KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.750

"Testing the Box-Cox Parameter for an Integrated Process"

Michael McAleer

December 2010

KYOTO UNIVERSITY

KYOTO, JAPAN

Testing the Box-Cox Parameter for an Integrated Process*

Jian Huang
Guangdong University of Finance
PRC

Masahito Kobayashi
Faculty of Economics
Yokohama National University
Japan

Michael McAleer
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Institute of Economic Research
Kyoto University
Japan

Revised: December 2010

---

* The second author wishes to acknowledge the financial support of the Ministry of Education, Science, Culture and Sports of Japan, and the third author is most grateful for the financial support of the Australian Research Council, National Science Council, Taiwan, and the Japan Society for the Promotion of Science.

Corresponding author: Masahito Kobayashi, Faculty of Economics, Yokohama National University, Japan. Email:mkoba@ynu.ac.jp.

Abstract

This paper analyses the constant elasticity of volatility (CEV) model suggested by Chan et al. (1992). The CEV model without mean reversion is shown to be the inverse Box-Cox transformation of integrated processes asymptotically. It is demonstrated that the maximum likelihood estimator of the power parameter has a nonstandard asymptotic distribution, which is expressed as an integral of Brownian motions, when the data generating process is not mean reverting. However, it is shown that the t-ratio follows a standard normal distribution asymptotically, so that the use of the conventional t-test in analyzing the power parameter of the CEV model is justified even if there is no mean reversion, as is often the case in empirical research. The model may applied to ultra high frequency data

Keywords: Box-Cox transformation, Brownian Motion, Constant Elasticity of Volatility, Mean Reversion, Nonstandard distribution.

1. Introduction

The constant elasticity of volatility (CEV) model has been an important tool in analyzing short-time interest rates. The estimation and testing of the power parameter γ of the CEV model in the discrete form, namely:

$$y_t - y_{t-1} = \alpha + \beta y_{t-1} + y_{t-1}{}^{\gamma} e_t, \quad t = 1, \ldots, T, \quad e_t \sim NID(0, \sigma^2), \tag{1}$$

has been a focus of research over an extended period. It is well known that the differential equation in the continuous version of (1) can be solved analytically only for specific values of γ, namely 1/2 and 0, as proved by Cox, Ingersoll and Ross (1985) and Vasicek (1977). However, the actual estimate of γ is often larger than the theoretically permissible values. This puzzle, which was first suggested by Chan et al (1992), has attracted the attention of a number of researchers (see, for example, Brenner et al. (1996), Koedijk et al. (1997), and Bliss and Smith (1998)). Yu and Phillips (2001) estimated the continuous version of the model directly in order to avoid the bias arising through discretization. Further discussion of discretization errors in models of ultra high frequency data are given in McAleer (2005) and McAleer and Mederios (2008).

In empirical research, it is standard to assume the asymptotic normality of both the estimator and the t-ratio of $\hat{\gamma}$ in the CEV model, which requires that $-2 < \beta < 0$ in (1), so that the process y_t is mean reverting. However, the mean reversion of the process is rarely supported empirically, since it is often reported that the estimate of the mean reversion parameter β in (1) is not significantly different from zero. Rodrigues and Rubia (2004) considered this problem in detail through simulation. Even Chan et al. (1992) reported a t-statistic of 1.54 for $\hat{\beta}$, which is not statistically significant at any conventional level, assuming the asymptotic normality of the t-ratio for $\hat{\beta}$. Adkins and Krehbiel (1999) found that neither the 3-month nor 6-month LIBOR (London Inter-Bank Offered Rate) was mean reverting. Treepongkaruna and Gray (2003) reported that the mean reversion of daily short-term interest rate data was not statistically significant at the 5 percent level in any of the eight countries they examined.

The main implication of the apparent lack of mean reversion is that the asymptotic normality of the estimator of γ may no longer be guaranteed. In most empirical research, the hypothesized value of γ is tested under the assumption that the t-statistic for $\hat{\gamma}$ is distributed as asymptotic normal. However, statistical inference is no longer reliable if the process is not mean reverting, as most empirical research would seem to suggest.

