulated and actual exchange rate data, the feasibility of this novel approach is shown. The parametric setting is used to estimate the intra-day trend in the Euro/U.S. Dollar exchange rate.

Keywords: High frequency; integrated variation; intra-day; jump diffusions; microstructure noise; stochastic volatility; exchange rates.

JEL classification: C11, C14, D53, E44

*Address: C.S. Bos, VU/FEWEB, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands. Tel/fax: +31 20 598 60 23/+31 20 598 60 20. Email: cbos@feweb.vu.nl.

1 Introduction

1.1 Motivation

With the advent of available intra-day financial time series data, interest in analysing these high frequency series sparked over the last decade. These high frequency time series offer the opportunity to study in more detail the evolution of volatility over time, improving on earlier variance estimates.

Even on a daily (or lower) frequency, time varying volatility is a major characteristic of financial time series. Moving to hourly or higher frequencies, several other important aspects arise. For intra-day data, characteristic findings include foremost, the time varying volatility, of a stochastic nature. Apart from that, intra-daily there is also time varying volatility, of a deterministic, seasonal, nature, related to time-of-day effects and opening/closing of markets. There exists micro-structure noise on the price process, and further unexplained apparent jumps, which are both especially observed with higher frequency data. Finally, for higher frequencies of data one can usually not escape the non-equal time spacing between successive observations, as markets are closed over weekends or holidays, or no quotes are available occur for thinly traded stocks and assets.

Treatment of this full set of characteristics of financial data occurs in a strand of literature which looks at realised variance, integrated variance, and related concepts (see for an overview the articles in Shephard (2005); more details in subsequent sections). This leads to the development of non-parametric methods to estimate the daily (integrated) volatility, and to a test for the occurrence of jumps within a specific day.

Though jumps in financial series are often not immediately explicable, at times they might also seem to be connected with announcements made on macro-variables, like on the future of the interest rate or the state of the unemployment. The precise link between such macro-announcements and the probability of a jump cannot be investigated in a non-parametric setting. The ultimate holy grail of this article will be to start with a parametric setting which, in the end, would allow for inclusion of explanatory variables in the jump process. This in turn would allow measuring the effectiveness and importance of certain announcements, with respect to the probability of a jump in the series and the related jump size.

For this purpose, a parametric technique is applied. After a literature overview, in Section

1.2, a diffusion model and its discretisation at high frequency is presented in Section 2, together with several concepts from the non-parametric approach which can serve as a baseline for comparison.

After a first look at the data in Section 3, the implementation of the Markov chain is discussed in Section 4, with a simulation exercise to investigate the properties of the algorithm in Section 5.1. An application on Euro/U.S. Dollar exchange rates follows in Section 6. Section 7 summarises and outlines the range of possibilities to continue this research.

1.2 Literature overview

The literature related to this article is large, and growing at a high rate, especially during the last decade. Without any intention of giving a complete overview, this section identifies the most important strands of literature, and how they are interrelated. It ends with a description of the contribution of this article to the literature.

Clearly, one major topic is stochastic volatility (SV). As an alternative to the generalised autoregressive conditional heteroskedasticity (GARCH) approach of modelling time varying volatility, allowing a separate source of uncertainty, it was introduced into the econometrics literature by Harvey et al. (1994) and Jacquier et al. (1994), though the ideas can be traced back to an article by Taylor (1982). The GARCH approach, including jumps, is presented in Drost et al. (1998), but in its original form dates back to Engle (1982) and Bollerslev (1986). The comparison between GARCH and SV is made in Kim et al. (1998), and also in Bos et al. (2000). An overview over the history of the SV model, and present research, is given in Shephard (2005).

Most stochastic volatility models are set up with the daily (or lower) frequency in mind. For higher frequencies, very quickly the switch to continuous time diffusion models is made. And if a price process can be considered as a continuous time process, observed in discrete (preferably evenly spaced) time points, then measures like realised volatility or bipower variation can be used to estimate the daily stochastic volatility.

Both Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) use intra-day data to help in estimating SV. Other articles by the these and many other authors (Andersen et al., 2002; Barndorff-Nielsen and Shephard, 2004b, 2006a,b) extend the knowledge on using high frequency returns for calculating integrated variance on a daily basis, in a ‘model-free’

manner.

One important aspect of using continuous time models, and their approximation through realised variance-type measures is the effect of microstructure noise or the occurrence of jumps on the estimates. Apart from the previously mentioned articles, Aït-Sahalia et al. (2005a) and Aït-Sahalia et al. (2005b) discuss the possible effects of microstructure noise on estimates of (integrated) volatility. Extracting the jumps in a non-parametric fashion is the topic of articles by Andersen et al. (2007), Barndorff-Nielsen and Shephard (2004b, 2006a), Haung and Tauchen (2005), or Lee and Mykland (2008), among many others. Eraker et al. (2003) use a discretisation of the diffusion model to a daily frequency, allowing for jumps and stochastic volatility, in a parametric fashion.

Parametric estimation of intra-day time series is seldomly done. Elerian et al. (2001) and Andersen et al. (2002) use simulated intra-day data to estimate daily returns of either exchange rates or the S&P500 index. The model is a discretisation of a diffusion process with SV, while the latter paper extends the model to include jumps. The first authors apply a Bayesian simulation method, in the second paper a simulated method of moments approach is used. The Bayesian approach is described in more detail in Jacquier et al. (1994) and Chib et al. (2002, 2006), and is also the methodology used in this article. Alternatively, one can use simulated maximum likelihood using importance sampling, as e.g. in Sandmann and Koopman (1998). All these articles effectively estimate daily SV models, not using intra-day information.

An exception to the rule that SV models do not use intra-day information is the article by Koopman et al. (2005), where the intra-day realised variance is used as an explanatory variable for estimating a daily SV model, so the intra-day information is used in an indirect manner. In Jungbacker and Koopman (2006) one day of second-by-second data is estimated at a time using a model-based approach including microstructure noise and stochastic volatility. No link is made between subsequent days. More purely intra-day and parametric is Rydberg and Shephard (2003), who model a larger sample of tick-by-tick data. Here the price process is modelled through the (discrete) steps up or down taken at each trade. A different approach is taken by Aït-Sahalia and Kimmel (2007), who use either an implied volatility proxy or option prices implying a volatility estimate through a pricing formula, to get to a closed form maximum likelihood estimator.

Table 1 about here

The interest in these diffusion or stochastic volatility models is driven by the interest in measuring, with as much precision as possible, the volatility in the market, for risk management, pricing options, monetary policy evaluation and making, and many other causes. These models also improve our understanding of the price process, and the microstructure foundations of it. Lyons (2001) discusses this approach to foreign exchange rate modelling. With the advent of high frequency data sets and the growth in computational possibilities, we can model and test ever finer structures. A good application of this microstructure approach is found in Andersen et al. (2003). On a daily basis, within a multivariate setting, the effect of Central Bank interventions is measured parametrically in Beine et al. (2007); this is exactly the type of analysis which could be done with more detail if an intra-day evaluation would be possible. As an example, the application in this article revisits a finding in Ranaldo (2007), who finds indications of a time-of-day pattern in intra-day exchange rate returns using non-parametric methods.

Table 1 recaps the discussion of this section in a structured manner: It indicates for a subset of the articles mentioned here whether the approach can be called parametric, if intra-day data is used for estimation, whether the underlying model is a diffusion model, if there is attention for microstructure effects, if jumps are allowed in the model, and if there is explicit modelling of a stochastic volatility structure. A ‘+’ indicates a ‘yes’, a ‘-’ indicates ‘no’, and a ‘±’ stands for a ‘sometimes/partly’.

What is missing from the table is a row for an article taking the parametric approach for a high frequency dataset. In such a case, the model is best derived from a continuous time diffusion model. Any application should then allow for both the possibility of jumps, and for some microstructure noise, together with a flexible specification for the volatility of the process. And that is exactly what the contribution of this article will be: It will look into the possibility of putting an entry in the table with only ‘+’ everywhere. For this purpose, the next section will introduce the diffusion model, and its operationalisation in discrete time (Sections 2.1–2.3), comparing it to daily variance measures in Section 2.4. After an intermediate section on the evidence of jumps in exchange rate data, and the thickness of the tails of the return distribution in Section 3, implementation of the model is discussed in Section 4. Both a simulation exercise and an application follow in Sections 5.1–6, with

conclusions in Section 7.

2 Theory of diffusion processes and related literature

2.1 Jump-SV-Diffusion

A log-price process of a financial asset can be modelled through a diffusion, as is done in e.g. Barndorff-Nielsen and Shephard (2004a) and Andersen et al. (2007) through

$$p(t) = p^*(t) + \epsilon(t), \quad \epsilon(t) \sim \text{i.i.d.}(0, \sigma_\epsilon^2), \quad (1)$$

$$dp^*(t) = \mu(t)dt + \sigma(t)dW_1(t) + \kappa(t)dq(t), \quad 0 \leq t \leq T. \quad (2)$$

Such a diffusion process states that

- the underlying price $p^*(t)$ is observed with error $\epsilon(t)$ as $p(t)$;
- the underlying price evolves according to a trend component $\mu(t)$, often fixed at zero for simplicity. In Section 2.3, $\mu(t)$ introduces deterministic trend effects into the model;
- plus a volatility element $\sigma(t)dW_1(t)$, with Brownian motion $W_1(t)$; instantaneous volatility is $\sigma(t)$, which may contain intra-day or intraweek seasonality;
- plus a jump component; $q(t)$ is a counting process with intensity $\lambda(t)$, and $\kappa(t) \equiv p^*(t) - p^*(t-)$ the corresponding instantaneous jump size.

In the aforementioned articles, the volatility process $\sigma(t)$ is only restricted to be a strictly positive càdlàg sequence; that is all that is necessary in order to derive the integrated daily variance.

As here the object of interest is the full volatility sequence, also intra-day, more assumptions on the process are necessary. Here we assume an independent diffusion for the log-volatility $h(t)$, combined with a separate deterministic mean process $g(t)$ describing intra-day or intraweek seasonality of the variance:

$$\log \sigma^2(t) = h(t) + g(t), \quad (3)$$

$$dh(t) = \mu_h(h(t); \theta)dt + \sigma_\xi dW_2(t) \quad (4)$$

with $W_2(t)$ a second Brownian motion. The variance of the volatility diffusion, σ_ξ^2 , is assumed constant, and for the moment the price, volatility and jump processes are taken to be independent apart from the direct link between them in the price diffusion. At a later stage, this assumption can be relaxed further.

2.2 Discretisation through Euler expansion

To operationalise the model, the diffusion has to be linked to the (discrete) observations. Given observations p_i at times $T_n = \{0 = t_1, \dots, t_n = T\}$ with time distances $\delta_i = t_{i+1} - t_i$ between observations, an Euler approximation of the process gives

$$p_i = p_i^* + \epsilon_i, \tag{5}$$

$$p_{i+1}^* - p_i^* = \mu(p_i; \theta) + s_i \sqrt{\delta_i} \eta_i + \sum_{j=1}^{q_i} \kappa_{i,j}, \tag{6}$$

$$\log s_i^2 = h_i + g_i, \tag{7}$$

$$h_{i+1} - h_i = \mu_h(h_i; \theta) \delta_i + \sigma_\xi \sqrt{\delta_i} \xi_i \tag{8}$$

with

$$\epsilon_i, \eta_i, \xi_i \sim \text{i.i.d.}(0, 1),$$

$$\kappa_{i,j} \sim \mathcal{N}(0, \sigma_\kappa^2), \tag{9}$$

$$q_i \sim \text{Po}(\delta_i \lambda). \tag{10}$$

The choice of density for $\kappa_{i,j}$ for jump j in time interval $[t_i, t_i + \delta_i]$ is an approximation to the density proposed in e.g. Andersen et al. (2002), who use

$$\log(1 + \kappa_{i,j}) \sim \mathcal{N}(-0.5\sigma_\kappa^2, \sigma_\kappa^2). \tag{11}$$

For small variances, densities (9) and (11) correspond as Chib et al. (2002) show, while the normal specification for κ in (9) is more convenient computationally.

The number of jumps per time period is specified as a Poisson process. In practice not more than one jump is expected within a short time interval. Hence in most cases specifying a Bernoulli density with probability $\delta_i \lambda$ is sufficient, and considerably simpler in implementation.

Aït-Sahalia (2002) discusses other discretisations for diffusion processes, which are more precise and quicker than the Euler expansion. However, these methods are limited to a diffusion process without micro-structure noise or stochastic volatility. Secondly, as in the application the frequency of the data will be high, the approximation error on the volatility process can be expected to be relatively small.

