

Modeling foreign exchange rates with jumps

John M. Maheu and Thomas H. McCurdy*

October 2003
preliminary and incomplete

Abstract

This paper proposes a new discrete-time model of returns in which jumps capture persistence in the conditional variance. Jump arrival is governed by a heterogeneous Poisson process. The intensity is directed by a latent stochastic autoregressive process, while the jump-size distribution allows for conditional heteroskedasticity. Model evaluation focuses on the dynamics of the conditional distribution of returns. For example, using Bayesian simulation methods, this paper estimates predictive densities which provide a period-by-period comparison of the performance of this new specification relative to a conventional stochastic volatility model. Further, in contrast to previous studies on the importance of jumps, we utilize realized volatility to assess out-of-sample variance forecasts.

*Maheu (jmaheu@chass.utoronto.ca) is from the Department of Economics, University of Toronto, and McCurdy (tmccurdy@rotman.utoronto.ca) is from the Joseph L. Rotman School of Management, University of Toronto, and an Associated Fellow of CIRANO. We thank Olsen and Associates for providing the intraday foreign exchange quotes. We have benefited from discussions with Stephen Gordon.

1 Introduction

Measuring and forecasting the distribution of returns is important for many problems in finance. Pricing of financial securities, risk management decisions, and portfolio allocations all depend on the distributional features of returns. Possibly the most important feature of financial returns is the structure in the variance. As a result, a vast literature has sprung from the ARCH model of Engle (1982) and the stochastic volatility (SV) approach of Taylor (1986).

Current research has documented the importance of jump dynamics in combination with autoregressive stochastic volatility for modeling returns. Examples of this work include Andersen, Benzoni, and Lund (2002), Bates (2000), Chernov, Gallant, Ghysels, and Tauchen (2003), Chib, Nardari, and Shephard (2002), Eraker, Johannes, and Polson (2003), Jorion (1988), and Maheu and McCurdy (2004) among others. Jumps provide a useful addition to SV models by explaining occasional, large abrupt moves in financial markets, but they are generally not used to capture volatility clustering. As a result, jumps account for neglected structure, usually tail dynamics, that autoregressive SV cannot capture.¹

All financial data are measured in discrete time which suggests that jumps provide a natural framework to model price moves. Nevertheless, surprisingly few applications focus on the potential performance of a pure jump model. Recent research, including Das (2002), Johannes, Kumar, and Polson (1999), and Oomen (2002), suggests that jumps alone could provide a good specification for financial returns. Those applications all feature some form of dependence in the arrival rate of jumps. For example, Johannes, Kumar, and Polson (1999) allow jump arrival to depend on past jumps and the absolute value of returns. Indeed, the success of pure jump models will depend on whether the specification can explain dynamics of the conditional distribution, in particular, volatility clustering. However, the performance of existing models in this regard is unclear. For example, can models with only jumps produce good volatility forecasts? Are they competitive with standard volatility models? The purpose of this paper is to investigate these questions.

This paper proposes a new discrete-time model of returns in which jumps capture persistence in the conditional variance. The jump intensity is directed by a latent stochastic autoregressive process. Therefore, jump arrivals can cluster. We also allow the jump-size distribution to be conditionally heteroskedastic. Larger jumps occur during volatile periods and smaller ones during quiet periods.

We propose a Markov chain Monte Carlo (MCMC) approach to model estimation. We follow Johannes, Kumar, and Polson (1999) in treating unobserved state variables such as jump times and jump sizes as parameters. A byproduct of MCMC output is estimates of these quantities which incorporate parameter uncertainty.

¹Additional SV factors is another possible solution. For example, Chernov, Gallant, Ghysels, and Tauchen (2003) find that a multifactor loglinear SV specification is equivalent to an affine class of SV with jumps for equity data.

Model evaluation focuses on the dynamics of the conditional distribution of returns, as we compare this new specification to a conventional stochastic volatility model. Using Bayesian simulation methods, we estimate predictive likelihoods for models as suggested by Geweke (1995). This allows for a period-by-period comparison of the jump model and the SV model, and is particularly useful in identifying influential observations. From these calculations we report the cumulative model probability as a function of time. This provides insight into the performance of volatility forecasts.

In addition to conditional density evaluation, forecasts of volatility are compared. Andersen and Bollerslev (1998) show that accurate measures of ex post volatility can be constructed from high frequency intraday data. The sum of squared intraday returns is an efficient measure of daily volatility. Following Andersen, Bollerslev, Diebold, and Labys (2001), this estimator is often called realized volatility and is calculated based on 5-minute price data from the foreign exchange market for JPY-USD and DEM-USD rates. Our out-of-sample forecasts of volatility are assessed using these realized volatility measures.

Results — To be completed

This paper is organized as follows. The next section presents a heterogeneous jump model for foreign exchange returns, while Section 3 briefly discusses a benchmark SV model used for comparison purposes. Section 4 considers Bayesian estimation of the jump model. The use, and estimation of predictive densities for model comparison is reviewed in Section 5, while model forecasts are explained in Section 6. Data sources are found in Section 7, results in Section 8 and conclusions in Section 9. An Appendix contains detailed calculations for the estimation algorithms.

2 Heterogeneous Jump Parameterization

This section proposes a new discrete-time model which can capture the autoregressive pattern in the conditional variance of returns by allowing jumps to arrive according to a heterogeneous Poisson process. Our parameterization includes a latent autoregressive structure for the jump intensity, as well as a conditionally heteroskedastic variance for the jump-size distribution. The mean of the jump-size distribution can be significantly different from zero, allowing the specification to capture a skewed distribution of returns. Johannes, Kumar, and Polson (1999) consider a reversal effect through the jump-size mean for equity data, however, this is likely to be less important for FX rates.

The parameterization of the heterogeneous jump model is as follows:

$$r_t = \mu + \sigma z_t + J_t \xi_t, \quad z_t \sim N(0, 1) \quad (2.1)$$

$$\xi_t \sim N(\mu_J, \sigma_{J,t}^2), \quad J_t \in \{0, 1\} \quad (2.2)$$

$$P(J_t = 1|w_t) = \lambda_t \text{ and } P(J_t = 0|w_t) = 1 - \lambda_t \quad (2.3)$$

$$\lambda_t = \frac{\exp(w_t)}{1 + \exp(w_t)} \quad (2.4)$$

$$w_t = \gamma_0 + \gamma_1 w_{t-1} + u_t, \quad u_t \sim NID(0, 1), \quad |\gamma_1| < 1 \quad (2.5)$$

$$\sigma_{J,t}^2 = \eta_0 + \eta_1 X_{t-1}, \quad (2.6)$$

where r_t denotes returns, $t = 1, \dots, T$, μ is the mean of the returns, conditional on no jump, and ξ_t is the jump size and follows a normal distribution. J_t is an indicator that identifies when jumps occur. In particular, the set $\{J_t = 1\}_{t=1}^T$ denotes jump times. λ_t is the time-varying jump intensity or arrival process which is directed by the latent autoregressive process w_t . The logistic function ensures that w_t is mapped into a $(0, 1)$ interval for λ_t .

Besides capturing occasional large moves in returns, this specification can account for volatility clustering through persistence in w_t . One interpretation of w_t is that it represents the unobserved news flows into the market that causes trading activity. The variance of the jump-size distribution, $\sigma_{J,t}^2$, is allowed to be a function of weakly exogenous regressors X_{t-1} . In this paper we consider $X_{t-1} = |r_{t-1}|$. This permits the jump-size variance to be sensitive to recent market conditions. For example, if $\eta_1 > 0$ jumps will tend to be larger (smaller) in volatile (quiet) markets.

Imposing the restrictions $\lambda_t = \lambda, \forall t$ and $\eta_1 = 0$ obtains a simple jump model with an iid arrival of jumps and a homogeneous jump size distribution. Although not considered in this paper, it is straightforward to allow additional dynamics like a different jump-size distribution or time dependence in μ_J .

3 SV model

In this paper we consider a standard log-linear stochastic volatility (SV) model as a benchmark for comparison purposes. A large literature discusses estimation methods for this model.² Surveys of the SV literature are provided by Ghysels, Harvey, and Renault (1996), Shephard (1996) and Taylor (1994).