In this paper it is shown that, when the data generating process is not mean reverting, the CEV model (1) can be expressed as the inverse Box-Cox (1964) transformation of integrated series, and that the maximum likelihood estimator of γ in (1) has a nonstandard asymptotic distribution. It is also demonstrated that the t-ratio for $\hat{\gamma}$ follows an asymptotic standard normal distribution under the assumption that the variance of innovations is small in relation to the levels. Therefore, it follows that the use of the conventional t-test in analyzing the CEV model is justified even for processes that are not mean reverting.

The results in this paper are a useful application of small σ-asymptotics for purposes of deriving the asymptotic distribution of a nonlinear transformation of integrated series. This simple approach is particularly helpful in analyzing high-frequency financial time series data, where the variance of innovations is small in relation to the levels. A rigorous and general approach to nonlinear transformations of integrated series can be found in Park and Phillips (1999, 2001).

The plan of the remainder of the paper is as follows. Section 2 shows that the constant elasticity of volatility process without mean reversion can be approximated by the inverse Box-Cox transformation of a random walk. Maximum likelihood estimation and testing of the CEV model are discussed in Section 3. A Monte Carlo experiment is presented in Section 4, and some concluding remarks are given in Section 5. The algebraic details are given in the Appendix.

2. Approximation of the CEV Process Without Mean Reversion

We first show that the CEV process (1) without mean reversion can be approximated by the inverse Box-Cox transformation of a random walk. It is assumed that the variance of innovations of (1) converges to zero as the sample size increases, namely:

$$\sigma^2 = \omega^2 T^{-1-d} = O(T^{-1-d}) \tag{2}$$

for some $0 < d < 1$. Statistical analysis using the assumption of small variance, which is referred to as small-σ expansion, has been a powerful tool in statistics. The dependence of the parameter value on the sample size T is not essential in this assumption as it simply means that the parameter value is small in relation to the levels. Similar ideas can be found in the Pitman drift, the near-unit root process, and weak instrumental variables by Staiger and Stock (1997), in that the unknown parameter converges to zero as the sample size increases in these models. The small-σ expansion was used, for example, in analyzing the Box-Cox model (1964) by Bickel and Doksum (1981), and in testing linear and logarithmic transformations of the random walk process by Kobayashi and McAleer (1999). This assumption is justifiable in analyzing financial time series because the sampling frequency is often very high, so that the variance of the innovations is very small in comparison with the levels.

It is also assumed that the actual data generating process has no mean reversion, namely $\alpha = 0$ and $\beta = 0$ in (1), but the model is estimated without the restriction $\alpha = 0$ and $\beta = 0$. This procedure reflects the practice of estimating the CEV model (1) under the assumption of mean reversion even if the estimated β parameter is not statistically significant.

We first show that the discretized CEV model (1) can be approximated by the inverse Box-Cox transformation of the random walk, as defined by:

$$\tilde{y}_t \equiv (1+(1-\gamma)z_t)^{\frac{1}{1-\gamma}}, \tag{3}$$

$$z_t \equiv e_1 + \cdots + e_t, \ \ \mathrm{var}(e_t) \equiv \sigma^2 = \omega^2 n^{-1-d} \ \text{ for some } 0 < d < 1.$$

The model given above is a generalization of that used in Kobayashi and McAleer (1999), where the logarithmic and linear transformations of the random walk processes were tested against each other. It follows from (3) that:

$$z_t = \frac{\tilde{y}_t^{1-\gamma} - 1}{1-\gamma},$$

which is the Box-Cox transformation. The assumption of small variance of e_t ensures that, as n increases, the random walk process, z_t, is bounded stochastically, because it follows that:

$$E(z_t) = 0, \ \ \mathrm{var}(z_t) = \mathrm{var}(e_1) + \cdots + \mathrm{var}(e_t) = c \times n \times n^{-1-d},$$

$$z_t = O_p(n^{-d/2}),$$

so that $\tilde{y}_t$ is bounded in the neighborhood of 1, namely:

$$\tilde{y}_t - 1 \equiv z_t + \frac{1}{2}\frac{\gamma}{1-\gamma} z_t^2 + \cdots = O_p(n^{-d/2}). \tag{4}$$

Under this assumption, the series, $\tilde{y}_t$, as defined by (3), expresses the CEV model asymptotically when $\alpha = 0$ and $\beta = 0$. Therefore, the expression given by:

$$\tilde{y}_t - \tilde{y}_{t-1} = \tilde{y}_{t-1}{}^{\gamma} e_t + O_p(n^{-1-d}) = O_p(n^{-1/2-d/2})$$

follows from the formal Taylor expansion, namely:

$$\tilde{y}_t = \tilde{y}_{t-1} + \tilde{y}_{t-1}{}^{\gamma} e_t + \tfrac{1}{2}\gamma \tilde{y}_{t-1}^{2\gamma-1} e_t{}^2 + \cdots$$

and the order condition given by:

$$e_t = O_p(n^{-1/2-d/2}).$$

In the following, we use the notation $y_1, ..., y_T$ and $\tilde{y}_1, ..., \tilde{y}_T$ interchangeably when there is no fear of ambiguity.