2.3 Refinements and final specification

The above model is still missing the definitions of the drift components, and the specific distributional choices of $\epsilon_i, \eta_i, \xi_i$. To start with the last, take

$$\begin{aligned}\epsilon_i, \xi_i &\sim \text{i.i.d. } \mathcal{N}(0, 1), \\ \eta_i &\sim \text{i.i.d. } \sqrt{\frac{\nu - 2}{\nu}} t_\nu.\end{aligned}$$

A fully Gaussian choice would allow for the model to be cast into a (non-linear) state space framework (Harvey, 1989; Durbin and Koopman, 2001), which lends itself to straightforward implementation of differing time distances δ_i . However, as a generalisation a Student- t density (standardised to unit variance) is implemented for the transition equation. This delivers extra flexibility as returns might display a heavier tail than the normal, without these returns being so extreme that they should count as true jumps.

From the original (Gaussian) diffusion model, a Student- t density in cannot be derived through discretisation. The introduction of the heavy tails here is done on pragmatic grounds. In the application it will be found that Gaussian densities throughout the model are not sufficient to distinguish ‘large’ jumps from more standard behaviour of the price process. One could consider the Student- t process however as a high frequency jump component with relatively low variance, and consider the jump component κ_i as a low-frequency high variance jump effect.

As the fully (conditionally) Gaussian state space model is more convenient for implementing the model, the unobserved states can be augmented by an inverse Gamma random variable z_i . This z_i is used as the standard deviation of the disturbance η_i in the transition

equation, as in

$$z_i \sim \text{IG-1} \left(\alpha_z = \frac{\nu}{2}, \beta_z = \frac{2}{\nu - 2} \right),$$

$$\eta_i | z_i \sim \mathcal{N}(0, z_i^2).$$

Bauwens et al. (1999, Theorem A.7) show how this corresponds (unconditionally) to the rescaled Student- t density for η_i as defined before.

Chib et al. (2002) extend a stochastic volatility model to allow for Student- t errors and jumps, though only at a daily frequency. Their findings indicate that, on this daily frequency, the inclusion of Student- t errors obviates the need for the inclusion of jumps, either with or without the Student- t disturbances. On the intra-daily frequency used here, it may become even harder to clearly distinguish between the occurrence of a jump or an observation from the heavy tail of a Student density.

The first drift component, of the evolution of prices, is simple: In the applied section attention is paid to exchange rates, and the log-exchange rate is not supposed to drift (disregarding any interest rate disparities), so a logical choice would be to take

$$\mu(p_i; \theta) \equiv 0.$$

To test for an intra-day deterministic drift in exchange rates, one could specify alternatively

$$\mu(p_i; \theta) \equiv X_i \beta \tag{12}$$

with X_i a vector of explanatory variables measuring the trending effect.

For the volatility specification, a reasonable drift could be

$$\mu_h(h_i; \theta) \equiv (\phi - 1)h_i \tag{13}$$

which for time difference $\delta_i \equiv 1$ leads to

$$h_{i+1} = \phi h_i + \sigma_\xi \xi_i, \tag{8'}$$

the standard AR(1) specification for the stochastic volatility equation.

For general time difference δ_i , the Euler expansion gives

$$h_{i+1} = \left(\phi - 1 + \frac{1}{\delta_i} \right) h_i + \sigma_\xi \sqrt{\delta_i} \xi_i, \tag{8''}$$

see Andersen et al. (2004, Section 3.1.2). This compares with the multi-step AR(1) equation of

$$h_{i+1} = \phi^\delta h_i + \sigma_\xi \sqrt{\sum_{j=0}^{\delta_i-1} \phi^{2j}} \xi_i, \quad (8''')$$

which is a refinement of the Euler approximation along the lines of the augmentation of the grid in Elerian et al. (2001).

For the deterministic part of volatility, $g_i \equiv M_i \gamma$, a periodic spline (Poirier, 1973; Stoer and Bulirsch, 1991) is taken, allowing smooth variation of the volatility throughout the week.

The full model, combining all previous equations and the final notation, is repeated below. For clarity, each equation is labelled with a term indicating the concept to which the equation relates; these terms will be used extensively in later sections.

$$p_i = \log(P_i) = p_i^* + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad [\text{Observations}] \quad (14)$$

$$p_{i+1}^* = p_i^* + b_i + \kappa_i + \sigma_i \eta_i, \quad \eta_i \sim \mathcal{N}(0, z_i^2), \quad [\text{Transition}] \quad (15)$$

$$\sigma_i^2 \equiv \delta_i \exp(h_i + g_i), \quad z_i \sim \text{IG-1} \left(\alpha_z = \frac{\nu}{2}, \beta_z = \frac{2}{\nu - 2} \right), \quad [\text{Variance}] \quad (16)$$

$$\kappa_i \sim \mathcal{N}(0, q_i \sigma_\kappa^2), \quad q_i \sim \text{Po}(\delta_i \lambda), \quad [\text{Jumps}] \quad (17)$$

$$b_i = X_i \beta, \quad g_i = M_i \gamma, \quad [\text{Trend \& Season}] \quad (18)$$

$$h_{i+1} = \phi^{\delta_i} h_i + \xi_i, \quad \xi_i \sim \mathcal{N} \left(0, \sigma_\xi^2 \sum_{j=0}^{\delta_i-1} \phi^{2j} \right), \quad [\text{Volatility}] \quad (19)$$

for N observations, time difference $\delta_i = t_{i+1} - t_i$ between observations, and $i = 1, \dots, N$. M_i is one row of matrix M containing the periodic spline. M is of size $N \times k$, with k the number of knots in the spline. Similarly, X_i is a row of matrix X with explanatory variables governing the trending behaviour of the exchange rate. Estimation is done using a Markov chain Monte Carlo (MCMC) method. Details are to be found in Section 4.

2.4 Daily variation

In financial markets, the current daily variation is an important input for gauging the risk of portfolios. Though an estimate from a GARCH or SV model on daily data can be used as a first impression of the variation at the daily level, the introduction of intra-day data in the measures of non-parametric realised variation, led to a large increase in precision of

the estimates. This theory was developed concurrently and independently in Andersen and Bollerslev (1998), Comte and Renault (1998) and Barndorff-Nielsen and Shephard (2001).

If there are a maximum of D observations within a day, such that $[t_i, t_{i+D}]$ is an interval of 24 hours, assuming that there is no microstructure noise, then

$$[r, r]_{t_i} = \int_{t_i}^{t_{i+D}} \sigma^2(s) ds + \sum_{t_i < s \leq t_{i+D}} \kappa^2(s)$$

is the quadratic variation of the price process, consisting of the sum of the integrated variance and the squared jumps occurring during the day.¹ This

Using realised variation (RV), the quadratic variation at the day starting at t_i , over D observations, can be estimated as

$$\text{RV}_{t_i}(D) = \sum_{j=0}^{D-1} r_{i+j}^2 \rightarrow \int_{t_i}^{t_{i+D}} \sigma^2(s) ds + \sum_{t_i < s \leq t_{i+D}} \kappa^2(s),$$

whereas the bipower variation (BPV) can be used to estimate only the integrated variance, excluding the jump process:

$$\text{BPV}_{t_i}(D) = \sum_{j=1}^{D-1} |r_{i+j}| |r_{i+j-1}| / \mu^2 \rightarrow \int_{t_i}^{t_{i+D}} \sigma^2(s) ds, \quad \mu \equiv \sqrt{2/\pi}.$$

Hence, the integrated/summed quadratic jump process can be estimated as the difference between the RV and the BPV, following Barndorff-Nielsen and Shephard (2004b) to get rid of negative squared jumps as

$$\text{J}_{t_i}(D) \equiv \max(\text{RV}_{t_i}(D) - \text{BPV}_{t_i}(D), 0). \quad (20)$$

To test whether a jump is significant, Barndorff-Nielsen and Shephard (2006a, BN-S for short) use (among others) a test statistic which they call the feasible linear jump statistic $\hat{G}_{t_i}(D)$,

$$\hat{G}_{t_i}(D) = \frac{\text{BPV}_{t_i}(D) - \text{RV}_{t_i}(D)}{\sqrt{\delta \theta \text{QPV}}} \rightarrow \mathcal{N}(0, 1), \quad (21)$$

$$\text{QPV}_{t_i}(D) = \frac{1}{\delta} \sum_{j=4}^{D-4} |r_{i+j}| |r_{i+j-1}| |r_{i+j-2}| |r_{i+j-3}| / \mu^4 \rightarrow \int_{t_i}^{t_{i+D}} \sigma^4(s) ds,$$

with QPV the realised quadpower variation and $\theta \equiv \pi^2/4 + \pi - 5$.

¹For simplicity of presentation, this section assumes that there are exactly D observations each day; if not, the returns have to be scaled accordingly, with negative effects on the precision of the estimates of quadratic variation.

A second option to test for jumps is the statistic introduced by Lee and Mykland (2008, L-M, with a small change in notation for consistency with previous results). This test adjusts a return by an estimate of the instantaneous standard deviation, the latter calculated through the square root of the average BPV over the previous D observations, as

$$\mathcal{L}_{t_i}(D) = \frac{r_i}{\hat{s}_{i-1}}, \quad \hat{s}_{i-1}^2 \equiv \frac{1}{D} \text{BPV}_{t_{i-1}}(D).$$

By using the BPV over D previous observations, the variance estimate is robust for jumps, especially for a jump at time t_i as this return itself is not included in the calculation of the variance.

The statistic itself $\mathcal{L}_{t_i}(D)$, in the absence of jumps, follows a $\mathcal{N}(0, 1)$ distribution. It enables one to test for jumps at each moment in the sample. To check whether a specific time interval, say a day, contains a jump, a max statistic $\mathcal{L}_{t_i}^m(D)$ is derived, with

$$\mathcal{L}_{t_i}^m(D) = \frac{\max_{j \in [i, \dots, i+D-1]} |\mathcal{L}_{t_j}(D)| - C_n}{S_n} \underset{a}{\sim} \text{Gumbel}(0, 1), \quad (22)$$

$$C_n = \sqrt{2 \log D} - \frac{\log \pi + \log \log D}{2\sqrt{2 \log D}}, \quad S_n = \frac{1}{\sqrt{2 \log D}}.$$

The asymptotic density of the test statistic is the standard Extreme Value-I or Gumbel density.

The theory of realised variation and bipower variation, and also the test of Lee and Mykland (2008), assumes that the price process is observed without error from possible microstructure effects. Alternatively, the multiscale realised volatility measures developed by Zhang et al. (2005) and Zhang (2006) allow for estimation of the integrated variance robust to microstructure effects. These measures, however, do not explicitly take jumps into account.

Based on the model of Section 2.3, alternative estimators can be constructed integrating directly over the estimated intra-day variances and jumps. The expected integrated model variation is

$$\text{IMV}_{t_i}(D) = \mathbb{E} \sum_{j=0}^{D-1} \exp(h_{i+j} + g_{i+j}) \approx \frac{1}{S} \sum_{s=1}^S \sum_{j=0}^{D-1} \exp(h_{i+j}^{(s)} + g_{i+j}^{(s)})$$

where a posterior sample of size S is assumed. Likewise, the posterior variance of the IMV can be evaluated, or other functions of the posterior density.

Such an IMV measure is directly comparable to the bipower variation, at least in theory. Differences between the two are errors in the approximation of the BPV to the true inte-

grated variance on the one hand (among them, the effect of microstructure noise), and model misspecification and estimation errors on the part of the IMV measure.

3 Data description and evidence of jumps

In the next section the performance of the model is analysed using data on the 2005 Euro/US Dollar exchange rate. Before moving to this analysis, this section takes a first look at the characteristics of the data and the type of jumps the model might find.

The data was provided by DiskTrading, who constructed the 5-minute data from tick-by-tick data. As observation at time t the price of the first trade in the 5-minute interval preceding t is reported. The prices P_i are transformed into $p_i = 100 \times \log(P_i)$, with (percentage) returns $r_i = p_{i+1} - p_i$. No further cleaning or transforming of the data has taken place. All observations between 2005/1/1, 0:00h and 2005/12/31, 23:55h are taken, using GMT+1 as the timing. This time period contains effectively 76,339 prices. No trading took place between Saturday 0:00h and Sunday 21:45h.

