INFERENCE IN STOCHASTIC VOLATILITY MODELS FOR GAUSSIAN AND T DATA

NAN ZHENG

Inference in Stochastic Volatility Models for Gaussian and t Data

by

©Nan Zheng

A thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirement for the Degree of

Master of Science in Statistics

Department of Mathematics and Statistics

Memorial University of Newfoundland

St. John's Newfoundland and Labrador, Canada July 2013

Abstract

Two competing analytical approaches, namely, the generalized method of moments (GMM) and quasi-maximum likelihood (QML) are widely used in statistics and econometrics literature for inferences in stochastic volatility models (SVMs). Alternative numerical approaches such as Markov chain Monte Carlo (MCMC), simulated maximum likelihood (SML) and Bayesian approaches are also available. All these later approaches are, however, based on simulations. Tagore (2010) revisited the analytical estimation approaches and proposed simpler and more efficient method of moments (MM) and approximate GQL (AGQL) inferences for the estimation of the volatility parameters. However, Tagore (2010) did not consider the estimation of the intercept parameter (γ_0) in the SV model, and also the model was confined to the normal based errors only.

In this thesis, we first extend Tagore's MM and AGQL approaches (Tagore 2010) to the estimation of all parameters of the SV model including the so-called intercept

parameter γ_0. Second, we modify the existing QML approach and unlike Tagore (2010) include this approach in the simulation study. Furthermore, all three approaches are applied to analyze a real life dataset.

Next, we consider a t-distribution based SV model, and apply the aforementioned estimation approaches for all parameters including a new degrees of freedom parameter of the t-distribution. Simulation studies are conducted to examine the relative performances of the estimation approaches. We also compute the kurtosis of the t-distribution based SV models and make an exact comparison with those of normal distribution based SV models. The estimation effect of parameters on the kurtosis is given for a special case.

Acknowledgements

I would like to express my deep gratitude to my supervisor Brajendra Sutradhar for highly thoughtful comments, remarks and engagements through the learning process of this thesis, including the writing of the thesis.

Thank the Department of Mathematics and Statistics of the Memorial University of Newfoundland for the facilities and scholarship.

I would like to thank my family for their constant love and support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background of the Problem

For many financial time series data, it is more important to study the variation in responses over time as opposed to studying the changes in the mean. For example, in financial problems dealing with exchange rates and stock returns, data may exhibit high variation in one time range, but low variation in another time range, and so on. This makes the variance of the responses non-stationary over time. To make prediction of the responses at a given time or future time, it is, therefore, important to understand the time series dynamics in variances. This non-stationary variation problem is referred to as the volatility problem. For the purpose, the variance of a response y_t at a desired time t conditional on its history of data is referred to as a volatility parameter. We will denote this conditional variance as σ_t^2. The modeling of the relationship between σ_t^2 and the variances from the past such as $\sigma_{t-\ell}^2$, ℓ being a suitable lag, is however not so easy. Many authors, such as Taylor (1986), Melino and Turnbull (1990), Taylor (1994), Harvey, Ruiz and Shephard (1994), Jacquier, Polson and Rossi (1994), Ruiz (1994), Harvey et.al (1994), Anderson and Sorensen (1996), and Mills (1999, p.127-128) have used a simple Gaussian type AR(1) relationship to model such non-stationary variation. To be specific, this simple model can be written as

$$y_t = \sigma_t \, \epsilon_t, \qquad t = 1, \ldots, T, \tag{1.1}$$

$$\log(\sigma_t^2) \equiv h_t = \gamma_0 + \gamma_1 \, h_{t-1} + \eta_t, \qquad t = 2, \ldots, T, \tag{1.2}$$

where in (1.1), ϵ_t's are independently and identically (iid) distributed with mean zero and variance one, i.e., $\epsilon_t \overset{iid}{\sim} (0,1)$. This non-stationary variance model (1.1) - (1.2) is referred to as the stochastic volatility (SV) model, where γ_0 is the intercept parameter, γ_1 is the volatility persistent parameter, and $\eta_t \overset{iid}{\sim} N(0, \sigma_\eta^2)$ with σ_η^2 being the measure of uncertainty about future volatility. Furthermore, ϵ_t's in (1.1) and η_t's in (1.2) are assumed to be independent. As a result, ϵ_t and σ_t are independent, and y_t should have zero mean (or shifted to zero). The initial variance σ_1^2 at time $t = 1$

can be reasonably modeled as

$$\log(\sigma_1^2) = h_1 \overset{iid}{\sim} N\left(\frac{\gamma_0}{1-\gamma_1}, \frac{\sigma_\eta^2}{1-\gamma_1^2}\right). \tag{1.3}$$