3. Estimation and Testing

The model (1) can be rewritten as

$$y_t - y_{t-1} = \delta + \beta(y_{t-1} - 1) + y_{t-1}{}^{\gamma} e_t,$$

where $\alpha \equiv -\beta + \delta$. This transformation is necessary to avoid the degeneracy of the asymptotic distribution of the estimators, as will be shown below. The maximum likelihood estimators, say $\hat{\delta}, \hat{\beta}, \hat{\gamma}, \hat{\omega}^2$, are defined as the solutions of the following equations:

$$\frac{\partial L}{\partial \delta} = 0, \ \frac{\partial L}{\partial \beta} = 0, \ \frac{\partial L}{\partial \gamma} = 0, \ \frac{\partial L}{\partial \omega^2} = 0, \qquad (5)$$

where the log-likelihood function is defined by:

$$L \equiv \log f(y_1, \cdots, y_T)$$

$$= -\frac{T}{2}\log(2\pi\sigma^2) - \frac{\sum [y_t - \delta - y_{t-1} - \beta(y_{t-1} - 1)]^2 y_{t-1}{}^{-2\gamma}}{2\sigma^2} - \gamma \sum \log y_{t-1}.$$

The distributions of the maximum likelihood estimator are obtained under the assumption that the data are generated by the model (2), which is an approximation of the CEV process without mean reversion. Assuming the consistency of the estimators, we can invert the Taylor expansion of (5) as follows:

$$\left(\hat{\delta} - \delta, \quad \hat{\beta} - \beta, \quad \hat{\gamma} - \gamma, \quad \hat{\omega}^2 - \omega^2\right)' = -K^{-1}\ell + \cdots, \tag{6}$$

where l and K are defined by:

$$\ell \equiv \begin{pmatrix} \partial L / \partial \delta \\ \partial L / \partial \beta \\ \partial L / \partial \gamma \\ \partial L / \partial \omega^2 \end{pmatrix}, \quad K \equiv \begin{pmatrix} \frac{\partial^2 L}{\partial \delta^2} & \frac{\partial^2 L}{\partial \delta \partial \beta} & \frac{\partial^2 L}{\partial \delta \partial \gamma} & \frac{\partial^2 L}{\partial \delta \partial \omega^2} \\ \frac{\partial^2 L}{\partial \beta \partial \delta} & \frac{\partial^2 L}{\partial \beta^2} & \frac{\partial^2 L}{\partial \beta \partial \gamma} & \frac{\partial^2 L}{\partial \beta \partial \omega^2} \\ \frac{\partial^2 L}{\partial \gamma \partial \delta} & \frac{\partial^2 L}{\partial \gamma \partial \beta} & \frac{\partial^2 L}{\partial \gamma^2} & \frac{\partial^2 L}{\partial \gamma \partial \omega^2} \\ \frac{\partial^2 L}{\partial \omega^2 \partial \delta} & \frac{\partial^2 L}{\partial \omega^2 \partial \beta} & \frac{\partial^2 L}{\partial \omega^2 \partial \gamma} & \frac{\partial^2 L}{\partial (\omega^2)^2} \end{pmatrix}, \tag{7}$$

which are evaluated at the true parameter values. After some algebra given in the Appendix, it can be shown that:

$$Q^{-1}\begin{pmatrix} \partial L/\partial\beta \\ \partial L/\partial\delta \\ \partial L/\partial\gamma \\ \partial L/\partial\omega^2 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \omega^{-1}B_1(1) \\ \int B_1(r)dB_1(r) \\ \sqrt{2}\omega\int B_1(r)dB_2(r) \\ \dfrac{1}{\omega^2\sqrt{2}}B_2(1) \end{pmatrix} \tag{8}$$