Table 2 about here

The moments of the percentage returns are presented in Table 2. As usual with financial data, there is virtually no mean return. There is some negative skewness, and a huge kurtosis, indicating the thick tails of the return distribution. Figure 1 presents the percentage returns in the first panel. Several large returns are apparent. The usual shape of high versus low volatility periods is less obvious from the return plot. The second panel displays the autocorrelation of the squared returns. This panel clearly indicates the intra-day volatility movement, with high correlation in volatility at the daily frequency. The last panel compares the return distribution to the Student- t density. The QQ-plot indicates that for values in the .1%-99.9% quantile of the returns (between -8.3 and 8.3 on the y -axis in the last panel) the Student- t with $\nu = 3.5$ degrees of freedom fits the data.

Figure 1 about here

Figure 2 about here

Figure 2 displays the exchange rate over the full year 2005, the month of January, the

third week, and day 12/1², with increasing frequency (indicated between parentheses within the graphs). What is clear from such a figure is that on a yearly basis using only daily data, one hardly notices the missing data in the weekends and there are no apparent sudden shifts. With data concerning one month, the missing weekends are starting to become a problem, as the gaps are relatively large compared to the available information. One might wonder about sudden shifts in the data, but they are not fully clear yet.

When a week of hourly data is used, one does notice a sharp increase on January 12, but it cannot be judged clearly whether it was an increase over a number of hours or a sudden jump. Only in the bottom right panel with intra-day data at the 5-minute frequency, it is seen that in the interval from 14:30h-14:35h the exchange rate jumped suddenly by 0.7%. Though a 0.7% increase sounds small, on a daily basis (multiplying by $\sqrt{288}$, to adapt for the number of observations within a day) it would correspond to a sudden 12% increase in exchange rate, which is considerable.

From the plot of the returns, and the QQ-plot of returns versus the Student- t density, it was found that there could be many jumps in this data set. Applying the feasible jump test statistic (20) or the version of Lee and Mykland (2008), (22), on the 5-minute data for the above series, indeed leads to numerous jumps being detected. On a standard significance level of $\alpha = 5\%$, the tests indicate between 43 and 86% of the days as containing (one or more) jumps. Decreasing the significance level, even up to $\alpha = .005\%$ still leads to 10–35% of days rejecting the hypothesis of the absence of jumps. Table 3 presents an overview of significance levels and rejections.

Table 3 about here

When studying these results on the jump test statistics, one should keep in mind that these tests do not (explicitly) allow for microstructure effects. In the application below, the goal is to study what a parametric approach can tell us about the possible jump process. When the model is parametrised fully, the timing of the jumps comes out as a side-product of the estimation procedure, and it would be possible to relate jumps to external effects like announcements at a later stage.

²This date is chosen as having the largest 5-minute absolute return of all days in 2005

4 Implementation of the Markov chain

In Section 2.3 the full parametric model of the log-price process is given in Equations (14)–(19). Implementing a MCMC algorithm for this model is not straightforward. Several issues arise when moving from daily data to high frequency data. This section describes first the general approach, in Section 4.1, with further details on the initialisation in Section 4.2.

4.1 General procedure

The model referred to above is an unobserved components model (Harvey, 1989; Durbin and Koopman, 2001), linking log-prices p to unobserved components p^*, q, κ, z and h (for the underlying price, number of jumps, jump size, Student- t effect and stochastic volatility, respectively), and parameters $\theta = (\sigma_\epsilon, \sigma_\kappa, \lambda, \beta, \gamma, \phi, \sigma_h, \nu)$, with the long-run standard deviation $\sigma_h \equiv \sigma_\xi / \sqrt{1 - \phi^2}$.

The unobserved components model is convenient in this situation as it is relatively straightforward to use in a situation with missing observations, or, according to the notation in (14)–(19), when the time difference $\delta_i = t_{i+1} - t_i$ between successive observations is not constant. In the notation below, and in subsequent sections, dependence of the parameters or states on this fixed vector δ of time differences, and likewise on the matrices X and M describing the trend in mean and the periodic spline for the volatility, is suppressed for clarity.

The parameters are estimated using a Bayesian approach applying data augmentation (Tanner and Wong, 1987; Gelfand and Smith, 1990) for the unobserved components p^*, q, κ, h and z . Conditional on parameters β and γ , the trend $b = X\beta$ and seasonality $g = M\gamma$ are fixed, so these are not counted among the unobserved components.

In order to use a Bayesian method, prior densities for the parameters have to be specified. For the parameters conjugate priors are taken as far as possible, choosing IG-1 densities for the parameters $\sigma_\epsilon, \sigma_\kappa$ and σ_h , normal for β, γ , and Gamma for the jump frequency parameter λ . For ϕ , a Beta density is most logical, also to limit prior mass to the range $[0, 1]$. The degrees-of-freedom parameter ν has to be larger than 2 for the variance to exist. The prior density for $\nu - 2$ is conveniently taken to be exponentially distributed.

Furthermore, the states h_0 and p_0^* need initial distributions. The initial value of h_0 is taken to be distributed according to $h_0 \sim \mathcal{N}(0, \sigma_h^2)$, its long-run distribution, whereas p_0^* is initialised uninformatively as $p_0^* \sim \mathcal{N}(0, K), K \rightarrow \infty$. The IG-1 density specified for z in

Equation (16) can be considered its prior distribution. This distribution is combined with the likelihood to derive the posterior density of z .

The full MCMC algorithm consists of the following steps:

- i. Initialise states p^*, q, κ, h, z and parameters θ , see Section 4.2, and set the simulation counter $s = 0$.
- ii. Sample new states p^*, q, κ, z, h successively, using
 - (a) the simulation smoother (De Jong and Shephard, 1995; Durbin and Koopman, 2002) for $p^*|p, q, \kappa, h, z, \theta$;
 - (b) a mixed continuous-discrete sampling step for drawing $q_i, \kappa_i, z_i|h_i, \nabla p_i^*, \theta$ for $i = 1, \dots, N - 1$, with $\nabla p_i^* \equiv p_{i+1}^* - p_i^*$. Only few jumps can occur per time interval, usually $q_i \in \{0, 1\}$, though the implementation does allow for more jumps. See below for further details;
 - (c) a Metropolis-Hastings (MH) step (Metropolis et al., 1953; Chib and Greenberg, 1995) for sampling $h_i|h_{i-1}, h_{i+1}, \nabla p_i^*, q_i, \kappa_i, z_i, \theta$, again for $i = 1, \dots, N - 1$, as in Jacquier et al. (1994) and Kim et al. (1998).
- iii. Sample new parameters θ , using the following substeps:
 - (a) Draw $\sigma_\epsilon|p, p^*$ from a IG-1 density;
 - (b) Sample $\gamma|\nabla p^*, \kappa, h, z, \beta$ from its full conditional posterior. Take e as the vector of residuals of the state transition equation (15). Then this posterior can be approximated in a point $\tilde{\gamma}$ (e.g. the last sampled value of γ) by a normal density

$$p_a(\gamma) \sim \mathcal{N}(\mu_a, \Sigma_a)$$

with

$$\begin{aligned}\mu_a &= -\mathbf{H}^{-1}(\mathbf{g} - \mathbf{H}h^*), & \Sigma_a &= -\mathbf{H}^{-1}, \\ \mathbf{g} &\equiv \left. \frac{\partial \log \mathcal{L}(e; \dots)}{\partial \gamma} \right|_{\gamma=\tilde{\gamma}}, & \mathbf{H} &\equiv \left. \frac{\partial^2 \log \mathcal{L}(e; \dots)}{\partial \gamma^2} \right|_{\gamma=\tilde{\gamma}}, \\ \log \mathcal{L}(e; \dots) &= C - \frac{1}{2} \sum \left(h + M\gamma + \frac{e^2}{z_i^2 \delta_i \exp(h + M\gamma)} \right), \\ \frac{\partial \log \mathcal{L}(e; \dots)}{\partial \gamma} &= -\frac{1}{2} M' \left(1 - \frac{e^2}{z_i^2 \delta_i \exp(h + M\gamma)} \right), \\ \frac{\partial^2 \log \mathcal{L}(e; \dots)}{\partial \gamma \partial \gamma'} &= -\frac{1}{2} M' \text{diag} \left(\frac{e_i^2}{z_i^2 \delta_i \exp(h_i + M_i \gamma)} \right) M,\end{aligned}$$

for $e \equiv \nabla p_i^* - X_i \beta - \kappa_i$ and some constant C . A short Metropolis-Hastings chain is used to adapt for the difference between the approximate normal density $p_a(\gamma)$ and the target density of γ ;

- (c) Sample $\beta|p, \kappa, g, h, z$ from a Normal density;
- (d) Sample $\lambda|q$, from a Gamma density;
- (e) Draw $\sigma_\kappa|q, \kappa, h, \theta$. This density is effectively IG-1, depending on the number and size of the sampled jumps;
- (f) Draw ϕ, σ_h conditionally on the standardised disturbances u of the SV process by
 1. filtering

$$\begin{aligned}u_i &\equiv \frac{h_{i+1} - \phi_i^\delta h_i}{\sigma_i}, \\ \sigma_i^2 &\equiv (1 - \phi^2) \sigma_h^2 \sum_{j=0}^{\delta_i - 1} \phi^{2j},\end{aligned}$$

to get the standardised disturbances u ;

2. sampling $\phi, \sigma_h|p, p^*, u, q, \sigma_\epsilon, \sigma_\kappa, \beta, \gamma$ using a random walk MH step with a normal candidate density, conditional on those u ;
3. inverting filter iii(f)1. to reconstruct the corresponding series of $h|\phi, \sigma_h, u$, adapting for the change in SV parameters.

See Bos and Shephard (2006) for details on the conditioning on the disturbances. It is found that conditioning on h , as is normally done, does not lead to convergence in high frequency/large sample applications, as the volatility process h is too informative on the values of ϕ, σ_h , and the sampler mixes too slow if at all.

(g) For ν , the posterior density is not of a known form. However, if the number of observations in the sample is large, the posterior will be approximately normal, and a simple random walk Metropolis-Hastings step with a normal candidate density suffices;

- iv. Save sampled parameters $\theta^{(s)} = \theta$ and possibly the states, increase s , and repeat from ii. until sufficient parameter vectors are collected.

The general outline of the sampling scheme is not much different from earlier samplers in e.g. Kim et al. (1998) or Chib et al. (2002), apart from the introduction of the jumps and the unobserved underlying price process. What makes the true difference, however, is the level of detail with which the data is modelled. In standard SV models the return sequence immediately provides information on the volatility. Here the volatility sequence has to be estimated as the volatility of the increments of an unobserved process, after extracting jumps occurring very infrequently but at unobserved time periods. At the same time, one has to take care to adequately handle variation of volatility within the week and within the day.

The increase in frequency from daily data to intra-day data at the 5 minute frequency leads to a sample of around 75,000 observations for a full year of data. This implies that standard samplers for the SV parameters ϕ, σ_h no longer converge. Here the transformation from Bos and Shephard (2006) to condition on the SV disturbances is found to be essential.

In the above algorithm, it is not specified in detail how to sample jointly from q_i, κ_i, z_i . How can the algorithm decide for one observation whether the return came far from the tail of the Student- t density, or whether there has been a jump? And if there was a jump, what size could it have been? If the model would be simplified to have normal disturbances, then the size of the jump can be integrated out of the model, and the marginal density of $\nabla p^*|q$ is

$$p(\nabla p_i^*|q, \beta) = \int_{\kappa} p(\nabla p_i^*|q_i, \beta, \kappa) \pi(\kappa) d\kappa = \mathcal{N}(X_i \beta, \delta_i \exp(h_i + g_i) + q_i \sigma_{\kappa}^2).$$

The discrete distribution of $q_i|\nabla p_i^*, h_i, \theta$ would be easily constructed from the above marginal density and the prior $\pi(q_i)$.

A draw from this approximate distribution of $q_i|\nabla p_i^*, h_i, \theta$ can be supplemented by a draw from $k_i|q_i, \nabla p_i^*, h_i, \theta$, again using the normal assumption. To correct for the error in the sampling distribution, a Metropolis-Hastings step is used to weigh in the effect of the

Student- t density.³ The resulting sample from $q_i, k_i | \nabla p_i^*, h_i, \theta$ is then augmented by a draw from $z_i | \nabla p_i^*, q_i, k_i, h_i, \theta$, which is IG-1, such that in the end a sample from the correct density is obtained.

Improved mixing of the chain can be achieved by repeating each step which uses an approximating density with MH a number of times before continuing. This lowers the correlation induced by possibly low acceptance rates of the MH sequence. The repetition of sampling σ_ϵ is done in a loop together with the sampling of the underlying price process p^* , as these two are heavily correlated. By blocking them together, and repeating the sampling of $[p^* | p, \sigma_\epsilon, \dots]$ and $[\sigma_\epsilon | p, p^*, \dots]$ together, correlation in the chain is found to drop considerably.

