Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility †

Tim Bollerslev
and
Hao Zhou

Department of Economics
Duke University
Durham, NC 27708-0097

February 2000

†We would like to thank Ravi Bansal, Neil Shephard, George Tauchen, and the participants in the Duke financial econometrics lunch group for many valuable comments and discussions. Financial support from an NSF grant to the NBER (Bollerslev) is gratefully acknowledged. An updated version of this paper can be downloaded from the following website `http://www.duke.edu/~hz/paper/dnld.html`. For questions and comments, please contact Tim Bollerslev, Department of Economics, Duke University, Box 90097, Durham NC 27708 USA, Email boller@econ.duke.edu, Phone 919-660-1846, Fax 919-684-8974; or Hao Zhou, Department of Economics, Duke University, Box 90097, Durham NC 27708 USA, Email zhou@econ.duke.edu, Phone 919-660-1800, Fax 919-684-8974

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 integrated volatility, which is effectively approximated by the quadratic variation of the process. We successfully implement the resulting GMM estimator with high-frequency five-minute foreign exchange and equity index returns. Our simulation evidence and actual empirical results indicate that the method is very reliable and accurate. The computational speed of the procedure compares very favorably to other existing estimation methods in the literature.

JEL Classification: C13, C22.

Keywords: Stochastic Volatility Diffusions; Integrated Volatility; High-Frequency Data; GMM Estimation.

1 Introduction

Continuous time methods and no-arbitrage arguments figure prominently in the theoretical asset pricing literature. However, some of the most influential contributions have been based upon fairly simple and restrictive assumptions concerning the process for the underlying state variable(s)—leading examples include the celebrated Black-Scholes option pricing formula, which assumes that the true process for the underlying asset follows a geometric Brownian motion; and the CIR model for the term structure of interest rates, which is derived under the assumption of a square-root process for the short rate. Meanwhile, the burgeon empirical literature on discrete time ARCH and stochastic volatility models (see Bollerslev et al., 1994; Ghysels et al., 1996), have called into question the empirical validity of a time invariant diffusion, or a single state variable, as a reasonable assumption for most speculative rate of return series. In response to this, several recent studies have utilized more realistic continuous time models, explicitly allowing for time varying volatility in the state variables. The Hull and White (1987) and Heston (1993) stochastic volatility option pricing formula, and the exponential-affine class of term structure models in Duffie and Kan (1996) and Dai and Singleton (2000), are all notable examples.

Aside from a few special cases, estimation of these continuous time stochastic volatility models are complicated by the lack of a closed form expression for the transition density function for the corresponding discretely sampled observations, and numerous competing estimation strategies have been proposed in the literature. An incomplete list of these different techniques includes the Markov Chain Monte Carlo (MCMC) methods advanced by Jacquier et al. (1994), Eraker (1998), Kim et al. (1998) and Elerian et al. (1998); the simulated methods of moments approach in Duffie and Singleton (1993); the indirect inference procedure of Gourieroux et al. (1993) utilized by Engle and Lee (1997); the Efficient Methods of Moments (EMM) developed by Gallant and Tauchen (1996) and Gallant and Long (1997) and implemented by Andersen et al. (1999c); the infinitesimal moment generator underlying the GMM procedure in Hansen and Scheinkman (1995) and Conley et al. (1997); the non-parametric series expansions of the transition density function advocated by Aït-Sahalia (1996) and Stanton (1997) and the related kernel estimator in Bandi and Phillips (1999); the approximation method to the likelihood function building on the Kolmogorov forward equations in Lo (1988) and Aït-Sahalia (1998); and the spectral GMM estimator utilizing the empirical characteristic function in Chacko and Viceira (1999) and Singleton (1999). While all of these procedures yield consistent, and in most 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. The basic idea is straight forward. Instead of integrating out the latent volatility, as it is implicitly done in the estimation procedures in the extant literature, the strategy proposed here utilizes high-frequency data for explicitly measuring the realized volatility.

High-frequency, or tick-by-tick, data have recently become available for a host of different financial instruments. Following the work of Merton (1980) and Nelson (1992), such data could in principle be used to construct point-wise consistent filtering measurements for the instantaneous volatility. Unfortunately, the optimal filter weights depend in complicated ways on the particular model structure (Nelson and Foster, 1994), and in practice the continuous record asymptotics underlying the theoretical arguments are corrupted by inherent discreteness, time-of-day effects, bid-ask spreads, and other market microstructure frictions.[1] Meanwhile, it is possible to construct *model-free* unbiased estimates of the integrated volatility over a fixed time interval, say one day, by simply summing the squared returns over the relevant time-period. Moreover, by the theory of quadratic variation, the sum of the squared inter-period returns afford increasingly more accurate ex-post volatility measurements as the length of the return-horizon decreases.[2] Motivated by this idea, Andersen et al. (1999b, 2000b) offer a detailed descriptive characterization of the salient distributional features of daily *realized* foreign exchange and individual stock return volatilities constructed from high-frequency five-minute returns.

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 concreteness we restrict the analysis in the paper to the square-root, or affine, class of stochastic volatility models. This particular class of models arguably constitutes the leading case in the literature, but the method is general. In particular Barndorff-Nielsen and Shephard (1999) present analytical expressions for the moments of the integrated volatility for a general class of continuous

[1] In a related context, Brandt and Santa-Clara (1999) and Ledoit and Santa-Clara (1999) suggest using Black-Scholes implied volatilities for short-lived at-the-money options to estimate the instantaneous volatility. In practice, this implicitly assumes that the volatility is constant over the remaining life of the option, and that the volatility risk is not priced.

[2] Alternatively, it is possible to extract information about the forward integrated volatility from the high-low range of the discretely sampled data as in Gallant et al. (1999). Also, Alizadeh et al. (1999) have recently proposed using the high-low range as a volatility proxy in a Gaussian quasi-maximum likelihood estimation procedure for a simple stochastic volatility model.

time stochastic volatility models, in which the instantaneous variance is defined by the sum of multiple Ornstein-Uhlenbeck processes, each of which is driven by a homogeneous Levy process.

The rest of the paper is organized as follows. The next section demonstrates how the population moments for the integrated volatility 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 highlight, that the method works very well in empirically realistic finite sample settings, and that the efficiency of the parameter estimates compares favorably to that of a non-feasible QML procedure treating the instantaneous volatility as observable. The statistical inference concerning the true values of the individual parameters and the overall fit of the model are generally also very reliable. The only caveat is a negligible upward bias in the estimates of the variance-of-variance parameter. This is directly attributable to the measurement error in the quadratic variation as a proxy for the integrated volatility, and we show how a simple adjustment term in the moment conditions is effectively able to eliminate this bias. Section 4 gives the empirical results from applying the new estimation procedure to a set of high-frequency five-minute foreign exchange rates and Japanese equity index returns. Section 5 concludes. Mathematical details regarding the derivation of the moment conditions for the integrated volatility are relegated to a technical appendix.

2 Estimating Stochastic Volatility Diffusion

The basic estimation strategy builds on the usual 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 almost sure convergence of the quadratic variation to the integrated volatility of the process. We begin with a general discussion of the main idea, and then proceed to a concrete illustrative example.

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

$$\begin{aligned} dp_t &= \mu(p_t, V_t, t; \xi)dt + \upsilon(p_t, V_t, t; \xi)dB_t, \\ dV_t &= \kappa(p_t, V_t, t; \xi)dt + \sigma(p_t, V_t, t; \xi)dW_t, \end{aligned} \tag{1}$$

where V_t denotes a vector of latent volatility factors, dB_t and dW_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). Moreover, 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 dW_t through both V_t and $\mathrm{corr}(dB_t, dW_t)$ may be redundant. Also, for concreteness, in the subsequent empirical analysis we will normalize the unit time interval to correspond to one day.

The exact form of the drift function, $\mu(p_t, V_t, t; \xi)$, is generally irrelevant for the consistent estimation of the parameters entering the diffusion functions. Meanwhile, the estimation of these parameters based on discretely sample observations for the p_t process are complicated by the V_t process being latent, and the lack of a closed form expression for the corresponding transition density. As noted in the introduction, this in turn has spurred the development of several alternative computationally demanding estimation procedures. However, by the theory of quadratic variation

$$\lim_{N\to\infty} \sum_{i=1}^{2^N} \left[p_{t+\frac{i}{2^N}(T-t)} - p_{t+\frac{i-1}{2^N}(T-t)} \right]^2 \xrightarrow{a.s.} \int_t^T \upsilon(p_s, V_s, s; \xi) ds \equiv \mathcal{V}_{t,T}, \tag{2}$$

where $\mathcal{V}_{t,T}$ denotes for 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 is effectively observable.[3]

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, in practice continuously sampled observations are not available, so that the integrated volatility is not actually observable. However, 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 validity of this assumption is obviously an empirical question.