and

$$Q^{-1}KQ^{-1}\xrightarrow{p} -\begin{pmatrix} \omega^{-2} & \omega^{-1}\int B_1(r)dr & 0 & 0 \\ \omega^{-1}\int B_1(r)dr & \int B_1^2(r)dr & 0 & 0 \\ 0 & 0 & 2\omega^2\int B_1(r)^2dr & \omega^{-1}\int B_1(r)dr \\ 0 & 0 & \omega^{-1}\int B_1(r)dr & \dfrac{1}{2\omega^4} \end{pmatrix}, \tag{9}$$

where the standardizing matrix is defined by

$$Q=\begin{bmatrix} T^{1+d/2} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T^{1/2-d/2} & 0 \\ 0 & 0 & 0 & T^{1/2} \end{bmatrix}.$$

Noting that $Q^{-1}KQ^{-1}$ converges to a block diagonal matrix asymptotically, the estimator of β is expressed as

$$T(\hat{\beta}-\beta)\xrightarrow{p}\omega\frac{\int\left(B_1(r)-\int B_1(r)dr\right)dB_1(r)}{\int\left(B_1(r)-\int B_1(r)dr\right)^2dr}.$$

The asymptotic distribution of the statistic given above is nonstandard as its expression is identical to that of the Dickey-Fuller statistic with a constant term.

We now show that $\hat{\gamma}$ has a nonstandard distribution but that its t-ratio is distributed as asymptotic normal. First, equations (8) and (9) yield the following asymptotic expressions:

$$T^{1/2-d/2}(\hat{\gamma}-\gamma) \xrightarrow{p} \frac{1}{\omega\sqrt{2}} \frac{\int\left(B_1(r)-\int B_1(r)dr\right)dB_2(r)}{\int\left(B_1(r)-\int B_1(r)dr\right)^2 dr} \tag{11}$$

and

$$t_\gamma \equiv \frac{\hat{\gamma}-\gamma}{\sqrt{-K^{33}}} \xrightarrow{p} \frac{\int\left(B_1(r)-\int B_1(r)dr\right)dB_2(r)}{\sqrt{\int\left(B_1(r)-\int B_1(r)dr\right)^2 dr}}, \tag{12}$$

where K^{33} denotes the (3,3)'th element of the inverse of the second derivative matrix K. It can be shown that the t-ratio of $\hat{\gamma}$ is asymptotically normally distributed, with zero mean and unit variance, since the conditional distribution of

$$\int\left[B_1(r)-\int B_1(r)dr\right]dB_2(r),$$

given $B_1(r)$, is asymptotically normal with zero mean and variance given by:

$$\int\left[B_1(r)-\int B_1(r)dr\right]^2 dr.$$

Therefore, the conditional distribution of

$$\frac{\int \left[B_1(r) - \int B_1(r)dr \right] dB_2(r)}{\sqrt{\int \left[B_1(r) - \int B_1(r)dr \right]^2 dr}},$$

given $B_1(r)$, is asymptotically normal with zero mean and unit variance, and hence is distributed independently of $B_1(r)$. The asymptotic distribution of $T^{1/2-d/2}(\hat{\gamma} - \gamma)$ is nonnormal, since the distribution of

$$\sqrt{\int \left[B_1(r) - \int B_1(r)dr \right]^2 dr}$$

on the left-hand side of

$$T^{1/2-d/2}(\hat{\gamma} - \gamma) \xrightarrow{p} t_\gamma \sqrt{\int \left[B_1(r) - \int B_1(r)dr \right]^2 dr}$$

is asymptotically nonstandard.

4. Monte Carlo Experiment

In a small Monte Carlo experiment, 500 series of artificial data with $T = 5000$ are generated using the data generating process (1), with $\alpha = \beta = 0$, $\gamma = 0.5, 1.0$, $\sigma^2 = 0.02$. The unknown parameters are estimated using the maximum likelihood method. Normality is tested using the Jarque-Bera (1990) Lagrange multiplier (LM) test statistic, namely

$$\frac{N}{6}\left[skewness^2 + \frac{(kurtosis - 3)^2}{4} \right],$$

which is distributed asymptotically as $\chi^2(2)$ under the null hypothesis of normality.

