GMM Estimation for Long Memory Latent Variable Volatility and Duration Models

Willa Chen and Rohit S. Deo*
Texas A&M University and New York University

January 14, 2005

Abstract: We study the rate of convergence of moment conditions that have been commonly used in the literature for Generalised Method of Moments (GMM) estimation of short memory latent variable volatility models. We show that when the latent variable possesses long memory, these moment conditions have an $n^{1/2-d}$ rate of convergence where $0 < d < 0.5$ is the memory parameter. The resulting GMM estimators will thus not be $\sqrt{n}$ consistent. We then provide an alternative set of moment conditions that are $\sqrt{n}$ consistent and asymptotically normal under long memory in the latent variable, thus allowing for $\sqrt{n}$ consistent GMM estimation.

JEL: C4, C5

*Corresponding author. rdeo@stern.nyu.edu, Tel: (212) 998-0469, Stern School of Business, New York University, 8-57 KMEC, 44 W. 4th Street, New York NY 10012, USA.

1 Introduction:

The stochastic volatility (SV) model is one of the popular models used in the literature (see, for example, Taylor (1986), Harvey (1998)) to model the conditional heteroscedasticity in returns on financial assets. The SV model is given by

$$r_t = \exp\left(Y_t/2\right) v_t, \tag{1}$$

where v_t is a sequence of *i.i.d.* $\left(0, \sigma^2\right)$ random variables independent of Y_t, and Y_t is a stationary Gaussian process. When the v_t are assumed to be an *i.i.d.* series of random variables with positive support, independent of Y_t, the model is referred to as the stochastic conditional duration model (SCD) and can be used to model financial durations. See Bauwens and Veredas (2004). When the SV model (1), assuming an AR(1) for Y_t, is fit to returns on financial assets, the estimated autoregressive parameter is generally found to be close to unity, suggesting very strong dependence in squared returns. This has prompted researchers (See Breidt, Crato and de Lima 1998, Harvey 1998) to consider a long memory process for Y_t, such as an Autoregressive Fractionally Integrated Moving Average Process (ARFIMA), which has slow power law decay in its autocorrelations. We will refer to an SV model where Y_t is a long memory Gaussian process as a Long Memory SV (LMSV) model. Though it has not been proposed in the literature so far, a long memory SCD (LMSCD) model would also seem to be a suitable candidate for tick-by-tick durations of trades. This is due to the fact that durations exhibit the same strong dependence that is seen in squared returns and that SCD models with AR(1) models for Y_t when fit to the data yield estimated coefficients very close to unity. See Bauwens and Veredas (2004). Thus, efficient estimation of LMSV and LMSCD models is an important issue.

The estimation of SV/SCD models is made difficult by the fact that the dependence is modelled non-linearly through an unobserved latent process. Though this use of a latent process allows one to obtain a wide range of theoretical properties of the model easily, it is impossible to write the exact likelihood of the SV/SCD model analytically. Hence, various alternative procedures have been proposed in the literature for estimation, including quasi maximum likelihood (QML) estimation and generalised method of moments (GMM) estimation. As Andersen and Sorensen (1996) note, the procedures other than QML and GMM are computationally intensive.

The QML procedure for the SV model exploits the fact that the transformed process $\log r_t^2$ may be written as a sum of a Gaussian stationary time series and an independent white noise series. The QML estimates of the model parameters are obtained by maximising the Gaussian likelihood of the $\log r_t^2$, even though this series is not Gaussian. Deo (1995) has shown that the QML estimators of the LMSV model (1) based on the $\log r_t^2$ are $\sqrt{n}$ consistent and asymptotically normal. Deo's (1995) result can easily be shown to hold for the LMSCD model too.

Unlike QML estimation, there is currently no known $\sqrt{n}$ consistent GMM estimation procedure of LMSV/SCD models. For GMM estimation, one specifies a set of sample moments denoted by $M_n = (M_{1n}, ..., M_{qn})$, where $M_{in} = \sum_{t=j+1}^{n} g_i(r_t, r_{t-j}) / (n-j)$, j is the maximum lag being used, g_i is some smooth function and q, the number of selected moments, is at least as large as the dimension of the parameter vector θ to be estimated. The GMM estimator, $\hat{\theta}$, minimises the distance $(M_n - M(\theta))' \Lambda^{-1} (M_n - M(\theta))$, where $M(\theta) = E_\theta(M_n)$ and Λ is some suitably chosen weight matrix. Under suitable regularity conditions, $\hat{\theta}$ is $\sqrt{n}$ consistent and asymptotically normal (Hansen 1982). These suitable conditions include the requirement that the vector of moments M_n be a $\sqrt{n}$ consistent estimator of $M(\theta)$. In the literature on short memory SV models(see, for example, Andersen and Sorensen 1996 and Jacquier, Polson and Rossi 1994), the moment conditions that have generally been used are obtained by using functions g_i of the form