[3] Andersen and Bollerslev (1998a) provide simulation evidence in support of this idea, and argue for the practical use of the quadratic variation as a meaningful measure of the ex-post *realized* volatility.

The next section details the derivation of the first two population moments for a particular class of stochastic volatility models. For concreteness, we shall focus on the square-root volatility model, or the single factor affine diffusion, analyzed by Heston (1993) among others. But, the same basic approach employed in the next section could in principle be extended to any multifactor stochastic volatility processes 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 class of stochastic differential equations popularized by Duffie and Kan (1996), and Dai and Singleton (2000), as well as the quadratic stochastic volatility class of models (see, e.g., Kloeden and Platen, 1992).

2.2 Conditional Moments of Integrated Volatility

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

$$\begin{aligned} dp_t &= \mu_t dt + \sqrt{V_t} dB_t, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t, \end{aligned} \tag{3}$$

where V_t is a scalar latent volatility process. While this first-order parameterization is obviously somewhat restrictive, it is nonetheless rich enough to illustrate the general idea, and it has in fact been widely used in the literature. In this parameterization, θ determines the long-run (unconditional) mean, κ is the mean reversion parameter, and σ denotes the local variance (volatility-of-volatility) parameter. For the process to be well defined, the parameters must satisfy: $\theta > 0$ (non-negativity), $\kappa > 0$ (stationary in mean), and $\sigma^2 \leq 2\kappa\theta$ (stationary in volatility). Note that the drift of the asset returns, μ_t, can be any linear or nonlinear function of the state variables, p_t, V_t, or even other unobservable factor(s), without impeding the estimation of the stochastic volatility component.

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

[4] Although the procedure implemented here hinges on the matching of selected population and sample moments for the integrated volatility, in situations where analytical expressions for the population moments $\mathcal{V}_{t,T}$ are not directly available, these could easily be evaluated by simulations, and the underlying parameters estimated by simulated methods of moments (Duffie and Singleton, 1993).

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 result in Cox et al. (1985) that,

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

where α_{T-t} and β_{T-t} are functions of the structural parameters and the horizon of the forecast, $T-t$ (see Appendix A for details). The second step is to express the conditional mean of the integrated volatility as a (linear) function of the point-in-time volatility by interchanging the integration operators

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

where a_{T-t} and b_{T-t} denote other explicit functions of the drift parameters and the sampling interval.

Now, by iteratively substituting the above two results, the conditional mean of the integrated volatility for the one-day horizon, given the finer information set $\mathcal{F}_t$, is readily expressed as (see Appendix A),

$$E\left(\mathcal{V}_{t+1,t+2}|\mathcal{F}_t\right) = \alpha E\left(\mathcal{V}_{t,t+1}|\mathcal{F}_t\right) + \beta,$$

where for notational simplicity we omit the subscript for the daily horizon; i.e. $\alpha \equiv \alpha_1$ and $\beta \equiv \beta_1$. Using the Law of Iterated Expectation, the above relationship can be conditioned on the coarser information set $\mathcal{G}_t$, yielding

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

The first order moments for the multi-period integrated volatility may be derived by similar reasoning.

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 detailed derivations).

Now combining the conditional variance formula in (8) and the conditional mean formula in (5), we can derive the second moment of the integrated volatility conditional on the finer information set $\mathcal{F}_t$. In particular, for the one-day horizon this takes the form

$$E(\mathcal{V}_{t,t+1}^2|\mathcal{F}_t) = Var(\mathcal{V}_{t,t+1}|\mathcal{F}_t) + [E(\mathcal{V}_{t,t+1}|\mathcal{F}_t)]^2 = a^2V_t^2 + (2ab + A)V_t + (b^2 + B), \tag{9}$$

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, it follows that

$$E[E(\mathcal{V}_{t+1,t+2}^2|\mathcal{F}_t)|\mathcal{G}_t] = E(\mathcal{V}_{t+1,t+2}^2|\mathcal{G}_t) = HE(\mathcal{V}_{t,t+1}^2|\mathcal{G}_t) + IE(\mathcal{V}_{t,t+1}|\mathcal{G}_t) + J, \tag{10}$$

where the functions H, I, and J are again defined in the Appendix. Corresponding moment conditions for the squared multi-period integrated volatility follow by analogous arguments.

2.2.3 Conditional Moment Restrictions

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

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

By construction $E[f_t(\xi_0)|\mathcal{G}_t] = 0$, 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 distributed asymptotically as a chi-square distribution with three degrees of freedom, 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 equation (11) 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 below, we used a heteroskedasticity and autocorrelation consistent robust covariance matrix estimator with a Bartlett-kernel and a lag length of five (Newey and West, 1987).[5] The next section details the results from a Monte Carlo study designed to investigate the finite sample performance of this particular GMM estimator.

3 Monte Carlo Study

One important aspect in evaluating econometric methods for estimating continuous time process concerns their finite sample performance. With strong temporal dependence and/or conditional heteroskedasticity in the data generating process, asymptotically sound estimators have been shown to exhibit very slow convergence rates (see, e.g., Pritsker, 1998). This section qualifies the small sample efficiency of our GMM estimator, along with the resulting omnibus specification test, and Wald based parameter inference.

3.1 Experimental Design

we presents the results for three benchmark specifications. Scenario A ($\kappa = 0.03$, $\theta = 0.25$, $\sigma = 0.10$) features a highly persistent volatility process (nearly unit-root); Scenario B ($\kappa = 0.10$, $\theta = 0.25$, $\sigma = 0.10$) has a more stationary 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 ($\sigma^2 > 2\kappa\theta$).

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

In simulating the data, we utilize a first order Euler scheme with 82 artificial "five-minute" intervals per day, further partitioning each five-minute interval into 10 segments.[6] The quadratic variation formula (2) is employed to approximate the integrated volatility series. To check the standard "long-span" asymptotics, the econometric sample sizes are chosen as T = 1,000 and T = 4,000. Since the true "continuous time" record is known inside the simulations, we compare the GMM estimator using the "five-minute" quadratic variation with the corresponding non-feasible estimator based on the true integrated volatility. Lastly, we also compare the results for the GMM estimator with a QML estimator based on the "daily" point-in-time volatility assuming the process to be Gaussian.[7] Of cause, this latter estimator is not feasible in practice either. The results are summarized in Table 1 and Figures 1 and 2.

3.2 Parameter Estimate and Efficiency

First, the finite sample results not only corroborate the moment conditions derived earlier for the integrated volatility, but also indicate that the 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. The root-mean-squared-errors (RMSEs) of the drift parameters, κ and θ, decrease roughly at the rate of $\sqrt{4}$ as the sample size increases from 1,000 to 4,000 "days". Meanwhile, the mean-reversion parameter κ is upward biased, and the long-run mean parameter θ exhibits a small downward bias.

Second, 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 when the sample size goes from 1,000 to 4,000, the rate is not always $\sqrt{4}$. 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. This confirms the theoretical arguments that the diffusion parameter can be estimated exactly with continuous sampling (see, Merton, 1980; Lo, 1988; Nelson, 1992).

Third, when the process is close to a unit-root (Scenario A), the variance parameter seems to converge at a faster rate than $\sqrt{T}$ (Table 1 Panel A and Figure 1). Also when the variance-of-variance parameter is large (Scenario C in Figure 1), the finite sample biases of

[6] Most US financial markets are open between six-and-a-half to seven hours, corresponding to 78-84 five-minute intervals.

[7] This estimator is closely related to the ideas in Fisher and Gilles (1996), who propose a Quasi-Maximum Likelihood estimator for Affine diffusion process, using closed form solutions for the conditional mean and variance.

the drift parameter estimates are larger than for the more stationary case (Scenario B in Figure 1). Basically the GMM estimator is not able to distinguish between a very persistent yet stationary process and a non-stationary, near unit-root process in "small" samples.

Lastly, the GMM estimates of the local variance parameter, σ, are systematically upward biased in all three scenarios. 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. We will investigate this issue further in Section 3.4.