The numerical results show that the actual distribution of $\hat{\gamma}$ is nonnormal, but the actual distribution of the t-ratio of $\hat{\gamma}$ can be regarded as normal. For the data generating process with $\gamma = 1.0$, the Jarque-Bera LM test statistic is 298.5 for $\hat{\gamma}$, with p-value of 0.0, so that the normality of $\hat{\gamma}$ is clearly rejected. On the other hand, the skewness and kurtosis of the t-ratio for $\hat{\gamma}$ are both near zero. The normality of the t-ratio for $\hat{\gamma}$ cannot be rejected, with the value of the Jarque-Bera LM test statistic as 1.045 and a p-value of 0.593. The results for the data generating process with $\gamma = 0.5$ are essentially the same, in that the actual distribution of $\hat{\gamma}$ is nonnormal, but the actual distribution of the t-ratio of $\hat{\gamma}$ can be regarded as normal.

5. Concluding Remarks

This paper analyzed the constant elasticity of volatility (CEV) model that was first suggested by Chan et al. (1992). The CEV model without mean reversion was shown to be the inverse Box-Cox transformation of integrated processes asymptotically. It was demonstrated that the maximum likelihood estimator of the power parameter of the CEV model had a nonstandard asymptotic distribution, which was expressed as an integral of Brownian motions, when the data generating process was not mean reverting. It was also shown that the t-ratio followed a standard normal distribution asymptotically. Therefore, the use of the conventional t-test in analyzing the power parameter of the CEV model can be justified even in the absence of mean reversion, as is often found in empirical research. The model may be applied to ultra high frequency data.

Table 1. Distribution of the Estimator and t-ratio of $\hat{\gamma}$

$\gamma = 1.0$

Estimators	Mean	SD	Skewness	Kurtosis-3	Jarque-Bera LM test for normality (p-value)
$\hat{\gamma}$	0.996	0.100	-0.522	3.618	298.5 (0.0)
t-ratio	-0.028	1.059	-0.111	-0.004	1.045 (0.593)

$\gamma = 0.5$

Estimators	Mean	SD	Skewness	Kurtosis-3	Jarque-Bera LM test for normality (p-value)
$\hat{\gamma}$	0.501	0.090	-0.191	1.568	54.58 (0.0)
t-ratio	0.025	0.997	-0.088	-0.122	0.958 (0.619)

Appendix: Algebraic Details

In this section we derive the first and second derivatives of the log-likelihood function:

$$L \equiv \log f(y_1, \cdots, y_T)$$

$$= -\frac{T}{2}\log(2\pi\sigma^2) - \frac{\sum [y_t - \delta - y_{t-1} - \beta(y_{t-1} - 1)]^2 y_{t-1}^{-2\gamma}}{2\sigma^2} - \gamma \sum \log y_{t-1}.$$

given in (8) and (9) for the case where $0 < \gamma < 1$. The algebraic derivations in the case where $\gamma = 1$ are analogous. First, note that

$$T^{-1/2} z_t / \sigma = T^{-1/2} \left(e_1 + \cdots + e_t\right) / \sigma \overset{p}{\rightarrow} B_1(r), \quad t = [rT] \tag{13}$$

and

$$T^{-1/2} \left((e_1^2 / \sigma^2 - 1) + \cdots + (e_t^2 / \sigma^2 - 1)\right) / \sqrt{2} \overset{p}{\rightarrow} B_2(r), \ \ t = [rT] \tag{14}$$

as $T \to \infty$, where $[rT]$ denotes the largest integer not greater than rT. It is straightforward to show that

$$B_1(r) \text{ and } B_2(r), \quad 0 \le r \le 1,$$

are independent standard Brownian motions, because e_t and $\frac{1}{\sqrt{2}}(e_t^2 / \sigma^2 - 1)$ are mutually uncorrelated and serially independent, with mean zero and unit variance.

We will show that

$$T^{-1-d/2}\frac{\partial L}{\partial\delta}=T^{-1/2}\frac{1}{\omega}\frac{\sum e_t y_{t-1}^{-\gamma}}{\sigma}\xrightarrow{p}\frac{1}{\omega}B_1(1)\,. \tag{15}$$