The SV model is parameterized as,

$$r_t = \mu + \exp(h_t/2)z_t, \quad z_t \sim N(0, 1) \quad (3.1)$$

$$h_t = \rho_0 + \rho_1 h_{t-1} + \sigma_v v_t, \quad v_t \sim N(0, 1). \quad (3.2)$$

²Important contributions to the theory of estimation include: Melino and Turnbull (1990) Jacquier, Polson, and Rossi (1994), Andersen and Sorensen (1996), Kim, Shephard, and Chib (1998), Danielsson and Richard (1993), Gouriou, Monfort, and Renault (1993), Gallant, Hsieh, and Tauchen (1997), and Alizadeh, Brandt, and Diebold (2002).

Given h_t , the conditional variance of returns is $\exp(h_t)$.

Compared to GARCH models, estimation is more difficult due to the latent volatility which must be integrated out of the likelihood. Bayesian methods rely on Markov Chain Monte Carlo (MCMC) sampling to estimate SV models. The properties of the estimator compare favorably with other approaches. It is straightforward to obtain smoothed estimates of volatility from MCMC output. In addition, these estimates of volatility take parameter uncertainty into account.

4 Posterior Inference

Johannes and Polson (2003) provide a good introduction to Bayesian methods for financial models including a simple jump model. Eraker, Johannes, and Polson (2003) and Johannes, Kumar, and Polson (1999) discuss a data augmentation approach to deal with jump times and jump sizes.

From a Bayesian perspective, inference regarding parameters takes place through the posterior which incorporates both the prior and likelihood function. Let the history of data be denoted as $\Phi_t = \{r_1, \dots, r_t\}$. In the case of the Jump model we augment the parameters $\theta = \{\mu, \sigma^2, \mu_J, \eta, \gamma\}$ where $\eta = \{\eta_0, \eta_1\}$, and $\gamma = \{\gamma_0, \gamma_1\}$, with the unobserved state vectors $\omega = \{\omega_1, \dots, \omega_T\}$, jump times $J = \{J_1, \dots, J_T\}$, and jump sizes $\xi = \{\xi_1, \dots, \xi_T\}$ and treat these as parameters. For the Jump model, Bayes rule gives us

$$p(\theta, \omega, J, \xi | \Phi_T) \propto p(r | \theta, \omega, J, \xi, \Phi_T) p(\omega, J, \xi | \theta) p(\theta) \quad (4.1)$$

where $r = \{r_1, \dots, r_T\}$. $p(r | \theta, \omega, J, \xi)$ is the joint density of returns conditional on the state variables ω, J, ξ , $p(\omega, J, \xi | \theta)$ is the density of the state variables and $p(\theta)$ is the prior. In practice, analytical results are not available and we use MCMC methods to draw samples from the posterior. Recent surveys of MCMC methods include Chib (2001), Geweke (1997) and Robert and Casella (1999).

MCMC theory allows valid draws from the posterior to be obtained by sampling from a series of conditional distributions. It is often much easier to work with the conditional distributions. In the limit, draws converge to samples from the posterior. A consistent estimate of any function of the parameter vector can be constructed from sample averages. For instance, if we have $\{\mu^{(i)}\}_{i=1}^N$ draws of μ from the posterior, and assuming the integral of $g(\mu)$ with respect to the marginal posterior exists, we can estimate $E[g(\mu)]$ as $\frac{1}{N} \sum_{i=1}^N g(\mu^{(i)})$.³ Further assumptions on the integrability of $g(\mu)^2$ permit consistent estimation of the asymptotic standard error of the estimate, based on conventional time series methods.⁴

In the following we denote the vector θ excluding the k th element θ_k , as $\theta_{-\theta_k}$, the subvector $\{\omega_t, \dots, \omega_\tau\}$ as $\omega_{(t,\tau)}$ and ω excluding $\omega_{(t,\tau)}$ as $\omega_{-(t,\tau)}$. Sampling is based on

³For example, to compute $E\mu$, set $g(\mu) = \mu$, to compute $E\mu^2$ and the variance, set $g(\mu) = \mu^2$, etc. See Tierney (1994) for technical details.

⁴For instance, see Geweke (1992).

Gibbs and Metropolis-Hasting (MH) routines. Draws from the posterior $\psi = \{\theta, \omega, J, \xi\}$ are obtained by cycling over the following steps.

1. sample $\mu|\theta_{-\mu}, \omega, J, \xi, r$
2. sample $\sigma^2|\theta_{-\sigma^2}, \omega, J, \xi, r$
3. sample $\mu_J|\theta_{-\mu_J}, \omega, J, \xi, r$
4. sample $\eta|\theta_{-\eta}, \omega, J, \xi, r$
5. sample blocks $\omega_{(t,\tau)}|\theta, \omega_{-(t,\tau)}, J, \xi, r, t = 1, \dots, T$.
6. sample $\gamma|\theta_{-\gamma}, \omega, J, \xi, r$
7. sample $\xi|\theta, \omega, J, r$
8. sample $J|\theta, \omega, \xi, r$
9. goto 1

A pass through 1 - 8 provides a draw from the posterior. We repeat this several thousand times and collect these draws after an initial burn-in period. There does not appear to be an agreed upon method to establish convergence of the MCMC samples. It is often useful to view time series plots of the parameter draws versus iterations, for multiple starts of the chain, in order to assess convergence. Note that the parameters $\eta, \gamma, J, \xi,$ and ω are sampled as blocks which may contribute to better mixing of the MCMC output. Detailed steps of the algorithm are collected in the Appendix.

5 Model Comparison

The key ingredient in Bayesian model comparison is the marginal likelihood which can be used to form Bayes factors or model probabilities. However, a drawback of any statistical approach that summarizes a model's performance with a single number is understanding why and when a model performs well or poorly. Geweke (1995) suggests the use of a predictive density decomposition of the marginal likelihood. Estimating the predictive density allows us to compare models on an observation by observation basis. This may be useful in identifying influential observations, or periods that make a large contribution to Bayes factors. Applications of this idea include Gordon (1997), and Min and Zellner (1993). In addition, the calculation of the predictive density for model comparison is also useful if we are interested in other features of the predictive density such as variance forecasts.

Consider a model with parameter vector Θ . The predictive density for observation y_{t+t} based on the information set Φ_t is,

$$p(y_{t+1}|\Phi_t) = \int p(y_{t+1}|\Phi_t, \Theta)p(\Theta|\Phi_t)d\Theta \quad (5.1)$$

where $p(\Theta|\Phi_t)$ is the posterior, and $p(y_{t+1}|\Phi_t, \Theta)$ is the conditional distribution. When it is clear we suppress conditioning on a model for notational convenience. Evaluating (5.1) at the realized \tilde{y}_{t+1} gives the predictive likelihood,

$$\hat{p}_t^{t+1} = p(\tilde{y}_{t+1}|\Phi_t) = \int p(\tilde{y}_{t+1}|\Phi_t, \Theta)p(\Theta|\Phi_t)d\Theta \quad (5.2)$$

which can be estimated from MCMC output. Models that have a larger predictive likelihood are preferred to ones with a smaller value, as they are more likely to have generated the data. The predictive likelihood for observations $\tilde{y}_u, \dots, \tilde{y}_v$, $u < v$ is

$$\hat{p}_{u-1}^v = p(\tilde{y}_u, \dots, \tilde{y}_v|\Phi_{u-1}) = \int p(\tilde{y}_u, \dots, \tilde{y}_v|\Phi_{u-1}, \Theta)p(\Theta|\Phi_{u-1})d\Theta = \prod_{i=u-1}^{v-1} \hat{p}_i^{i+1}, \quad (5.3)$$

see Geweke (1995) for details. If $u = 1$ and $v = T$ then (5.3) provides a full decomposition of the marginal likelihood. In practice we will use a *training sample* that we condition on for all models. This initial sample of observations $1, 2, \dots, u - 1$, combined with the likelihood and prior, forms a new prior, $p(\Theta|\Phi_{u-1})$, on which all calculations are based. If u is large then $p(\Theta|\Phi_{u-1})$ will be dominated by the likelihood function and the original prior $p(\Theta)$ will have a minimal contribution to model comparison exercises. Note that conditional on this training sample, the log predictive Bayes factor in favor of model j versus k for the data $\tilde{y}_u, \dots, \tilde{y}_v$ is