In terms of relative efficiency, the GMM estimator using the "five-minute" realized volatility actually performs better than the non-feasible QML estimator using the unobservable point-in-time volatility for the drift parameters in all three scenarios, and better for the variance parameter in all but the stationary scenario. The middle rows in Table 1 suggest that the RMSEs of the GMM estimator using the true integrated volatility process are much smaller than those of the QML estimator. However, going to the "continuous time" limit does not necessarily improve the efficiencies of the GMM drift parameters, but it does increase the convergence rate of the diffusion parameter.

3.3 Statistical Inference

In practice, inference concerning the individual model parameters and the overall specification of the model will have to 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 Figure 1 clearly indicate that the GMM estimator based on the "five-minute" quadratic variation is close to normal for both 1,000 and 4,000 "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. This is consistent with the earlier explanation of the non-dissipating measurement error in the second moment condition.[8]

Turning to Figure 2 and the GMM tests of overidentifying restrictions, it follows that except for the near unit-root case (Scenario A in Figure 2), the test performs very well. Moreover, the slight over-rejection and under-rejection biases largely vanishes as the sample size increases from 1,000 to 4,000.[9]

[8]The corresponding t-tests for the non-feasible QML estimator based on the point-in-time volatility are generally much 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.

[9]Overrejection biases of GMM omnibus tests are widely reported in the literature (Andersen and Sørenson,

3.4 Measurement Error Adjustment

By construction the quadratic variation based on the simulated "five-minute" returns provides an unbiased estimate for the true integrated volatility. At the same time, the squared quadratic variation for any fixed sampling interval yields a biased estimate of the true squared integrated volatility. Consequently, while the linear expectations operator washes out the measurement errors in the first conditional moment and the corresponding two augmented moments in equation (11), the three moment conditions involving the squared quadratic variation will entail a non-zero measurement error.[10] Although the exact form of the measurement error is not known, it follows 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). In order to conveniently approximate this 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, $\mathcal{V}^2_{t+1,t+2}$, by $\mathcal{V}^2_{t+1,t+2} + \gamma$.

Not surprisingly, from the results reported in Table 2 and Figure 3, the parameter estimates for the two drift parameters and the corresponding t-statistics are largely unaffected by the estimation of this additional nuisance parameter. More important, the pervasive finite sample biases for the local variance parameter estimates have completely disappeared.[11] Moreover, the rejection frequencies for the GMM specification test for the overidentifying restrictions appear marginally closer to their nominal sizes. Thus, all in all, the inclusion of a simple additive correction term for the squared quadratic variation has effectively eliminated the only notable statistical bias in the procedure. Of course, it is possible that more advanced measurement error adjustment procedures could result in further improvements, especially for more complicated models. However, for the square-root volatility diffusion in equation (3), the GMM estimation procedure proposed here works very well in realistic fixed-interval finite sample settings.

1996; Hansen et al., 1996), whereas underrejection biases often occur when lag instruments are used to form the moment conditions (Tauchen, 1986).

[10] Andersen and Bollerslev (1998a) provide some limited simulation evidence on the size of this measurement error as a function of the sampling frequency. Andersen et al. (2000c) 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.

[11] Meanwhile, the RMSEs for σ have increased somewhat.

4 Empirical Illustration

This section provides an empirical illustration of the new estimation procedure using actual high-frequency date. For ex positional purposes, we will focus on the estimation results for the simple scalar affine diffusion analyzed in the previous two sections. To illustrate the applicability of the procedure across different markets and institutional arrangements, we present the results for two separate data sets: spot foreign exchange rates, and Japanese equity index returns. In line with the simulations in the preceding section, we partition the trading day for each of the markets into five-minute intervals, incorporating an additive nuisance parameter to correct the inherent measurement error in the resulting five-minute quadratic variation measures.

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 Deutsche Mark/U.S. Dollar (DM/$) and Japanese Yen/U.S. Dollar (Yen/$) 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. (1999b), we are left with a total of 2,445 trading days, each of which consists of 288 five-minute returns over the 24-hour trading cycle.

The intraday data for the Nikkei 225 composite stock market index were provided by Nihon Keizei Shimbun Inc. The five-minute returns for the Nikkei 225 covers the period from January 2, 1994 through December 31, 1997. Excluding days on which the Japanese equity market was closed results in a total of 984 trading days. The Tokyo Stock Exchange opens at 9:00 a.m., closes for lunch from 11:00 to 12:30, and closes for the day at 15:00 p.m.. Omitting the first five-minute interval of the day associated with the special Itayose batched trading process at the opening, leaves us with 53 five-minute returns per day. In contrast to the very actively traded foreign exchange rates, the five-minute returns for the Nikkei 225 cash index is plagued by important non-synchronous trading effects (see, e.g., Lo and MacKinlay, 1990; Chan et al., 1991, for a discussion of non-synchronous trading effects in equity index returns). While the resulting autocorrelation in the high-frequency returns does not formally affect the continuous record asymptotics underlying the GMM estimator, any mean dependencies in the discretely sampled returns will systematically bias the quadratic variation as an estimate for the true latent integrated volatility. In order to minimize this bias we pre-whitened the returns by a first order autoregressive model, treating the residuals from this model as the actual five-minute return series.[12] For a more detailed discussion of the pertinent institutional arrangements and the pre-whitened five-minute Nikkei 225 returns, we refer to Andersen et al. (2000a).

Next we transform the three five-minute return series into daily time series of integrated volatilities, as approximated by the quadratic variations in equation (2). Table 3 provides the standard set of summary statistics for each series. The means of the integrated volatility for the two exchange rates imply an annualized standard deviation of approximately 11.5 percent, whereas the annualized volatility for the Japanese stock market equals 14.6 percent.[13] The standard deviations of the integrated volatilities are close to the mean for all three markets. The higher order moments indicate extremely heavy tails and, most notably in the case of the Yen/$ spot exchange rate, important skewness to the right. These distributional features are confirmed by visual inspection of the time series plots in Figure 5. Each of the panels also reveals a high degree of serial correlation in the integrated volatility series. The next subsection presents the estimation results from the stochastic volatility model in equation (3) explicitly designed to capture this volatility clustering effect.

4.2 Estimation Results

Before proceeding to the actual estimation results, we caution that the scalar volatility diffusion in equation (3) is too simplistic to fully account for the complex dynamic dependencies in the high-frequency return series. In particular, there are sound theoretical reasons to expect there to be at least two factors affecting the exchange rates (Bansal, 1997). Also, a number of recent studies have argued for the empirical relevance of including multiple factors and/or jump components when modeling equity index returns (e.g., Chacko and Viceira, 1999; Andersen et al., 1999a; Chernov et al., 1999). Moreover, the model in equation (3) does not incorporate the strong periodic dependencies in the volatility within the day documented in several recent studies (see e.g., Andersen and Bollerslev, 1997).[14] In spite of these deficiencies, we feel that the square-root stochastic volatility model is rich enough

[12]The estimated AR(1) coefficient for the raw Nikkei 225 five-minute returns equals 0.1429. Details concerning the model estimates based on the un-adjusted five-minute Nikkei 225 returns are available on request. The resulting daily integrated volatility series is slightly smoother, and the parameter estimates are marginally lower in level, persistence, and variance.

[13]The annualized standard deviation is obtained by multiplying the mean of the daily integrated volatility by 250 and taking the square-root.

[14]Interestingly, the forecasting result in Andersen and Bollerslev (1998b) and Andersen et al. (2000a) suggest that the influences of the intraday periodicities are effectively eliminated in the daily integrated volatility measures utilized in the GMM estimation.

to offer a first meaningful empirical illustration of the applicability of the new estimation procedure.

The parameter estimates for the three series are reported in Table 4.[15] With the exception of the slightly higher values for σ, the estimates are almost identical to the ones reported here. As expected the estimates for the long-run means, or θ, are all fairly close to the sample means for the three integrated volatility series reported in Table 3. Also, not surprisingly, the estimates of the variance-of-variance parameter, or σ, have the largest standard errors among all of the parameters. Meanwhile, the estimated mean reversion parameters, κ, are on the high side relative to the values reported in the extant literature using more complicated discrete time ARCH and stochastic volatility type formulations. Even though the GMM omnibus test only rejects the model for the DM/$ exchange rate, the one-factor model is obviously an oversimplification of the true dynamic dependencies for all three markets. However, from an overall perspective, the estimation results in Table 4 are generally in line with the simulation evidence reported in the previous section, and clearly suggest that the new estimation procedure could effectively be employed in the empirical estimation of more complicated continuous time diffusions.