From the expansion

$$y_{t-1}{}^{-\gamma}=(1+(1-\gamma)z_{t-1})^{-\gamma/(1-\gamma)}=1-\gamma z_{t-1}+\cdots \tag{16}$$

and the definition $\sigma=\omega T^{-1/2-\varepsilon/2}$, it follows that

$$\omega T^{-1-d/2}\sum e_t\, y_{t-1}^{-\gamma}/\sigma^2=\sigma T^{-1/2}\sum e_t\,(1+(1-\gamma)z_{t-1})^{-\gamma/(1-\gamma)}/\sigma^2$$

$$=\sigma T^{-1/2}\sum e_t/\sigma^2-\gamma\sigma T^{-1/2}\sum e_t z_{t-1}/\sigma^2+\cdots$$

$$\xrightarrow{p} B_1(1)+o_p(1),$$

as the second term

$$\sigma T^{-1/2}\sum e_t z_{t-1}/\sigma^2\xrightarrow{p}\omega T^{-\varepsilon/2}\int B_1(r)dB_1(r)=O_p(T^{-d/2})$$

is of smaller order than the first term $\sigma T^{-1/2}\sum e_t/\sigma^2=O_p(1)$, and hence is negligible.

Next, we show that

$$\begin{aligned}T^{-1}\frac{\partial L}{\partial\beta}&=T^{-1}\frac{\sum e_t(y_{t-1}-1)y_{t-1}^{-\gamma}}{\sigma^2}\\&=T^{-1}\sum(e_t z_{t-1}+\frac{1}{2}\frac{\gamma}{1-\gamma}e_t z_{t-1}^2+\cdots)(1-\gamma z_{t-1}+\cdots)/\sigma^2\xrightarrow{p}\int B_1(r)dB_1(r)\end{aligned} \tag{17}$$

upon substituting (16) and

$$y_{t-1} - 1 = z_{t-1} + \frac{1}{2}\frac{\gamma}{1-\gamma} z_{t-1}^2 + \cdots \tag{18}$$

Then we show that

$$\begin{aligned} T^{-1/2+d/2}\frac{\partial L}{\partial \gamma} &= \omega \sum \left(e_t^2 - \sigma^2\right) / \sigma^2 \log y_{t-1} / \sigma \\ &= \omega\sqrt{2} T^{-1} \sum \left(\frac{e_t^2 - \sigma^2}{\sqrt{2}\sigma^2} \right) \log y_{t-1} / \sigma \xrightarrow{p} \omega \sqrt{2} \int B_1(r) dB_2(r) \end{aligned} \tag{19}$$

upon substituting the expansion

$$\log y_{t-1} = \log(1 + z_{t-1} + \frac{1}{2}\frac{\gamma}{1-\gamma} z_{t-1}^2 + \cdots) = z_{t-1} + \frac{1}{2} z_{t-1}^2 + \frac{1}{2}\frac{\gamma}{1-\gamma} z_{t-1}^2 + \cdots \tag{20}$$

Noting that

$$\frac{\partial L}{\partial \omega^2} = T^{-1-d} \frac{\partial L}{\partial \sigma^2}, \tag{21}$$

it follows that

$$\begin{aligned} T^{-1/2}\frac{\partial L}{\partial \omega^2} &= T^{-1/2} T^{-1-d} \frac{\partial L}{\partial \sigma^2} = T^{-1/2} T^{-1-d} \frac{1}{2} \frac{\sum \left(e_t^2 - \sigma^2\right)}{\sigma^4} \\ &= \frac{1}{2\omega^2} \frac{T^{-1/2} \sum \left(e_t^2 - \sigma^2\right)}{\sigma^2} \xrightarrow{p} \frac{1}{\omega^2 \sqrt{2}} B_2(1). \end{aligned} \tag{22}$$

The limit of the first derivative vector (8) from (15), (17), (19) and (22) is then given as

$$Q^{-1}\begin{pmatrix} \partial L/\partial\delta \\ \partial L/\partial\beta \\ \partial L/\partial\gamma \\ \partial L/\partial\omega^2 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \omega^{-1}B_1(1) \\ \int B_1(r)dB_1(r) \\ \sqrt{2}\omega\int B_1(r)dB_2(r) \\ \dfrac{1}{\omega^2\sqrt{2}}B_2(1) \end{pmatrix},$$

where

$$Q = \begin{bmatrix} T^{1+d/2} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T^{1/2-d/2} & 0 \\ 0 & 0 & 0 & T^{1/2} \end{bmatrix}.$$