Kim et al. (1998) discuss the importance of the method used in sampling a new vector of h . Their suggestion of using a mixture density in an approximate parametric model is not easily applicable to the present situation, due to the effects of the jumps on the squared return sequence. This mixture density also applies an extra unobserved indicator for the mixture, implying further cause for correlations in the chain. An alternative to step iic. is to implement a multi-move sampler (Shephard and Pitt, 1997), where a block of (h_i, \dots, h_{i+K}) are drawn jointly, conditional on the values h_{i-1} and h_{i+K+1} at the end points. This method does not have the disadvantage of applying an extra unobserved component, but has not proven itself on high frequency data sets.

4.2 Initialisation of the chain

The initialisation of the parameters and states is important, as the MCMC algorithm itself may mix slowly. This may be particularly acute when initialised from bad starting values.

The initialisation for the parameters is based on a rough quasi maximum likelihood approach. First, parameters β for the trend can be estimated by OLS, disregarding microstructure effects and stochastic volatility all together. Though not efficient, such estimates should be consistent.

³When no jump occurs, there is no need to sample k_i and the Metropolis-Hastings step can be circumvented. When a jump is found, the Metropolis-Hastings step adapting the candidate Gaussian density towards the Student- t density theoretically has bad properties, as the candidate has thinner tails than the target. In practice, this implies that the effective density implemented here of the returns conditional on a jump of size k has thinner tails than the theoretical t_ν .

Secondly, to estimate the volatility spline, one can use

$$\begin{aligned} \nabla p_i^* &\sim \mathcal{N}(X_i\beta, \delta_i \exp(W_i\gamma)), & \Leftrightarrow & \quad \log(\nabla p_i^* - X_i\beta)^2 = W_i\gamma + \xi_i, \\ \xi_i &\sim \text{i.i.d.}(-1.27, \pi^2/2), \end{aligned}$$

see Harvey et al. (1994). Using the the observed log squared returns (adapting for the trend and the time distance between successive observations, disregarding jumps and microstructure noise), a first estimate of γ can be computed using OLS. As this is far from efficient, it can be followed by an optimisation of γ over the approximate likelihood of the returns themselves.

Given the trend and volatility spline, still disregarding the microstructure effects, daily realised volatility measures can be extracted for each day in the sample. Using the feasible jump test (20) of Section 2.4, days are selected at which a jump is found using a significance level of $\alpha = .1\%$. At these days, the largest return is initially flagged as a jump, with size equal to the return. The fraction of jumps λ is initialised according to the number of jumps found, and the jump standard deviation σ_κ is extracted likewise.

For the stochastic volatility parameters, the original quasi maximum likelihood method of Harvey et al. (1994) is used, adapting for the trend, jumps and time distance δ between observations. This method also provides a first estimate of the intra-day volatility sequence h .

Conditionally on the volatility, a fully Gaussian state space model can be estimated to extract the variance σ_ϵ^2 of the microstructure effects in observation equation (14). From the Kalman filter equations, a filtered estimate of the efficient price p^* is extracted. The degrees of freedom parameter ν is estimated from the increments ∇p_i^* , adapted for SV, time distances δ , jumps κ and volatility seasonality $W\gamma$. This estimate for the degrees of freedom parameter will be positively biased in general, as the returns in the price process are extracted under the assumption of normal disturbances. It remains a good proxy for a first estimate. The Student- t tail component z is initialised at 1, as it is not too influential for the start-up of the Gibbs algorithm of Section 4.1.

5 Performance in a simulated setting

5.1 Simulating intra-day returns with jumps and diffusion

As the model is fully specified, a choice of parameters and initial conditions is sufficient to fix a DGP. For simplicity, the parameters and their prior choices are collected in Table 4. These parameters correspond to values that could be found in practice. Some further discussion of the parameter values and sample sizes, relating to the simulation results, is postponed to Section 5.3.

Table 4 indicates, for each parameter, what setting is chosen for the DGP, followed by the prior density, the parameters of the prior density, and the mean and standard deviation of the prior. Note that for the degrees-of-freedom parameter, the prior for $\nu - 2$ is specified as $\exp(.25)$, with expectation $E(\nu - 2) = 4$. The value used in the DGP is $\nu = 6$ degrees-of-freedom. Furthermore, in this simulation, the specification for the trend is not used at all, so β is fixed at 0 both in generating the data and simulating the parameters.

Parameters are presented at the daily frequency, e.g. a value of $\lambda = .1$ indicates on average one jump every 10 days, even though multiple observations per day are generated. Likewise, the autocorrelation of the SV component is $\phi = 0.97$ per day, implying a correlation of $\phi^{1/D} = 0.99989$ between two successive 5-minute observations. All innovations in the price increments ∇p_i^* have been transformed by multiplying them by \sqrt{D} , to effectively measure volatility at the daily frequency. This implies that the jump standard deviation of $\sigma_\kappa = 5$ also refers to this daily frequency.

Table 4 about here

The spline used for creating a seasonality effect within the volatility is given knots throughout the working days at every six hours, repeating the daily shape from Monday until Friday. In the weekend, only a knot in the middle of the weekend is specified. The level of the spline is modelled after the volatilities found in the next section, with exchange rate data.

From this DGP a data set of 365 days, with a maximum of 288 observations per day, is generated. As in a real data set, some of the observations will be missing. Throughout the weekend, no observations are preserved, while randomly around 95% of the observations throughout the working days are retained. This resulted in a dataset of 71,225 generated prices. Of these prices, 45 (slightly more than the expected 36) contain a jump. Through the

specification of the process, 13 of those jumps fell within the (missing) weekend, such that their effect might be harder to recognise.

5.2 Discussion of simulation results

A sample from the posterior distribution is collected as described in Section 4. After initialising, a burn-in sample of size 5,000 of states is drawn repeatedly, until all means of the parameters in the first and last third of the sample lie within 1.96 standard deviations of each other. Subsequently a final sample of 50,000 parameter vectors is collected. Within each iteration, each MH sampling step is repeated 5 times to further lower correlation in the chain. Such a setup takes around 12 hours of time to complete using a combination of Ox (Doornik, 2006), SsfPack (Koopman et al., 1999) and C, on a 2.4Ghz AMD64 machine.

Table 5 about here

Table 5 reports statistics on the posterior distribution for the model parameters. Missing in the table are the remaining intra-week spline parameters, as their posterior is very similar to the density of $\gamma_{\text{Mon } 0:00\text{h}}$. From the table it can be seen that estimation of these parameters is not easy. Most clearly this is seen from the inefficiency measures⁴ in the last column. Especially for the SV parameters ϕ and σ_h , correlation within the chain is still rather high. On the other hand, these inefficiencies are a large improvement over a sampler without the transformation to SV disturbances mentioned in Step iiif. of the algorithm in Section 4.1 (results not reported here). Without the transformation, inefficiencies surge to levels around 2,000, for the parameters ϕ and σ_h .

Apart from the inefficiencies, the table also reports the parameters according to the DGP, and the posterior mode, mean, standard deviation, and the lower and upper 2.5% quantile, which give a good approximation to the 95% highest posterior density (HPD) region. Apart from the parameter σ_h , the values of the DGP lie within the interquantile ranges. The latter parameter is however still estimated reasonably close to the value according to the DGP.

⁴The inefficiency measure compares the posterior sample variance to the variance adapted for autocorrelation in the chain, using a window width of (in this case) 2,000 observations. A value around 1 indicates a fully efficient estimator, with effectively independent draws from the posterior. High values indicate that the posterior sample is less informative due to high correlation; see also Shephard and Pitt (1997), Kim et al. (1998) and Bos and Shephard (2006).

Figure 3 about here

Figure 3 displays the posterior density of the main parameters together with their priors, and the value of the DGP. It is seen that most parameters are retrieved from the data remarkably well, though the uncertainty on the jump parameters λ and σ_κ remains large. The jump process is hard to recognise when jumps are this infrequent, in a model with heavy tails. More on the uncertainty related to these parameters follows below in Section 5.3. Note however that the data is informative on these parameters, in the sense that the posterior clearly moves away from the prior.

Figure 4 about here

When plotting the sampled parameters against the index (not reproduced in the article), the parameters do not trend. Different starting values, longer or shorter samples, all lead to the same posterior densities. The sampler clearly converged without much of a problem for the data of this DGP.

Figure 4 studies the relation between estimated jumps and the true jumps from the DGP. In the first panel, for each 5-minute return, the posterior probability of a jump occurring in that time interval is given, for all 75,000 intervals. Most of the returns of those intervals are deemed to come from the standard Student- t distribution, with a few intervals which are deemed highly suspect. There is also a range of observations with a posterior probability of a jump of around 0.2. These are the observations just before the weekend, where the combined uncertainty of the whole weekend without observations makes it plausible that a jump could have occurred.

The second panel plots the absolute value of corresponding expected jump sizes, with one standard deviation error bands, for those intervals with a jump probability of at least 0.75. Without error bands, the true jumps of the DGP are plotted. For some jumps, the estimated jump sizes are correct, but smaller jumps and those occurring in the weekend are not detected.

The third panel displays the relation between the true jump size, horizontally, and the estimated jump probability, split out between weekdays (for which most neighbouring observations are available) and just before the weekends (with many missing observations, making it harder to recognise jumps). Of the 45 jumps generated throughout the year, 13 intra-week

jumps are recognised correctly; jumps below a size of 3 tend not to be recognised at all, neither are jumps in the weekends.

Figure 5 about here

Given that the important jumps can be found, one can wonder whether the model-based estimation can reconstruct the integrated variation over the day. Figure 5 displays the estimates of integrated variation using realised variance (RV), bipower variation (BPV) and the integrated model variation (IMV), together with the underlying integrated variation from the DGP. The left panel displays all measures against time, whereas the right panel compares either BPV or IMV to the underlying integrated variation.

From the figure, little is learned, apart from the fact that all measures appear to differ little from the underlying integrated variation. To formalise this, Table 6 reports results on the Mincer-Zarnowitz (Mincer and Zarnowitz, 1969; Andersen et al., 2005) regressions explaining IV from one of the other measures.

Table 6 about here

The MZ-regression indicates the validity of the IMV approach. The IMV approach results in the highest R^2 . When explaining the integrated variation using the IMV, the estimated parameters β_0, β_1 are not significantly different from 0 and 1, respectively. For the BPV, a t -test does indicate a difference of the parameters from their theoretical values. The RV measure does not account for jumps, and hence is not supposed to deliver an estimate of the IV better than the BPV in the present setting. It is only included for comparison.

5.3 Considerations on choices for the DGP

Several remarks could have been made before when the settings of the DGP were chosen, in Section 5.1, but these are better related to the estimation results of the previous section. First, there is the length of the data sample. Most articles estimating SV models use between 500-3,000 daily or weekly observations. Moving to 5-minute data, the year of data used both in the DGP above, and in the application below, should provide sufficient information for estimating the (weekly) spline. Shortening the sample to e.g. a quarter could be problematic for these parameters.

The jump process can be harder to estimate than the spline parameters, for two reasons.

First, the number of jumps in an applied data set is expected to be low. Andersen et al. (2007) report between 0.083-0.137 jumps per day (at a significance level of $\alpha = 0.1\%$ resp. 0.01% ; this corresponds roughly to the fraction of jumps found using the test of BN-S in Table 3). Over the year of data sampled here, 45 jumps were generated, and from those the parameters λ and σ_κ of the jump process should be estimated. Clearly, the amount of information on these parameters is going to be limited. Increasing the frequency of the data set could help in distinguishing the jumps from the non-jumps, as the jump variance is not dependent on the frequency of the data set, but the non-jump return variance is. Aït-Sahalia (2004) discusses this effect. Even though this article settled on the 5-minute frequency, initial estimations indicate that there are few problems to move towards 1-minute or even tick-by-tick data.

Secondly, it can be hard to distinguish the jumps from the standard returns, especially in combination with a Student- t density for the disturbance term. Financial returns are often found to be heavy tailed, even in the absence of jumps. For daily observations on S&P500 data, Chib et al. (2002) report degrees-of-freedom parameters ν between 12-15, though they also refer to other articles (Sandmann and Koopman, 1998; Jacquier et al., 1995) with lower values of 7.9 or 11, using different time periods of the same series. Given that the daily returns are composed of a large number of intra-day returns, by a central limit theorem, these daily returns can be expected to be ‘more normal’ than the intra-day returns themselves. The resulting parameter ν for the intra-day data would end up even (considerably) below 7.9, assuming the exchange rate data used below is comparable to the tail behaviour of the S&P 500. In the DGP, the value of $\nu = 6$ was chosen, even though the description of the data in Section 3 indicates a possibly lower degrees-of-freedom parameter.