$$g_i(r_t, r_{t-j}) = |r_t|^{a_i} |r_{t-j}|^{b_i} \tag{2}$$

for some integer valued non-negative a_i, b_i. When Y_t is assumed to follow an AR(1) process, Andersen and Sorensen (1997) report that GMM estimators based on moments of the form (2) perform more poorly than QML estimation when the AR(1) coefficient is close to the unit root. i.e. when the persistence in the volatility is high. Bauwens and Veredas (2004) also report in their simulations that the sample moments based on functions of the form (2) converge very slowly to the population analogues in an SCD model with a strongly persistent AR(1) process for Y_t. These two observations indicate that the convergence of the sample moments of the form (2) will also be very slow if Y_t were a genuine long memory process instead of a near unit root AR(1), thus yielding poor GMM estimators. In the next section, we show that this is indeed the case. More specifically, we show that the rate of convergence of the moments (2) is slower than $\sqrt{n}$ and is a decreasing function of the memory parameter. Furthermore, this rate can be arbitrarily close to a constant. We than provide an alternative set of moment conditions and prove that the new conditions are indeed $\sqrt{n}$ consistent and asymptotically normal. The proofs of all of our results are in the Appendix at the end of the paper.

2 GMM estimation for the LMSV/SCD model:

We will assume that the spectral density $f(\cdot)$ of Y_t in (1) is of the form

$$f_Y(\lambda) = \lambda^{-2d} g(\lambda) \tag{3}$$

for some $d \in (0, 0.5)$, where $g(\cdot)$ is a differentiable function on $[-\pi, \pi]$. The parameter d is called the memory parameter of the process and controls the rate of decay of the correlations of the process Y_t. Processes such as the well known ARFIMA models have spectral densities that satisfy (3). It is well known that under (3) the correlations of Y_t at lag j, $\gamma_h(j)$, decay hyperbolically in j at a rate given by j^{2d-1}. Furthermore, it can be shown that the correlations of the transformed processes $|r_t|^c$ and $\log r_t^2$ display the same hyperbolic decay for any $c > 0$. Thus, the process r_t displays very strong conditional heteroscedasticity.

The following theorem establishes the asymptotic behaviour of sample moments of the form (2) under LMSV/SCD models.

Theorem 1 *For any integer valued non-negative a, b and under the conditions (1) and (3), assuming that v_t has all the required moments,*

$$n^{1/2-d}\left(\frac{1}{n}\sum_{t=1}^{n}|r_t|^a |r_{t-j}|^b - E\left(|r_t|^a |r_{t-j}|^b\right)\right) \xrightarrow{D} X,$$

where X is a zero mean Gaussian random variable.

It is clear from theorem 1 that the rate of convergence of the sample moments is slower than $\sqrt{n}$. Furthermore, this rate gets worse as d approaches $1/2$. This is particularly of concern since numerous studies (see for example Andersen et al. 2001) have found that high frequency returns tend to yield estimated values of d which are around 0.3 to 0.45. This problem may be exarcebated further by the fact that Deo and Hurvich (2001, 2002) have shown that semi-parametric estimation of d for LMSV models can be negatively biased indicating that the real values of d may be even greater than the values obtained in the studies.

The particular form of the SV/SCD model can however be exploited to get a set of moment conditions that retain a $\sqrt{n}$ rate of convergence. Using (1), we can write the transformed series $Z_t = \log r_t^2$ as $Z_t = \mu + Y_t + u_t$ where $u_t = \log v_t^2 - E(\log v_t^2)$ and $\mu = E\left(\log v_t^2\right)$. Since Y_t

and u_t are independent, we get a signal plus noise representation for Z_t. QML estimation of the SV/SCD model is based on precisely this transformation. The transformation Z_t has also been used in the literature to suggest GMM estimators for the LMSV model. Wright (1999) has proposed using the sample covariances of Z_t as the moment conditions to estimate the model parameters. However, Wright (1999) shows that these moment conditions are $\sqrt{n}$ consistent only when the memory parameter d satisfies $d < 0.25$. When $d > 0.25$, the sample covariances of Z_t can be shown (Hosking, 1996) to be slower than $\sqrt{n}$ consistent. Since, as argued above, the interval (0.25,0.5) constitutes the more empirically relevant range for d, it is crucial to have moments which will retain the $\sqrt{n}$ convergence rate over the entire parameter space of d. The following theorem provides precisely such a set of moment conditions.