5 Concluding Remarks

Exploiting closed form analytic expressions for the conditional moments of integrated volatility coupled with highly accurate empirical quadratic variation measures constructed from high-frequency data, we proposed a new class of GMM-type estimators for stochastic volatility diffusions. In contrast to other computationally demanding estimation procedures routinely used in the literature, such as the simulation based EMM and MCMC methods, 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 inferences in realistic finite samples. In implementing the new estimator with actual five-minute rates of return, our results confirm prior evidence in the literature concerning the existence of strong volatility clustering at the inter-daily level.

It would be interesting to extend the estimator developed here to more complicated continuous time jump-diffusion and multi-factor diffusion processes. More ambitious empirical applications might also entail the estimation of multivariate diffusions, which in turn would

[15] Details regarding the parameter estimates without the additive measurement error term are available upon request.

require vector versions of the integrated volatility and quadratic variation measures exploited here. Another interesting extension, would be to use the distributional features of the integrated volatility in pricing financial options, although this would necessitate additional assumptions about the price of volatility risk. We leave further work along these lines for future research.

A Conditional Moments of Integrated Volatilities

A.1 Conditional Mean of Integrated Volatility

Because of the linear drift specification of the stochastic volatility, the conditional mean of the integrated volatility can be shown as a linear function of the point-in-time volatility

$$\begin{aligned} E\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right) &= E\left(\left.\int_t^T V_s ds\right| \mathcal{F}_t\right) \\ &= \int_t^T E(V_s|\mathcal{F}_t)ds \\ &= \int_t^T \left[V_t e^{-\kappa(s-t)} + \theta\left(1-e^{-\kappa(s-t)}\right)\right] ds \\ &= V_t\frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right) + \theta(T-t) - \frac{\theta}{\kappa}\left(1-e^{-\kappa(T-t)}\right) \\ &= a_{T-t}V_t + b_{T-t}, \end{aligned} \tag{A1}$$

where a_{T-t} and b_{T-t} are functions of the drift parameters and the time difference $(T-t)$. For notational simplicity we denote the parameters for the daily horizon, or $T-t=1$, by $a \equiv \frac{1}{\kappa}\left(1-e^{-\kappa}\right)$ and $b \equiv \theta - \frac{\theta}{\kappa}\left(1-e^{-\kappa}\right)$. The above derivation explicitly uses the conditional mean of the point-in-time volatility

$$E(V_T|\mathcal{F}_t) = V_t e^{-\kappa(T-t)} + \theta\left(1-e^{-\kappa(T-t)}\right) = \alpha_{T-t}V_t + \beta_{T-t}, \tag{A2}$$

where α_{T-t} and β_{T-t} are also functions of the drift parameters and the time difference $(T-t)$. Again for $T-t=1$, we define $\alpha \equiv e^{-\kappa}$ and $\beta \equiv \theta\left(1-e^{-\kappa}\right)$.

Focusing on the one-day horizon with $E\left(\mathcal{V}_{t+1,t+2}|\mathcal{F}_{t+1}\right) = aV_{t+1} + b$, and $E(V_{t+1}|\mathcal{F}_t) = \alpha V_t + \beta$, it follows that

$$\begin{aligned} E\left[E\left(\mathcal{V}_{t+1,t+2}|\mathcal{F}_{t+1}\right)|\mathcal{F}_t\right] &= aE(V_{t+1}|\mathcal{F}_t) + b \\ &= a(\alpha V_t + \beta) + b \\ &= \alpha[E\left(\mathcal{V}_{t,t+1}|\mathcal{F}_t\right) - b] + a\beta + b, \end{aligned}$$

which simplifies as

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

Finally, by the Law of Iterated Expectations,

$$E\left[E\left(\mathcal{V}_{t+1,t+2}|\mathcal{F}_t\right)|\mathcal{G}_t\right] = E\left(\mathcal{V}_{t+1,t+2}|\mathcal{G}_t\right) = \alpha E\left(\mathcal{V}_{t,t+1}|\mathcal{G}_t\right) + \beta. \tag{A3}$$

A.2 Conditional Variance of Integrated Volatility

By definition $\mathcal{V}_{t,T} = \int_t^T V_s ds$, and from equation (A1) $E\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right) = a_{T-t}V_t + b_{T-t}$. The stochastic differential equation (SDE) for $E\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right)$ may therefore be generated as a function of V_t by applying Itô's formula to the affine diffusion in equation (3),[16]

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

which may be further simplified as

$$dE\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right) = -V_t dt + a_{T-t}\sigma\sqrt{V_t}dW_t. \tag{A4}$$

Now fix the upper limit T, and let the lower limit t be time varying. The Itô integral implied by SDE (A4) then takes the form

$$E\left(\mathcal{V}_{T,T}|\mathcal{F}_T\right) = E\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right) + \int_t^T (-V_s)ds + \int_t^T a_{T-s}\sigma\sqrt{V_s}dW_s.$$

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

$$\mathcal{V}_{t,T} - E\left(\mathcal{V}_{t,T}|\mathcal{F}_t\right) = \int_t^T a_{T-s}\sigma\sqrt{V_s}dW_s.$$

Using standard arguments in stochastic calculus, it follows from the substitution of equation (A2) that

$$\begin{aligned}
Var(\mathcal{V}_{t,T}|\mathcal{F}_t) &= E[(\mathcal{V}_{t,T} - E\left(\mathcal{V}_{t,T}|\,\mathcal{F}_t\right))^2|\mathcal{F}_t] \\
&= E\left\{\left[\int_t^T a_{T-s}\sigma\sqrt{V_s}dW_s\right]^2\middle|\mathcal{F}_t\right\} \\
&= \int_t^T a_{T-s}^2\sigma^2 E(V_s|\mathcal{F}_t)ds \\
&= \int_t^T a_{T-s}^2\sigma^2[\alpha_{s-t}V_t + \beta_{s-t}]ds \\
&= A_{T-t}V_t + B_{T-t},
\end{aligned} \tag{A5}$$

where

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

[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

$$df(V_t,t,T) = [f_V(V_t,t,T)\mu(V_t,t) + f_t(V_t,t,T) + \frac{1}{2}f_{VV}(V_t,t,T)\upsilon^2(V_t,t)]dt + f_V(V_t,t,T)\upsilon(V_t,t)dW_t,$$

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

$$
\begin{aligned}
B_{T-t} &= \frac{\sigma^2}{\kappa^2}\left[\theta(T-t)\left(1+2e^{-\kappa(T-t)}\right)-\frac{3\theta}{\kappa}\left(1-e^{-\kappa(T-t)}\right)+\frac{\theta}{2\kappa}\left(1-e^{-\kappa(T-t)}\right)^2\right] \\
&= \frac{\sigma^2}{\kappa^2}\left[\theta(T-t)\left(1+2e^{-\kappa(T-t)}\right)+\frac{\theta}{2\kappa}\left(e^{-\kappa(T-t)}+5\right)\left(e^{-\kappa(T-t)}-1\right)\right].
\end{aligned}
$$

In particular, the conditional variance of the integrated volatility is a linear function of the point-in-time volatility. It follows also from Cox et al. (1985) and equation (A2) above that,

$$
\begin{aligned}
E(V_T^2|\mathcal{F}_t) &= Var(V_T|\mathcal{F}_t)+[E(V_T|\mathcal{F}_t)]^2 \\
&= V_t\frac{\sigma^2}{\kappa}\left(e^{-\kappa(T-t)}-e^{-2\kappa(T-t)}\right)+\frac{\sigma^2\theta}{2\kappa}\left(1-e^{-\kappa(T-t)}\right)^2+[\alpha_{T-t}V_t+\beta_{T-t}]^2 \\
&= C_{T-t}V_t+D_{T-t}+\alpha_{T-t}^2V_t^2+\beta_{T-t}^2+2\alpha_{T-t}\beta_{T-t}V_t \\
&= \alpha_{T-t}^2V_t^2+[C_{T-t}+2\alpha_{T-t}\beta_{T-t}]V_t+[D_{T-t}+\beta_{T-t}^2] \qquad \text{(A6)}
\end{aligned}
$$

where

$$
\begin{aligned}
C_{T-t} &= \frac{\sigma^2}{\kappa}\left(e^{-\kappa(T-t)}-e^{-2\kappa(T-t)}\right), \\
D_{T-t} &= \frac{\sigma^2\theta}{2\kappa}\left(1-e^{-\kappa(T-t)}\right)^2.
\end{aligned}
$$