$$\log B_{j,k,u-1}^v = \sum_{i=u-1}^{v-1} \log \frac{\hat{p}_{j,i}^{i+1}}{\hat{p}_{k,i}^{i+1}} \quad (5.4)$$

where the predictive likelihood is now indexed by the model j and model k . Model probabilities associated with the predictive likelihood are calculated as,

$$p(M_i|\tilde{y}_u, \dots, \tilde{y}_v, \Phi_{u-1}) = \frac{p(\tilde{y}_u, \dots, \tilde{y}_v|M^i, \Phi_{u-1})p(M^i|\Phi_{u-1})}{\sum_{k=1}^K p(\tilde{y}_u, \dots, \tilde{y}_v|M^k, \Phi_{u-1})p(M^k|\Phi_{u-1})}, \quad i = 1, \dots, K \quad (5.5)$$

where there are K models, and M^k denotes model k . In all calculations equal prior model probabilities are used. Models with a high predictive likelihood will be assigned a high model probability. Calculating (5.5) for each $v = u + 1, \dots, T$, provides a cumulative assessment of the evidence for model i as more observations are used.

5.1 Calculations

5.1.1 Jump Model

The predictive likelihood can be estimated from the MCMC output. For the Jump model, with $y_{t+1} = r_{t+1}$, and $\Theta = (\theta, \omega)$ the predictive likelihood is,

$$\hat{p}_t^{t+1} = \int p(r_{t+1}|\theta, \omega_t, \Phi_t)p(\theta, \omega_t|\Phi_t)d\theta d\omega_t \quad (5.6)$$

$$= \int p(r_{t+1}|\theta, \omega_{t+1}, \Phi_t)p(\omega_{t+1}|\omega_t, \theta)p(\theta, \omega_t|\Phi_t)d\theta d\omega_t d\omega_{t+1} \quad (5.7)$$

$$\approx \frac{1}{N} \sum_{i=1}^N \frac{1}{R} \sum_{j=1}^R p(r_{t+1}|\theta^{(i)}, \omega_{t+1}^{(j)}, \omega_t^{(i)}, \Phi_t) \quad (5.8)$$

where i denotes the i th draw from the posterior $p(\theta, \omega|\Phi_t)$, $i = 1, \dots, N$, j indexes simulated values of w_{t+1} , and

$$p(r_{t+1}|\theta^{(i)}, \omega_{t+1}^{(j)}, \omega_t^{(i)}, \Phi_t) = \lambda_{t+1}^{(j)} \phi(r_{t+1}|\mu^{(i)} + \mu_J^{(i)}, \sigma^{2(i)} + \sigma_{J,t+1}^{2(i)}) \quad (5.9)$$

$$+ (1 - \lambda_{t+1}^{(j)}) \phi(r_{t+1}|\mu^{(i)}, \sigma^{2(i)}) \quad (5.10)$$

$$\lambda_{t+1}^{(j)} = \frac{\exp(w_{t+1}^{(j)})}{1 + \exp(w_{t+1}^{(j)})} \quad (5.11)$$

$$w_{t+1}^{(j)} = \gamma_0^{(i)} + \gamma_1^{(i)} w_t^{(i)} + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim NID(0, 1). \quad (5.12)$$

$\phi(x|\mu, \sigma^2)$ denotes the normal density function evaluated at x with mean μ and variance σ^2 . The following steps summarize the estimation of \hat{p}_t^{t+1} .

1. Set model parameters to the i th draw from the posterior $\{\theta^{(i)}, \omega^{(i)}\}$.
2. Generate $j = 1, \dots, R$, values of $w_{t+1}^{(j)}$ according to (5.12) and calculate the average of $p(r_{t+1}|\theta^{(i)}, \omega_{t+1}^{(j)}, \omega_t^{(i)}, \Phi_t)$ for these values. Save the result.
3. If $i < N$ then set $i = i + 1$ and goto 1
4. Calculate the average of the N values obtained in step 2.

Standard errors for this estimate can be calculated as usual from MCMC output (for example Geweke (1992)).

5.1.2 SV Model

Very similar calculations are used for the SV model.

$$\hat{p}_t^{t+1} = \int p(r_{t+1}|\theta, h_t, \Phi_t)p(\theta, h_t|\Phi_t)d\theta dh_t \quad (5.13)$$

$$= \int p(r_{t+1}|\theta, h_{t+1}, \Phi_t)p(h_{t+1}|\omega_t, \theta)p(\theta, h_t|\Phi_t)d\theta dh_t dh_{t+1} \quad (5.14)$$

$$\approx \frac{1}{N} \sum_{i=1}^N \frac{1}{R} \sum_{j=1}^R p(r_{t+1}|\theta^{(i)}, h_{t+1}^{(j)}, h_t^{(i)}, \Phi_t). \quad (5.15)$$

To summarize,

1. Set model parameters to the i th draw from the posterior $\{\theta^{(i)}, h^{(i)}\}$.
2. Generate $j = 1, \dots, R$, values of $h_{t+1}^{(j)}$ according to

$$h_{t+1}^{(j)} = \rho_0^{(i)} + \rho_1^{(i)} h_t^{(i)} + \sigma_\nu^{(i)} \nu_{t+1}, \quad \nu_{t+1} \sim NID(0, 1) \quad (5.16)$$

and calculate the average of $p(r_{t+1}|\theta^{(i)}, h_{t+1}^{(j)}, h_t^{(i)}, \Phi_t) = \phi(r_{t+1}|\mu^{(i)}, \exp(h_{t+1}^{(j)}))$ for these values. Save the result.

3. If $i < N$ then set $i = i + 1$ and goto 1
4. Calculate the average of the N values obtained in step 2.

6 Volatility Forecasts

Consider a generic model with parameter vector Θ . Moments of y_{t+1} (assuming they exist), based on time t information, can be calculated as

$$E[y_{t+1}^s|\Phi_t] = \int y_{t+1}^s p(y_{t+1}|\Phi_t) dy_{t+1} \quad (6.1)$$

$$= \int y_{t+1}^s \int p(y_{t+1}|\Phi_t, \Theta)p(\Theta|\Phi_t)d\Theta dy_{t+1} \quad (6.2)$$

$$= \int E[y_{t+1}^s|\Phi_t, \Theta]p(\Theta|\Phi_t)d\Theta, \quad s = 1, 2, \dots \quad (6.3)$$

which can be approximated from MCMC output. Note that the posterior $p(\Theta|\Phi_t)$ was used in the last section for calculating the predictive likelihood. Therefore, very little additional computation is needed to obtain out-of-sample forecasts. Conditional variance forecasts are $E[y_{t+1}^2|\Phi_t] - E[y_{t+1}|\Phi_t]^2$.

6.1 Calculations

6.1.1 Jump model

$$E[r_{t+1}^s | \Phi_t] = \int E[r_{t+1}^s | \Phi_t, \theta, \omega_t] p(\theta, \omega_t | \Phi_t) d\theta d\omega_t \quad (6.4)$$

$$\approx \frac{1}{N} \sum_{i=1}^N E[r_{t+1}^s | \Phi_t, \theta^{(i)}, \omega_t^{(i)}], \quad s = 1, 2, \quad (6.5)$$

where $\{\theta^{(i)}, \omega_t^{(i)}\}$ is the i th draw from the posterior distribution, $p(\theta, \omega | \Phi_t)$. For the variance we require the first two moments

$$E[r_{t+1} | \Phi_t, \theta^i, \omega_t^i] = \mu^{(i)} + \mu_J^{(i)} E[J_{t+1} | \theta^{(i)}, \omega_t^{(i)}] \quad (6.6)$$

$$E[r_{t+1}^2 | \Phi_t, \theta^{(i)}, \omega_t^{(i)}] = \mu^{(i)2} + \sigma^{(i)2} + (\mu_J^{(i)2} + \sigma_J^{(i)2}) E[J_{t+1} | \theta^{(i)}, \omega_t^{(i)}] \quad (6.7)$$

$$+ 2\mu^{(i)} \mu_J^{(i)} E[J_{t+1} | \theta^{(i)}, \omega_t^{(i)}]. \quad (6.8)$$