To see how hard it can be to distinguish a jump from a standard return, simplify the model to

$$\begin{aligned}
 p_{i+1}^* &= p_i^* + q_i \kappa_i + \eta_i, & \eta_i &\sim t\left(0, \nu, \frac{\exp(\gamma)}{D}\right), \\
 q_i &\sim \text{Bern}(\lambda), & \kappa_i &\sim \mathcal{N}\left(0, \frac{\sigma_\kappa^2}{D}\right),
 \end{aligned}$$

with $\lambda = 0.1$ jumps per day, with $D = 288$ 5-minute returns ∇p^* , $\nu = 6$ and $\gamma = -1.4$ fixing the variance of the Student- t density. For the jump process, use $\sigma_\kappa = 5$ as in the DGP. In

such a case, the following probabilities for observing a large return are easily calculated:

$$\begin{aligned}
P\left(|\nabla p_i^*| > \sigma_\kappa/\sqrt{D} | q_i = 0\right) &= 2\left(1 - F_t\left(\frac{\sigma_\kappa}{\exp(\gamma)}\sqrt{\frac{\nu}{\nu-2}}\right)\right) = 0.017\% \equiv A, \\
P\left(|k_i| > \sigma_\kappa/\sqrt{D} | q_i = 1\right) &= 2(1 - F_N(1)) = 317\% \equiv B, \\
P\left(|\nabla p_i^*| > \sigma_\kappa/\sqrt{D}\right) &\approx \left(1 - \frac{\lambda}{D}\right)A + \frac{\lambda}{D}B = 0.128\%.
\end{aligned}$$

Unless normality can be assumed or the degrees-of-freedom parameter of the Student- t density is large, the jump-process standard deviation σ_κ needs to be factors of magnitude larger than the non-jump standard deviation, to clearly distinguish between true jumps and non-jumps.

Even so, it was found that a data series including microstructure noise, infrequent jumps and heavy tails can be estimated using the present setup. The model-based approach even delivers a closer approximation to the true integrated variation than either bipower variation or realised variation. This provides a good starting point for the next section, looking into the possibility of using this same model on a time series concerning the Euro/U.S. Dollar exchange rate.

6 An empirical illustration: Exchange rate and jumps

6.1 Data and setup

In Section 3, the exchange rate of the EUR/USD over the year 2005 was presented. In this section the discretised jump-diffusion-SV model with microstructure noise, (14)–(19), is estimated on this exchange rate using 5 minute frequency for the observations.

The model at hand gives a full parametric specification for the evolution of prices and volatility within the day. To put these to use, consider the findings in Rinaldo (2007). This author found that exchange rates tend to trend within the day according to the market which is more active. Roughly, the Dollar appreciates against the Euro between 9-13h GMT+1, during European active trading hours. Vice versa, the Dollar is found to depreciate again from 17-21h GMT+1, when only the U.S. market is open. Testing for such an effect can be done simply by introducing two dummies in X (see Equation (12)), multiplying the trend parameters in β for these different periods of the day.

In order to start the analysis, first the priors on the parameters have to be specified. The DGP in the previous Section 5.1 was set up in such a manner that the same priors from

Table 4 can be used for the present situation. The data used are the log-prices p_i presented in Section 3. A sample of size 50,000 is drawn from the Markov chain after initialising the parameters and the states as before.

6.2 On convergence with FX data

When using the simulated data in Section 5.1, the Markov chain converged without problems, and a decent sample was collected after the burn-in period. When the same algorithm is run on the FX data, initially a sample is found where the parameters ϕ , σ_h and ν seem to trend throughout the 50,000 iterations of the chain. A decrease in ϕ seems to be offset by increases in σ_h and ν . The movement of the parameters is however relatively slow, with ϕ decreasing from an initial value of 0.96 to 0.83 at the end the sample. Continuing the chain for another 50,000 or even 150,000 iterations does not lead to convergence either, with ϕ decreasing ever further towards zero.

The behaviour of a low ϕ , high σ_h , and increasing ν indicates some short-run bursts of volatility, modelled through a temporary increase or decrease of the SV component instead of through the Student- t tail of the density. In the extreme case, with $\phi \downarrow 0$, the instantaneous extra volatility originating from the SV is indistinguishable from the tail effects of the Student- t density (cf. in Section 2.3 where the Student- t density is implemented using a mixture of a Gaussian density and an instantaneous random variance component z_i^2).

When estimating the model on simulated data the algorithm was stable, converging rapidly towards the parameter values used in the DGP. In the present application the model is the same, the sample size is the same, and most of the parameter estimates will result to be similar as well. This leaves two options for the lack of convergence in the present situation. Either the model is misspecified for the data at hand, or there is a difference in parameter settings which makes the model unidentified in the present case. The main difference between the parameters used before for the DGP in Section 5.1 and present preliminary estimates is found in the degrees-of-freedom parameter. Instead of $\nu = 6$, in the DGP, here a value of ν between 3.2 and 3.4 is found.

A similar convergence problem estimating this FX data is found when ν is fixed at 3.4, when a Gaussian transition density (corresponding to $\nu \rightarrow \infty$) is used, or for a model where the SV process is forced to be integrated with $\phi \equiv 1$. In these three cases the standard

deviation of the SV process very slowly increases during the iterations of the sample.

These results point in a similar direction. The long run variance of the volatility is not well identified from the data. Given a degrees-of-freedom parameter estimated around 3.4, this does not have to be a surprise. For the Student- t density, such a value of ν implies that the fourth moment of the (unobserved) returns does not exist. Correspondingly, the second moment of the volatility, related to its long-run variance σ_h^2 , is not defined as well.

As the long-run variance cannot be estimated from the intra-day data, the SV parameters ϕ, σ_h are fixed at a combination of plausible values, related to estimates at a daily frequency. The estimated volatility sequence h_i effectively will be a weighted average of neighbouring observations, where the weights are fixed through the parameters of ϕ and σ_h . This however still allows for more freedom than the non-parametric approach, where volatility is estimated using equal weights throughout a day, weight zero outside the specific day.

For the SV parameters values of $\phi = 0.98$ and $\sigma_h = 0.5$ are chosen. Such values correspond with common findings for estimates of SV models on daily data, relate back to the values found during the initialisation stage, and also correspond to values found for these parameters when estimating daily EUR/USD exchange rates over a longer time period.

6.3 Resulting posterior

A collection of 50,000 vectors of parameters was sampled, while retaining information on the sampled price, volatility and jump processes. Before moving on to the results in terms of the jumps and integrated model variation, some statistics on the posterior densities of the parameters are presented in Table 7. The densities of the main parameters are plotted in Figure 6, together with their priors.

Table 7 about here

Figure 6 about here

From the densities in Figure 6, it is seen that the posteriors are smooth. This is a first indication of a well-behaved sample from the posterior. The data is informative on the parameters, as the posteriors are concentrated relative to the priors. A plot of the parameters against the index (not reproduced here) confirms that the sampler has converged.⁵ Also the

⁵To check the robustness of the results against the choice of the SV parameters, ϕ was fixed alternatively

inefficiency measures, in the last column of Table 7 indicate little remaining autocorrelation in the chain.

Concerning the size of the parameters, σ_ϵ is estimated around 0.01. This implies that the difference between $p_i = 100 \times \log P_i$ and the underlying p_i^* has a standard deviation of 0.01, which can be read as a 0.01% difference in the prices corresponding to one standard deviation. For the number of jumps, a value of $\hat{\lambda} = 0.42$ results. Some jump occurs a posteriori about every 2.4 days, or 690 observations. The standard deviation of the jump size is estimated at $\hat{\sigma}_\kappa = 3.4$, again measured on a daily scale, implying that the intra-day standard deviation is $\sigma_\kappa/\sqrt{288}$. That is to say, a one-standard deviation jump would correspond on a daily scale to a jump in the daily underlying price of around 3.4%, slightly smaller than the value assumed in the simulation exercise. Linked to the frequency of the jumps and the jump size is the degrees-of-freedom parameter. The estimate of $\hat{\nu} = 3.4$ implies far heavier tails than found by Sandmann and Koopman (1998) or Chib et al. (2002) for daily S&P 500 data. It corresponds closely to the findings in Section 3, where a first look at the data was taken. Such an estimate implies that the third moment of the returns still exists, but the kurtosis of the Student- t density is infinite. The trend parameters β , governing the trending behaviour within the trading day, are estimated very precisely and with high efficiency. A discussion on these parameters is postponed to Section 6.6.

Figure 7 about here

The volatility spline is modelled through the parameters γ , fixing the height of the periodic spline at the knots. Estimating $\gamma(\text{Sun}, 0:00\text{h})$ is difficult, judging from the wide bounds on the 95% interquantile range in Table 7. Intra-week γ 's are estimated more precisely, with similarly low inefficiencies. Instead of tabulating the values for all the knots, Figure 7 plots the estimates at each of the knots (indicated by the vertical lines) and the resulting spline, together with a one standard deviation error band. Over the weekend, volatility appears to be low but imprecisely measured. Throughout the weekdays, volatility is highest during at 0.97 or 0.99. In the latter case, all remaining parameters were estimated at very similar values as the ones presented here, and the same jumps (see below) were found. For $\phi \equiv 0.97$, the SV process becomes more flexible, picking up a larger part of the jumps. Estimates for the parameters are $\hat{\lambda} \approx 0.07$, $\hat{\sigma}_\kappa \approx 5.5$, $\hat{\nu} \approx 3.18$ when $\phi \equiv 0.97$. Only the most severe jumps are found in this case, other larger returns are accounted for by the tails of the Student- t density.

the (European) afternoons, and can be estimated far more accurately. There are some small differences between the weekdays. The expected variance of the percentage price increase moves between $\exp(-2.2) \approx 0.11\%$ at night and $\exp(-0.7) \approx 0.5\%$ during daytime, again measured on a daily scale.

6.4 Jumps and sizes

The object of this article is to see whether it is feasible to construct a model-based estimator for the occurrence of jumps in a high-frequency diffusion-type model. The jumps in this model are governed by the unobserved vector of q , the number of jumps occurring at each time point.

Figure 8 about here

Figure 8 displays the jumps, with in the first panel the posterior mean of q_i , the probability of a jump occurring within a specific five minute interval. Clearly a large number of intervals could have had a minor jump, resulting in a low value of q_i . On the other hand, there are a reasonable number of observations where $E(q_i) \approx 1$, such that a jump is indicated by the algorithm to have occurred with high probability.

A distinction is made in the panel for days with possible jumps during the week (indicated with a plus sign), and jumps over the weekend (the crosses). The observations around $E(q_i) = 0.38$, in the left-most panel, almost entirely stem from observations around Friday 23:55h, when the last observation before the weekend comes in. The model has a hard time deciding over the weekend whether there might have been a jump (among the almost 2×288 missing observations) or whether the change was caused by the normal diffusion process. Hence, an estimate of $E(q_i) \approx 0.38$ results in many of these cases.

The second panel plots the absolute expected jump size, for those jumps with $E(q_i) > .75$. The jump at January 12 2005 stands out, with a size corresponding to the value found in Section 3. Many other jumps are detected by the algorithm as well.

The final panel of the figure relates posterior jump probability to estimated (absolute) jump size (as the underlying true jump sizes are not available, in contrast to the case of the simulation exercise). Obviously, larger jumps are more easily detectable. The distinction between intra-week and weekend data is significant. If a possible jump is detected at the start of the weekend, its size is relatively small, below 3% in all cases. Intra-week jumps are

clearer both in estimated posterior probability and jump size.

6.5 Comparison with non-parametric measures

From the model, measures for the intra-day integrated variation including jumps, for the variation excluding jumps, and for the integrated squared jump process itself, can be constructed. These are comparable to the realised variation, bipower variation, and the difference between the two.