We can obtain the diagonal elements of the second derivative matrix as

$$T^{-2-d}\frac{\partial^2 L}{\partial\delta^2} = -T^{-2-d}\frac{\sum y_{t-1}^{-2\gamma}}{\sigma^2} \xrightarrow{p} -\omega^{-2}$$

$$T^{-2}\frac{\partial^2 L}{\partial\beta^2} = -T^{-2}\frac{\sum (y_{t-1}-1)^2 y_{t-1}^{-2\gamma}}{\sigma^2} \xrightarrow{p} -\int B_1^2(r)dr\,,$$

$$T^{-1+d}\frac{\partial^2 L}{\partial\gamma^2} = -2\omega^2 T^{-2}\frac{\sum e_t^2(\log y_{t-1})^2}{\sigma^4} \xrightarrow{p} -2\omega^2\int B_1(r)^2 dr\,,$$

$$T^{-1}\frac{\partial^2 L}{\partial(\omega^2)^2} = \left(\frac{T}{2\sigma^4} - \frac{\sum e_t^2}{\sigma^6}\right)T^{-3-2d} = \frac{1}{T^{-2-2d}\omega^4}\left(\frac{T}{2} - \frac{\sum e_t^2}{\sigma^2}\right)T^{-3-2d} \xrightarrow{p} -\frac{1}{2\omega^4}\,,$$

upon substituting (16), (18), and (21). Analogously, the off-diagonal elements are given as

$$T^{-1-d/2}\frac{\partial^2 L}{\partial\delta\partial\beta}=-T^{-1-d/2}\sum z_t y_{t-1}^{-2\gamma}/\sigma^2$$

$$=-\omega^{-1}T^{1/2+d/2}T^{-3/2}\sum z_t y_{t-1}^{-2\gamma}/\sigma \xrightarrow{p} -\omega^{-1}\int B_1(r)dr,$$

$$T^{-1+d/2}\frac{\partial^2 L}{\partial\beta\partial\gamma}=-2T^{-1+d/2}\frac{\sum e_t z_{t-1} y_{t-1}^{-\gamma}\log y_{t-1}}{\sigma^2}\xrightarrow{p} -2\omega\int B_1(r)^2 dB_1(r)\ ,$$

$$T^{-1}\frac{\partial^2 L}{\partial\delta\partial\gamma}=-2T^{-1}\frac{\sum e_t y_{t-1}^{-\gamma}\log y_{t-1}}{\sigma^2}\xrightarrow{p} -2\int B_1(r)dB_1(r),$$

$$T^{-1-d/2}\frac{\partial^2 L}{\partial\delta\partial\omega^2}=-T^{-1-d/2}T^{-1-d}\frac{\sum e_t y_{t-1}^{-\gamma}}{\sigma^4}=-T^{-1-d/2}T^{-1-d}\frac{T^{1/2}T^{-1/2}}{\omega^3 T^{-3/2-3d/2}}\frac{\sum e_t y_{t-1}^{-\gamma}}{\sigma}$$

$$\xrightarrow{p} -\omega^{-3}B_1(1),$$

$$T^{-1}\frac{\partial^2 L}{\partial\beta\partial\omega^2}=-T^{-1}T^{-1-d}\frac{\sum e_t(y_{t-1}-1)y_{t-1}^{-\gamma}}{\sigma^4}=-T^{-1}\frac{T^{-1-d}}{\omega^2 T^{-1-d}}\frac{\sum e_t(y_{t-1}-1)y_{t-1}^{-\gamma}}{\sigma^2}$$

$$\xrightarrow{p} -\omega^{-2}\int B_1(r)dB_1(r),$$

$$T^{-1+d/2}\frac{\partial^2 L}{\partial\gamma\partial\omega}=-T^{-1+d/2}\frac{\sum e_t^2\log y_{t-1}}{\sigma^4}T^{-1-d}$$

$$=-T^{-1+d/2}\omega^{-1}T^{1/2+d/2}\frac{\sum e_t^2\log y_{t-1}}{\sigma^3}T^{-1-d}$$

$$\xrightarrow{p} -\omega^{-1}\int B_1(r)dr,$$

$$T^{-1}\frac{\partial^2 L}{\partial\beta\partial\omega^2} = -T^{-1}T^{-1-d}\frac{\sum e_t(y_{t-1}-1)y_{t-1}^{-\gamma}}{\sigma^4} = -T^{-1}\frac{T^{-1-d}}{\omega^2 T^{-1-d}}\frac{\sum e_t(y_{t-1}-1)y_{t-1}^{-\gamma}}{\sigma^2}$$