Focusing on the one-day horizon, the conditional variance formula (A5), $Var(\mathcal{V}_{t,t+1}|\mathcal{F}_t) = AV_t + B$, and the corresponding one-day conditional mean formula (A1), $E\left(\mathcal{V}_{t,t+1}|\mathcal{F}_t\right) = aV_t + b$, implies that

$$
E(\mathcal{V}_{t,t+1}^2|\mathcal{F}_t)=Var(\mathcal{V}_{t,t+1}|\mathcal{F}_t)+[E(\mathcal{V}_{t,t+1}|\mathcal{F}_t)]^2=a^2V_t^2+(2ab+A)V_t+(b^2+B). \qquad \text{(A7)}
$$

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

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

Now substitute $E(V_{t+1}|\mathcal{F}_t)$ by (A2) and $E(V_{t+1}^2|\mathcal{F}_t)$ by (A6), and reversely substitute out V_t^2 by (A7) and V_t by (A1), it is clear that

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

$$
\begin{aligned}
& +[a^2(D+\beta^2)+\beta(2ab+A)+(b^2+B)] \\
= \ & \alpha^2 E(\mathcal{V}^2_{t,t+1}|\mathcal{F}_t) \\
& +[a^2(C+2\alpha\beta)+(\alpha-\alpha^2)(2ab+A)]\frac{1}{a}[E(\mathcal{V}_{t,t+1}|\mathcal{F}_t)-b] \\
& +[a^2(D+\beta^2)+\beta(2ab+A)+(1-\alpha^2)(b^2+B)] \\
= \ & \alpha^2 E(\mathcal{V}^2_{t,t+1}|\mathcal{F}_t) \\
& +\frac{1}{a}[a^2(C+2\alpha\beta)+(\alpha-\alpha^2)(2ab+A)]E(\mathcal{V}_{t,t+1}|\mathcal{F}_t) \\
& -\frac{b}{a}[a^2(C+2\alpha\beta)+(\alpha-\alpha^2)(2ab+A)] \\
& +[a^2(D+\beta^2)+\beta(2ab+A)+(1-\alpha^2)(b^2+B)] \qquad \text{(A8)}
\end{aligned}
$$

Lastly, applying the Law of Iterated Expectations to (A8) and changing the information set, we have

$$
\begin{aligned}
E[E(\mathcal{V}^2_{t+1,t+2}|\mathcal{F}_t)|\mathcal{G}_t] &= E(\mathcal{V}^2_{t+1,t+2}|\mathcal{G}_t) \\
&= \alpha^2 E(\mathcal{V}^2_{t,t+1}|\mathcal{G}_t) \\
&\quad +\frac{1}{a}[a^2(C+2\alpha\beta)+(\alpha-\alpha^2)(2ab+A)]E(\mathcal{V}_{t,t+1}|\mathcal{G}_t) \\
&\quad -\frac{b}{a}[a^2(C+2\alpha\beta)+(\alpha-\alpha^2)(2ab+A)] \\
&\quad +[a^2(D+\beta^2)+\beta(2ab+A)+(1-\alpha^2)(b^2+B)] \\
&= HE(\mathcal{V}^2_{t,t+1}|\mathcal{G}_t)+IE(\mathcal{V}_{t,t+1}|\mathcal{G}_t)+J, \qquad \text{(A9)}
\end{aligned}
$$

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

References

Aït-Sahalia, Yacine (1996), "Nonparametric Pricing of Interest Rate Derivatives," *Econometrica*, vol. 64, 527–560.

Aït-Sahalia, Yacine (1998), "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach," *Working Paper*, Department of Economics, Princeton University.

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold (1999), "Range-Based Estimation of Stochastic Volatility Models," *Working Paper*, Department of Finance, NYU.

Andersen, Torben G., Luca Benzoni, and Jesper Lund (1999a), "Estimating Jump-Diffusions for Equity Returns," *Working Paper*, Kellogg Graduate School of Management, Northwestern University.

Andersen, Torben G. and Tim Bollerslev (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, vol. 4, 115–158.

Andersen, Torben G. and Tim Bollerslev (1998a), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forcasts," *International Economic Review*, vol. 39, 885–905.

Andersen, Torben G. and Tim Bollerslev (1998b), "DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies," *Journal of Finance*, vol. 53, 219–265.

Andersen, Torben G., Tim Bollerslev, and Jun Cai (2000a), "Intraday and Interday Volatility in the Nikkei 225 Index," *Journal of International Financial Markets, Institutions & Money*, forthcoming.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens (2000b), "The Distribution of Stock Return Volatility," *Working Paper*, Department of Economics, Duke University.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys (1999b), "The Distribution of Exchange Rate Volatility," *NBER Working Paper*, No. 6961.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys (2000c), "Microstructure Bias and Volatility Signatures," *Work in Progress*, Department of Economics, Duke University.

Andersen, Torben G., Hyung-Jin Chung, and Bent E. Sørenson (1999c), "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study," *Journal of Econometrics*, vol. 91, 61–87.

Andersen, Torben G. and Bent E. Sørenson (1996), "GMM estimation of a stochastic volatility model: a monte Carlo study," *Journal of Business and Economic Statistics*, vol. 14, 328–352.

Bai, Xuezheng, Jeffrey R. Russell, and George C. Tiao (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.

Bandi, Federico M. and Peter C. B. Phillips (1999), "Econometric Estimation of Diffusion Models," *Working Paper*, Department of Economics, Yale University.

Bansal, Ravi (1997), "An Exploration of the Forward Premium Puzzle in Currency Market," *The Review of Financial Studies*, vol. 10, 369–403.

Barndorff-Nielsen, Ole and Neil Shephard (1999), "Non-Gaussian OU Based Models and Some of Their Uses in Financial Economics," *Working Paper*, Nuffield College, Oxford University.

Bollerslev, Tim, Robert F. Engle, and Daniel B. Nelson (1994), "ARCH Models," in "Handbook of Econometrics," (edited by Engle, Robert F. and Daniel L. McFadden), vol. IV, Amsterdam: North-Holland.

Brandt, Michael W. and Pedro Santa-Clara (1999), "Simulated Likelihood Estimation of Multivariate Diffusions with an Application to Interest Rates and Exchange Rates with Stochastic Volatility," *Working Paper*, Wharton School, University of Pennsylvania.

Chacko, George and Luis M. Viceira (1999), "Spectral GMM Estimation of Continuous-Time Processes," *Working Paper*, Graduate School of Business, Harvard University.

Chan, Kalok, K. C. Chan, and G. Andrew Karolyi (1991), "Intraday Volatility in the Stock Index and Stock Index Futures Markets," *Review of Financial Studies*, vol. 4, 1161–1187.

Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels, and George Tauchen (1999), "A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation," *Working Paper*, Department of Economics, Duke University.

Conley, Tim, Lars Peter Hansen, Erzo Luttmer, and Jose Scheinkman (1997), "Short Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, vol. 10, 525–578.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica*, vol. 53, 385–407.

Dai, Qiang and Kenneth J. Singleton (2000), "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, forthcoming.

Duffie, Darrell and Rui Kan (1996), "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, vol. 6, 379–406.

Duffie, Darrell and Kenneth Singleton (1993), "Simulated moments estimation of Markov models of asset prices," *Econometrica*, vol. 61, 929–52.

Elerian, Ola, Siddhartha Chib, and Neil Shephard (1998), "Likelihood Inference for Discretely Observed Non-Linear Diffusions," *Working Paper*, Nuffield College, Oxford University.

Engle, Robert F. and Gary G. J. Lee (1997), "Estimating Diffusion Models of Stochastic Volatility," in "Modeling Stock Market Volatility: Bridging the Gap to Continuous Time," (edited by Rossi, Peter E.), Academic Press, New York.

Eraker, Bjorn (1998), "MCMC Analysis of Diffusion Models with Application to Finance," *Working Paper*, Graduate School of Business, University of Chicago.

Fisher, Mark and Christian Gilles (1996), "Estimating Exponential Affine Models of the Term Structure," *Working Paper*.

Gallant, A. Ronald, Chien-Te Hsu, and George Tauchen (1999), "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance," *Review of Economics and Statistics*, vol. 81, 617–631.

Gallant, A. Ronald and Jonathan R. Long (1997), "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square," *Biometrika*, vol. 84, 125–141.

Gallant, A. Ronald and George E. Tauchen (1996), "Which Moment to Match?" *Econometric Theory*, vol. 12, 657–681.