Figure 9 about here

Figure 9 plots these measures, both against time (top panels) and against each other (bottom panels, with the non-parametric measures on the vertical axes). In the left column, realised variation is compared with the model-based daily integrated variation+jumps measure, followed by the bipower variation versus integrated model variance in the middle column, and the jump measure in the third column of plots. Though the correspondence is not 1-on-1, clearly the measures are strongly related. If pressed one could find that the model-based variation measures tend to underestimate the non-parametric measures. This is partly due to the definition of variance in the model. The variance is comprised of the elements $\delta_i \exp(h_i + g_i)$, and does not account for variability due to the Student- t component. Likewise, these smaller jumps which the model attributes to the non-Gaussian transition density are depicted by the non-parametric measures as possible jumps. The righthand panels consequently display more NP- than model-based jumps.

6.6 Intra-day behaviour

Modelling intra-day returns in a fully parametric manner allows for a straightforward test of the hypothesis that the exchange rate trends within the day. Instead of measuring only the returns over fixed periods in the day and taking averages of those, in the present model the trend in the exchange rate can be allowed to change between the morning hours (fixed here at 9-13h GMT+1, when the European markets are most active) or the late afternoon period of 17-21h GMT+1, when most trading occurs on U.S. markets. Outside these periods, the standard assumption of a random walk without trend for the exchange rate is kept.

Introducing the trend through the parametrisation of Equations (15)–(18), a posterior density for the parameters β as in Figure 10 is found (see also the posterior statistics at the

bottom of Table 7). Both for the morning and afternoon period, a clear non-zero trend is found. For the morning-parameter, 99.2% of the posterior mass lies below zero, giving a very clear signal of the direction of the trend. For the afternoon, the signal of a positive trend is also clear, with 96.6% of the posterior mass at positive values of β .

Figure 10 about here

Over the 48 intervals of 5-minutes during morning and afternoon hours, in total the price drifted approximately $0.0102 \times 48/\sqrt{D} \approx 0.029\%$ downwards in the morning, and $0.0077 \times 48/\sqrt{D} \approx 0.022\%$ upwards in the afternoons. Such shifts are hardly discernible on a specific day. Figure 11 displays the exchange rate, trend and volatility components at a tranquil date halfway the sample, at June 30, 2005. The exchange rate drifted as in the left panel of Figure 11, where both the (log-) exchange rate p and the unobserved price process p^* are drawn. The difference between the two, due to the microstructure noise, is minimal. The cumulative trend line resulting from the intra-day trending is also drawn at the same scale. Though significant in the long run, on a specific day the effect is very small indeed.

Figure 11 about here

The right hand panel of the figure displays the components h_i (for the stochastic volatility, on the left axis) and $h_i + g_i$ (for the SV and spline jointly right axis) of the variance $\sigma_i^2 \equiv \delta_i \exp(h_i + g_i)$. The stochastic part moves slowly (for this day, over a range of 0.1). Including the seasonality the volatility moves over a range of 1.8 points in total. This occurs on most other days as well, that deterministic movements in volatility are far larger than the stochastic part of the variability.

6.7 Overview of FX results

The model-based approach to estimating the evolution over time of exchange rate prices and volatilities applied to high frequency data is new. It allows for a wealth of different and detailed viewpoints on the estimation results. This sections intends to summarise the main findings concerning the estimation of the model on exchange rate data.

The section started off finding problematic convergence for the parameters when the model was estimated without any constraints. This was caused by the thickness of the tails of the return distribution, with a degrees-of-freedom parameter ν dropping below 4. As the

kurtosis did not exist for these returns, the long-run variance of the SV component driven by parameters σ_h and ϕ had to be fixed at values linked to estimates at the daily frequency. Remaining parameters could be estimated well, with reasonable efficiency of the sample, given the refined MCMC algorithm applied here.

The parametric approach was able up to detect the jumps in the financial series. This of course does depend on the model assumptions. Allowing more flexibility in the SV process (by lowering ϕ or increasing σ_h) lowers the number of jumps detected, whereas assuming Gaussianity of the returns process would increase the number of detected jumps strongly (detailed results not reported here, but available upon request). With the variance parameters fixed at present values, integrated volatility measures comparable to the non-parametric realised variance and bipower variation resulted.

An advantage of a model-based approach is the ability to estimate jointly intra-day effects, like the trend-effect in exchange rates here. Day-to-day the effect is minimal, but like in Ranaldo (2007) a clear trending effect is found differing between morning and late afternoon trading hours. With the present approach, it is straightforward to include further dummies, checking for a trend component in each separate hour of the day.

Lastly, and possibly most importantly, this application shows that it is doable to extract intra-day volatility estimates in a sensible manner. These could eventually be used as input into financial models on portfolio optimisation, related to option prices etc. The extra level of detail in modelling can open up many other areas of research.

7 Conclusions and future work

This article studied the feasibility of a model incorporating micro-structure effects and jumps to approximate a jump-diffusion process. By modelling these aspects of the data at high frequency, statistics similar to the realised variation, bipower variation and related jump measures can be derived without having to take the non-parametric route. In both a simulation exercise and in an application this feasibility was found.

In the case of a simulated example, the underlying integrated variation process could be estimated from the data, even though a large number of observations were missing from the data set throughout the weekends. The model-based estimates delivered a better R^2 than non-parametric estimates in rebuilding the simulated integrated variation. Only with the

parametric approach did the parameter estimates not differ significantly from 0 and 1 for the intercept and the variance measure, respectively.

Apart from the Gaussian disturbance process with jumps for returns, it was found necessary to introduce also the possibility of a heavy-tailed returns process. The Student- t density that is implemented accounts for shocks of intermediate size and higher frequency than the jumps related to the separate jump process. In the simulation, with $\nu = 6$ according to the DGP, all parameters could be recovered from the data set quite precisely. The exchange rate data at the 5-minutes frequency however indicated a value for the degrees-of-freedom parameter of $\hat{\nu} < 4$. Such a value implies that the kurtosis of the return process does not exist, and indeed it was found that the model was not able to estimate the volatility of volatility as measured by SV parameters ϕ and σ_h . Instead, these had to be fixed at values derived from estimates of SV at a daily frequency, using a longer series of EUR/USD exchange rates. The implication of the non-existence of the kurtosis for the non-parametric measures of variance remains to be investigated.

In the application, only a rather simple use of the parametric setup for detecting the intraday trends was made. Combining this approach with e.g. data on macro-economic announcements would allow to measure the effect of surprises in the market on financial time series to a far higher precision. One of the driving factors for pursuing the present line of research were findings in Beine et al. (2007), where the impact of central bank interventions could only be modelled at a daily frequency. An intra-day analysis of a similar data set could shed further light on such issues.

The latter article uses a multivariate approach to modelling the time series. Such an extension of the present model is conceptually straightforward, delivering joint estimation of the volatilities of exchange rates or assets in a portfolio. The volatilities extracted here were constructed conditional on the full data set. Instead they could be filtered using an approach as the particle filter (Pitt and Shephard, 1999; Godsill et al., 2004). This in turn would allow to track intra-day, online, the risk of a portfolio of assets, and could lead to an alternative approach for pricing options.

Acknowledgements

This paper has benefited greatly from discussion with, among others, Drew Creal, Siem Jan Koopman, Sébastien Laurent, Bert Menkveld, Chris Neely, Angelo Ranaldo, and seminar participants at Aarhus university, the Tinbergen Institute and the ESEM'07 conference.

References

- Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica* 70, 223–262.
- Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics* 74, 487–528.
- Aït-Sahalia, Y. and R. Kimmel (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics* 83, 413–452.
- Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005a). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* 18, 351–416.
- Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005b). Ultra high frequency volatility estimation with dependent microstructure noise. Working paper 11380, NBER.
- Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239–1284.
- Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885–905.
- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2004). Parametric and nonparametric volatility measurement. In Y. Aït-Sahalia and L. Hansen (Eds.), *Handbook of Financial Econometrics*. North-Holland: Elsevier.

- Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* 89, 701–720.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review* 93, 38–62.
- Andersen, T. G., T. Bollerslev, and N. Meddahi (2005). Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica* 73, 279–296.
- Barndorff-Nielsen, O. E. and N. Shephard (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B* 63, 167–241.
- Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253–280.
- Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realised covariation: High frequency based covariance, regression and correlation in financial economics. *Econometrica* 72, 885–925.
- Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* 2, 1–48.
- Barndorff-Nielsen, O. E. and N. Shephard (2006a). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1–30.
- Barndorff-Nielsen, O. E. and N. Shephard (2006b). Impact of jumps on returns and realised variances: Econometric analysis of time-deformed Lévy processes. *Journal of Econometrics* 131, 217–252.

- Bauwens, L., M. Lubrano, and J.-F. Richard (1999). *Bayesian Inference in Dynamic Econometric Models*. Advanced Texts in Econometrics. Oxford: Oxford University Press.
- Beine, M., C. S. Bos, and S. Laurent (2007). The impact of central bank FX interventions on currency components. *Journal of Financial Econometrics* 5, 154–183.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bos, C. S., R. J. Mahieu, and H. K. Van Dijk (2000). Daily exchange rate behaviour and hedging of currency risk. *Journal of Applied Econometrics* 15, 671–696.
- Bos, C. S. and N. Shephard (2006). Inference for adaptive time series models: Stochastic volatility and conditionally Gaussian state space form. *Econometric Reviews* 25, 219–244.
- Chib, S. and E. Greenberg (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician* 49, 327–335.
- Chib, S., F. Nardari, and N. Shephard (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* 108, 281–316.
- Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134, 341–371.
- Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance* 8, 291–323.
- De Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika* 82, 339–350.
- Doornik, J. A. (2006). *Ox: An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Ltd.
- Drost, F. C., T. E. Nijman, and B. J. M. Werker (1998). Estimation and testing in models containing both jumps and conditional heteroscedasticity. *Journal of Business and Economic Statistics* 16, 237–243.
- Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

- Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika* 89, 603–615.
- Elerian, O., S. Chib, and N. Shephard (2001). Likelihood inference for discretely observed non-linear diffusions. *Econometrica* 69, 959–993.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica* 50, 987–1008.
- Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in volatility and returns. *Journal of Finance* 58, 1269–1300.
- Gelfand, A. E. and A. F. M. Smith (1990, June). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85, 398–409.
- Godsill, S. J., A. Doucet, and M. West (2004). Monte Carlo smoothing for nonlinear time series. *Journal of the American Statistical Association* 99, 156–168.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61, 247–264.
- Haung, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics* 3, 456–499.
- Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics* 12, 371–417.
- Jacquier, E., N. G. Polson, and P. E. Rossi (1995). Models and priors for multivariate stochastic volatility. CIRANO Working Paper 95s-18, CIRANO.
- Jungbacker, B. and S. J. Koopman (2006). Model-based measurement of actual volatility in high frequency data. In T. B. Fomby and D. Terrell (Eds.), *Advances in Econometrics*. New York: JAI Press. Volume 20, part A.
- Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65, 361–393.

- Koopman, S. J., B. Jungbacker, and E. Hol (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance* 12, 445–475.
- Koopman, S. J., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal* 2, 107–160.
- Lee, S. S. and P. A. Mykland (2008). Jumps in financial markets: A new non-parametric test and jump dynamics. *Review of Financial Studies*. Advance access, doi:10.1093/rfs/hhm056.
- Lyons, R. K. (2001). *The Microstructure Approach to Exchange Rates*. MIT Press.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21, 1087–1091.
- Mincer, J. and V. Zarnowitz (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic Forecasts and Expectations*. New York: NBER.
- Pitt, M. K. and N. Shephard (1999). Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association* 94, 590–599.
- Poirier, D. J. (1973). Piecewise regression using cubic spline. *Journal of the American Statistical Association* 68, 515–524.
- Rinaldo, A. (2007). Segmentation and time-of-day patterns in foreign exchange markets. Working paper 2007/3, Swiss National Bank.
- Rydberg, T. H. and N. Shephard (2003). Dynamics of trade-by-trade price movements: Decomposition and models. *Journal of Financial Economics* 1, 2–25.
- Sandmann, G. and S. J. Koopman (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics* 87, 271–301.
- Shephard, N. (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

- Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* 84, 653–667.
- Stoer, J. and R. Bulirsch (1991). *Introduction to Numerical Analysis* (Second ed.). Texts in Applied Mathematics 12. Berlin: Springer-Verlag.
- Tanner, M. A. and W. H. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82, 528–550.
- Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes, a study of daily sugar prices 1961-79. In O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice 1*, pp. 203–226. Amsterdam: North-Holland. Reprinted in Shephard (2005).
- Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* 12, 1019–1043.
- Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394–1411.

