Estimating stochastic volatility diffusion using conditional moments of integrated volatility

Tim Bollerslev[a], Hao Zhou[b,*]

[a]*Department of Economics, Duke University, Durham, NC 27708-0097, USA*
[b]*Division of Research and Statistics, Federal Reserve Board, Washington, DC 20551, USA*

Abstract

We exploit the distributional information contained in high-frequency intraday data in constructing a simple conditional moment estimator for stochastic volatility diffusions. The estimator is based on the analytical solutions of the first two conditional moments for the latent integrated volatility, the realization of which is effectively approximated by the sum of the squared high-frequency increments of the process. Our simulation evidence indicates that the resulting GMM estimator is highly reliable and accurate. Our empirical implementation based on high-frequency five-minute foreign exchange returns suggests the presence of multiple latent stochastic volatility factors and possible jumps.© 2002 Elsevier Science B.V. All rights reserved.

1. Introduction

Continuous-time stochastic volatility models have figured very prominently in the recent econometrics and empirical finance literature (see e.g., Ghysels et al., 1996). Aside from a few special cases, the estimation of such models are complicated by two distinct challenges. First, since the volatility is latent, any estimation method must either adopt an imperfect proxy for the unobserved volatility, typically involving the inversion of a specific asset pricing formula, or alternatively integrate out this latent variable from the model. Second, even if the volatility process was directly observable at discrete points in time, closed form expressions for the corresponding transition

---

*Corresponding author.
E-mail address:* hao.zhou@frb.gov (H. Zhou).

density functions are generally not available for the continuous time models commonly employed in the literature. Using high-frequency data, this paper tries to meet both of these challenges by making the latent volatility effectively observable at discrete points in time through its representation in quadratic variation form, in turn relying on simple moment conditions of this measure that involve the underlying structural parameters for estimation purposes.

Our work complements numerous estimation strategies that have been proposed in the literature for dealing with the above mentioned two complications, either separately or jointly. Important contributions include the quasi-maximum likelihood (QML) estimator for the discrete-time stochastic volatility model in Ruiz (1994) and Harvey and Shephard (1994); the Markov Chain Monte Carlo (MCMC) methods advanced by Jacquier et al. (1994), Eraker (2001), and Kim et al. (1998); the simulated methods of moments approach in Duffie and Singleton (1993); the indirect inference procedure of Gourieroux et al. (1993); the efficient methods of moments (EMM) developed by Gallant and Tauchen (1996) and Gallant and Long (1997); the infinitesimal moment generator underlying the GMM procedure in Hansen and Scheinkman (1995) and Conley et al. (1997); the approximation method to the likelihood function building on the Kolmogorov forward equations in Lo (1988) and Aït-Sahalia (2001); and the spectral GMM estimator utilizing the empirical characteristic function in Chacko and Viceira (1999), Jiang and Knight (2001), and Singleton (2001). While all of these procedures yield consistent, and in many cases also asymptotically efficient, parameter estimates for the various model specifications, they are all computationally demanding and cumbersome to implement in practice. In the present paper we propose a new, much easier to compute, estimation procedure for stochastic volatility diffusions outside the realm of any specific asset pricing framework. The basic idea is straight forward. Instead of integrating out the latent volatility, the strategy proposed here utilizes high-frequency data for explicitly measuring the *realized volatility*.[1]

High-frequency, or tick-by-tick, data have recently become available for a host of different financial instruments and markets.[2] Following the work of Merton (1980) and Nelson (1992), such high-frequency data could in principle be used to construct point-wise consistent filtering measurements for the instantaneous volatility. However, the optimal filter weights depend in complicated ways on the particular model structure. More importantly, in practice the continuous record asymptotics underlying the theoretical arguments are corrupted by inherent discreteness, time-of-day effects,

---

[1] Alternatively, within the context of a specific asset pricing framework using options data and/or term structure variables, it may be possible to solve for the point-in-time implied volatilities based on the inversion of the pricing formula; see e.g., Dai and Singleton (2000), Duffie et al. (2000), and Pan (2001) for a thorough analysis of affine stochastic volatility models.

[2] An incomplete list of these includes quote-by-quote foreign exchange rates for all of the major currencies provided by Olsen & Associates; trade and quote data for almost all exchanges listed US equities as provided by the TAQ database; tick prices for the derivatives contracts on the Chicago Mercantile Exchange (CME) and the Chicago Board of Trade (CBOT); complete order books for the Paris Bourse and all Australian equities; prices and best quotes from the Deutsche Termine-Borse (DTB) and the London International Financial Futures and Options Exchange (LIFFE) (see also the discussion of high-frequency data sources in Goodhart and O'Hara, 1997).

bid–ask spreads, and other market microstructure frictions. Meanwhile, by the theory of quadratic variation, it is possible to construct increasingly more accurate *model-free* realized variability, or integrated volatility, measures over a fixed time interval, say one day, by simply summing increasingly finer sampled squared high-frequency returns over the relevant time-interval. Importantly, the empirical findings in Andersen and Bollerslev (1998) suggest that the same market microstructure frictions which plague the standard continuous record asymptotics, including the strong intraday volatility patterns, do not materially affect these interday volatility measures (see also the forecasting results in Martens, 2001). Motivated by these results, Andersen et al. (2001a, b) offer a detailed characterization of the salient distributional features of daily *realized* foreign exchange and individual stock return volatilities constructed from high-frequency data.

Here, we go one step further, and show that by matching the sample moments of the realized volatility to the population moments of the integrated volatility implied by a particular continuous-time model structure, a standard, and easy-to-compute, GMM-type estimator for the underlying model parameters is immediately applicable. For illustrative purposes we restrict the initial analysis in the paper to the single-factor square-root, or affine, class of stochastic volatility models. However, we subsequently suggest generalizations to incorporate jumps, asymmetries or leverage effects, as well as multiple volatility factors. This set of models arguably constitute the leading cases in the literature, but the method is general and could be extended to many other continuous time models, including the constant elasticity of variance, or SR-SARV, class of models analyzed by Meddahi and Renault (2000), and the non-Gaussian Lévy based Ornstein–Uhlenbeck processes in Barndorff-Nielsen and Shephard (2001b). [3]

The rest of the paper is organized as follows. The next section demonstrates how the population moments of the integrated volatility for the baseline one-factor square-root model may be derived from the moments for the point-in-time volatility. This section also briefly discusses the basic GMM setup employed in the estimation. The Monte Carlo simulations in Section 3 show that the method works very well in empirically realistic finite sample settings, the only caveat being the inclusion of an additive adjustment term to account for the measurement error in the squared realized volatility constructed from the high-frequency "five-minute" returns. Moreover, the efficiency of the parameter estimates compares favorably to that of a non-feasible QML procedure explicitly treating the "daily" instantaneous volatility as observable. Section 4 details the empirical results from applying the new estimation procedure to a data set of high-frequency five-minute foreign exchange rates spanning an entire decade. Consistent with other recent findings in the literature, the results suggest the presence of multiple latent volatility factors and/or jumps. Section 5 concludes. Mathematical details regarding the derivation of the various moment conditions are relegated to two technical appendices.

---

[3] In concurrent work, Barndorff-Nielsen and Shephard (2001a) have recently proposed a QML estimator for the Lévy driven OU processes based on Kalman filtering and a state space representation of the latent integrated volatility.

2. Estimating stochastic volatility diffusion

The basic estimation strategy builds on the conventional asymptotic theory of GMM assuming an increasing number of discretely sampled observations (Hansen, 1982). However, the construction of the sample moments explicitly relies on the availability of high-frequency data, and the corresponding almost sure convergence of the quadratic variation. We begin with a general discussion of the main idea, followed by a detailed discussion of the popular one-factor square-root, or affine, class of stochastic volatility models. Generalizations of this model to include jumps, leverage effects, or asymmetries, and multiple latent volatility factors are considered in the technical appendix.

2.1. Integrated volatility and GMM estimation

To set out the main idea, let p_t denote the time t logarithmic price for some asset. The generic continuous time stochastic volatility model may then be written as

$$\mathrm{d}p_t = \mu(p_t, V_t, t; \xi)\,\mathrm{d}t + \upsilon(p_t, V_t, t; \xi)\,\mathrm{d}B_t,$$

$$\mathrm{d}V_t = \kappa(p_t, V_t, t; \xi)\,\mathrm{d}t + \sigma(p_t, V_t, t; \xi)\,\mathrm{d}W_t, \tag{1}$$

where V_t denotes a vector of latent volatility factors, $\mathrm{d}B_t$ and $\mathrm{d}W_t$ denote compatible, possibly correlated, Brownian motions, and the drift and diffusions functions are assumed to be sufficiently regular to guarantee the existence of a unique strong solution (see, e.g., Karatzas and Shreve, 1997). The parameters, ξ, are restricted to lie within some compact set, Ξ, containing the true parameters of the process, say ξ_0. Of course, the dependence of p_t on $\mathrm{d}W_t$ through both V_t and $\mathrm{corr}(\mathrm{d}B_t, \mathrm{d}W_t)$ may be redundant. Also, for concreteness, in the subsequent empirical analysis we will normalize the unit time interval to correspond to one day.

In order to focus the analysis, we shall fix the drift in the log price process at zero, concentrating exclusively on the estimation of the time-varying volatility of the process, i.e., $\mu(p_t, V_t, t; \xi) = 0$. This is consistent with existing empirical evidence, which indicates very little predictable variation in the mean of high-frequency asset returns (e.g., Andersen and Bollerslev, 1997). Moreover, unless the full model exhibits direct feedback effects from the log price process p_t to the volatility process V_t, leverage effects or asymmetries, and/or jump components, the exact form of the drift function is generally irrelevant for the consistent estimation of the parameters entering the diffusion functions.

Meanwhile, the estimation of the volatility parameters based on discretely sample observations for the p_t process is complicated by the V_t process being latent, and the lack of a closed form expression for the corresponding transition density function. However, by the theory of quadratic variation,

$$\lim_{N \to \infty} \sum_{i=1}^{2^N} [p_{t+i/2^N(T-t)} - p_{t+(i-1)/2^N(T-t)}]^2 \xrightarrow{\text{a.s.}} \int_t^T \upsilon^2(p_s, V_s, s; \xi)\,\mathrm{d}s \equiv \mathscr{V}_{t,T}, \tag{2}$$

where $\mathscr{V}_{t,T}$ denotes the integrated volatility from time t to T. Thus, while the point-in-time volatility, $\upsilon(p_t, V_t, t; \xi)$, is generally unobservable, by summing increasingly finer

sampled squared high-frequency returns, it is possible to obtain increasingly more accurate estimates of the integrated volatility of the process. Importantly, in the limit the integrated volatility becomes effectively observable.

Explicitly treating the integrated volatility as observable, in turn permits the implementation of a standard GMM-type estimator for the underlying model parameters, by minimizing the weighted distance between the sample moments and the corresponding population moments of $\mathcal{V}_{t,T}$ implied by the particular model structure. Of course, as discussed further below in practice continuously sampled observations are not available, so that the integrated volatility is not truly observable. Still, the same GMM estimation strategy may be formally justified under the additional assumption that the number of observations employed in the construction of the sample moments converges to infinity at a slower rate than the almost sure convergence rate of $1/2^N$ for the quadratic variation. The operative of this assumption is obviously a practical empirical question.

For illustrative purposes, we restrict the discussion in the following section to the simple one-factor square-root model. However, the same basic approach could in principle be applied in the estimation of any stochastic volatility model for which the conditional mean and conditional variance of the point-in-time volatility have tractable analytical expressions. [4] This latter class is quite general, including the affine jump-diffusion class of models popularized by Duffie and Kan (1996), Duffie et al. (2000), and Dai and Singleton (2000), the constant elasticity of variance class of models studied by Meddahi and Renault (2000), as well as the quadratic stochastic volatility class of models (see, e.g., Kloeden and Platen, 1992).

2.2. Solutions for the baseline stochastic volatility model

The square-root volatility model, or scalar affine diffusion, is succinctly defined by

$$\mathrm{d}p_t = \sqrt{V_t}\,\mathrm{d}B_t,$$
$$\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t, \qquad (3)$$

where V_t is a scalar latent volatility process. [5] This model was popularized by Heston (1993), and it has been widely used in the empirical finance literature. While the first-order parameterization is obviously somewhat restrictive, it is nonetheless rich enough to illustrate the general idea. In this parameterization, θ determines the long-run (unconditional) mean of the volatility, κ corresponds to the mean reversion parameter, while σ denotes the local variance (volatility-of-volatility) parameter. For the process to be well defined, the parameters must satisfy standard restrictions (see e.g., Feller (1951) for details): $\theta > 0$ (non-negativity), $\kappa > 0$ (stationary in mean), and $\sigma^2 \leqslant 2\kappa\theta$ (stationary in volatility).

---

[4] Although the procedure implemented here hinges on analytical expressions for the population moments of $\mathcal{V}_{t,T}$, these could alternatively be evaluated by simulations, and the underlying parameters estimated by simulated methods of moments (Duffie and Singleton, 1993).