Ghysels, Eric, Andrew Harvey, and Eric Renault (1996), "Stochastic Volatility," in "Handbook of Statistics Vol 14., Statistical Method in Finance," (edited by Maddala, G. S.), Amsterdam: North-Holland.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993), "Indirect Inference," *Journal of Applied Econometrics*, vol. 8, s85–s118.

Hansen, Lars Peter (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, vol. 50, 1029–1054.

Hansen, Lars Peter (1985), "A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators," *Journal of Econometrics*, vol. 30, 203–238.

Hansen, Lars Peter, John Heaton, and Amir Yaron (1996), "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics*, vol. 14, 262–280.

Hansen, Lars Peter, John C. Heaton, and Masao Ogaki (1988), "Efficiency Bounds Implied by Multiperiod Conditional Moment Restrictions," *Journal of the American Statistical Association*, vol. 83, 863–871.

Hansen, Lars Peter and Jose Alexandre Scheinkman (1995), "Back to the Future: Generalized Moment Implications for Continuous Time Markov Process," *Econometrica*, vol. 63, 767–804.

Heston, Steven (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Finanacial Studies*, vol. 6, 327–343.

Hull, John and Alan White (1987), "The pricing of Options on Assets with Stochastic Volatility," *Journal of Finance*, vol. 42, 381–340.

Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, vol. 12, 371–389.

Karatzas, Ioannis and Steven E. Shreve (1997), *Brownian Motion and Stochastic Calculus*, Springer-Verlag.

Kim, S., Neil Shephard, and Siddhartha Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, vol. 65, 361–394.

Kloeden, Perter E. and Eckhard Platen (1992), *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics, Springer-Verlag.

Ledoit, Olivier and Pedro Santa-Clara (1999), "Relative Pricing of Options with Stochastic Volatility," *Working Paper*, Anderson Graduate School of Management, UCLA.

Lo, Andrew W. (1988), "Maximum Likelihood Estimation of Generalized Itô Process with Discretely Sampled Data," *Econometric Theory*, vol. 4, 231–247.

Lo, Andrew W. and A. Craig MacKinlay (1990), "An Econometric Analysis of Nonsynchronous Trading," *Journal of Econometrics*, vol. 45, 181–212.

Merton, Robert C. (1980), "On Estimating the Expected Return on the Market," *Journal of Financial Economics*, vol. 8, 323–361.

Nelson, Daniel B. (1992), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model," *Journal of Econometrics*, vol. 52, 61–90.

Nelson, Daniel B. and Dean P. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH Models," *Econometrica*, vol. 62, 1–41.

Newey, Whitney K. and Kenneth D. West (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, vol. 55, 703–708.

Pritsker, Matt G. (1998), "Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models," *Review of Financial Studies*, vol. 11, 449–487.

Protter, Philip (1992), *Stochastic Integration and Differential Equations: A New Approach*, Springer-Verlag.

Singleton, Kenneth (1999), "Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function," *Working Paper*, Graduate School of Business, Stanford University.

Stanton, Richard (1997), "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, vol. 52, 1973–2002.

Tauchen, George (1986), "Statistical properties of generalized methods-of-moments estimators of structural parameters obtained from financial market data," *Journal of Business and Economic Statistics*, vol. 4, 397–416.

B Tables and Figures

Table 1

Monte Carlo Experiment

Panel A

True Value	Mean		Median		RMSE	
GMM with Quadratic Variation from High-Frequency Return						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.03$	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						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.03$	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						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.03$	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

Note: 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 equation (3). The total number of Monte Carlo replications is 1,000.

Table 1 cont.

Panel B

True Value	Mean		Median		RMSE	
GMM with Quadratic Variation from High-Frequency Return						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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						
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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

Table 1 cont.

Panel C

True Value	Mean		Median		RMSE	
	GMM with Quadratic Variation from High-Frequency Return					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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

Table 2

Monte Carlo Experiment with Measurement Error Correction

True Value	Mean		Median		RMSE	
	Scenario A: GMM with Quadratic Variation					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.03$	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					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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					
	T = 1000	T = 4000	T = 1000	T = 4000	T = 1000	T = 4000
$\kappa = 0.10$	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 1,000 Monte Carlo replications.

Table 3

Summary Statistics for Daily Integrated Volatility

Statistics	DM/$ Rate	Yen/$ Rate	Nikkei 225
Mean	0.5290	0.5383	0.8511
Std. Dev.	0.4839	0.5217	0.7757
Skewness	3.7083	5.5713	3.0203
Kurtosis	24.0505	66.6545	18.1780
Minimum	0.0517	0.0280	0.0309
5% Quant.	0.1384	0.1382	0.1494
25% Quant.	0.2542	0.2533	0.3681
Medium	0.3990	0.4008	0.6479
75% Quant.	0.6252	0.6317	1.0782
95% Quant.	1.3450	1.3598	2.2491
Maximum	5.2453	10.0971	7.5651
Num. of Obs.	2445	2445	984

Note: The daily integrated volatilities are approximated by the quadratic variations constructed from five-minute returns.

Table 4

GMM Estimation of Stochastic Volatility Model

Parameter	DM/$ Rate	Yen/$ Rate	Nikkei 225
Mean Reversion κ	0.1464	0.2472	0.1236
(Standard Error)	(0.0387)	(0.0463)	(0.0492)
Long-run Mean θ	0.5172	0.5190	0.8312
(Standard Error)	(0.0342)	(0.0240)	(0.0950)
Local Variance σ	0.5789	0.4242	0.1909
(Standard Error)	(0.0580)	(0.1804)	(0.3992)
GMM Specification Test			
Chi-Square (2)	12.1476	3.6182	0.8040
p-Value	0.0023	0.1638	0.6690

Note: The GMM estimator and the specification test are described in Section 2. The daily integrated volatilities are approximated by the quadratic variations from five-minute returns. The variance-covariance matrix is estimated using a Newey-West weighting scheme with a lag-length of five.

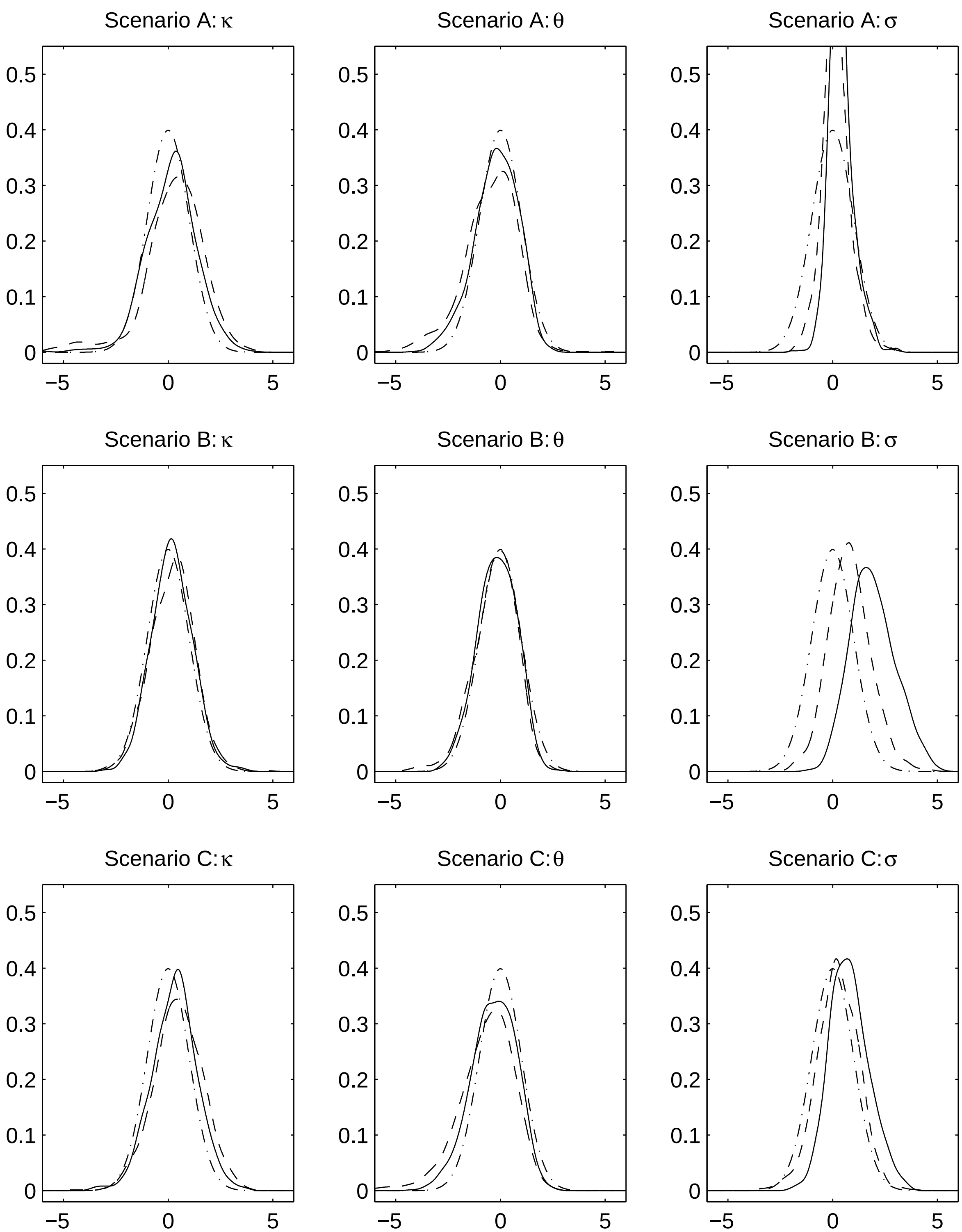

Figure 1: t-test Distributions. "- - -" t-statistics with 1000 observations; "—" t-statistics with 4000 observations; "-.-.-" Normal (0,1) reference density.