These moments are substituted into (6.5) for each draw of $\{\theta^{(i)}, \omega_t^{(i)}\}$. Since λ_{t+1} depends on ω_{t+1} we approximate each conditional expectation of J_{t+1} as

$$E[J_{t+1} | \theta^{(i)}, \omega_t^{(i)}] \approx \frac{1}{R} \sum_{j=1}^R \frac{\exp(\omega_{t+1}^{(j)})}{1 + \exp(\omega_{t+1}^{(j)})} \quad (6.9)$$

$$(6.10)$$

where $\omega_{t+1}^{(j)}$ is generated from

$$\omega_{t+1}^{(j)} = \gamma_0^{(i)} + \gamma_1^{(i)} \omega_t^{(i)} + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim NID(0, 1), \quad j = 1, \dots, R. \quad (6.11)$$

6.1.2 SV model

For the SV model we have,

$$E[r_{t+1}^s | \Phi_t] = \int E[r_{t+1}^s | \Phi_t, \theta, h_t] p(\theta, h_t | \Phi_t) d\theta dh_t \quad (6.12)$$

$$\approx \frac{1}{N} \sum_{i=1}^N E[r_{t+1}^s | \Phi_t, \theta^i, h_t^i], \quad s = 1, 2, \quad (6.13)$$

The conditional moments are

$$E[r_{t+1} | \Phi_t, \theta^{(i)}, h_t^{(i)}] = \mu^{(i)} \quad (6.14)$$

$$E[r_{t+1}^2 | \Phi_t, \theta^{(i)}, h_t^{(i)}] = \mu^{(i)2} + \exp(\rho_0^{(i)} + \rho_1^{(i)} h_t^{(i)} + \sigma_\nu^{(i)2}/2). \quad (6.15)$$

7 Data

Five-minute intraday FX quote data (bid and ask) were obtained from Olsen and Associates for every day from 1986/2/1 - 2002/12/31 for both DEM-USD and JPY-USD exchange rates.⁵ With a few exceptions, the construction of daily returns and realized volatility closely follow Andersen, Bollerslev, Diebold, and Labys (2001) and Maheu and McCurdy (2002).

Currencies trade 24 hours a day, 7 days a week. The raw data included a number of missing observations. Missing quotes on the 5-minute grid were linearly interpolated from the nearest available quote. Prices were constructed as the midpoint of the bid and ask quote. This resulted in 1.8 million, 5-minute price observations. A day was defined as beginning at 00:05 GMT and ending 24:00 GMT.⁶ Continuously compounded 5-minute returns (in percent) were constructed from the price data. Following Andersen, Bollerslev, Diebold, and Labys (2001), all weekends (Saturday and Sunday) were removed as well as the following slow trading days: December 24-26, 31, January 1,2. In addition, the moving holidays: Good Friday, Easter Monday, Memorial Day, July Fourth, Labor Day, Thanksgiving and the day after, as well as any days in which more than half (144) of the day's quotes were missing, were removed. The remaining data were linearly filtered by an MA(q) to remove autocorrelation which may be due to the discrete nature of bid/ask quotes and market microstructure effects. For the JPY-USD, $q = 4$ and for DEM-USD, $q = 10$.

From these filtered data, daily realized volatility was constructed as

$$RV_t = \sum_{j=1}^{288} r_{t,j}^2 \quad (7.1)$$

where $r_{t,j}$ is the j th 5-minute return in day t . Daily returns were constructed as the sum of the intraday 5-minute returns, $r_t = \sum_{j=1}^{288} r_{t,j}$.

In a series of papers, Andersen, Bollerslev, Diebold and co-authors, Barndorff-Nielson and Shephard, among others, have documented various features of realized volatility and derived their asymptotic properties as the sampling frequency of prices increases over a fixed-time interval. See Andersen, Bollerslev, and Diebold (2003) and Barndorff-Nielsen, Graversen, and Shephard (2003) for an overview of the results.

Table 1 reports our summary statistics for both currencies. Notice that the JPY-USD return is the more volatile of the two currencies. In addition, JPY-USD returns display larger kurtosis, and vary over a larger range.

⁵From 1999/1/1 on, we converted the EUR-USD rate to an implied DEM-USD rate based on the final fixed exchange rate of 1.95583 DEM/EUR at 1998/12/31.

⁶Data characteristics using the start time 21:05 and end time 21:00 were very similar.

8 Results

The data sample for JPY-USD is 1986/12/16 to 2002/12/31 (4001 observations), and 1986/11/04 to 2002/12/31 (4025 observations) for the DEM-USD. Tables 2 and 3 contain full-sample model estimates based on those JPY-USD and DEM-USD FX rates. For example, estimates of the SV specification in Table 2 imply an unconditional variance of .617 which is close to the value in Table 1. For each model a total of 90000 MCMC iterations were performed. The first 10000 draws were discarded to minimize the influence of starting values. Thus, $N = 80000$ samples from the posterior distribution are used to calculate posterior moments.

Figure 1 shows various features of the jump model while Figures 2 and 3 display characteristics of the out-of-sample performance for the Jump and SV specification for the JPY-USD data. Out-of-sample calculations are based on a training sample of the first 3000 observations. Therefore predictive likelihood estimates and model forecasts appearing in Table 5 are based on the remaining 1001 observations for the JPY-USD, and 1026 observations for the DEM-USD. Predictive likelihood estimates are displayed in Figure 3.

Table 2 contains estimates of the Jump model and a SV specification. Note that Bayesian methods provide exact finite sample performance conditional on the prior specification. Reported are various features of the posterior distribution. Based on the posterior mean the unconditional expectation for the autoregressive latent variable, ω_t , driving jump arrival is -2.62. This implies that jumps are infrequent. For instance, the empirical average of λ_t is .07. This is consistent with previous studies which combine SV with simple jumps. However, in our case, the jump probability shows clear time dependencies.

From the model estimates it is seen that the process governing jump arrival is very persistent. A 95% confidence interval for γ_1 is (.937, 971). The posterior mean of η_1 is .4532 which indicates that lagged absolute returns are important in affecting the jump-size variance. Notice that the jump-size variance is 3-4 times larger than the normal innovation variance σ^2 .

Features of the jump model are displayed in Figure 1. Panel A is returns, B is the inferred jump probability λ_t , C the jump size and D the estimated jump times over the full sample. Clearly this model identifies large moves in returns as jumps. Panels C and D suggest that some jumps are isolated events while others cluster and lead to more jumps and higher volatility. This is an important feature of the model.

Predictive likelihood estimates appear in Table 4. The estimates were obtained by re-estimating the posterior for each observation from 3001 on. For each run, we collected $N = 20000$ (after discarding the first 10000) samples from the posterior simulator and set $R = 100$. Kass and Raftery (1995) recommend considering twice the logarithm of the Bayes factor for model comparison, as it has the same scaling as the likelihood ratio statistic. Based on this, the evidence in favor of the Jump model is positive.⁷

⁷ $2 \log BF = 2(-1103.760 + 1105.4758) = 3.431$. Kass and Raftery (1995) suggest the following

Model probabilities based on the predictive likelihood estimates are calculated as in Equation 5.5. The probability in favor of the jump specification is .85 while it is .15 for the SV model.

Although Bayes factors are an assessment of all distributional features it is interesting to focus on volatility forecasts. Out-of-sample forecasts for both models are assessed using realized volatility. Table 4 contains Mincer and Zarnowitz (1969) forecast regressions of realized volatility regressed on a model's conditional variance forecast. Reported is regression R^2 and mean absolute error (MAE). Both measures favor the Jump model although the differences appear to be small on average. Panel A of Figure 2 displays realized volatility for the out-of-sample period. Panel B plots the model forecasts. Notice that the model forecasts are much less variable than realized volatility. During high volatility periods both models produce similar forecasts, but during low periods the Jump model's variance is lower. During these times the Jump variance is essentially flat, a time-series pattern similar to that of realized volatility, while the SV forecast appears to be trending upward.