[5] The drift of the log price process is fixed at zero for simplicity. However, a non-trivial drift function, say $\mu_t\,\mathrm{d}t$, could be included in the first equation without impeding the estimation of the three volatility parameters.

In deriving the conditional moments for the integrated volatility, it is useful to distinguish between two information sets—the continuous sigma-algebra $\mathscr{F}_t = \sigma\{V_s; s \leqslant t\}$, generated by the point-in-time volatility process, and the discrete sigma-algebra $\mathscr{G}_t = \sigma\{\mathscr{V}_{t-s-1,t-s}; s = 0, 1, 2, \ldots, \infty\}$, generated by the integrated volatility series. Obviously, the coarser filtration is nested in the finer filtration (i.e., $\mathscr{G}_t \subset \mathscr{F}_t$), and by the Law of Iterated Expectations, $\mathrm{E}[\mathrm{E}(\cdot|\mathscr{F}_t)|\mathscr{G}_t] = \mathrm{E}(\cdot|\mathscr{G}_t)$.

We begin with a detailed discussion of the logic behind the derivation of the operational moment conditions for the conditional mean of the integrated volatility. The derivation of the second moment proceeds in a similar fashion, but it is more involved and the technical details are referred to Appendix A.

2.2.1. Conditional mean

In deriving the conditional mean for the integrated volatility, it is useful to start with the conditional mean of the point-in-time volatility. In particular, it follows from the results in Cox et al. (1985) that

$$\mathrm{E}(V_T|\mathscr{F}_t) = \alpha_{T-t} V_t + \beta_{T-t}, \tag{4}$$

where $\alpha_{T-t} = \mathrm{e}^{-\kappa(T-t)}$ and $\beta_{T-t} = \theta(1 - \mathrm{e}^{-\kappa(T-t)})$ are functions of the structural parameters and the horizon of the forecast, $T - t$. Now, expressing the conditional mean of the integrated volatility as a (linear) function of the point-in-time volatility and interchanging the integration operators, it follows that

$$\mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t) = \mathrm{E}\left(\int_t^T V_s \, \mathrm{d}s \middle| \mathscr{F}_t\right) = a_{T-t} V_t + b_{T-t}, \tag{5}$$

where $a_{T-t} = 1/\kappa(1 - \mathrm{e}^{-\kappa(T-t)})$ and $b_{T-t} = \theta(T - t) - (\theta/\kappa)(1 - \mathrm{e}^{-\kappa(T-t)})$ denote explicit functions of the drift parameters and the sampling interval (see Appendix A for details).

Focusing on the one-day horizon (we omit the subscript, i.e., $a \equiv a_1$, $b \equiv b_1$, $\alpha \equiv \alpha_1$ and $\beta \equiv \beta_1$), repeated use of Eqs. (4) and (5) then yields

$$\mathrm{E}[\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_{t+1})|\mathscr{F}_t] = a\,\mathrm{E}(V_{t+1}|\mathscr{F}_t) + b$$

$$= a(\alpha V_t + \beta) + b$$

$$= \alpha[\mathrm{E}(\mathscr{V}_{t,\,t+1}|\mathscr{F}_t) - b] + a\beta + b,$$

which simplifies as

$$\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_t) = \alpha\,\mathrm{E}(\mathscr{V}_{t,t+1}|\mathscr{F}_t) + \beta.$$

This latter equation only involves the moments of the integrated volatility, and by the Law of Iterated Expectation, it may be conditioned on the coarser information set $\mathscr{G}_t$, yielding

$$\mathrm{E}[\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_t)|\mathscr{G}_t] = \mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{G}_t) = \alpha\,\mathrm{E}(\mathscr{V}_{t,t+1}|\mathscr{G}_t) + \beta. \tag{6}$$

This last equation establishes the required link between the first moment of the integrated volatility and lagged integrated volatility.

2.2.2. Conditional second moment

Analogous to the derivation of the conditional first moment above, it is convenient to start from the expression for the conditional variance for the point-in-time volatility. Again, following Cox et al. (1985), we have

$$E(V_T^2|\mathcal{F}_t) = Var(V_T|\mathcal{F}_t) + [E(V_T|\mathcal{F}_t)]^2$$

$$= C_{T-t}V_t + D_{T-t} + [\alpha_{T-t}V_t + \beta_{T-t}]^2, \tag{7}$$

where C_{T-t} and D_{T-t} are functionally dependent on the structural parameters and the sampling interval. Now by expressing the conditional variance of the integrated volatility as a linear function of the point-in-time volatility and by exploiting Itô's Lemma, it is possible to show that

$$Var(\mathcal{V}_{t,T}|\mathcal{F}_t) = A_{T-t}V_t + B_{T-t}, \tag{8}$$

where A_{T-t} and B_{T-t} represent other functionals of the parameters (see Appendix A for details).

Combining the conditional variance formula in (8) and the conditional mean formula in (5), it follows that for the one-day horizon,

$$E(\mathcal{V}_{t,t+1}^2|\mathcal{F}_t) = Var(\mathcal{V}_{t,t+1}|\mathcal{F}_t) + [E(\mathcal{V}_{t,t+1}|\mathcal{F}_t)]^2$$

$$= a^2 V_t^2 + (2ab + A)V_t + (b^2 + B), \tag{9}$$

where we have omitted the "daily" subscript "1" on a, b, A and B for notational convenience. Finally by repeatedly applying the Law of Iterated Expectation on different information sets and substituting expressions between integrated volatility and point-in-time volatility,

$$E[E(\mathcal{V}_{t+1,t+2}^2|\mathcal{F}_t)|\mathcal{G}_t] = E(\mathcal{V}_{t+1,t+2}^2|\mathcal{G}_t)$$

$$= H\,E(\mathcal{V}_{t,t+1}^2|\mathcal{G}_t) + I\,E(\mathcal{V}_{t,t+1}|\mathcal{G}_t) + J, \tag{10}$$

where the functions H, I, and J are again defined in the Appendix. Corresponding moment conditions for the squared multi-period integrated volatility may be derived by similar arguments. [6]

2.3. Conditional moment restrictions

The analytical solutions for the conditional first and second moments in Eqs. (6) and (10), set the stage for the construction of a standard GMM type estimator. Of course, the efficiency of the resulting estimator defined from these conditions will depend upon the particular choice of instruments (see Hansen (1985); Hansen et al. (1988) and Gallant and Tauchen (1996) for additional discussion and formal results along these lines). In the implementation pursued here, we simply augment the two

---

[6] Following the initial draft of the present paper, Lewis (2001) has recently derived a combinatorial algorithm for numerically calculating all of the higher order moments of the integrated volatility for the one-factor square-root model.

basic moments derived above with their own lag-one and lag-one squared counterparts, resulting in the following six moments:

$$f_t(\xi) \equiv \begin{bmatrix} \mathrm{E}[\mathscr{V}_{t+1,t+2}|\mathscr{G}_t] - \mathscr{V}_{t+1,t+2} \\ \mathrm{E}[\mathscr{V}^2_{t+1,t+2}|\mathscr{G}_t] - \mathscr{V}^2_{t+1,t+2} \\ \mathrm{E}[\mathscr{V}_{t+1,t+2}\mathscr{V}_{t-1,t}|\mathscr{G}_t] - \mathscr{V}_{t+1,t+2}\mathscr{V}_{t-1,t} \\ \mathrm{E}[\mathscr{V}^2_{t+1,t+2}\mathscr{V}_{t-1,t}|\mathscr{G}_t] - \mathscr{V}^2_{t+1,t+2}\mathscr{V}_{t-1,t} \\ \mathrm{E}[\mathscr{V}_{t+1,t+2}\mathscr{V}^2_{t-1,t}|\mathscr{G}_t] - \mathscr{V}_{t+1,t+2}\mathscr{V}^2_{t-1,t} \\ \mathrm{E}[\mathscr{V}^2_{t+1,t+2}\mathscr{V}^2_{t-1,t}|\mathscr{G}_t] - \mathscr{V}^2_{t+1,t+2}\mathscr{V}^2_{t-1,t} \end{bmatrix}. \tag{11}$$

By construction $\mathrm{E}[f_t(\xi_0)|\mathscr{G}_t] = 0$, and the corresponding GMM, or minimum chi-square, estimator is defined by $\hat{\xi}_T = \arg\min g_T(\xi)' W g_T(\xi)$, where $g_T(\xi)$ refers to the sample mean of the moment conditions, $g_T(\xi) \equiv 1/T \sum_{t=1}^{T-2} f_t(\xi)$, and W denotes the asymptotic covariance matrix of $g_T(\xi_0)$ (Hansen, 1982). Under standard regularity conditions, the minimized value of the objective function multiplied by the sample size is asymptotically chi-square distributed, which allows for an omnibus test of the overidentifying restrictions. Moreover, inference concerning the individual parameters is readily available from the standard formula for the asymptotic covariance matrix, $(\partial f_t(\xi)/\partial \xi' W \partial f_t(\xi)/\partial \xi)/T$.

The one-period lag in the moment conditions in Eqs. (6) and (10) implies an MA(1) error structure. However, in order to avoid any finite sample problems with the sample analogue of W not being positive definite, in the simulations and the actual empirical estimates reported on below, we used a heteroscedasticity and autocorrelation consistent robust covariance matrix estimator with a Bartlett-kernel and a lag length of five (Newey and West, 1987).[7]

Note that in contrast to the GMM-based estimators in Chacko and Viceira (1999), Jiang and Knight (2001), and Singleton (2001), which rely on discretely sampled, say daily, data, the moments employed here explicitly depend on high-frequency, intraday, observations. As documented in the next section, this wider information set, coupled with the ad hoc moment conditions in Eq. (11), easily translates into more efficient finite-sample parameter estimates when compared to (non-feasible) QML based on discretely sampled point-in-time volatility observations.

3. Monte Carlo study

The presence of strong temporal dependence and/or conditional heteroscedasticity have been shown to result in very slow convergence rates and associated poor finite sample performance of asymptotically justified econometric estimation procedures for continuous-time processes (see, e.g., Pritsker, 1998). This section qualifies the small sample behavior of the parameter estimates based on the GMM procedure outlined

---

[7] We also experimented with other lag lengths. All of the results were very similar to the ones reported here, and are available upon request.

above, along with the corresponding Wald-based inference and omnibus test for general model misspecification.

3.1. Experimental design

To keep the simulation results manageable, we concentrate on the simple model in Eq. (3) and the three parameter configurations: Scenario A ($\kappa = 0.03$, $\theta = 0.25$, $\sigma = 0.10$) features a highly persistent volatility process (near unit-root); Scenario B ($\kappa = 0.10$, $\theta = 0.25$, $\sigma = 0.10$) has a less persistent variance process; Scenario C ($\kappa = 0.10$, $\theta = 0.25$, $\sigma = 0.20$) has a higher variance-of-variance and is close to the non-stationary region. In simulating the data, we utilized a first order Euler scheme with 82 artificial "five-minute" intervals per day, further partitioning each five-minute interval into 10 smaller segments for the "continuous-record".[8] The quadratic variation formula (2) is employed to approximate the integrated volatility series. To check the standard "long-span" asymptotics, we consider the two sample sizes $T = 1000$ and $T = 4000$, corresponding to roughly 4 and 16 years of daily observations, respectively.

Since the true "continuous-time" record is known inside the simulations, we are able to compare the GMM estimator using the "five-minute" quadratic variation with the corresponding non-feasible estimator using the true integrated volatility. We also compare the results to a QML estimator based on the "daily" point-in-time volatility assuming the process to be Gaussian.[9] Although this estimator is not feasible in practice either, it provides a useful benchmark for assessing the performance of the GMM procedure and the additional informational content provided by the high-frequency data. In particular, it follows from the Monte Carlo evidence in Zhou (2001) that in empirically realistic situations involving highly persistent and/or variable volatility processes, approximate MLE for the baseline square-root model tend to perform relatively poorly compared to this much simpler QML procedure.[10] The results for all of the different models and estimators are summarized in Table 1 and Figs. 1 and 2.

3.2. Parameter estimates and efficiency

The finite sample results indicate that the feasible GMM estimator fairs well (if not better) than the other two non-feasible alternatives—using "unobserved" point-in-time volatility or the "continuous time" record integrated volatility. The GMM estimates for the mean-reversion parameter κ are only slightly upward biased, while the long-run mean parameter estimates for θ exhibit a small downward bias. At the same time, the root-mean-squared-errors (RMSEs) for both of the two drift parameters, κ and θ, decrease roughly at the rate of $\sqrt{4}$ as the sample size goes from 1000 to 4000 "days".

---

[8] This particular choice was dictated by the fact that most US financial markets are open between six-and-a-half to seven hours per day, corresponding to 78–84 five-minute intervals per day.

[9] This estimator is closely related to the procedure in Fisher and Gilles (1996), who propose a QML estimator for Affine diffusion process, using closed form solutions for the conditional mean and variance.

[10] The exact likelihood function has to be numerically approximated, and the quality of this approximation deteriorates quickly when the parameter is close to the boundary region, i.e., when the process is either very persistent or highly volatile.