Table 1: Literature and topics

Article	Parametric	Intra-day	Diffusion	Micro	Jumps	SV
Aït-Sahalia (2002), Aït-Sahalia et al. (2005a)	-	+	+	+	-	-
Aït-Sahalia et al. (2005b)	-	+	+	+	-	±
Aït-Sahalia and Kimmel (2007)	+	-	+	+	-	+
Andersen et al. (2003)	+	+	-	+	±	-
Andersen et al. (2007)	-	+	+	+	+	+
Barndorff-Nielsen and Shephard (2002)	±	±	+	-	-	+
Barndorff-Nielsen and Shephard (2004b) a.o.	-	±	+	-	+	+
Drost et al. (1998)	+	-	-	-	+	-
Elerian et al. (2001)	+	-	+	-	-	-
Eraker et al. (2003)	+	-	+	-	+	+
Harvey et al. (1994), Jacquier et al. (1994)	+	-	-	-	-	+
Jungbacker and Koopman (2006)	+	+	+	+	-	+
Lyons (2001)	±	+	+	+	±	±
Rydberg and Shephard (2003)	+	+	-	+	-	-

Table 2: Moments of exchange rate returns

Moment	
N	76,339
Mean	-0.000
Standard deviation	0.035
Skewness	-0.150
Kurtosis	17.410

Table 3: Non-parametric jump test statistics

α	BN-S	L-M	% BN-S	% L-M
5.000	111	224	42.7	86.2
1.000	86	184	33.1	70.8
0.500	77	172	29.6	66.2
0.100	57	145	21.9	55.8
0.050	52	130	20.0	50.0
0.010	31	101	11.9	38.9
0.005	26	92	10.0	35.4

The table reports, for different significance levels, the number of days (out of a total of 260 days with sufficient observations) where a jump was detected either by the feasible jump test (20) or the L-M max test (22), indicating in the latter columns the percentages of days.

Table 4: DGP parameters and prior settings

Parameter	DGP	Prior	a	b	μ_π	σ_π
σ_ϵ	.01	IG-1	2	7500	.01	.005
λ	.1	Gamma	1	10	.1	.1
σ_κ	5	IG-1	2	.03	5	2.7
ϕ	.97	Beta	32	3.5	.9	.05
σ_h	.5	IG-1	2	3	.5	.27
$\nu - 2$	4	Exp	.25		4	4
β	0	\mathcal{N}	0	1	0	1
$\gamma_{\text{Sun 0:00h}}$	-4	\mathcal{N}	0	4	0	2
$\gamma_{\text{Wk 0:00h}}$	-2	\mathcal{N}	0	4	0	2
$\gamma_{\text{Wk 6:00h}}$	-1.5	\mathcal{N}	0	4	0	2
$\gamma_{\text{Wk 12:00h}}$	-1	\mathcal{N}	0	4	0	2
$\gamma_{\text{Wk 18:00h}}$	-1.2	\mathcal{N}	0	4	0	2

Table 5: Posterior statistics using simulated data

Parameter	DGP	Mode	Mean	σ	$[Q_{2.5\%}, Q_{97.5\%}]$	Inefficiency
σ_ϵ	.01	0.0101	0.0101	(0.000)	[0.010, 0.010]	29.702
λ	.1	0.1192	0.1254	(0.029)	[0.075, 0.188]	109.843
σ_κ	5	4.3766	4.5779	(0.662)	[3.517, 6.085]	39.357
ϕ	.97	0.9712	0.9711	(0.002)	[0.968, 0.974]	414.572
σ_h	.5	0.4721	0.4725	(0.012)	[0.449, 0.497]	450.896
ν	6	5.9879	5.9880	(0.164)	[5.678, 6.318]	33.803
$\gamma(\text{Sun}, 00:00\text{h})$	-4	-1.2913	-1.1539	(2.136)	[-5.281, 3.063]	5.253
$\gamma(\text{Mon}, 00:00\text{h})$	-2	-2.0007	-2.0001	(0.077)	[-2.153, -1.850]	37.339

The table reports the values of the parameters in the DGP, the posterior mode, mean, standard deviation, 95% interquartile range and a measure of inefficiency.

Table 6: Mincer-Zarnowitz regression results

	β_0	β_1	$[R^2]$
RV	0.0175 (0.009)	0.7768 (0.029)	[0.742]
BPV	-0.0376 (0.007)	1.0678 (0.026)	[0.871]
IMV	0.0073 (0.005)	0.9786 (0.020)	[0.906]

The table reports the regression results, both the estimated parameters β and their standard deviations (between parentheses) together with the R^2 , regressing the integrated variation on a constant and either realised variance, bipower variation or the integrated model variation.

Table 7: Posterior statistics using exchange rate data

Parameter	Mode	Mean	σ	$[Q_{2.5\%}, Q_{97.5\%}]$	Inefficiency
σ_ϵ	0.0103	0.0103	(0.000)	[0.010, 0.011]	34.676
λ_K	0.4178	0.4251	(0.057)	[0.322, 0.544]	272.174
σ_K	3.4213	3.4962	(0.256)	[3.041, 4.047]	75.778
ϕ	0.98			(fixed)	
σ_h	0.5			(fixed)	
ν	3.3920	3.3900	(0.061)	[3.272, 3.512]	89.788
$\beta(9-13h)$	-0.0102	-0.0105	(0.004)	[-0.019,-0.002]	5.026
$\beta(17-21h)$	0.0077	0.0077	(0.004)	[-0.001, 0.016]	1.185
$\gamma(\text{Sun},00:00h)$	-3.1609	-3.1700	(1.980)	[-7.104, 0.669]	37.911
$\gamma(\text{Mon},00:00h)$	-2.2163	-2.2130	(0.054)	[-2.319,-2.108]	36.477

See Table 5 for a description of the entries in the table.

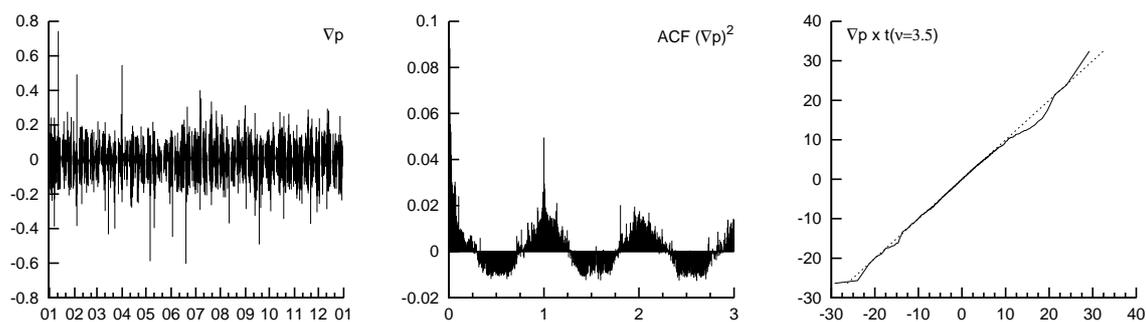


Figure 1: Returns through time, autocorrelation of squared returns, and QQ-plot of returns vs. Student- $t_{\nu=3.5}$ density.

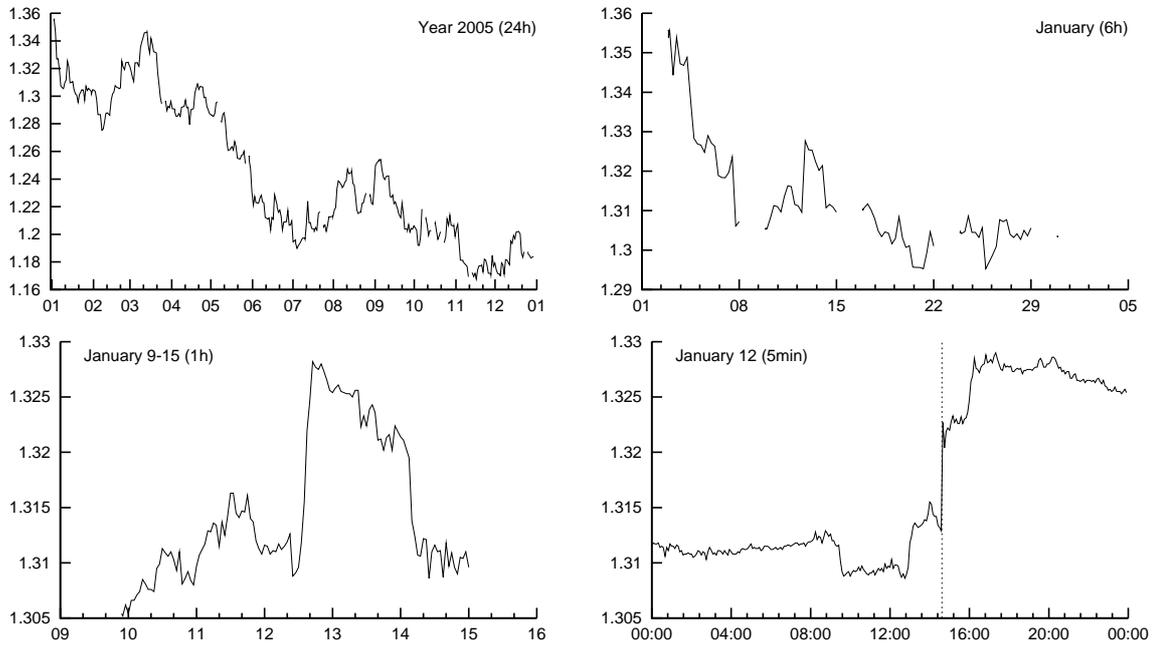


Figure 2: Euro/US Dollar exchange rate throughout 2005, for different periods (frequencies indicated between parentheses).

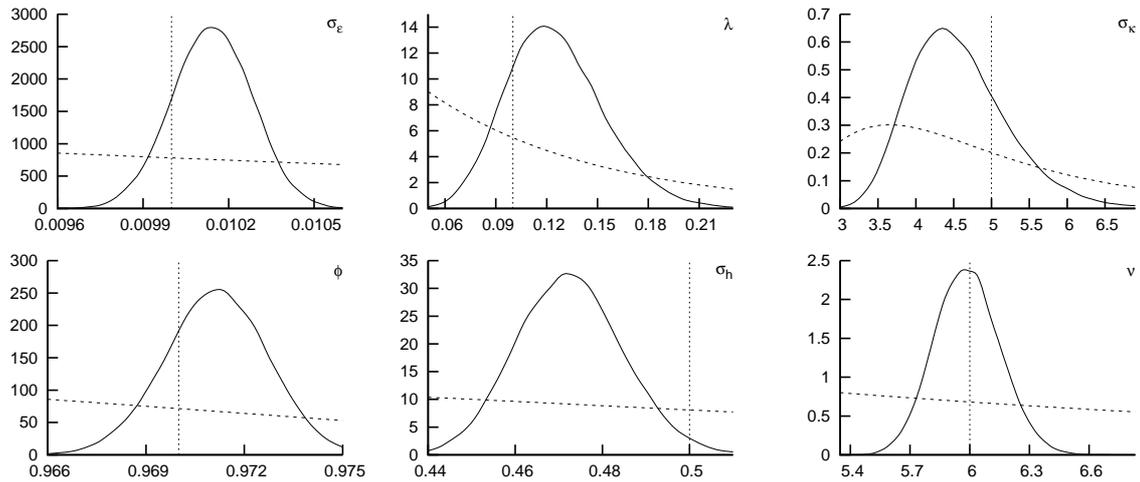


Figure 3: Prior and posterior density of parameters σ_ϵ (standard deviation of microstructure noise), λ (jump intensity), σ_κ (jump standard deviation), ϕ (SV autocorrelation), σ_h (SV long run standard deviation) and ν (Student- t degrees-of-freedom parameter) for simulated data, indicating the DGP value.

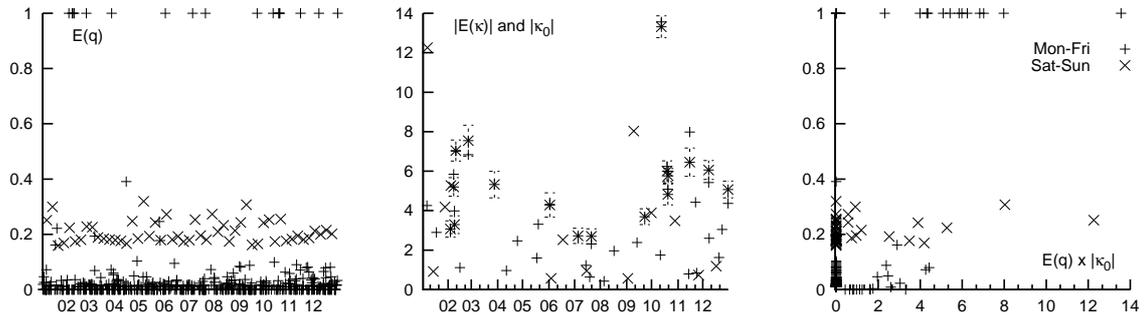


Figure 4: Posterior mean of jump probability $E(q)$ (first panel, against time), absolute expected jump size $|E(\kappa)|$ with one standard deviation error bounds and true jump sizes $|\kappa_0|$ without error bounds (second panel, against time), and a scatterplot of the posterior mean of jump probability vs. absolute true jump size (third panel), for the simulated data set. A '+' indicates an intra-week observation, a 'x' an observation falling just before the weekend.