Differences in the log predictive likelihood in favor of the Jump model are found in panel B of Figure 3. A positive (negative) value occurs when an observation is more likely under the Jump (SV) model. There are several influential observations favoring the Jump model that appear to be high volatility episodes. The influential observations show up as spikes in the model probability in panel C. It is interesting to note that there are several upward trending periods in Figure 3C which are not from tail occurrences in returns, for instance from the middle of 2000 to 2001. The SV model appears to perform best at the start of the sample, 1999-2000.

The differences in the volatility forecasts become clear when we compare Figure 2B to 3C. Periods when the Jump variance is lower is exactly when the probability for the Jump model is increasing (just before 2001 and around 2002). This suggests that state dependent volatility clustering may be important, that is, periods of normal homoskedasticity along with periods of high volatility that clusters. The Jump model is able to capture these dynamics relatively well.

Preliminary evidence reveals that the Jump model is quite competitive with a traditional SV parameterization. Further, out-of-sample predictive likelihoods reveal periods when the heteroskedastic jump structure appears to do a better job at capture the dynamics of realized volatility.

To be completed - DEM-USD results.

9 Conclusions

To be completed

interpretation of $2 \log BF$: 0 to 2 not worth more than a bare mention, 2 to 6 positive, 6 to 10 strong, and greater than 10 as very strong.

10 Appendix

Below we provide the details of the posterior simulator for the Jump specification. For steps 1 - 3 and 6 we use standard conjugate results for the linear regression model (see Koop (2003)).

1. $\mu|\theta_{-\mu}, \omega, J, \xi, r$. If the prior is normal, $p(\mu) \sim N(a, A^{-1})$

$$p(\mu|\theta_{-\mu}, \omega, J, \xi, r) \propto p(r|\theta, \omega, J, \xi)p(\mu) \quad (10.1)$$

$$\sim N(m, V^{-1}) \quad (10.2)$$

where $V = \sigma^{-2}T + A$ and $m = V^{-1}(\sigma^{-2}\sum_{t=1}^T(r_t - \xi_t J_t) + Aa)$.

2. $\sigma^2|\theta_{-\sigma^2}, \omega, J, \xi, r$. With an inverse gamma prior, $\sigma^2 \sim IG(\nu_{\sigma^2}/2, s_{\sigma^2}/2)$ we have⁸

$$p(\sigma^2|\theta_{-\sigma^2}, \omega, J, \xi, r) \propto p(r|\theta, \omega, J, \xi)p(\sigma^2) \quad (10.3)$$

$$\sim IG\left(\frac{T + \nu_{\sigma^2}}{2}, \frac{\sum_{t=1}^T(r_t - \xi_t J_t - \mu)^2 + s_{\sigma^2}}{2}\right) \quad (10.4)$$

3. $\mu_J|\theta_{-\mu_J}, \omega, J, \xi, r$. If $p(\mu_J) \sim N(b, B^{-1})$

$$p(\mu_J|\theta_{-\mu_J}, \omega, J, \xi, r) \propto p(\xi|\theta)p(\mu_J) \quad (10.5)$$

$$\sim N(m, V^{-1}) \quad (10.6)$$

where $V = \sigma_J^{-2}T + B$, and $m = V^{-1}(\sigma_J^{-2}\sum_{t=1}^T \xi_t + Bb)$.

4. $\eta|\theta_{-\eta}, \omega, J, \xi, r$. We use independent inverse gamma priors for both parameters, $p(\eta_0) \sim IG(\nu_{\eta_0}/2, s_{\eta_0}/2)$, $p(\eta_1) \sim IG(\nu_{\eta_1}/2, s_{\eta_1}/2)$. This ensures that the jump size variance is always positive. The target distribution is

$$p(\eta|\theta_{-\eta}, \omega, J, \xi, r) \propto p(\xi|\theta)p(\eta_0)p(\eta_1) \quad (10.7)$$

$$\propto p(\eta_0)p(\eta_1) \prod_{t=1}^T (\eta_0 + \eta_1 X_{t-1})^{-1/2} \exp\left(-\frac{1}{2} \frac{(\xi_t - \mu_J)^2}{(\eta_0 + \eta_1 X_{t-1})}\right)$$

which is a nonstandard distribution. We use a MH algorithm to sample this parameter using a random walk proposal. The proposal distribution $q(x|\eta^{i-1})$, is a fat-tailed mixture of normals where the covariance matrix is calibrated so approximately 50% of candidate draws are accepted. The mixture specification is the same as what is used in step 5 below. If $\eta^i \sim q(x|\eta^{i-1})$ is a draw from the proposal distribution, it is accepted with probability

$$\min \left\{ \frac{p(\eta^i|\theta_{-\eta}, \omega, J, \xi, r)}{p(\eta^{i-1}|\theta_{-\eta}, \omega, J, \xi, r)}, 1 \right\} \quad (10.8)$$

and otherwise $\eta^i = \eta^{i-1}$.

⁸If $x \sim IG(\alpha, \beta)$ then $p(x) \propto x^{-(\alpha+1)} \exp(-\beta/x)$.

5. $\omega_{(t,\tau)}|\theta, \omega_{-(t,\tau)}, J, \xi, r$, $t < \tau$, $t = 1, \dots, T$. The conditional posterior is

$$p(\omega_{(t,\tau)}|\theta, \omega_{-(t,\tau)}, J) \propto p(J_{(t,\tau)}|\omega_{(t,\tau)})p(\omega_{(t,\tau)}|\omega_{-(t,\tau)}, \theta) \quad (10.9)$$

$$\propto p(\omega_{\tau+1}|\omega_{\tau}, \theta) \prod_{i=t}^{\tau} p(J_i|\omega_i)p(\omega_i|\omega_{i-1}, \theta) \quad (10.10)$$

which is a nonstandard distribution. To improve the mixing properties of the MCMC output we adapted the blocking procedure that Fleming and Kirby (2003) use for SV models. Specifically, we approximate the first-difference of ω_t as a constant over the interval (t, τ) , and sample a block using an independent Metropolis routine. If ω_t is strongly autocorrelated this provides a good proposal density. The proposal distribution is a fat-tailed multivariate mixture⁹ of normals,

$$q(x|\omega_{-(t,\tau)}) \sim \begin{cases} N(m, V) & \text{with probability } p \\ N(m, 10V) & \text{with probability } 1 - p \end{cases} \quad (10.11)$$

where $p = .9$. The mean vector and variance-covariance matrix are,

$$m_l = \omega_{t-1} + \frac{l}{k+1}(\omega_{t-1} + \omega_{\tau+1}) \quad l = 1, \dots, k, \quad (10.12)$$

$$V_{lm} = \min(l, m) - \frac{lm}{k+1}, \quad l = 1, \dots, k, \quad m = 1, \dots, k \quad (10.13)$$

where $k = \tau - t + 1$ is the block length and is chosen randomly from a Poisson distribution with parameter 15. V^{-1} has a very convenient tridiagonal form in which $V_{ii}^{-1} = 2$ and $V_{ij}^{-1} = -1$, $1 \leq i \leq k$, $j = i - 1, i + 1$, and otherwise 0. A new draw $\omega_{(t,\tau)}^{(i)}$ from $q(x|\omega_{-(t,\tau)}^{(i-1)})$, is accepted with probability

$$\min \left\{ \frac{p(\omega_{(t,\tau)}^{(i)}|\theta, \omega_{-(t,\tau)}^{(i-1)}, J)/q(\omega_{(t,\tau)}^{(i)}|\omega_{-(t,\tau)}^{(i-1)})}{p(\omega_{(t,\tau)}^{(i-1)}|\theta, \omega_{-(t,\tau)}^{(i-1)}, J)/q(\omega_{(t,\tau)}^{(i-1)}|\omega_{-(t,\tau)}^{(i-1)})}, 1 \right\} \quad (10.14)$$

and otherwise $\omega_{(t,\tau)}^{(i)} = \omega_{(t,\tau)}^{(i-1)}$.

6. $\gamma|\theta_{-\gamma}, \omega, J, \xi, r$. The conjugate prior for γ is a bivariate normal $p(\gamma) \sim N(m, V)$ which results in a conditional distribution that is also normal. See Koop for details.