Table 1
Monte Carlo experiment

True value	Mean		Median		RMSE	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Panel A						
GMM with quadratic variation from high-frequency return						
$\kappa = 0.03$	0.0352	0.0313	0.0340	0.0310	0.0130	0.0054
$\theta = 0.25$	0.2430	0.2487	0.2355	0.2460	0.0523	0.0258
$\sigma = 0.10$	0.1016	0.1030	0.1018	0.1030	0.0080	0.0050
GMM with integrated volatility						
$\kappa = 0.03$	0.0382	0.0323	0.0374	0.0319	0.0139	0.0055
$\theta = 0.25$	0.2338	0.2456	0.2273	0.2437	0.0521	0.0257
$\sigma = 0.10$	0.0992	0.0999	0.0992	0.0998	0.0044	0.0020
QML with point-in-time volatility						
$\kappa = 0.03$	0.0446	0.0360	0.0434	0.0361	0.0195	0.0095
$\theta = 0.25$	0.2327	0.2441	0.2271	0.2410	0.0537	0.0290
$\sigma = 0.10$	0.1012	0.1014	0.0999	0.1011	0.0095	0.0052
Panel B						
GMM with quadratic variation from high-frequency return						
$\kappa = 0.10$	0.1057	0.1023	0.1048	0.1016	0.0214	0.0100
$\theta = 0.25$	0.2478	0.2491	0.2474	0.2489	0.0158	0.0078
$\sigma = 0.10$	0.1059	0.1073	0.1061	0.1072	0.0093	0.0082
GMM with integrated volatility						
$\kappa = 0.10$	0.1102	0.1032	0.1090	0.1027	0.0214	0.0091
$\theta = 0.25$	0.2460	0.2486	0.2459	0.2483	0.0163	0.0078
$\sigma = 0.10$	0.0994	0.1000	0.0995	0.0998	0.0042	0.0020
QML with point-in-time volatility						
$\kappa = 0.10$	0.1136	0.1040	0.1134	0.1048	0.0259	0.0138
$\theta = 0.25$	0.2497	0.2517	0.2480	0.2510	0.0196	0.0097
$\sigma = 0.10$	0.0967	0.0956	0.0967	0.0958	0.0059	0.0054
Panel C						
GMM with quadratic variation from high-frequency return						
$\kappa = 0.10$	0.1113	0.1035	0.1091	0.1035	0.0253	0.0111
$\theta = 0.25$	0.2389	0.2468	0.2364	0.2463	0.0326	0.0158
$\sigma = 0.20$	0.2031	0.2051	0.2030	0.2049	0.0122	0.0078
GMM with integrated volatility						
$\kappa = 0.10$	0.1153	0.1048	0.1131	0.1047	0.0270	0.0114
$\theta = 0.25$	0.2346	0.2455	0.2319	0.2449	0.0341	0.0160
$\sigma = 0.20$	0.1984	0.1997	0.1982	0.1995	0.0097	0.0046
QML with point-in-time volatility						
$\kappa = 0.10$	0.1257	0.1093	0.1242	0.1107	0.0390	0.0208
$\theta = 0.25$	0.2459	0.2537	0.2432	0.2520	0.0336	0.0199
$\sigma = 0.20$	0.1977	0.1960	0.1966	0.1958	0.0135	0.0084

The table reports the simulation results for the GMM and QML procedures discussed in the main text applied in estimating the stochastic volatility diffusion in Eq. (3). The total number of Monte Carlo replications is 1000.
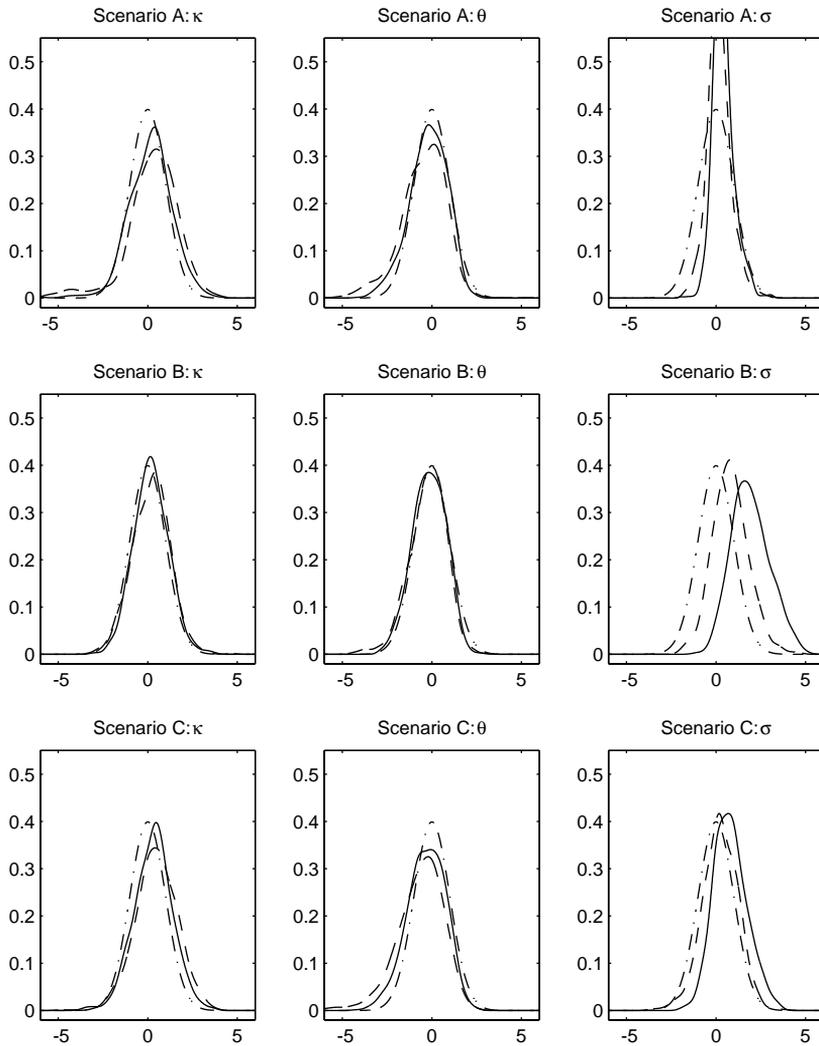
Fig. 1. Distributions of *t*-tests. "- - -" *t*-statistics with 1000 observations; "——" *t*-statistics with 4000 observations; "-.-.-" Normal (0,1) reference density.

Meanwhile, the accuracy of the local variance parameter estimates is affected by both the long-span asymptotics and the fill-in asymptotics. Although the RMSE of σ does decrease with the sample size, the rate is not always close to $\sqrt{4}$. Also, while the drift parameter estimates are almost unaffected by the fill-in asymptotics, the RMSE of σ clearly diminishes when the sampling frequency increases from "five-minute" to the "continuous-time" limit.
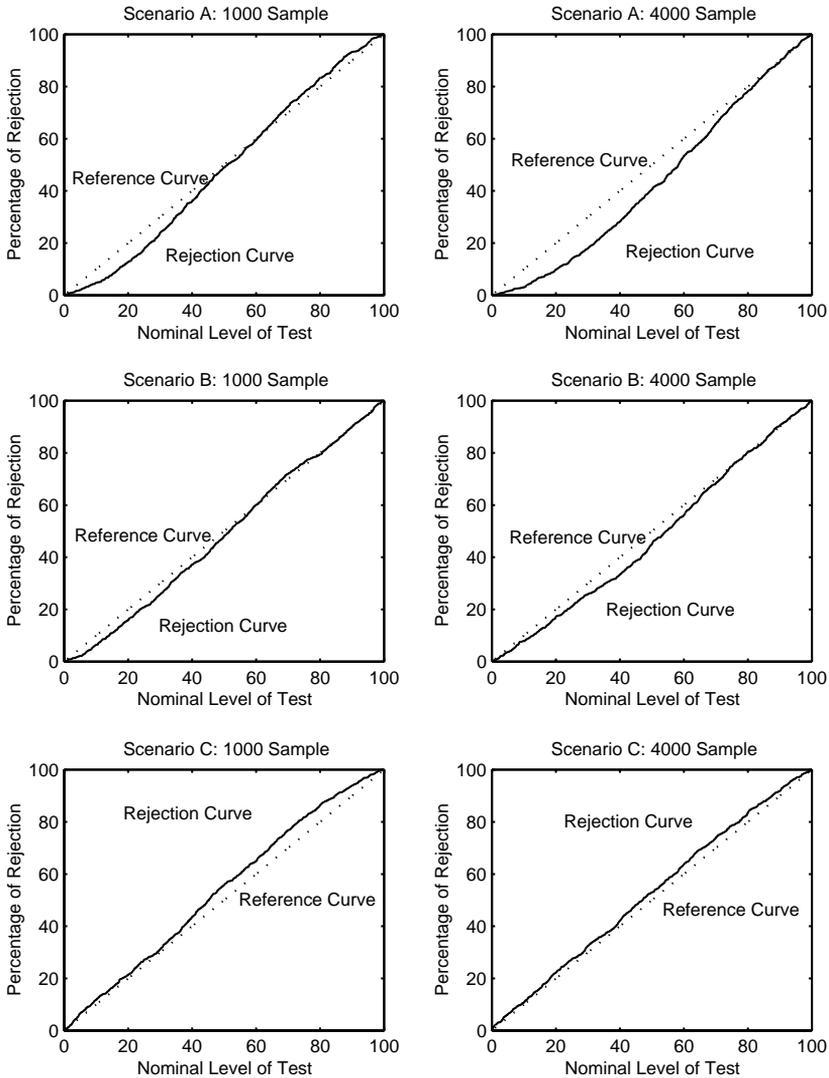
Fig. 2. GMM specification test of overidentifying restrictions.

It is also noteworthy that for the approximate unit root process (Scenario A), the variance parameter seems to converge at a faster rate than $\sqrt{T}$. Also when the variance-of-variance parameter is large (Scenario C), the finite sample biases of the drift parameter estimates are larger than for the less persistent case (Scenario B). Essentially, the estimator has difficulties in distinguishing between a very persistent yet stationary process and a non-stationary, near unit-root process in finite, or "small", samples.

The GMM estimates of the local variance parameter, σ, are systematically upward biased, albeit not by much. Interestingly, this bias completely vanishes when the true

integrated volatility is used in place of the "five-minute" quadratic variation. While the measurement error from using the quadratic variation to approximate the integrated volatility process is averaged out in the first moment condition, the second moment condition depends non-linearly on the measurement error. As discussed further in Section 3.4 below, this readily explains the small bias.

In terms of relative efficiency, the GMM estimates for the two drift parameters using the "five-minute" realized volatility uniformly perform better than the non-feasible QML estimates using the unobservable "daily" point-in-time volatility. The RMSEs for the variance parameter are also better in all but the stationary case (Scenario B). This is particularly noteworthy insofar as the results in Zhou (2001) suggest that the QML estimates analyzed here dominate the approximate (and computationally much more demanding) MLE based on "daily" data.

3.3. Statistical inference

In practice, inference concerning the individual model parameters and the overall specification of the model must be based on the standard GMM type test statistics discussed in Section 2.2.3. In this regard, the t-statistics for the drift parameters in Fig. 1 clearly indicate that the feasible GMM estimator based on the "five-minute" quadratic variation is close to normal for both of the 1000 and 4000 "daily" sample sizes analyzed here. Meanwhile for the diffusion parameter, the use of "five-minute" realized volatility in the GMM estimation gives rise to a systematic upward bias in the t-statistics. As analyzed further below, this is consistent with the earlier explanation of the non-dissipating measurement error in the second order moment condition. [11]

Turning to Fig. 2 and the results for the standard GMM J-tests of overidentifying restrictions, it follows that except for the near unit-root case (Scenario A), the test performs very well. Moreover, the slight over-rejection and under-rejection biases largely vanishes as the sample size increases from 1000 to 4000.

3.4. Measurement error adjustment

By construction the quadratic variation based on the simulated "five-minute" returns provides an unbiased estimate of the true integrated volatility. At the same time, the squared quadratic variation for any fixed sampling interval yields an upward biased estimate of the true squared integrated volatility. [12] Consequently, while the linear expectations operator washes out the measurement errors in the first conditional moment and the corresponding augmented moment conditions, the moment conditions

---

[11] The corresponding t-tests for the non-feasible QML estimator based on the point-in-time volatility are even more distorted, while the t-tests for the GMM estimates using the true integrated volatility are all extremely close to normal. These results are available upon request.

[12] Andersen and Bollerslev (1998) provide some limited simulation evidence related to the size of this measurement error, while asymptotically based analytical expressions for certain models have recently been derived by Barndorff-Nielsen and Shephard (2001a) and Meddahi (2001). Also, Andersen et al. (2000) and Bai et al. (1999) discuss practical considerations related to the inherent market microstructure frictions and the choice of the sampling frequency with actual high-frequency data.