$$\xrightarrow{p} -\omega^{-2}\int B_1(r)dB_1(r).$$

The second derivative matrix of the log-likelihood is asymptotically block-diagonal, as $T^{-3/2}\frac{\partial^2 L}{\partial\beta\partial\omega^2}$, $T^{-3/2-d/2}\frac{\partial^2 L}{\partial\delta\partial\omega^2}$, $T^{-3/2+d/2}\frac{\partial^2 L}{\partial\beta\partial\gamma}$ and $T^{-3/2}\frac{\partial^2 L}{\partial\delta\partial\gamma}$ converge to zero, and hence are negligible in $Q^{-1}KQ^{-1}$. Then, we have the probability limit of the second derivative matrix of the log-likelihood as

.

$$Q^{-1}KQ^{-1} \xrightarrow{p} -\begin{pmatrix} \omega^{-2} & \omega^{-1}\int B_1(r)dr & 0 & 0 \\ \omega^{-1}\int B_1(r)dr & \int B_1^2(r)dr & 0 & 0 \\ 0 & 0 & 2\omega^2\int B_1(r)^2 dr & \omega^{-1}\int B_1(r)dr \\ 0 & 0 & \omega^{-1}\int B_1(r)dr & \frac{1}{2\omega^4} \end{pmatrix}.$$

References

Adkins, L.C. and T. Krehbiel (1999), "Mean Reversion and Volatility of Short-term London Interbank Offer Rates: An Empirical Comparison of Competing Models", *International Review of Economics and Finance*, 8, 45-54.

Bickel, P.J. and K.A. Doksum (1981), "An Analysis of Transformations Revisited", *Journal of the American Statistical Association*, 76, 296-310.

Box, G.E.P. and D.R. Cox (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society* B, 26, 211-243.

Brenner, R.J.R., H. Harjes and K.F. Kroner (1996), "Another Look at Models of the Short-term Interest Rate", *Journal of Financial and Quantitative Analysis*, 31, 85-107.

Bliss, R.R. and D.C. Smith (1998), "The Elasticity of Interest Rate Volatility: Chan, Karolyi, Longstaff and Sanders Revisited", *Journal of Risk*, 1, 21-46.

Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders (1992), "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate", *Journal of Finance*, 47, 1209-1227.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-407.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74, 427-431.

Jarque, C. and A. Bera (1980), "Efficient Tests for Normality, Homoskedasticity, and Serial Independence of Regression Residuals", *Economics Letters*, 6, 255-259.

Kobayashi, M. and M. McAleer (1999), "Tests of Linear and Logarithmic Transformations for Integrated Processes", *Journal of the American Statistical Association*, 94, 860-868.

Koedijk, K.G., F.G.J. Nissen, P.C. Schotman and C.C.P. Wolfe (1997), "The Dynamics of Short-Term Interest Rate Volatility Reconsidered", *European Finance Review*, 1, 105-130.

McAleer, M. (2005), "Automated Inference and Learning in Modeling Financial Volatility", *Econometric Theory*, 21, 232-261.

McAleer, M. and M.C. Medeiros (2008), "Realized Volatility: A Review", *Econometric Reviews*, 27, 10-45.

Park, J.Y. and P.C.B. Phillips (1999), “Asymptotics for Nonlinear Transformations of Integrated Time Series”, *Econometric Theory*, 15, 269-298.

Park, J.Y. and P.C.B. Phillips (2001), “Nonlinear Regressions with Integrated Time Series”, *Econometrica*, 69, 117-161.

Rodrigues, P. and A. Rubia (2004), “Testing Non-stationarity in Short-term Interest Rates”, Technical Report, Faculty of Economics, University of Algarve.

Staiger, D. and J.H. Stock (1997), “Instrumental Variables Regression with Weak Instruments”, *Econometrica*, 65, 557-586.

Treepongkaruna, S. and S.F. Gray (2003), “On the Robustness of Short-rate Models”, *Accounting and Finance*, 43, 87-121.

Vasicek, O. (1977), “An Equilibrium Characterization of the Term Structure”, *Journal of Financial Economics*, 5, 177-158.

Yu, J. and P.C.B. Phillips (2001), “Gaussian Approach for Continuous Time Models of the Short Term Interest Rate”, *Econometrics Journal*, 4, 210-224.

.