7. $\xi|\theta, \omega, J, r$. The jump size can be sampled in a block by using the conditional independence of each ξ_t . The conditional distribution of ξ_t is,

$$p(\xi_t|\theta, \omega, J, r) \propto p(r_t|\theta, \omega_t, J_t, \xi_t)p(\xi_t|\theta) \sim N(c, C^{-1}) \quad (10.15)$$

where $C = \sigma^{-2}J_t + \sigma_{J,t}^{-2}$, and $c = C^{-1}(\sigma^{-2}J_t(r_t - \mu) + \sigma_{J,t}^{-2}\mu_J)$

⁹In practice a fat-tailed proposal was found to be important when comparing the accuracy of the block version with a single move sampler.

8. $J|\theta, \omega, \xi, r$. J_t is a Bernoulli random variable with parameter λ_t . To find the probability of $J_t = 0, 1$ note that,

$$\begin{aligned} p(J_t = 0|\theta, \omega, \xi, r) &\propto p(r_t|\theta, J_t, \xi_t)p(J_t = 0|\omega) \propto \exp(-.5\sigma^{-2}(r_t - \mu)^2)(1 - \lambda_t) \\ p(J_t = 1|\theta, \omega, \xi, r) &\propto p(r_t|\theta, J_t, \xi_t)p(J_t = 1|\omega) \propto \exp(-.5\sigma^{-2}(r_t - \mu - \xi_t)^2)\lambda_t \end{aligned}$$

which allows calculation of the normalizing constant and hence a draw of J_t .

9. goto 1

The priors used for this model are $\mu \sim N(0, 1000)$, $\mu_J \sim N(0, 100)$, $\sigma^2 \sim IG(3, .02)$, $\eta_0 \sim IG(2.5, 1)$, $\eta_1 \sim IG(2.5, 1)$, $\gamma_0 \sim N(0, 100)$, $\gamma_1 \sim N(0, 100)I_{|\gamma_1| < 1}$. The priors selected for this model are for the most part, non-informative. Note that ω_t is restricted to be stationary. The priors on η_0 and η_1 reflect a reasonable range for the conditional variance of the jump-size distribution. For example, the probability that these parameters lie in the interval $(.1, 2)$ is approximately .96.

SV model - to be completed

References

- ALIZADEH, S., M. BRANDT, AND F. X. DIEBOLD (2002): “Range-Based Estimation of Stochastic Volatility Models,” *Journal of Finance*, 57, 1047–1092.
- ANDERSEN, T., T. BOLLERSLEV, AND F. X. DIEBOLD (2003): “Parametric and Non-parametric Volatility Measurement,” to appear in the Handbook of Financial Econometrics, L.P. Hansen and Y. Ait-Sahalia (eds.).
- ANDERSEN, T., AND B. SORENSEN (1996): “GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study,” *Journal of Business & Economic Statistics*, 14, 328–352.
- ANDERSEN, T. G., L. BENZONI, AND J. LUND (2002): “An Empirical Investigation of Continuous-Time Equity Return Models,” *Journal of Finance*, 62, 1239–1284.
- ANDERSEN, T. G., AND T. BOLLERSLEV (1998): “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts,” *International Economic Review*, 39(4), 885–905.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2001): “The Distribution of Exchange Rate Volatility,” *Journal of the American Statistical Association*, 96, 42–55.
- BARNDORFF-NIELSEN, O. E., S. E. GRAVERSEN, AND N. SHEPHARD (2003): “Power Variation and Stochastic Volatility: a Review and some New Results,” Nuffield College, University of Oxford.
- BATES, D. S. (2000): “Post-’87 Crash Fears in the S&P 500 Futures Option Market,” *Journal of Econometrics*, 94, 181–238.
- CHERNOV, M., R. A. GALLANT, E. GHYSELS, AND G. TAUCHEN (2003): “Alternative models for stock price dynamics,” *Journal of Econometrics*, 116, 225–257.
- CHIB, S. (2001): “Markov Chain Monte Carlo Methods: Computation and Inference,” in *Handbook of Econometrics*, ed. by Heckman, and Leamer. Elsevier Science.
- CHIB, S., F. NARDARI, AND N. SHEPHARD (2002): “Markov Chain Monte Carlo Methods for Stochastic Volatility Models,” *Journal of Econometrics*, 108, 281–316.
- DANIELSSON, J., AND J. F. RICHARD (1993): “Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models,” *Journal of Applied Econometrics*, 8, S153–S173.
- DAS, S. R. (2002): “The Surprise Element: Jumps in Interest Rate,” *Journal of Econometrics*, 106, 27–65.

- ENGLE, R. F. (1982): “Autoregressive Conditional Heteroskedasticity with Estimates of the UK inflation,” *Econometrica*, 50, 987–1008.
- ERAKER, B., M. S. JOHANNES, AND N. G. POLSON (2003): “The Impact of Jumps in Volatility and Returns,” *Journal of Finance*, 58(3), 1269–1300.
- FLEMING, J., AND C. KIRBY (2003): “A Closer Look at the Relationship between GARCH and Stochastic Autoregressive Volatility,” Jones Graduate School, Rice University.
- GALLANT, A. R., D. HSIEH, AND G. TAUCHEN (1997): “Estimation of Stochastic Volatility Models with Diagnostics,” *Journal of Econometrics*, 81, 159–192.
- GEWEKE, J. (1992): “Evaluating the Accuracy of Sampling Based Approaches to the Calculation of Posterior Moments,” in *Bayesian Statistics*, ed. by Bernardo, Berger, Dawid, and Smith, vol. 4. Oxford: Clarendon Press.
- (1995): “Bayesian Comparison of Econometric Models,” University of Iowa.
- (1997): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 18(1), 1–73.
- GHYSELS, E., A. C. HARVEY, AND E. RENAULT (1996): *Stochastic Volatility* chap. 5, Handbook of Statistics vol. 14. Elsevier Science.
- GORDON, S. (1997): “Stochastic Trends, Deterministic Trends, and Business Cycle Turning Points,” *Journal of Applied Econometrics*, 12, 411–434.
- GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): “Indirect Inference,” *Journal of Applied Econometrics*, 8, S85–S118.
- JACQUIER, E., N. G. POLSON, AND P. E. ROSSI (1994): “Bayesian Analysis of Stochastic Volatility Models,” *Journal of Business & Economic Statistics*, 12, 371–417.
- JOHANNES, M., R. KUMAR, AND N. G. POLSON (1999): “State Dependent Jump Models: How do U.S. Equity Indices Jump?,” Working Paper, Graduate School of Business, University of Chicago.
- JOHANNES, M., AND N. POLSON (2003): “MCMC Methods for Financial Econometrics,” to appear in the Handbook of Financial Econometrics.
- JORION, P. (1988): “On Jump Processes in the Foreign Exchange and Stock Markets,” *Review of Financial Studies*, 1(4), 427–445.
- KASS, R. E., AND A. E. RAFTERY (1995): “Bayes Factors,” *Journal of the American Statistical Association*, 90(420), 773–795.

- KIM, S., N. SHEPHARD, AND S. CHIB (1998): “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models,” *Review of Economic Studies*, 65, 361–393.
- KOOP, G. (2003): *Bayesian Econometrics*. Wiley, Chichester, England.
- MAHEU, J. M., AND T. H. MCCURDY (2002): “Nonlinear Features of FX Realized Volatility,” *Review of Economics and Statistics*, 84(4), 668–681.
- (2004): “News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns,” *Journal of Finance*, 59(2).
- MELINO, A., AND M. TURNBULL (1990): “Pricing Foreign Currency Options with Stochastic Volatility,” *Journal of Econometrics*, 45, 239–265.
- MIN, C., AND A. ZELLNER (1993): “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates,” *Journal of Econometrics*, 56, 89–118.
- MINCER, J., AND V. ZARNOWITZ (1969): “The Evaluation of Economic Forecasts and Expectations,” in *Economic Forecasts and Expectations*, ed. by Mincer. New York, National Bureau of Economic Research.
- OOMEN, R. C. A. (2002): “Statistical Models for High Frequency Security Prices,” Warwick Business School, University of Warwick.
- ROBERT, C. P., AND G. CASELLA (1999): *Monte Carlo Statistical Methods*. Springer, New York.
- SHEPHARD, N. (1996): “Statistical Aspects of ARCH and Stochastic Volatility,” in *Time Series Models in Econometrics, Finance and Other Fields*, ed. by D. V. H. D. R. Cox, and O. E. Barndorff-Nielsen. Chapman and Hall.
- TAYLOR, S. J. (1986): *Modeling Financial Time Series*. John Wiley.
- TAYLOR, S. L. (1994): “Modelling Stochastic Volatility,” *Mathematical Finance*, 4(183–204).
- TIERNEY, L. (1994): “Markov Chains for Exploring Posterior Distributions,” *The Annals of Statistics*, 21, 1701–1762.