Table 2
Monte Carlo experiment with measurement error correction

True value	Mean		Median		RMSE	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Scenario A: *GMM with quadratic variation*						
$\kappa = 0.03$	0.0364	0.0317	0.0354	0.0315	0.0138	0.0056
$\theta = 0.25$	0.2456	0.2491	0.2384	0.2464	0.0520	0.0257
$\sigma = 0.10$	0.0909	0.0994	0.0905	0.0983	0.0230	0.0127
γ	0.0007	0.0004	0.0006	0.0004	0.0008	0.0005
Scenario B: *GMM with quadratic variation*						
$\kappa = 0.10$	0.1067	0.1027	0.1061	0.1023	0.0219	0.0104
$\theta = 0.25$	0.2489	0.2494	0.2484	0.2492	0.0157	0.0078
$\sigma = 0.10$	0.0990	0.1049	0.0986	0.1046	0.0214	0.0121
γ	0.0007	0.0004	0.0006	0.0003	0.0009	0.0005
Scenario C: *GMM with quadratic variation*						
$\kappa = 0.10$	0.1133	0.1042	0.1109	0.1043	0.0274	0.0119
$\theta = 0.25$	0.2435	0.2481	0.2400	0.2473	0.0314	0.0157
$\sigma = 0.20$	0.1893	0.1999	0.1884	0.1987	0.0303	0.0162
γ	0.0017	0.0010	0.0015	0.0009	0.0019	0.0013

Note. The table reports the GMM estimation results obtained by including an additive measurement error correction term, γ, in the moment conditions involving the squared integrated volatility. The RMSE column for γ gives the sample standard deviation across the 1000 Monte Carlo replications.

involving the squared quadratic variation inevitably entail a non-zero measurement error. However, it follows quite generally by the almost sure convergence of the quadratic variation, that the expectation of the squared error term is bounded by the local maximum of the continuous local martingale process (see e.g., Karatzas and Shreve, 1997; Protter, 1992). [13]

Thus, in order to conveniently account for this error term, we simply included an additive nuisance parameter, γ, in each of the three second order moment conditions, replacing the squared "five-minute" quadratic variation, $\mathscr{V}^2_{t+1,t+2}$, by $\mathscr{V}^2_{t+1,t+2} + \gamma$. [14] Not surprisingly, from the results reported in Table 2 and Fig. 3, the estimates for the two drift parameters and the corresponding t-statistics are largely unaffected by the addition of this nuisance parameter. However, while the RMSEs for σ are slightly

---

[13] In order to get a better idea about the population properties of the measurement error in the present context, we simulated a very long time series (more than four million "days") for each of the three scenarios. The correlation between the "five-minute" quadratic variation and the true latent integrated volatility for each of three parameterizations were 97.1%, 93.2%, and 97.3%, respectively, while the average absolute measurement errors equaled 12.5%, 12.5%, and 12.8%, respectively. Also, the R-squares in the regressions of the absolute measurement errors on V_t were 0.40, 0.20, and 0.40, respectively, while the squared measurement errors were more closely related to V_t^2 with "population" R-squares of 0.26, 0.16, and 0.25, respectively.

[14] Of course, this simple adjustment term ignores the state dependence in the measurement errors documented in the previous footnote and formally analyzed in Barndorff-Nielsen and Shephard (2001a) and Meddahi (2001). More complicated instrumental variables (IV) procedures could be employed to take account of this dependence.
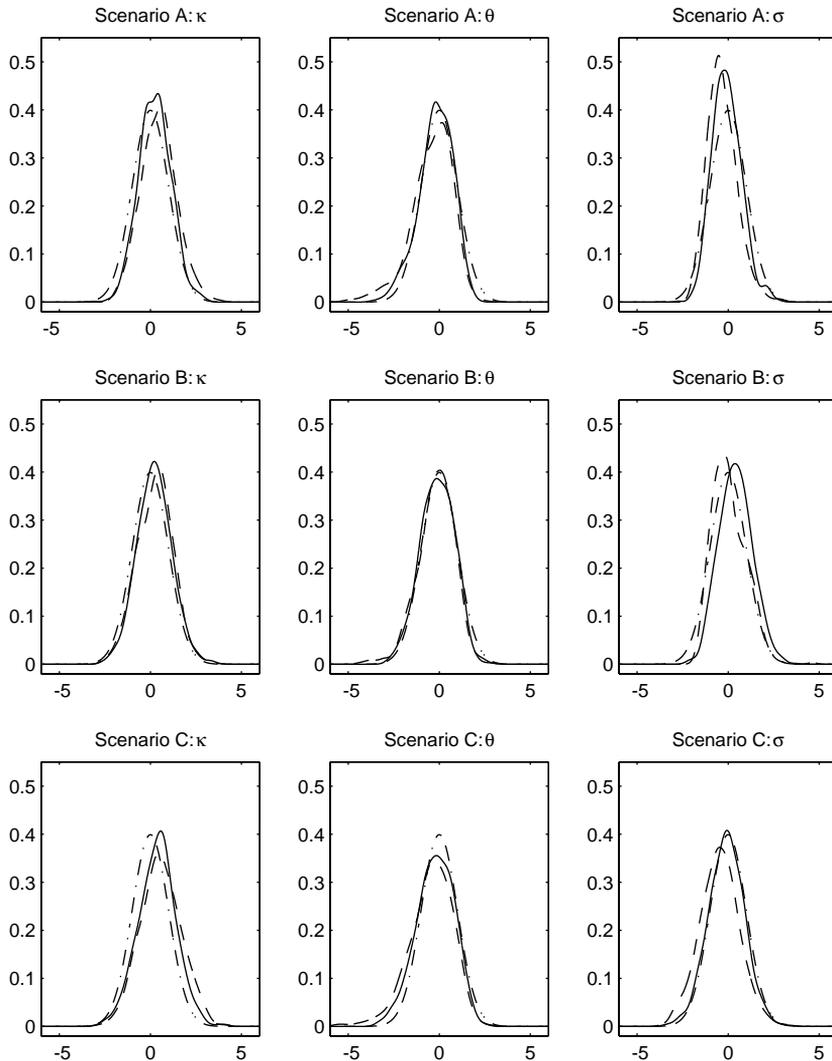
Fig. 3. Distributions of *t*-tests with measurement error correction. "- - -" *t*-statistics with 1000 observations; "——" *t*-statistics with 4000 observations; "-.-.-" Normal (0,1) reference density.

larger compared to the values in Table 1, the systematic finite sample biases for the local variance parameter estimates completely vanishes. The rejection frequencies for the GMM specification tests for the overidentifying restrictions in Fig. 4 also appear marginally closer to their nominal sizes. The empirical estimates in the next section further underscore the applicability of this simple measurement error adjustment procedure.
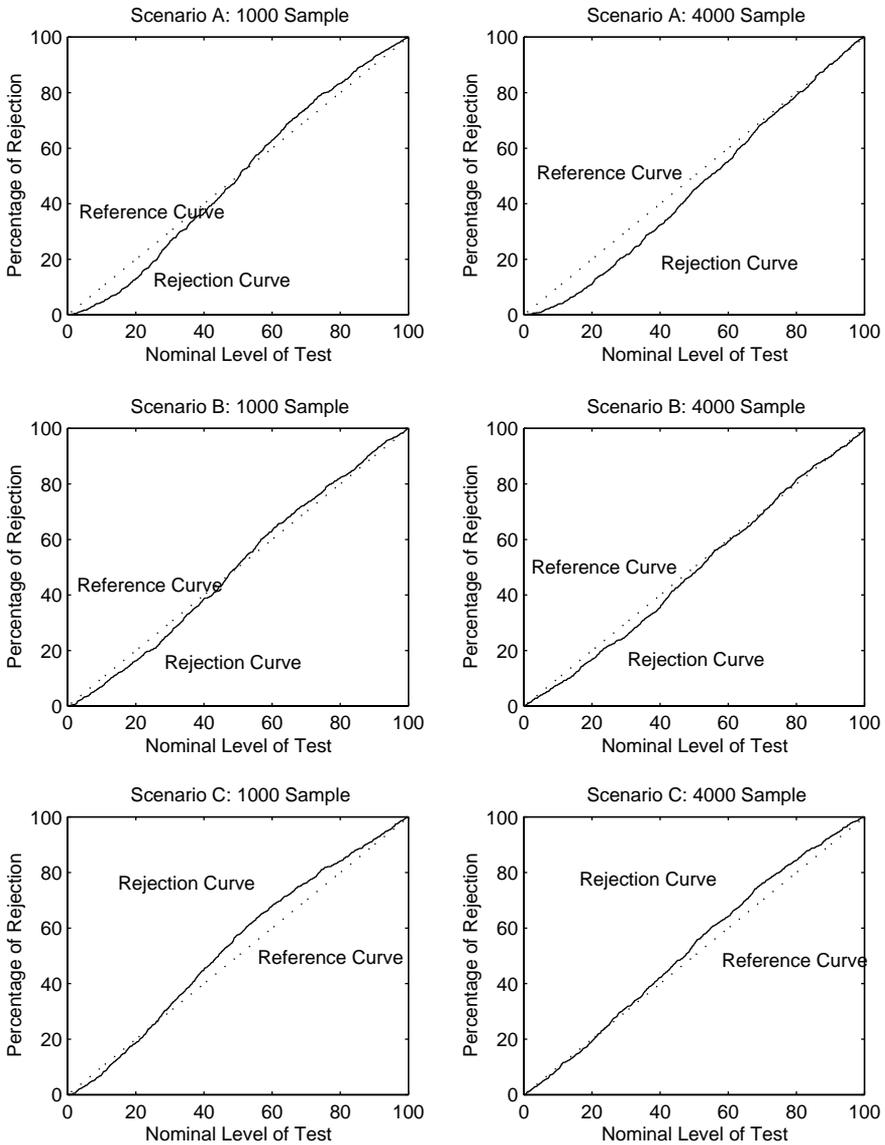
Fig. 4. GMM specification test of overidentifying restrictions with measurement error correction.

4. Empirical illustration

This section provides an empirical illustration using actual high-frequency five-minute returns for the foreign exchange market covering an entire decade. In addition to the estimates for the benchmark scalar affine diffusion, we also present estimates for

Table 3
Summary statistics for daily DM/\$ exchange rate volatility

Statistics	Quadratic variation	Squared return
Mean	0.5290	0.4889
SD	0.4839	1.0036
Skewness	3.7083	5.2285
Kurtosis	24.051	41.102
Minimum	0.0517	0.0000
5% Quant.	0.1384	0.0013
25% Quant.	0.2542	0.0324
Medium	0.3990	0.1513
75% Quant.	0.6252	0.4865
95% Quant.	1.3450	2.0478
Maximum	5.2453	12.2938

Note. The daily quadratic variation, or realized volatility, measure is constructed from the sum of 288 intraday five-minute squared percentage returns. The sample period extends from December 1, 1986 through December 1, 1996, for a total of 2445 observations.

models incorporating two stochastic volatility factors, leverage effects or asymmetries, as well as stochastic jumps. Appendix B gives a detailed description of each model, along with the requisite moment conditions. Motivated by the simulation results in the previous section, we include an additive nuisance parameter in all of the relevant moment conditions involving the squared five-minute-based quadratic variation measures. [15]

4.1. Data description

The data for the foreign exchange market were obtained from Olsen & Associates in Zürich, Switzerland, and consists of continuously recorded five-minute returns for the Deutsch Mark U.S. Dollar (DM/\$) spot exchange rates. The sample for the exchange rates spans the period from December 1, 1986 through December 1, 1996. After removing missing data, weekends, fixed holidays, and other calendar effects, as detailed in Andersen et al. (2001b), we are left with a total of 2445 trading days, each of which consists of 288 five-minute returns over the 24-hour trading cycle.

Table 3 provides a standard set of summary statistics for the corresponding quadratic variation, or realized volatility, measures based on the five-minute percentage returns, along with the same statistics for the daily squared percentage returns. The sample means for the two series imply an annualized volatility (standard deviation) of approximately $11.50 = (250 \times 0.5290)^{1/2}$ and $11.05 = (250 \times 0.4889)^{1/2}$ percent, respectively. Although both series afford unbiased estimates for the true quadratic variation on any given day, the squared returns are clearly much more noisy. This is

---

[15] Details regarding the parameter estimates for all of the models without the additive measurement error term are available upon request. With the exception of the slightly higher values for the variance-of-variance parameters, σ, the estimates are almost identical to the ones reported here.