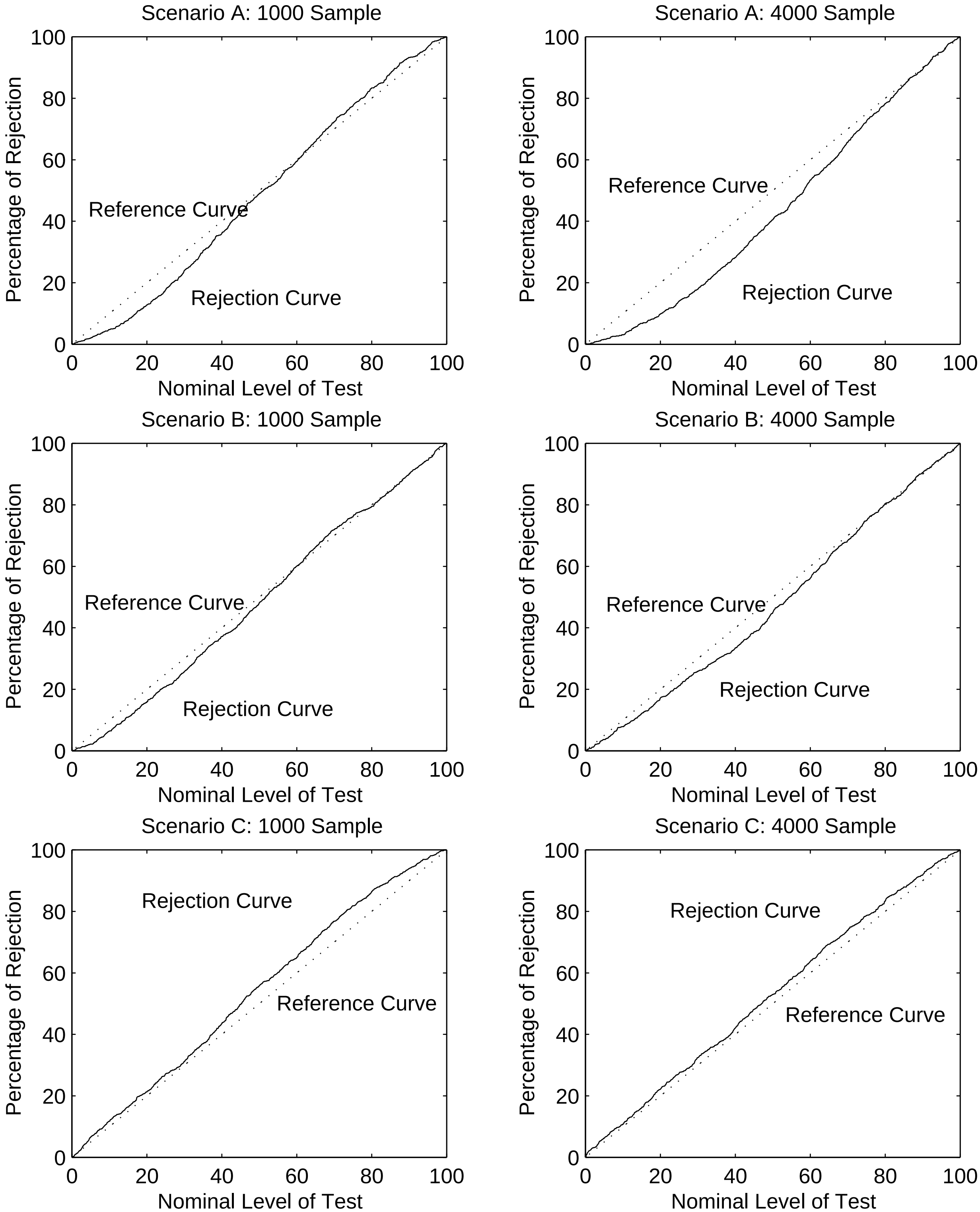

Figure 2: GMM Specification Test of Overidentifying Restrictions.

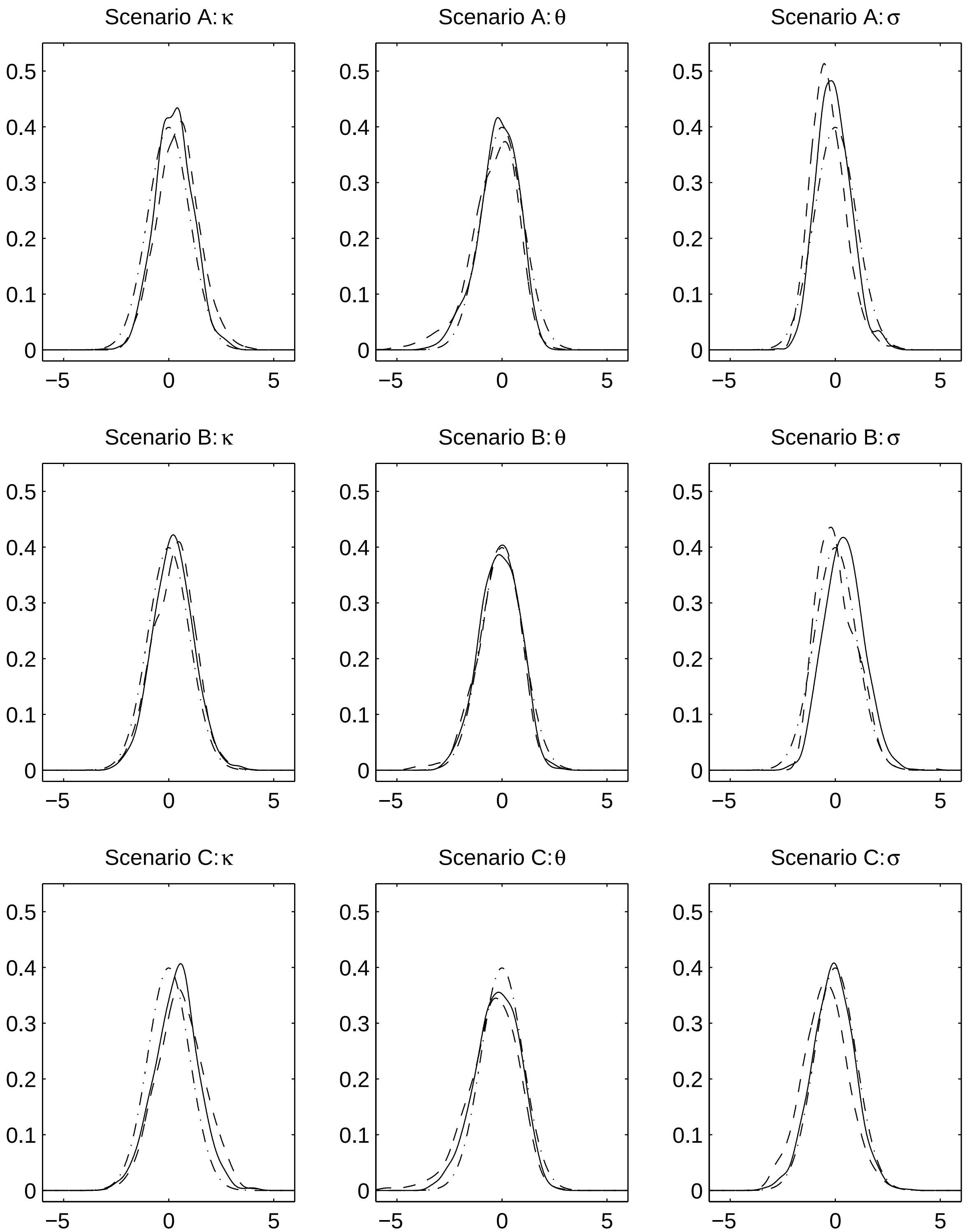

Figure 3: t-test Distributions with Measurement Error Correction. "- - -" t-statistics with 1000 observations; "—–" t-statistics with 4000 observations; "-.-.-" Normal (0,1) reference density.