Table 1: Summary Statistics

| Statistic | JPY-USD | | DEM-USD | |
|-----------|-----------|----------|------------|---------|
| | r_t | RV_t | r_t | RV_t |
| Mean | -4.566e-6 | 0.5992 | -0.2444e-8 | .5190 |
| Variance | 0.6352 | 0.62636 | 0.5447 | .2059 |
| Skewness | -0.3504 | 21.6655 | 0.0017 | 6.1888 |
| Kurtosis | 8.0587 | 854.2149 | 4.8476 | 89.3892 |
| Min | -6.9768 | 0.0290 | -3.7165 | .0369 |
| Max | 5.9780 | 34.3872 | 4.1043 | 10.8711 |
| Obs | 4001 | 4001 | 4026 | 4026 |

r_t is percent log differences of daily spot exchange rates, and RV_t is realized volatility, both from 1986/12/16 - 2002/12/31 JPY-USD, 1986/11/4 - 2002/12/31 DEM-USD.

Table 2: Model Estimates, JPY-USD

| Jump Model | | | | |
|------------|---------|--------|---------|-------------------|
| | mean | stdev | median | 95% CI |
| μ | 0.0307 | 0.0125 | 0.0307 | (0.0103, 0.0513) |
| σ^2 | 0.3226 | 0.0159 | 0.3223 | (0.2968, 0.3491) |
| μ_J | -0.1272 | 0.0510 | -0.1269 | (-0.2115,-0.0440) |
| η_0 | 0.9089 | 0.1238 | 0.8961 | (0.7273, 1.1276) |
| η_1 | 0.4532 | 0.1153 | 0.4443 | (0.2793, 0.6599) |
| γ_0 | -0.1169 | 0.0305 | -0.1158 | (-0.1688,-0.0686) |
| γ_1 | 0.9554 | 0.0103 | 0.9562 | (0.9372, 0.9709) |

| SV Model | | | | |
|--------------|---------|--------|---------|--------------------|
| | mean | stdev | median | 95% CI |
| μ | 0.0165 | 0.0105 | 0.0165 | (-0.0008, 0.03370) |
| ρ_0 | -0.0568 | 0.0125 | -0.0559 | (-0.0786,-0.03793) |
| ρ_1 | 0.9234 | 0.0148 | 0.9244 | (0.8974, 0.94599) |
| σ_v^2 | 0.0763 | 0.0160 | 0.0749 | (0.0524, 0.10433) |

The data is percent log differences of daily spot exchange rates from 1986/12/16 - 2002/12/31 JPY-USD, 1986/11/4 - 2002/12/31 DEM-USD. Tables report posterior mean, standard deviation, median and 95% confidence intervals for the models.

Table 3: Model Estimates, DEM-USD

Jump Model

| Table 4: Log Predictive Likelihood | | |
|------------------------------------|------------------------|---------|
| Model | JPY-USD | DEM-USD |
| Jump | -1103.7603 (0.3982) | |
| SV | -1105.4758 (0.2153) | |

This table reports estimates of log predictive likelihoods for observations 3001 - 4001 JPY-USD, and 3001 - 4025, DEM-USD. Numerical standard errors appear in parentheses.

Table 5: Out-of-Sample Forecast Performance

| | a | b | R^2 | MAE |
|---------|------------------|------------------|-------|-------|
| JPY-USD | | | | |
| SV | .0024 (.0334) | .8899 (.0509) | .2345 | .2505 |
| Jump | .2175 (.0218) | .7806 (.0423) | .2488 | .2351 |

This table reports Mincer and Zarnowitz (1969) forecast regressions of

$$RV_t = a + b\text{Var}_{t-1}(r_t) + \text{error}_t$$

where $\text{Var}(r_t)$ is a model forecast of the one period ahead conditional variance based on time $t - 1$ information, and RV_t is realized volatility for day t . Standard errors appear in parentheses. R^2 is the coefficient of determination and MAE is mean absolute error for $(RV_t - \text{Var}_{t-1}(r_t))$. The number of out-of-sample forecasts are 1001 JPY-USD, and 1027 DEM-USD.