DM/$ Exchange Rate Daily Quadratic Variation

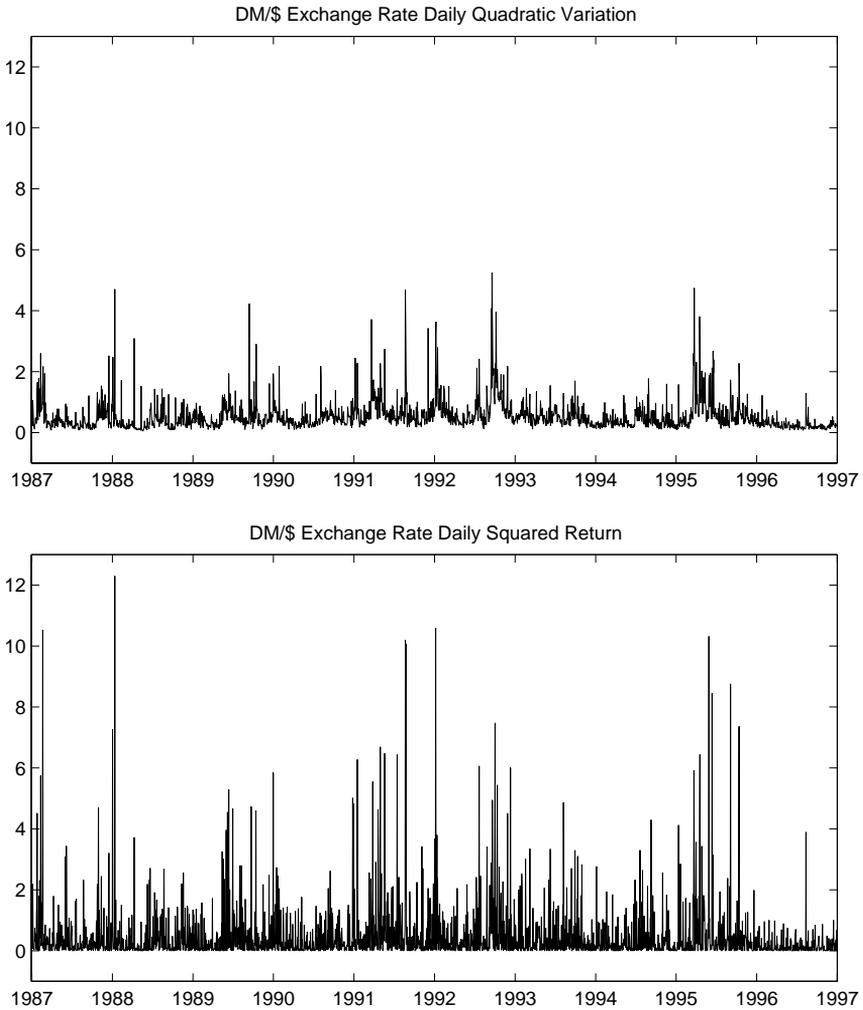DM/$ Exchange Rate Daily Squared Return

Fig. 5. Daily quadratic variation and squared return.

underscored by the sample standard deviation for the five-minute quadratic variation measure which is less than half of that for the daily squared returns. The higher order moments also indicate extremely heavy tails and significant skewness to the right in both of the unconditional distributions. Meanwhile, visual inspection of the corresponding time series plots in Fig. 5, clearly indicates a high degree of serial correlation in the underlying volatility process. The next section presents the estimation results from the different stochastic volatility models designed to capture these effects.

Table 4
Stochastic volatility model estimates

$$d\,p_t = \sqrt{V_{1t} + V_{2t}}\,dB_t + J\,dN(\lambda t),$$
$$dV_{1t} = \kappa_1(\theta_1 - V_{1t})\,dt + \sigma_1\sqrt{V_{1t}}\,dW_{1t},$$
$$dV_{2t} = \kappa_2(\theta_2 - V_{2t})\,dt + \sigma_2\sqrt{V_{2t}}\,dW_{2t},$$
where $\mathrm{corr}(dB_t, dW_{1t}) = \rho$ and $J \sim N(\mu_J, \sigma_J^2)$

	1FSV	1FSVJ	1FSVL	2FSV	2FSVJ
κ_1	0.1464(0.0387)	0.1389(0.0745)	0.1465(0.0375)	0.5708(0.0031)	0.5859(0.0032)
θ_1	0.5172(0.0342)	0.0415(0.2080)	0.5172(0.0355)	0.3257(0.0398)	0.3067(0.0179)
σ_1	0.5789(0.0580)	0.6109(0.0967)	0.5786(0.0578)	0.2286(0.0006)	0.2292(0.0001)
λ		52.6839(3.8436)			5.6930(0.6782)
μ_J		−0.0002(0.0003)			−0.0002(0.0026)
σ_J		0.0974(0.0167)			0.0807(0.0030)
ρ			−0.0243(0.0343)		
κ_2				0.0757(0.8984)	0.0734(0.0335)
θ_2				0.1786(0.0345)	0.1599(0.0168)
σ_2				0.1096(0.0063)	0.1098(0.00004)
GMM test of overidentifying restrictions					
$\mathscr{X}^2$	12.1476	18.5808	12.1526	6.0992	6.2216
d. o. f.	(2)	(1)	(2)	(3)	(2)
p-Value	0.0023	0.0001	0.0023	0.1069	0.0446

Note. The table reports the GMM estimates for the five different model specifications: one-factor stochastic volatility (1FSV), one-factor stochastic volatility with jumps (1FSVJ), one-factor stochastic volatility with leverage effect (1FSVL), two-factor stochastic volatility (2FSV), and two-factor stochastic volatility with jumps (2FSVJ). The sample period for the DM/\$ exchange rates extends from December 1, 1986 through December 1, 1996, for a total of 2445 daily observations. The daily quadratic variation measures are based on intraday five-minute returns. The actual moment conditions underlying the GMM estimates are detailed in Table 5. The asymptotic variance–covariance matrix for the parameter estimates is calculated using a Newey–West weighting scheme with a lag-length of five for the 1FSV, 1FSVJ, and 1FSVL models, and a lag-length of 60 for the 2FSV and 2FSVJ models.

4.2. Estimation results

The second column in Table 4 reports the estimates for the baseline one-factor stochastic volatility model (1FSV). As expected, the long-run mean parameter ($\theta = 0.517$) is fairly close to the sample mean of the underlying five-minute quadratic variation (0.529) reported in Table 3. Also, not surprisingly, the estimate for the variance-of-variance parameter, or σ, have the largest standard error of all the parameters. These results are generally in line with the simulation evidence reported in the previous section, and highlight the usefulness of the new estimation procedure in actual empirical applications. Meanwhile, the estimate for the mean reversion parameter, κ, is clearly on the high-side relative to the values reported in the extant literature using more complicated discrete time ARCH and stochastic volatility type models. The parameter estimates also violate the stationary condition $\sigma^2 \leqslant 2\kappa\theta$ in the 1FSV as well as the 1FSVJ and 1FSVL models. The GMM omnibus test for the overidentifying restrictions firmly rejects the 1FSV model.

The inability of the simple one-factor stochastic volatility model to fully account for the dynamic dependencies in the daily DM/\$ volatility is, of course, not surprising (see

e.g., Bates, 1996; Gallant and Long, 1997). A number of recent empirical studies have argued for the relevance of including multiple volatility factors, leverage effects, and/or jump components when modeling other speculative returns. The remaining columns in the table present the estimation results for various extensions of the baseline model specifically designed to accommodate each of these features.

The estimates for the one-factor jump-diffusion stochastic volatility (1FSVJ) model are reported in the third column. Compared to the 1FSV model, the estimate for the mean reversion parameter is slightly smaller, while the local variance parameter is slighter larger. The long-run mean parameter for the stochastic volatility factor is dramatically reduced to only 0.042. However, considering the expected jump volatility $\lambda(\mu_J^2 + \sigma_J^2) = 0.500$, the total expected unconditional volatility 0.541 is still very close to the sample mean. Unlike existing estimates for equity index returns, which tend to imply negative jumps on average, the average jump size for the DM/$ foreign exchange market is estimated to be close to zero. The daily jump rate of 52.7% is quite high, corresponding to roughly 100 jumps/yr. Meanwhile, the estimated standard deviation of the jumps is fairly small (one-tenth of a percent). The GMM test for the overidentifying restrictions also strongly rejects the 1FSVJ model.

The close correspondence between the estimates for κ, θ, and σ for the one-factor stochastic volatility model allowing for asymmetries or leverage effects (1FSVL) reported in the fourth column and the estimates for the baseline 1FSV model is to be expected. The first six moment conditions employed in the estimation are identical across the two models, and the last moment added to the 1FSVL is conditionally orthogonal to the other moments and only helps to identify the leverage coefficient, ρ. Also, not surprisingly in the present context the actual estimate of the leverage coefficient is insignificant and close to zero.

The fifth column in the table reports the estimates for the two-factor stochastic volatility (2FSV) model. One of the two volatility factors is strongly mean-reverting, and close to independent ($\kappa_1 = 0.571$); the other is highly persistent, and close to non-stationary ($\kappa_2 = 0.076$), although it is not very accurately estimated. Meanwhile, the sum of the estimates for the long-run mean parameters ($\theta_1 + \theta_2 = 0.504$) is fairly close to the sample mean of the five-minute quadratic variation (0.529). Both of the local variance parameters are also highly statistically significant, with the one associated with the first strongly mean-reverting volatility factor being roughly double of that for the more persistent second factor. These results are generally consistent with the recent findings in Alizadeh et al. (2001), who rely on the daily high-low range in estimating two-factor stochastic volatility models for several different exchange rates. It is noteworthy, that by including the second stochastic volatility factor, the GMM test for the overidentifying restrictions reported in Table 4 no longer rejects the model outright.

The last column in the table reports the estimation result for a two-factor stochastic volatility jump-diffusion model (2FSVJ). Consistent with the idea that the high jump-intensity (52.68%) for the 1FSVJ model is in part capturing the influence of the omitted less persistent volatility factor, the estimated jump rate for the 2FSVJ model is reduced to 5.69%, or roughly 15 jumps per year with a standard deviation of 8 basis

points. This finding of frequent small jumps is consistent with the results reported in Bates (1996) based on weekly foreign exchange options, and somewhat higher than the estimates for other financial assets recently reported in the literature. The parameter estimates for the two stochastic volatility factors are very close to the results for the 2FSV model (without jumps). Interestingly however, the estimates for the 2FSVJ model are more precisely determined. The implied steady state for the daily volatility, $\theta_1 + \theta_2 + \lambda(\mu_J^2 + \sigma_J^2) = 0.504$, also closely matches the sample mean of 0.529. Still, the addition of the two moment conditions used to identify the jump component results in a p-value of "only" 0.045 for the GMM omnibus test. In order to get a better idea about the fit of the various models, we, therefore, next report a series of specific conditional moment diagnostic tests.

4.3. Moment condition tests

The standard GMM omnibus tests reported in Table 4 are not particularly informative about the source(s) of model mis-specification. Following Tauchen (1985), it is relatively straightforward to construct a series of diagnostic t-tests associated with each of the moment conditions underlying the estimation, explicitly taking into account the variations resulting from the parameter estimation error uncertainty. In the notations of Section 2.3. above,

$$t = \text{diag}\{[\hat{W} - \hat{g}_\xi(\hat{g}'_\xi \hat{W}^{-1} \hat{g}_\xi)^{-1} \hat{g}'_\xi]\}^{-1/2} \sqrt{T} \hat{g},$$

where $\hat{W}$ refers to the estimated weighting matrix, $\hat{g}$ denotes the vector of moment conditions evaluated at the GMM parameter estimates, and $\hat{g}'_\xi$ gives the corresponding matrix of first derivatives with respect to each of the estimated parameters.

The resulting t-statistics for the different models are reported in Table 5. The apparent poor fit of the 1FSV model comes primarily from the moment involving the squared quadratic variation and the lagged squared quadratic variation (-4.459). This moment is most directly related to the dynamics of the volatility and the volatility-of-the-volatility. Allowing for stochastic jumps, as in the 1FSVJ model, the t-ratio associated with this same moment is reduced to -2.914. At the same time, the t-ratio for the moment involving the product of the squared quadratic variation with the lagged quadratic variation increases from -0.005 to -1.759. Since the moments are the same, and the cross moment appears to fit reasonably well, the t-ratios for the 1FSVL model are all very similar to the ones for the 1FSV model. Of course, all of these three models are rejected by the GMM omnibus tests in Table 4.