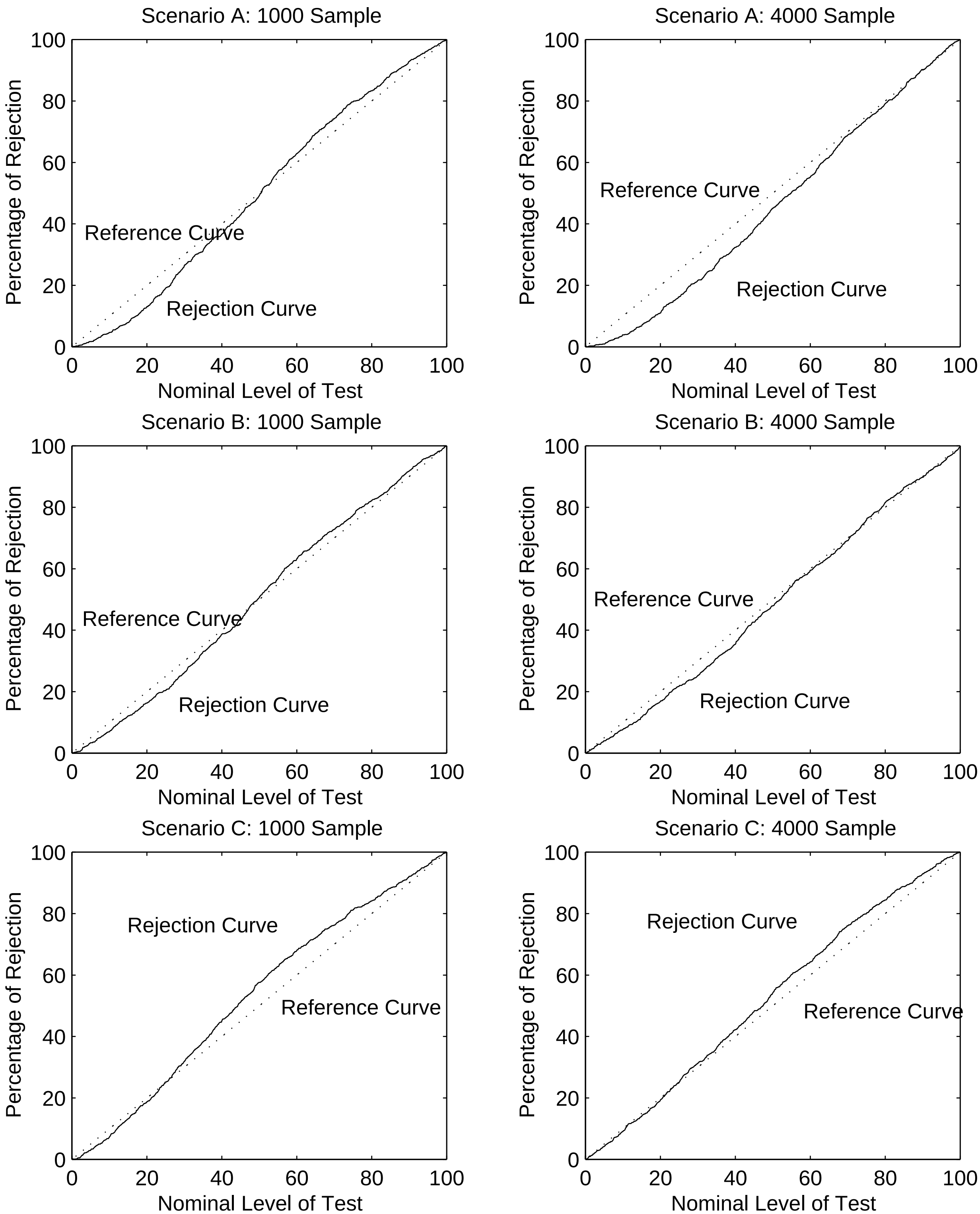

Figure 4: GMM Specification Test of Overidentifying Restrictions with Measurement Error Correction.

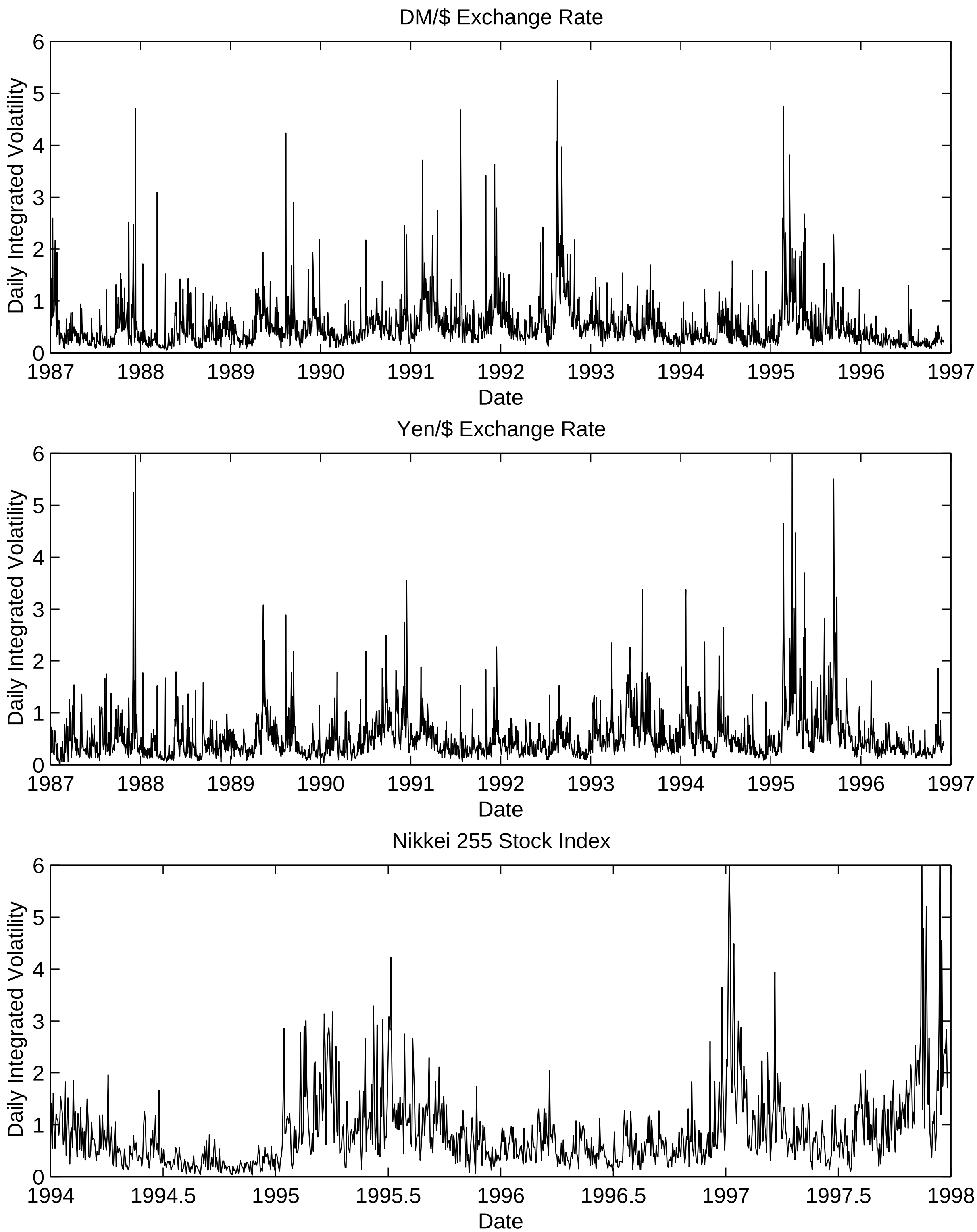

Figure 5: Daily Integrated Volatility on Financial Markets.