Theorem 2 *Assume the model given by (1) and (3) and that* $E\left\{u_t^8\right\} < \infty$. *Let*

$$\hat{\gamma}_j = (n-j)^{-1} \sum_{t=j+1}^{n} \left(Z_t - \bar{Z}\right)\left(Z_{t-j} - \bar{Z}\right),$$

$\gamma_j = Cov\left(Z_t, Z_{t-j}\right)$ *and* $W_j = \hat{\gamma}_0 - \hat{\gamma}_j - \left(\gamma_0 - \gamma_j\right)$. *Then for any integer* q

$$\sqrt{n}\mathbf{W} \xrightarrow{D} N(\mathbf{0}, \mathbf{\Sigma}),$$

where $\mathbf{W} = (w_1, ..., w_q)'$ *and* $\mathbf{\Sigma} = \mathbf{A}\mathbf{\Sigma}_1\mathbf{A}' + \mathbf{\Sigma}_2$,

the j^{th} *row of* $\mathbf{A}$ *is* $a_j' = (j/2, (j-1), (j-2), ..., 1, \underbrace{0, ..., 0}_{q-j \text{ terms}})$,

the $(j,k)^{th}$ *term of* $\mathbf{\Sigma}_1$ *is* $4\pi \int_{-\pi}^{\pi} \cos j\lambda \cos k\lambda \left|1 - \exp(i\lambda)\right|^4 f_Y^2(\lambda)\, d\lambda$,
and the $(j,k)^{th}$ *term of* $\mathbf{\Sigma}_2$ *is* $Eu_t^4 - \sigma_u^4 + \sigma_u^2 Cov(2Y_t - Y_{t-j} - Y_{t+j}, 2Y_t - Y_{t-k} - Y_{t+k})$.

Using standard Taylor series arguments, it ifollows from Theorem 2 that any differentiable transformation of the sample moments provided there will also be $\sqrt{n}$ consistent and asymptotically normal. Thus, one can choose any of these transformations of sample moments to construct the moment conditions to use for GMM estimation of the model. Needless to say, which moment conditions one chooses will dictate the efficiency of the resulting GMM estimator and a partial answer regarding this choice may be given by a detailed Monte Carlo simulation. We leave this issue for further research.

Appendix

Proof of Theorem 1: Let $\mu_1 = E\left\{\exp\left(0.5aY_t + 0.5bY_{t-j}\right)\right\}$ and $\mu_2 = E\left\{|v_t|^a |v_{t-j}|^b\right\}$. Let

$$T_1 = n^{-1}\sum_{t=1}^{n} |r_t|^a |r_{t-j}|^b - n^{-1}\sum_{t=1}^{n} \exp\left(0.5aY_t + 0.5bY_{t-j}\right)\mu_2$$

and

$$T_2 = n^{-1}\sum_{t=1}^{n} \exp\left(0.5aY_t + 0.5bY_{t-j}\right)\mu_2 - \mu_1\mu_2.$$

Then $n^{-1}\sum_{t=1}^{n} |r_t|^a |r_{t-j}|^b - E\left(|r_t|^a |r_{t-j}|^b\right) = T_1 + T_2$. Since v_t is an iid series, we get $Var\left(T_1\right) = O\left(n^{-1}\right)$ and hence $T_1 = o_p\left(n^{d-1/2}\right)$. The theorem is thus established if we prove that $n^{1/2-d}T_2$ is asymptotically normal. Let $X_t = 0.5aY_t + 0.5bY_{t-j}$. Then X_t is also a stationary long memory Gaussian series with a spectral density that satisfies (3) and $Corr\left(X_t, X_{t-s}\right) \tilde{} As^{2d-1}$ as $s \to \infty$ for some constant A. We have

$$T_2 = \mu_2 n^{-1}\sum_{t=1}^{n}\left(\exp\left(X_t\right) - \mu_1\right).$$

Furthermore, the function $G\left(x\right) = \exp\left(x\sigma\right) - \exp\left(\sigma^2/2\right)$, where $\sigma^2 = Var\left(X_t\right)$, has a Hermite rank of 1 as defined in Taqqu (1975). Hence, by Theorem 5.1 of Taqqu (1975), $n^{1/2-d}T_2 \xrightarrow{D} X$ where X is a zero mean Gaussian random variable.