Interestingly, while the omnibus tests for the 2FSV(J) models in Table 4 are fairly supportive, the corresponding moment diagnostics reported in the fifth and sixth columns in Table 5 indicate that in spite of closely matching the first eight moment conditions, the two-factor models are not able to explain the empirical correlation between the squared quadratic variation and the lag six and lag 12 squared quadratic variation. It is noteworthy, that while the addition of the jump component improves the fit of the sixth lagged squared moment, the fit involving the 12 lagged squared moment actually worsens. As such, these results point towards the existence of important long-run volatility dependencies not captured by the two-factor models analyzed

Table 5
Moment condition tests

Moment condition	1FSV	1FSVJ	1FSVL	2FSV	2FSVJ
$E[\mathcal{V}_{t+1,t+2}|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}$	0.0071	−0.0441	0.0071		
$E[\mathcal{V}_{t+1,t+2}^2|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}^2$	−0.0162	−0.0338	−0.0161		
$E[\mathcal{V}_{t+1,t+2}\mathcal{V}_{t-1,t}|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}\mathcal{V}_{t-1,t}$	−0.2288	−0.2228	−0.2250		
$E[\mathcal{V}_{t+1,t+2}^2\mathcal{V}_{t-1,t}|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}^2\mathcal{V}_{t-1,t}$	−0.0049	−1.7589	−0.0045		
$E[\mathcal{V}_{t+1,t+2}\mathcal{V}_{t-1,t}^2|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}\mathcal{V}_{t-1,t}^2$	−0.3301	−0.5222	−0.3264		
$E[\mathcal{V}_{t+1,t+2}^2\mathcal{V}_{t-1,t}^2|\mathcal{G}_t] - \mathcal{V}_{t+1,t+2}^2\mathcal{V}_{t-1,t}^2$	−4.4594	−2.9140	−4.4727		
$E[p_{t+1}|\mathcal{G}_t] - p_{t+1}$		0.0564			−0.0822
$E[p_{t+1}^2|\mathcal{G}_t] - p_{t+1}^2$		−1.3092			−0.4988
$E[p_{t+1}(\mathcal{V}_{t+1,t+2} - b)/a|\mathcal{G}_t] - p_{t+1}(\mathcal{V}_{t+1,t+2} - b)/a$			1.0244		
$E[\mathcal{V}_{t+5,t+6}|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}$				0.0024	0.0041
$E[\mathcal{V}_{t+5,t+6}\mathcal{V}_{t-1,t}|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}\mathcal{V}_{t-1,t}$				−0.0053	−0.0055
$E[\mathcal{V}_{t+5,t+6}\mathcal{V}_{t-7,t-6}|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}\mathcal{V}_{t-7,t-6}$				0.0187	0.0186
$E[\mathcal{V}_{t+5,t+6}\mathcal{V}_{t-1,t}^2|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}\mathcal{V}_{t-1,t}^2$				0.0362	0.0354
$E[\mathcal{V}_{t+5,t+6}\mathcal{V}_{t-7,t-6}^2|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}\mathcal{V}_{t-7,t-6}^2$				0.2242	0.2107
$E[\mathcal{V}_{t+5,t+6}^2|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}^2$				0.0170	0.0258
$E[\mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-1,t}|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-1,t}$				0.0654	0.0722
$E[\mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-7,t-6}|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-7,t-6}$				1.0436	1.0860
$E[\mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-1,t}^2|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-1,t}^2$				7.6160	4.3396
$E[\mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-7,t-6}^2|\mathcal{G}_t] - \mathcal{V}_{t+5,t+6}^2\mathcal{V}_{t-7,t-6}^2$				4.5445	11.0394

Note. The table reports the diagnostic *t*-statistics associated with each of the moment conditions underlying the model estimates reported in Table 4. For further discussion of the construction of the test statistics see Section 4.3.

here. This is consistent with recent empirical studies documenting long-memory volatility dependencies (see e.g., Andersen et al. (2001b), and the references therein). The formulation and estimation of more complicated continuous time stochastic volatility models explicitly incorporating such dependencies constitute an active area of current research.

5. Concluding remarks

Exploiting closed form analytic expressions for the conditional moments of integrated volatility coupled with empirical quadratic variation measures constructed from high-frequency intraday data, we proposed a new class of GMM-type estimators for stochastic volatility diffusions. In contrast to other computationally demanding estimation procedures routinely employed in the literature, the GMM estimator developed here is very easy to implement, requiring only the solution to a standard non-linear optimization problem. Our Monte Carlo evidence shows that the procedure results in highly accurate parameter estimates and reliable statistical inference in realistic finite sample settings. In implementing the new estimator with actual five-minute returns from the foreign exchange market, our results confirm prior evidence in the literature concerning the existence of strong volatility clustering at the inter-daily level, and the presence of multiple volatility factors, frequent symmetric jumps, and possible long-memory dependencies.

It would be interesting to extend the estimator developed here to more complicated continuous time processes, including the possibility of long-memory volatility dynamics. More ambitious empirical applications might also entail the estimation of multivariate diffusions, which in turn would require vector versions of the integrated volatility and quadratic variation measures exploited here. The distributional features of the integrated volatility coupled with additional assumptions about the price of volatility risk may also be used in the pricing of financial options and the joint estimation of objective and risk neutral dynamics (Garcia et al. (2001) offer some intriguing recent developments along these lines).

Acknowledgements

Appendix A. Conditional moments of the baseline model

A.1. Conditional mean of integrated volatility

The conditional mean of the point-in-time volatility satisfies,

$$\mathrm{E}(V_T|\mathscr{F}_t) = V_t \mathrm{e}^{-\kappa(T-t)} + \theta(1 - \mathrm{e}^{-\kappa(T-t)}) = \alpha_{T-t} V_t + \beta_{T-t}. \tag{A.1}$$

Hence,

$$
\begin{aligned}
\mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t) &= \mathrm{E}\left(\int_t^T V_s \,\mathrm{d}s \,\middle|\, \mathscr{F}_t\right) \\
&= \int_t^T \mathrm{E}(V_s|\mathscr{F}_t)\,\mathrm{d}s \\
&= \int_t^T [V_t \mathrm{e}^{-\kappa(s-t)} + \theta(1 - \mathrm{e}^{-\kappa(s-t)})]\,\mathrm{d}s \\
&= V_t \frac{1}{\kappa}(1 - \mathrm{e}^{-\kappa(T-t)}) + \theta(T-t) - \frac{\theta}{\kappa}(1 - \mathrm{e}^{-\kappa(T-t)}) \\
&= a_{T-t} V_t + b_{T-t}.
\end{aligned}
\tag{A.2}
$$

For notational simplicity, denote the parameters for the daily horizon, or $T - t = 1$, by $a \equiv (1/\kappa)(1 - e^{-\kappa})$, $b \equiv \theta - (\theta/\kappa)(1 - e^{-\kappa})$, $\alpha \equiv e^{-\kappa}$ and $\beta \equiv \theta(1 - e^{-\kappa})$. Focusing on the one-day horizon, it follows that:

$$\mathrm{E}[\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_{t+1})|\mathscr{F}_t] = a\,\mathrm{E}(V_{t+1}|\mathscr{F}_t) + b$$

$$= a(\alpha V_t + \beta) + b$$

$$= \alpha[\mathrm{E}(\mathscr{V}_{t,t+1}|\mathscr{F}_t) - b] + a\beta + b,$$

which simplifies as

$$\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_t) = \alpha\,\mathrm{E}(\mathscr{V}_{t,t+1}|\mathscr{F}_t) + \beta.$$

Finally, by the Law of Iterated Expectations or reduction in information sets (Meddahi and Renault, 2000),

$$\mathrm{E}[\mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{F}_t)|\mathscr{G}_t] = \mathrm{E}(\mathscr{V}_{t+1,t+2}|\mathscr{G}_t) = \alpha\,\mathrm{E}(\mathscr{V}_{t,t+1}|\mathscr{G}_t) + \beta. \tag{A.3}$$

A.2. Conditional variance of integrated volatility

The stochastic differential equation (SDE) for $\mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t)$ may be generated as a function of V_t by applying Itô's formula to the affine diffusion in Eq. (3), [16]

$$\mathrm{dE}(\mathscr{V}_{t,T}|\mathscr{F}_t) = \left[a_{T-t}\kappa(\theta - V_t) + \frac{\partial a_{T-t}}{\partial t}V_t + \frac{\partial b_{T-t}}{\partial t}\right]\mathrm{d}t + a_{T-t}\sigma\sqrt{V_t}\,\mathrm{d}W_t,$$

which may be further simplified as

$$\mathrm{dE}(\mathscr{V}_{t,T}|\mathscr{F}_t) = -V_t\,\mathrm{d}t + a_{T-t}\sigma\sqrt{V_t}\,\mathrm{d}W_t. \tag{A.4}$$

Now fix the upper limit T, and let the lower limit t be time varying. The Itô integral implied by the SDE in Eq. (A.4) then takes the form

$$\mathrm{E}(\mathscr{V}_{T,T}|\mathscr{F}_T) = \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t) + \int_t^T (-V_s)\,\mathrm{d}s + \int_t^T a_{T-s}\,\sigma\sqrt{V_s}\,\mathrm{d}W_s.$$

However, $\mathrm{E}(\mathscr{V}_{T,T}|\mathscr{F}_T) = 0$, which implies that

$$\mathscr{V}_{t,T} - \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t) = \int_t^T a_{T-s}\,\sigma\sqrt{V_s}\,\mathrm{d}W_s.$$

By standard arguments and the substitution of Eq. (A.1)

$$Var(\mathscr{V}_{t,T}|\mathscr{F}_t) = \mathrm{E}[(\mathscr{V}_{t,T} - \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t))^2|\mathscr{F}_t]$$

$$= \mathrm{E}\left\{\left[\int_t^T a_{T-s}\,\sigma\sqrt{V_s}\,\mathrm{d}W_s\right]^2 \middle| \mathscr{F}_t\right\}$$

---

[16] The simple version of Itô's Lemma for a smooth function $f(V_t, t, T) \in C^2$ of a diffusion process V_t states that

$$\mathrm{d}f(V_t, t, T) = [f_V(V_t, t, T)\mu(V_t, t) + f_t(V_t, t, T) + \tfrac{1}{2}f_{VV}(V_t, t, T)\upsilon^2(V_t, t)]\mathrm{d}t + f_V(V_t, t, T)\upsilon(V_t, t)\mathrm{d}W_t,$$

where $\mu(V_t, t)$ and $\upsilon(V_t, t)$ denote the drift and diffusion functions, respectively, defining the V_t process.

$$= \int_t^T a_{T-s}^2 \, \sigma^2 \, \mathrm{E}(V_s | \mathscr{F}_t) \, ds$$

$$= \int_t^T a_{T-s}^2 \, \sigma^2 [\alpha_{s-t} V_t + \beta_{s-t}] \, ds$$

$$= A_{T-t} V_t + B_{T-t}, \tag{A.5}$$

where

$$A_{T-t} = \frac{\sigma^2}{\kappa^2} \left[\frac{1}{\kappa} - 2e^{-\kappa(T-t)}(T-t) - \frac{1}{\kappa} e^{-2\kappa(T-t)} \right],$$

$$B_{T-t} = \frac{\sigma^2}{\kappa^2} \left[\theta(T-t)(1 + 2e^{-\kappa(T-t)}) - \frac{3\theta}{\kappa}(1 - e^{-\kappa(T-t)}) + \frac{\theta}{2\kappa}(1 - e^{-\kappa(T-t)})^2 \right]$$

$$= \frac{\sigma^2}{\kappa^2} \left[\theta(T-t)\left(1 + 2e^{-\kappa(T-t)}\right) + \frac{\theta}{2\kappa}(e^{-\kappa(T-t)} + 5)(e^{-\kappa(T-t)} - 1) \right].$$

It follows also from Cox et al. (1985) and Eq. (A.1) above that,

$$\mathrm{E}(V_T^2 | \mathscr{F}_t) = Var(V_T | \mathscr{F}_t) + [\mathrm{E}(V_T | \mathscr{F}_t)]^2$$

$$= V_t \frac{\sigma^2}{\kappa}(e^{-\kappa(T-t)} - e^{-2\kappa(T-t)}) + \frac{\sigma^2 \theta}{2\kappa}(1 - e^{-\kappa(T-t)})^2$$

$$+ [\alpha_{T-t} V_t + \beta_{T-t}]^2$$

$$= C_{T-t} V_t + D_{T-t} + \alpha_{T-t}^2 V_t^2 + \beta_{T-t}^2 + 2\alpha_{T-t}\beta_{T-t} V_t$$

$$= \alpha_{T-t}^2 V_t^2 + [C_{T-t} + 2\alpha_{T-t}\beta_{T-t}]V_t + [D_{T-t} + \beta_{T-t}^2], \tag{A.6}$$

where

$$C_{T-t} = \frac{\sigma^2}{\kappa}(e^{-\kappa(T-t)} - e^{-2\kappa(T-t)}),$$

$$D_{T-t} = \frac{\sigma^2 \theta}{2\kappa}(1 - e^{-\kappa(T-t)})^2.$$

Focusing on the one-day horizon, (A.5) and (A.2) implies that

$$\mathrm{E}(\mathscr{V}_{t,t+1}^2 | \mathscr{F}_t) = Var(\mathscr{V}_{t,t+1} | \mathscr{F}_t) + [\mathrm{E}(\mathscr{V}_{t,t+1} | \mathscr{F}_t)]^2$$

$$= a^2 V_t^2 + (2ab + A)V_t + (b^2 + B). \tag{A.7}$$

Leading the arguments by one period and applying the Law of Iterated Expectation yields

$$\mathrm{E}[\mathrm{E}(\mathscr{V}_{t+1,t+2}^2 | \mathscr{F}_{t+1}) | \mathscr{F}_t] = a^2 \, \mathrm{E}(V_{t+1}^2 | \mathscr{F}_t) + (2ab + A) \, \mathrm{E}(V_{t+1} | \mathscr{F}_t) + (b^2 + B).$$