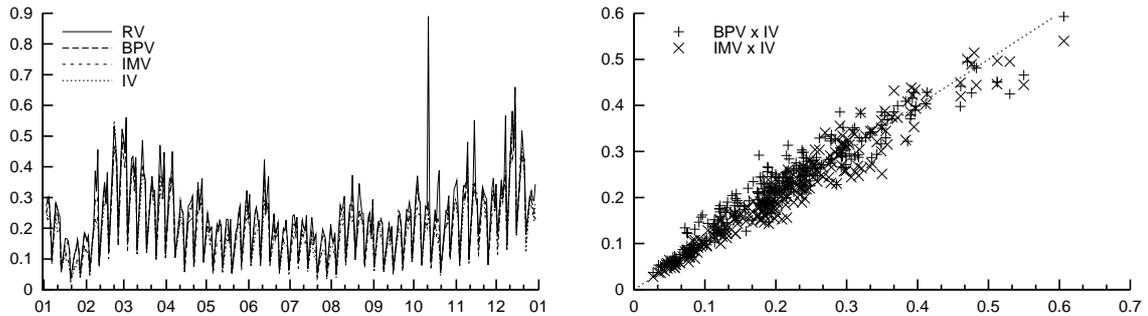


Figure 5: Integrated variation measures throughout the sample, against time (left panel) or BPV and IMV against IV (right panel), for simulated data.

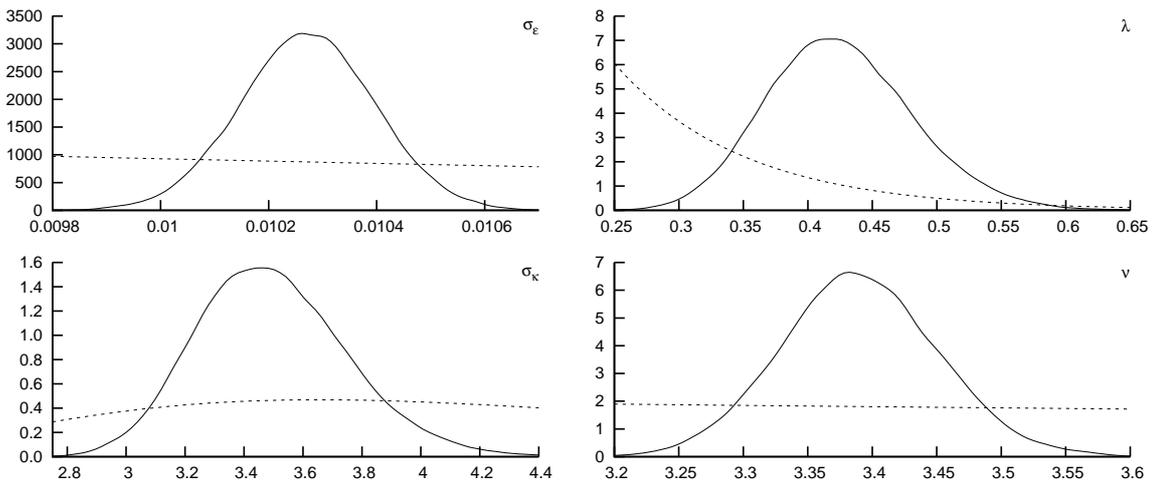


Figure 6: Prior and posterior density of parameters σ_ϵ (standard deviation of microstructure noise), λ (jump intensity), σ_κ (jump standard deviation) and ν (Student- t degrees-of-freedom parameter), using exchange rate data.

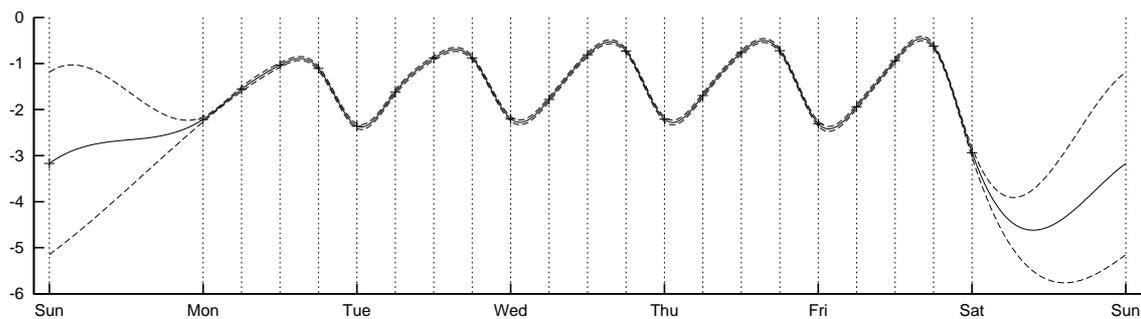


Figure 7: Volatility spline throughout the week for the exchange rate data, with one standard deviation error bands.

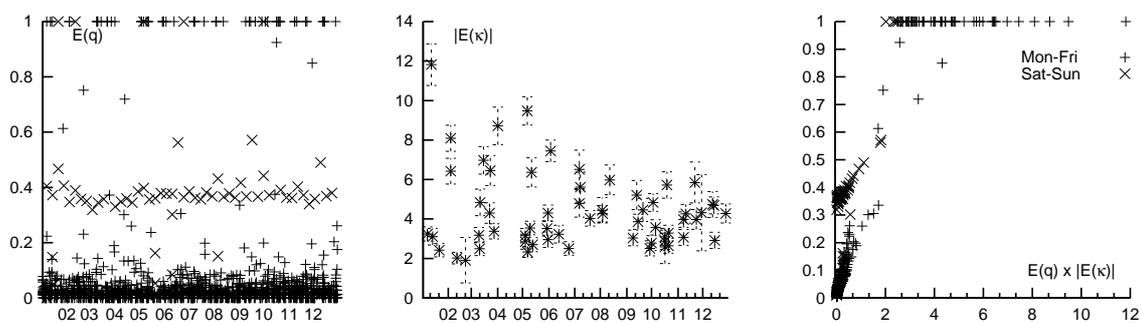


Figure 8: Posterior mean of jump probability $E(q)$ (first panel, against time), absolute expected jump size $|E(\kappa)|$ with one standard deviation error bounds (second panel, against time), and a scatterplot of the posterior mean of jump probability vs. absolute expected jump size (third panel), for the exchange rate data. A '+' indicates an intra-week observation, a 'x' an observation falling just before the weekend.

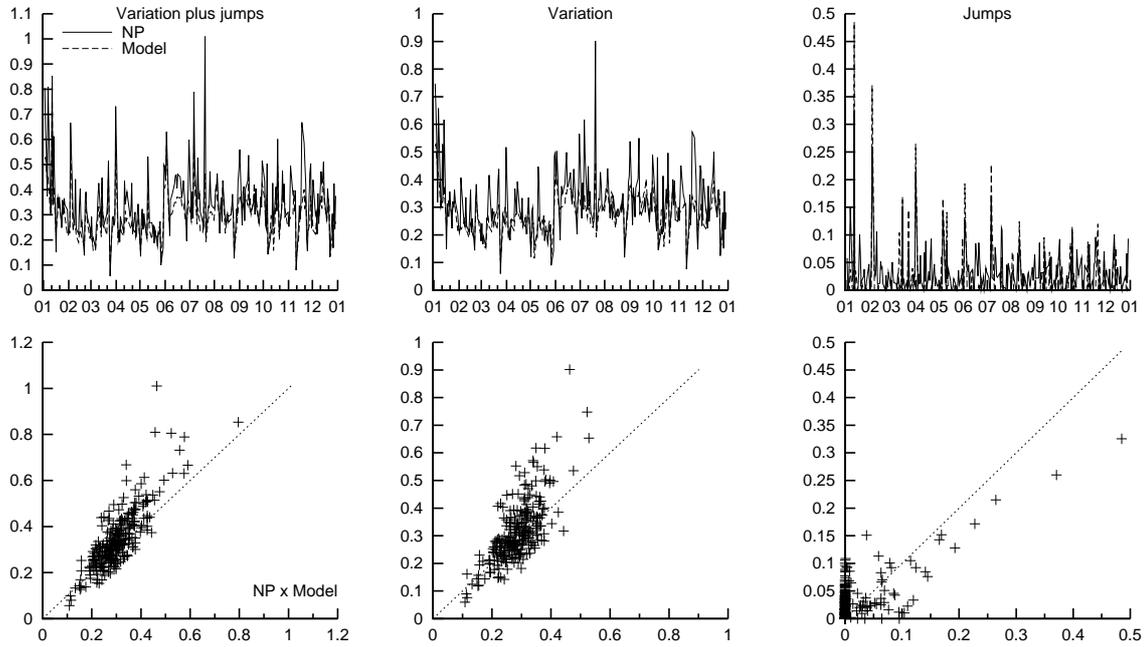


Figure 9: Non-parametric vs. model based measures of volatility and jumps, against time (top row) and against each other (bottom row, with model-based measures on the horizontal axis), for the exchange rate data. The first column displays realised variation against integrated model variation plus integrated jumps, the second contrasts bipower variation against integrated model variation, and the last column shows the jump measures.

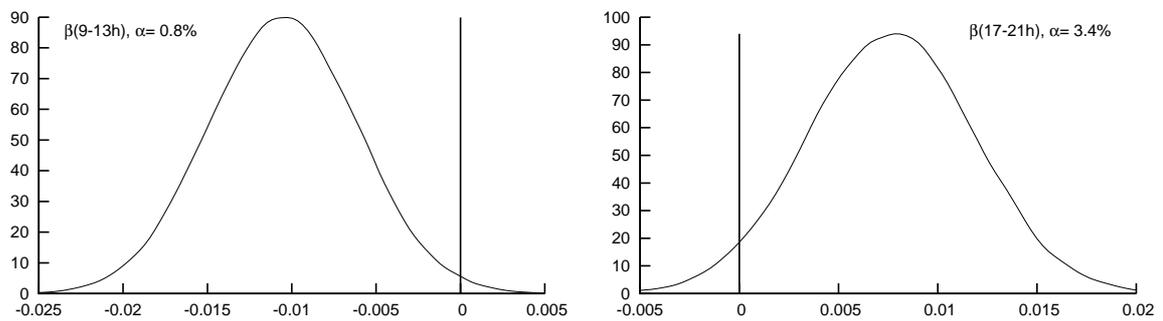


Figure 10: Intra-day trend in EUR/USD exchange rate during morning (9-13h, left panel) and late afternoon (17-21h, right panel).

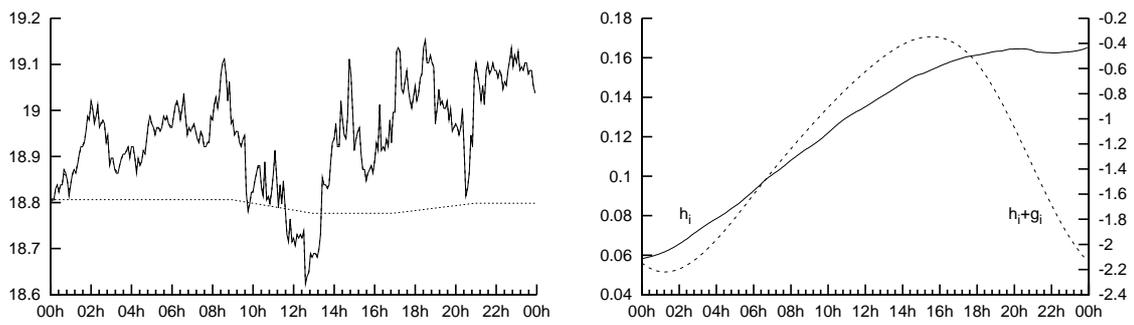


Figure 11: EUR/USD exchange rate on June 30, 2005 indicating the unobserved price and cumulative trend (left panel), and the variance components h_i (SV, right panel, against left axis) and $h_i + g_i$ (SV + spline, right axis).