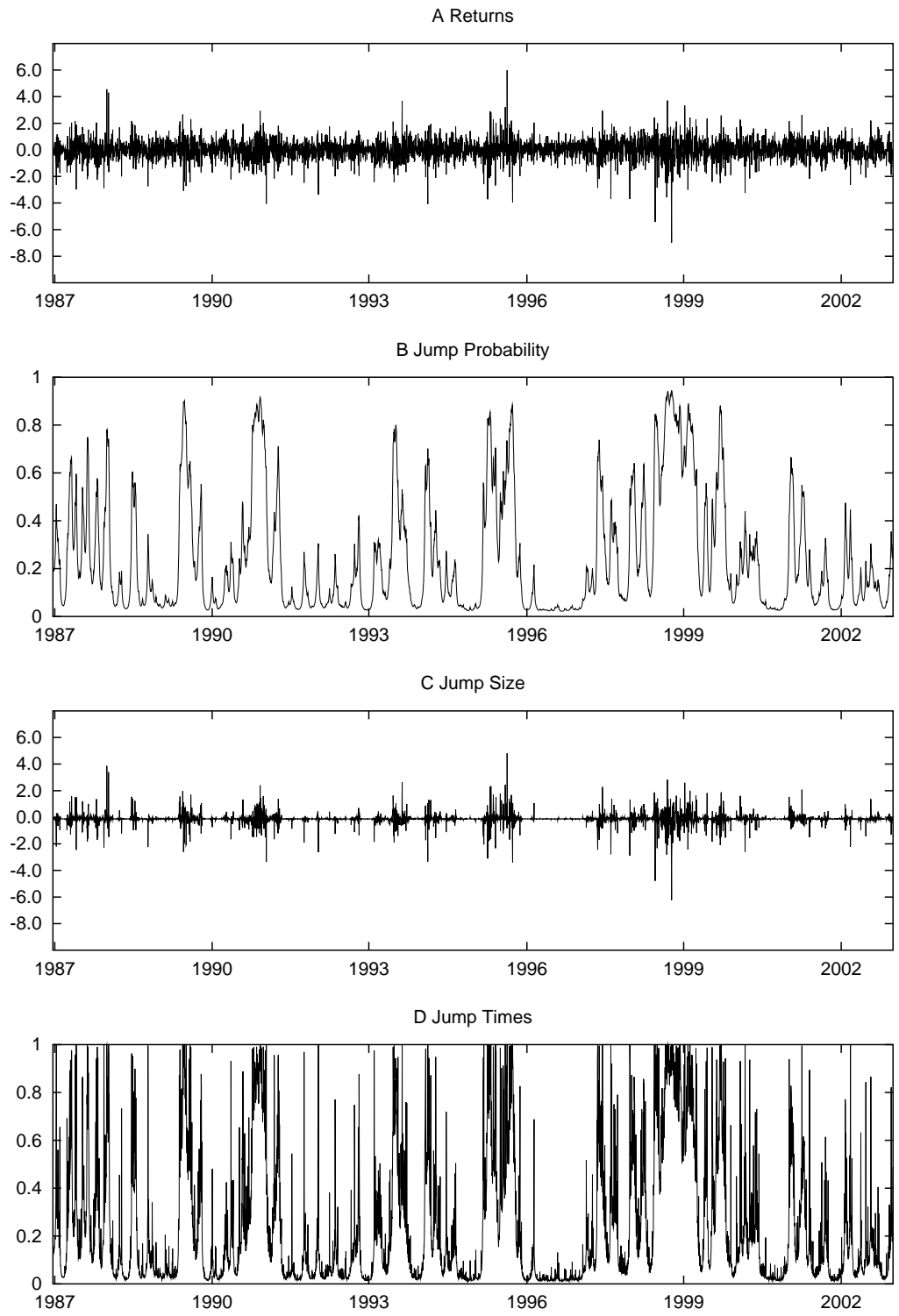


Figure 1: Jump model, JPY-USD

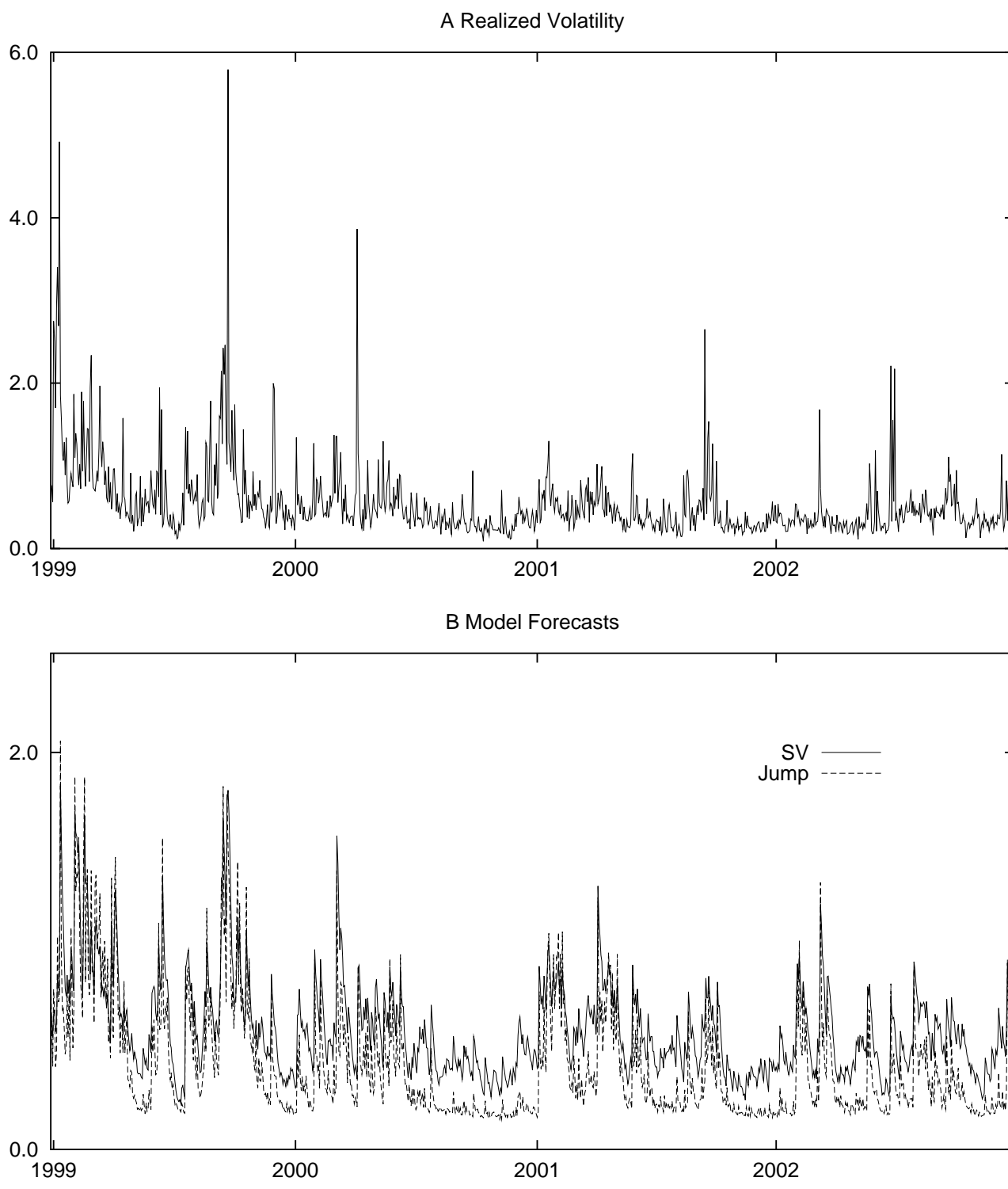


Figure 2: Volatility and Model Forecasts, JPY-USD

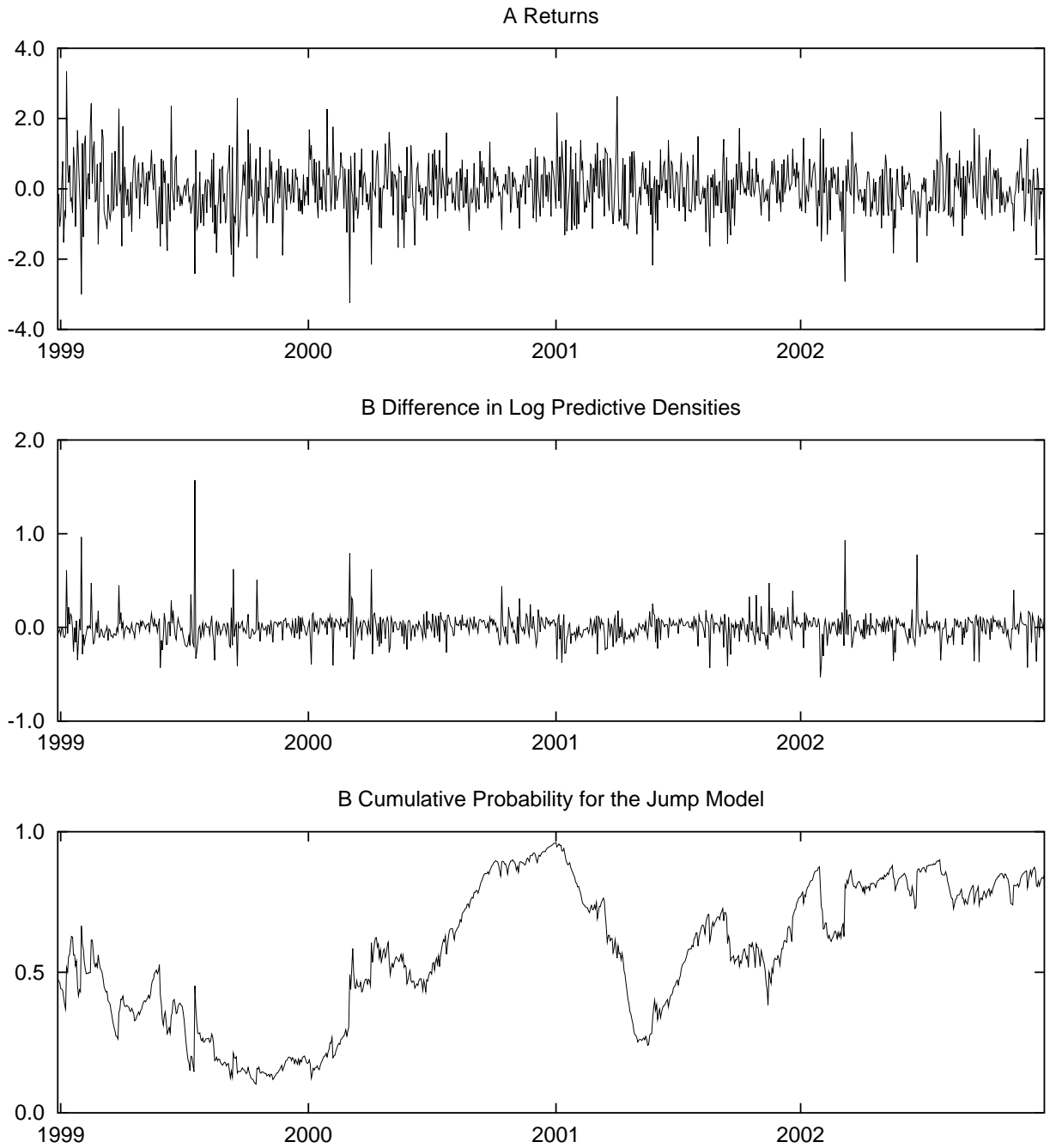


Figure 3: Model Probabilities based on Predictive Likelihood, JPY-USD