Now substituting for $E(V_{t+1}|\mathscr{F}_t)$ by Eq. (A.1) and $E(V_{t+1}^2|\mathscr{F}_t)$ by Eq. (A.6), and reversely substituting out V_t^2 by Eq. (A.7) and V_t by Eq. (A.2), it follows that

$$
\begin{aligned}
E(\mathscr{V}_{t+1,t+2}^2|\mathscr{F}_t) ={}& a^2[\alpha^2 V_t^2 + (C + 2\alpha\beta)V_t + (D + \beta^2)] \\
& + (2ab + A)(\alpha V_t + \beta) + (b^2 + B) \\
={}& \alpha^2 a^2 V_t^2 + [a^2(C + 2\alpha\beta) + \alpha(2ab + A)]V_t \\
& + [a^2(D + \beta^2) + \beta(2ab + A) + (b^2 + B)] \\
={}& \alpha^2[E(\mathscr{V}_{t,t+1}^2|\mathscr{F}_t) - (2ab + A)V_t - (b^2 + B)] \\
& + [a^2(C + 2\alpha\beta) + \alpha(2ab + A)]V_t \\
& + [a^2(D + \beta^2) + \beta(2ab + A) + (b^2 + B)] \\
={}& \alpha^2\, E(\mathscr{V}_{t,t+1}^2|\mathscr{F}_t) \\
& + [a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)]\frac{1}{a}[E(\mathscr{V}_{t,t+1}|\mathscr{F}_t) - b] \\
& + [a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B)] \\
={}& \alpha^2\, E(\mathscr{V}_{t,t+1}^2|\mathscr{F}_t) \\
& + \frac{1}{a}[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)]\, E(\mathscr{V}_{t,t+1}|\mathscr{F}_t) \\
& - \frac{b}{a}[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)] \\
& + [a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B)]. \qquad \text{(A.8)}
\end{aligned}
$$

Lastly, by the Law of Iterated Expectations,

$$
\begin{aligned}
E[E(\mathscr{V}_{t+1,t+2}^2|\mathscr{F}_t)|\mathscr{G}_t] ={}& E(\mathscr{V}_{t+1,t+2}^2|\mathscr{G}_t) \\
={}& \alpha^2\, E(\mathscr{V}_{t,t+1}^2|\mathscr{G}_t) \\
& + \frac{1}{a}[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)]\, E(\mathscr{V}_{t,t+1}|\mathscr{G}_t) \\
& - \frac{b}{a}[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)] \\
& + [a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B)] \\
={}& H\, E(\mathscr{V}_{t,t+1}^2|\mathscr{G}_t) + I\, E(\mathscr{V}_{t,t+1}|\mathscr{G}_t) + J, \qquad \text{(A.9)}
\end{aligned}
$$

where $H = \alpha^2$, $I = 1/a[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)]$, and $J = -b/a[a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)] + [a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B)]$.

Appendix B. Extensions of the baseline model

This appendix demonstrates how the moment conditions for the one-factor square-root volatility model may be extended to allow for multiple volatility factors, leverage effects or asymmetries, and jump components.

B.1. Multifactor stochastic volatility model

Consider the model,

$$\mathrm{d}p_t = \sqrt{V_{1t} + V_{2t}}\,\mathrm{d}B_t,$$

$$\mathrm{d}V_{1t} = \kappa_1(\theta_1 - V_{1t})\,\mathrm{d}t + \sigma_1\sqrt{V_{1t}}\,\mathrm{d}W_{1t},$$

$$\mathrm{d}V_{2t} = \kappa_2(\theta_2 - V_{2t})\,\mathrm{d}t + \sigma_2\sqrt{V_{2t}}\,\mathrm{d}W_{2t}. \tag{B.1}$$

Also, let $V_t = V_{1t} + V_{2t}$, and $\mathscr{V}_{t,T} \equiv \int_t^T V_s\,\mathrm{d}s = \mathscr{V}_{1t,T} + \mathscr{V}_{2t,T} \equiv \int_t^T V_{1s}\,\mathrm{d}s + \int_t^T V_{2s}\,\mathrm{d}s$.

B.1.1. First moment
From Appendix A.1,

$$\mathscr{V}_{1t+1,t+2} = \alpha_1\mathscr{V}_{1t,t+1} + \beta_1 + MA(1),$$

$$\mathscr{V}_{2t+1,t+2} = \alpha_2\mathscr{V}_{2t,t+1} + \beta_2 + MA(1), \tag{B.2}$$

where the symbolic $MA(1)$ representations for the error structures have the following "closed-from" expressions:

$$MA(1) = -\int_{t+1}^{t+2}\int_{t+1}^{s}\sigma_1\sqrt{V_{1u}}\,\mathrm{d}W_{1u}\,\mathrm{d}s + \alpha_1\int_{t}^{t+1}\int_{t}^{s}\sigma_1\sqrt{V_{1u}}\,\mathrm{d}W_{1u}\,\mathrm{d}s,$$

$$MA(1) = -\int_{t+1}^{t+2}\int_{t+1}^{s}\sigma_2\sqrt{V_{2u}}\,\mathrm{d}W_{2u}\,\mathrm{d}s + \alpha_2\int_{t}^{t+1}\int_{t}^{s}\sigma_2\sqrt{V_{2u}}\,\mathrm{d}W_{2u}\,\mathrm{d}s. \tag{B.3}$$

Using the standard argument that $ARMA(1,1) + ARMA(1,1) = ARMA(2,2)$ (Granger and Morris, 1976), it follows that

$$(1 - \alpha_1 L)(1 - \alpha_2 L)\mathscr{V}_{t+1,t+2} = [(1 - \alpha_2)\beta_1 + (1 - \alpha_1)\beta_2] + MA(2). \tag{B.4}$$

Thus, even though the autocovariance functions are not explicitly solvable (due to the unknown heteroscedasticity), additional moment conditions involving $\mathscr{V}_{t+1,t+2}$ are readily constructed through augmentation using appropriately lagged instruments.

B.1.2. Second moment
From Appendix A.2,

$$\mathscr{V}^2_{1t+1,t+2} = H_1\mathscr{V}^2_{1t,t+1} + I_1\mathscr{V}_{1t,t+1} + J_1 + MA(1),$$

$$\mathscr{V}^2_{2t+1,t+2} = H_2\mathscr{V}^2_{2t,t+1} + I_2\mathscr{V}_{2t,t+1} + J_2 + MA(1). \tag{B.5}$$

Of course, $\mathscr{V}^2_{t+1,t+2} = \mathscr{V}^2_{1t+1,t+2} + \mathscr{V}^2_{2t+1,t+2} + 2\mathscr{V}_{1t+1,t+2}\mathscr{V}_{2t+1,t+12}$. However, by multiplication of the two components in Eq. (B.2),

$$2\mathscr{V}_{1t+1,t+2}\mathscr{V}_{2t+1,t+2} = 2\alpha_1\alpha_2\mathscr{V}_{1t,t+1}\mathscr{V}_{2t,t+1} + 2\alpha_1\beta_2\mathscr{V}_{1t,t+1}$$

$$+ 2\alpha_2\beta_1\mathscr{V}_{2t,t+1} + 2\beta_1\beta_2 + MA(1), \tag{B.6}$$

where the $MA(1)$ error structure follows by the assumption that the two factors are independent. Now multiply the two equations in (B.5) by $(1 - H_2L)(1 - \alpha_1\alpha_2L)$ and $(1 - H_1L)(1 - \alpha_1\alpha_2L)$, respectively, multiply Eq. (B.6) by $(1 - H_1L)(1 - H_2L)$, and aggregate the results,

$$(1 - H_1L)(1 - H_2L)(1 - \alpha_1\alpha_2L)\mathscr{V}^2_{t+1,t+2}$$

$$= [(1 - H_2L)(1 - \alpha_1\alpha_2L)I_1 + 2\alpha_1\beta_2(1 - H_1L)(1 - H_2L)]\mathscr{V}_{1t,t+1}$$

$$+ [(1 - H_1L)(1 - \alpha_1\alpha_2L)I_2 + 2\alpha_2\beta_1(1 - H_1L)(1 - H_2L)]\mathscr{V}_{2t,t+1}$$

$$+ [(1 - H_2)(1 - \alpha_1\alpha_2)J_1 + (1 - H_1)(1 - \alpha_1\alpha_2)J_2$$

$$+ 2\beta_1\beta_2(1 - H_1)(1 - H_2)] + MA(3). \tag{B.7}$$

Finally, multiply Eq. (B.7) by $(1 - \alpha_1L)(1 - \alpha_2L)$, and each of the two components in Eq. (B.2) by $L(1 - \alpha_2L)[(1 - H_2L)(1 - \alpha_1\alpha_2L)I_1 + 2\alpha_1\beta_2(1 - H_1L)(1 - H_2L)]$ and $L(1 - \alpha_1L)[(1 - H_1L)(1 - \alpha_1\alpha_2L)I_2 + 2\alpha_2\beta_1(1 - H_1L)(1 - H_2L)]$, respectively, and it follows by aggregation that:

$$(1 - \alpha_1L)(1 - \alpha_2L)(1 - H_1L)(1 - H_2L)(1 - \alpha_1\alpha_2L)\mathscr{V}^2_{t+1,t+2}$$

$$= \beta_1(1 - \alpha_2)[(1 - H_2)(1 - \alpha_1\alpha_2)I_1 + 2\alpha_1\beta_2(1 - H_1)(1 - H_2)]$$

$$+ \beta_2(1 - \alpha_1)[(1 - H_1)(1 - \alpha_1\alpha_2)I_2 + 2\alpha_2\beta_1(1 - H_1)(1 - H_2)]$$

$$+ (1 - \alpha_1)(1 - \alpha_2)$$

$$\times [(1 - H_2)(1 - \alpha_1\alpha_2)J_1 + (1 - H_1)(1 - \alpha_1\alpha_2)J_2 + 2\beta_1\beta_2(1 - H_1)(1 - H_2)]$$

$$+ MA(5). \tag{B.8}$$

Additional moment conditions are readily constructed using instruments lagged six or more periods relative to $\mathscr{V}^2_{t+1,t+2}$.

B.2. Leverage effect

Consider the model,

$$\mathrm{d}p_t = \sqrt{V_t}\,\mathrm{d}B_t,$$

$$\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t,$$

$$\mathrm{corr}(\mathrm{d}B_t, \mathrm{d}W_t) = \rho. \tag{B.9}$$

Adding the log price p_t to the continuous filtration $\mathscr{F}_t = \sigma\{p_s, V_s; -\infty < s \leqslant t\}$, it follows that

$$\mathrm{E}(p_T V_T | \mathscr{F}_t) = p_t V_t \mathrm{e}^{-\kappa(T-t)} + (p_t \kappa + \rho\sigma)\theta(T-t)$$

$$+ \rho\sigma \left[\frac{V_t - \theta}{\kappa} (1 - \mathrm{e}^{-\kappa(T-t)}) \right]. \tag{B.10}$$

To derive this result, apply Itô's Lemma to $p_t V_t$, express the product as a stochastic differential equation, and take the conditional expectation of $p_T V_T$,

$$\mathrm{E}(p_T V_T | \mathscr{F}_t) = p_t V_t + \mathrm{E} \left\{ \int_t^T [p_u \kappa(\theta - V_u) + \rho\sigma V_u] \, \mathrm{d}u | \mathscr{F}_t \right\}.$$

Interchanging the two integration operators and taking derivatives of both sides with respect to the upper time limit, then yields the first-order linear ordinary differential equation

$$\frac{\mathrm{d}\mathrm{E}(p_s V_s | \mathscr{F}_t)}{\mathrm{d}s} = -\kappa \, \mathrm{E}(p_s V_s | \mathscr{F}_t) + \kappa\theta \, \mathrm{E}(p_s | \mathscr{F}_t) + \rho\sigma \, \mathrm{E}(V_s | \mathscr{F}_t),$$

from which Eq. (B.10) follows by existing solutions for $\mathrm{E}(p_T | \mathscr{F}_t)$ and $\mathrm{E}(V_T | \mathscr{F}_t)$. Focusing on the unit time-interval, or $T = t + 1$, and using the earlier result that $\mathrm{E}(\mathscr{V}_{t,t+1} | \mathscr{F}_t) = aV_t + b$, the left-hand side (LHS) of Eq. (B.10) may be written as

$$\mathrm{E}\{p_{t+1} \mathrm{E}[(\mathscr{V}_{t+1,t+2} - b)/a | \mathscr{F}_{t+1}] | \mathscr{F}_t\} = \mathrm{E}[p_{t+1}(\mathscr{V}_{t+1,t+2} - b)/a | \mathscr{F}_t].$$

Similarly, the right-hand side (RHS) of Eq. (B.10) may be expressed as

$$\mathrm{E}[p_t(\mathscr{V}_{t,t+1} - b)/a | \mathscr{F}_t]\mathrm{e}^{-\kappa} + (p_t \kappa + \rho\sigma)\theta$$

$$+ \frac{\rho\sigma}{\kappa} \mathrm{E}[(\mathscr{V}_{t,t+1} - b)/a | \mathscr{F}_t](1 - \mathrm{e}^{-\kappa}) - \frac{\rho\sigma\theta}{\kappa}(1 - \mathrm{e}^{-\kappa}).$$

Now taking conditional expectations of the LHS and RHS expressions above with respect to the extended discrete filtration $\mathscr{G}_t = \sigma\{p_{t-s}, \mathscr{V}_{t-s-1,t-s}; s = 0, 1, 2, \ldots, \infty\}$, the cross moment condition becomes

$$\mathrm{E}[p_{t+1}(\mathscr{V}_{t+1,t+2} - b)/a | \mathscr{G}_t] = \mathrm{E}[p_t(\mathscr{V}_{t,t+1} - b)/a | \mathscr{G}_t]\mathrm{e}^{-\kappa} + (p_t \kappa + \rho\sigma)\theta$$

$$+ (\rho\sigma/\kappa)\{\mathrm{E}[(\mathscr{V}_{t,t+1} - b)/a | \mathscr{G}_t] - \theta\}(1 - \mathrm{e}^{-\kappa}), \tag{B.11}$$

which is directly implementable. [17]

---

[17] As pointed out by Nour Meddahi, although the moment condition in (B.11) involves a non-linear function of the (non-stationary) price as opposed to the (stationary) return, the GMM estimator remains consistent (Kitamura and Phillips, 1997). This also holds true for the jump process analyzed below.