Proof of Theorem 2: We will demonstrate the proof only for W_j since the higher dimension case is obtained along similar lines by applying the Cramer Wold device. Letting $\hat{\gamma}_{Y,j} = n^{-1}\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)\left(Y_{t-j} - \bar{Y}\right)$ for $j \geq 0$ and using the fact that $\bar{Y} \xrightarrow{P} 0$, we get by simple algebra,

$$\begin{aligned}\hat{\gamma}_0 - \hat{\gamma}_j &= \hat{\gamma}_{Y,0} - \hat{\gamma}_{Y,j} + n^{-1}\sum_{t=1}^{n}\left(\xi_t^2 - \xi_t\xi_{t-j}\right) + n^{-1}\sum_{t=1}^{n} Y_t\left(2\xi_t - \xi_{t-j} - \xi_{t+j}\right) + o_p\left(n^{-1/2}\right)\\ &= \hat{\gamma}_{Y,0} - \hat{\gamma}_{Y,j} + n^{-1}\sum_{t=1}^{n}\left(\xi_t^2 - \xi_t\xi_{t-j}\right) + n^{-1}\sum_{t=1}^{n}\left(2Y_t - Y_{t-j} - Y_{t+j}\right)\xi_t + o_p\left(n^{-1/2}\right).\end{aligned} \tag{4}$$

From the proof of Theorem 5 of Hosking (1996), we see that

$$\hat{\gamma}_{Y,0} - \hat{\gamma}_{Y,j} = a_j'\mathbf{X} + O_p\left(n^{-1}\right), \tag{5}$$

where $\mathbf{X} = (\widetilde{\gamma}_0^{(1)} - \gamma_0^{(1)}, \widetilde{\gamma}_1^{(1)} - \gamma_1^{(1)}, ..., \widetilde{\gamma}_{q-1}^{(1)} - \gamma_{q-1}^{(1)})'$, $\widetilde{\gamma}_s^{(1)} = n^{-1}\sum_{t=1}^{n}(Y_t - Y_{t-1})(Y_{t-j} - Y_{t-1-j})$ and

$\gamma_s^{(1)} = E((Y_t - Y_{t-1})(Y_{t-j} - Y_{t-1-j}))$. From (4) and (5), we get

$$\hat{\gamma}_0 - \hat{\gamma}_j = a_j'\mathbf{X} + n^{-1}\sum_{t=1}^{n}\left(\xi_t^2 - \xi_t\xi_{t-j}\right) + n^{-1}\sum_{t=1}^{n}\left(2Y_t - Y_{t-j} - Y_{t+j}\right)\xi_t + o_p\left(n^{-1/2}\right).$$

The limiting distribution result now follows from Hannan (1976).

REFERENCES

Andersen, T. Bollerslev, T., Diebold, F.X. and Labys, P. (2001) "The Distribution of Realized Exchange Rate Volatility". *Journal of the American Statistical Association*, **96**, 42-55.

Andersen T. and Sorensen, B. (1996) "GMM estimation of a stochastic volatility model: A Monte Carlo study". *Journal of Business and Economic Statistics,* **14**, no. 3, 328-352.

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York.

Bauwens, L and Veredas, D. (2004) "The stochastic conditional duration model: a latent variable model for the analysis of financial durations". *Journal of Econometrics*, **119,** 381-412.

Breidt, F. J., Crato, N., and de Lima, P. (1998) "The detection and estimation of long memory in stochastic volatility". *Journal of Econometrics,* **83,** 325-348.

Brockwell, P.J. and Davis, R.A. (1987). Time series: Theory and Methods. (first edition) Springer, New York.

Deo, R. (1995) "Tests for unit roots in multivariate autoregressive processe". Ph.D. Dissertation, Iowa State University.

Deo, R. and Hurvich, C. (2001) "On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models". *Econometric Theory*, **17**, 686-710.

Deo, R. and Hurvich, C. (2002) "On long memory in stochastic volatility" Forthcoming.

Hansen, L. (1982) "Large sample properties of generalised methods of moments estimators". *Econometrica,* **50**, 1029-1054.

Jacquier, E., Polson, N and Rossi, P. (1994) "Bayesian analysis of stochastic volatility models". *Journal of Business and Economic Statistics,* **12**, 371-389.

Hannan, E. J. (1976) "The asymptotic distribution of serial covariances". *Annals of Statistics*, **4,** 396-399.

Harvey, A. (1998). Long memory in stochastoc volatility. In Forecasting Volatility in the Financial Markets. ed. J. Knight and S. Satchell. Butterworth and Heinemann.

Hosking, J. (1996) "Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series". *Journal of Econometrics,* **73**, no. 1, 261–284.

Taqqu, M. (1975) "Weak convergence to the fractional Brownian motion and to the Rosenblatt process". *Z. Wahrscheinlichkeitstheorie verw. Gebiete,* **31**, 287-302.

Wright, J. H. (1999). "A new estimator of the fractionally integrated stochastic volatility model". *Economics Letters*, **63**, 295-303