B.3. Jump component

Consider the model,

$$\mathrm{d}p_t = \sqrt{V_t}\,\mathrm{d}B_t + J\,\mathrm{d}N(\lambda t),$$
$$\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t, \tag{B.12}$$

where the Brownian motion $\mathrm{d}B_t$ and the Poisson process $J\,\mathrm{d}N(\lambda t)$ are independent, and the jumps arrive at the exponential rate of $\lambda\,\mathrm{d}t$ with the size of the jumps, J, determined by the normal distribution $\mathrm{N}(\mu_J, \sigma_J^2)$. Also, define the extended continuous and discrete filtrations, $\mathscr{F}_t = \sigma\{p_s, V_s; -\infty < s \leqslant t\}$ and $\mathscr{G}_t = \sigma\{p_{t-s}, \mathscr{V}_{t-s-1,t-s}; s = 0, 1, 2, \ldots, \infty\}$, respectively. The jump parameters may then be identified from the following two moment conditions:

$$\mathrm{E}(p_T|\mathscr{G}_t) = p_t + \lambda\mu_J(T-t),$$

$$\mathrm{E}(p_T^2|\mathscr{G}_t) = p_t^2 + 2\lambda\mu_J(T-t)p_t + 2\lambda^2\mu_J^2(T-t)$$
$$+ \lambda(\sigma_J^2 + \mu_J^2)(T-t) + \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{G}_t). \tag{B.13}$$

The first equation follows directly by noticing that

$$\mathrm{E}(p_T|\mathscr{F}_t) = p_t + \mathrm{E}\left[\int_t^T \lambda\,\mathrm{E}_J(J)|\mathscr{F}_t\right].$$

In deriving the second equation, first apply Itô's Lemma to p_t^2, and write p_T^2 as a stochastic integral starting at p_t^2. Then, taking conditional expectation of p_T^2 yields,

$$\mathrm{E}(p_T^2|\mathscr{F}_t) = p_t^2 + \mathrm{E}\left\{\int_t^T [V_s + \lambda\,\mathrm{E}_J((p_s + J)^2 - p_s^2)]\,\mathrm{d}s|\mathscr{F}_t\right\}$$

$$= p_t^2 + \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{F}_t) + \mathrm{E}\left\{\int_t^T [\lambda 2p_s E_J(J) + \lambda\,\mathrm{E}_J(J^2)]\,\mathrm{d}s|\mathscr{F}_t\right\}.$$

Substituting the first equation, evaluating the integral, and changing the information sets to $\mathscr{G}_t$, this reduces to the second expression in Eq. (B.13).

It is important to recognize, that in the presence of jumps the quadratic variation comprises two components, $QV_{t,T} = \int_t^T V_s\,\mathrm{d}s + \int_t^T J^2\,\mathrm{d}N(\lambda s)$. Hence, in implementing the moment conditions involving $\mathscr{V}_{t,T}$, the following substitutions are also required:

$$\mathrm{E}(\mathscr{V}_{t,T}|\mathscr{G}_t) = \mathrm{E}(QV_{t,T}|\mathscr{G}_t) - \lambda\,\mathrm{E}(J^2)(T-t),$$

$$\mathrm{E}(\mathscr{V}_{t,T}^2|\mathscr{G}_t) = \mathrm{E}(QV_{t,T}^2|\mathscr{G}_t) - 2\,\mathrm{E}(QV_{t,T}|\mathscr{G}_t)\lambda\,\mathrm{E}(J^2)(T-t) + \lambda\,\mathrm{E}(J^4)(T-t). \tag{B.14}$$

B.4. Multifactor model with jumps

Consider the model,

$$\mathrm{d}p_t = \sqrt{V_{1t} + V_{2t}}\,\mathrm{d}B_t + J\,\mathrm{d}N(\lambda t),$$

$$\mathrm{d}V_{1t} = \kappa_1(\theta_1 - V_{1t})\,\mathrm{d}t + \sigma_1\sqrt{V_{1t}}\,\mathrm{d}W_{1t},$$

$$\mathrm{d}V_{2t} = \kappa_2(\theta_2 - V_{2t})\,\mathrm{d}t + \sigma_2\sqrt{V_{2t}}\,\mathrm{d}W_{2t}, \tag{B.15}$$

where $\mathrm{d}N(\lambda t)$ is a Poisson jump arrival process and $J \sim \mathrm{N}(\mu_J, \sigma_J^2)$. Combining the arguments in Appendix B.1 and B.3 readily results in the following four conditional moments:

$$\mathrm{E}(p_T|\mathscr{G}_t) = p_t + \lambda\mu_J(T-t),$$

$$\mathrm{E}(p_T^2|\mathscr{G}_t) = p_t^2 + 2\lambda\mu_J(T-t)p_t + 2\lambda^2\mu_J^2(T-t)$$

$$\qquad\qquad + \lambda(\sigma_J^2 + \mu_J^2)(T-t) + \mathrm{E}(\mathscr{V}_{t,T}|\mathscr{G}_t),$$

$$(1-\alpha_1 L)(1-\alpha_2 L)\mathscr{V}_{t+1,t+2} = [(1-\alpha_2)\beta_1 + (1-\alpha_1)\beta_2] + MA(2),$$

$$(1-\alpha_1 L)(1-\alpha_2 L)(1-H_1 L)(1-H_2 L)(1-\alpha_1\alpha_2 L)\mathscr{V}_{t+1,t+2}^2$$

$$\quad = \beta_1(1-\alpha_2)[(1-H_2)(1-\alpha_1\alpha_2)I_1 + 2\alpha_1\beta_2(1-H_1)(1-H_2)]$$

$$\qquad + \beta_2(1-\alpha_1)[(1-H_1)(1-\alpha_1\alpha_2)I_2 + 2\alpha_2\beta_1(1-H_1)(1-H_2)]$$

$$\qquad + (1-\alpha_1)(1-\alpha_2)$$

$$\qquad \times[(1-H_2)(1-\alpha_1\alpha_2)J_1 + (1-H_1)(1-\alpha_1\alpha_2)J_2$$

$$\qquad + 2\beta_1\beta_2(1-H_1)(1-H_2)] + MA(5). \tag{B.16}$$

References

Aït-Sahalia, Y., 2001. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approach. Econometrica, forthcoming.

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2001. Range-based estimation of stochastic volatility models. Journal of Finance, forthcoming.

Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2000. Microstructure bias and volatility signatures. Department of Economics, Duke University, Work in progress.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001a. The distribution of realized stock return volatility. Journal of Financial Economics 61, 43–76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001b. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

Bai, X., Russell, J.R., Tiao, G.C., 1999. Beyond merton's utopia; effects of non-normality and dependence on the precision of variance estimates using high-frequency financial data. Working Paper Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O., Shephard, N., 2001a. Econometric analysis of realised volatility and its use in estimating lévy based non-gaussian ou type stochastic volatility models. Working Paper Nuffield College, Oxford University.

Barndorff-Nielsen, O., Shephard, N., 2001b. Non-gaussian ou based models and some of their uses in financial economics. Journal of the Royal Statistical Society, Series B 63, 167–241.

Bates, D.S., 1996. Jumps and stochastic volatility: exchange rate process implicit in deutsche mark options. The Review of Financial Studies 9, 69–107.

Chacko, G., Viceira, L.M., 1999. Spectral GMM estimation of continuous-time processes. Working Paper Graduate School of Business, Harvard University.

Conley, T., Hansen, L.P., Luttmer, E., Scheinkman, J., 1997. Short term interest rates as subordinated diffusions. Review of Financial Studies 10, 525–578.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–407.

Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. Journal of Finance 55, 1943–1978.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6, 379–406.

Duffie, D., Singleton, K., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.

Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.

Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. Journal of Business and Economic Statistics 19, 177–191.

Feller, W., 1951. Two singular diffusion problems. Annals of Mathematics 54, 173–182.

Fisher, M., Gilles, C., 1996. Estimating exponential affine models of the term structure. Working Paper.

Gallant, A.R., Long, J.R., 1997. Estimating stochastic differential equations efficiently by minimum chi-square. Biometrika 84, 125–141.

Gallant, A.R., Tauchen, G.E., 1996. Which Moment to Match?. Econometric Theory 12, 657–681.

Garcia, R., Lewis, M.-A., Renault, E., 2001. Estimation of objective and risk-neutral distributions based on moments of integrated volatility. Working Paper CRDE, Université de Montréal.

Ghysels, E., Harvey, A., Renault, E., 1996. Stochastic volatility. In: Maddala, G.S. (Ed.), Handbook of Statistics, Statistical Method in Finance, Vol. 14. North-Holland, Amsterdam.

Goodhart, C.A.E., O'Hara, M., 1997. High frequency data in financial markets: issues and applications. Journal of Empirical Finance 4, 73–114.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, s85–s118.

Granger, C.W.J., Morris, M.J., 1976. Time series modeling and interpretation. Journal of Royal Statistical Society, Series A 139, 246–257.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L.P., 1985. A method for calculating bounds on the asymptotic covariance matrices of generalized method of moments estimators. Journal of Econometrics 30, 203–238.

Hansen, L.P., Scheinkman, J.A., 1995. Back to the future: generalized moment implications for continuous time Markov process. Econometrica 63, 767–804.

Hansen, L.P., Heaton, J.C., Ogaki, M., 1988. Efficiency bounds implied by multiperiod conditional moment restrictions. Journal of the American Statistical Association 83, 863–871.

Harvey, A.C., Shephard, N., 1994. Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business and Economic Statistics 14, 429–434.

Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Finanacial Studies 6, 327–343.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371–389.

Jiang, G., Knight, J., 2001. Estimation of continuous time stochastic processes via the empirical characteristic function. Journal of Business and Economic Statistics, forthcoming.

Karatzas, I., Shreve, S.E., 1997. Brownian Motion and Stochastic Calculus. Springer, Berlin.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with arch models. Review of Economic Studies 65, 361–394.

Kitamura, Y., Phillips, P.C.B., 1997. Fully modified IV, GIVE and GMM estimation with possibly non-stationary regressors and instruments. Journal of Econometrics 80, 85–123.

Kloeden, P.E., Platen, E., 1992. Numerical Solution of Stochastic Differential Equations. Springer, Berlin.

Lewis, M.-A., 2001. Analytical expressions for the moments of the integrated volatility in affine stochastic volatility models. Working Paper Banque Nationale, Montréal and CIRANO.

Lo, A.W., 1988. Maximum likelihood estimation of generalized Itô process with discretely sampled data. Econometric Theory 4, 231–247.

Martens, M., 2001. Forecasting daily exchange rate volatility using intraday returns. Journal of International Money and Finance 20, 1–23.

Meddahi, N., 2001. A theoretical comparison between integrated and realized volatilities. Working Paper CRDE, Université de Montréal.

Meddahi, N., Renault, E., 2000. Temporal aggregation of volatility models. Working Paper CIRANO, Montréal.

Merton, R.C., 1980. On estimating the expected return on the market. Journal of Financial Economics 8, 323–361.

Nelson, D.B., 1992. Filtering and forecasting with misspecified arch models i: getting the right variance with the wrong model. Journal of Econometrics 52, 61–90.

Newey, W.K., West, K.D., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pan, J., 2001. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics, forthcoming.

Pritsker, M.G., 1998. Nonparametric density estimation and tests of continuous time interest rate models. Review of Financial Studies 11, 449–487.

Protter, P., 1992. Stochastic Integration and Differential Equations: a New Approach. Springer, Berlin.

Ruiz, E., 1994. Quasi-maximum likelihood estimation of stochastic volatility models. Journal of Econometrics 63, 289–306.

Singleton, K., 2001. Estimation of affine asset pricing models using the empirical characteristic function. Journal of Econometrics 102, 111–141.

Tauchen, G., 1985. Diagnostic testing and evaluation of maximum likelihood models. Journal of Econometrics 30, 415–443.

Zhou, H., 2001. Finite sample properties of EMM, GMM, QMLE, and MLE for a square-root interest rate diffusion model. Journal of Computational Finance, forthcoming.