[Lee and Koopman (2004, eqn(1.1c))]. Note that this simple volatility model (1.1) - (1.2) ensures that the kurtosis of the responses defined by

$$\kappa_t(\gamma_0,\ \gamma_1,\ \sigma_\eta^2) = \frac{E(Y_t^4)}{[E(Y_t^2)]^2}, \tag{1.4}$$

would be larger than Gaussian assumption based kurtosis. Further note that a larger kurtosis in practice helps to understand that there can be some larger or outlying responses present in the data which in turn helps to understand the variation in variances. This raises issues to know the kurtosis which requires the efficient and consistent estimates for the parameters γ_0, γ_1 and σ_η^2 under the model (1.2).

Note that the original work of Taylor (1986) was confined to the univariate case, and it was extended by Harvey et al. (1994) [see also Harvey (2013, P. 8)] as well as Jacquier et al. (1995) and Shephard (1996) to a multivariate SV setup. To be specific, Harvey et al. (1994) proposed a multivariate SV models where ϵ_t and η_t all become multivariate normal random vectors with constant covariance matrices. Jacquier et al. (1995) and Shephard (1996) have considered multivariate factor SV models, where an emphasis is given to construct a small number of factors when the multivariate observation at a given time has large dimension.

Traditionally for the model (1.1) - (1.2), it is assumed that ϵ_t's follow $N(0,1)$ [Ruiz (1994), Harvey et.al (1994), Anderson and Sorensen (1996), and Mills (1999, p.127-128)]. Some authors such as Nelson (1988), Harvey et al. (1994), Barndorff-Nielsen (1997), Gallant et al. (1997), Mahieu and Schotman (1998), Sandmann and Koopman (1998), Steel (1998), Liesenfeld and Jung (2000), Anderson (2001) and Watanabe and Asai (2001) have extended the normality assumption for ϵ_t to the heavy tailed distributions. More specifically, these authors have used $\epsilon_t \overset{iid}{\sim} t_\nu(0,1)$, $t_\nu(0,1)$ being a t-distribution with ν degrees of freedom. Note that this t-distribution assumption for ϵ_t makes the $\kappa_t()$ in (1.4) much larger than normality based kurtosis. In turn, this will accommodate much more volatility in the data.

In model (1.1) - (1.2), σ_t responds to negative and positive returns (y_t) symmetrically. But in some practical situations, σ_t can respond to negative and positive returns asymmetrically, which is referred to as the leverage effect [Black (1986)]. In order to incorporate the leverage effect, Harvey and Shephard (1996) proposed an extension to the basic SV models where ϵ_t and η_{t+1} are negatively correlated, and So et al. (2002) developed a threshold SV model where two sets of model parameter values are considered and the model can be switched between them based on the reactions of the individuals according to the rising and falling of the response y_t. To further allow for long memory persistence in conditional variance, Breidt et al. (1998) and

Harvey (1998) proposed independently the long-memory SV (LMSV) model.

Note that as opposed to the SV model (1.1) - (1.2) where variances were modeled, there exist another modeling approach where σ_t^2 is considered to be related to both σ_s^2's and y_s's with $s < t$. For example, Bollerslev (1986) proposed the generalized autoregressive conditional heteroscedastic (GARCH) model of order m and v (GARCH(m, v)) given by

$$y_t = \sigma_t \, \epsilon_t \tag{1.5}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i y_{t-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2, \tag{1.6}$$

where again $\epsilon_t \overset{iid}{\sim} (0, 1)$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(m,v)} (\alpha_i + \beta_i) < 1$ with $\alpha_i = 0$ for $i > m$ and $\beta_j = 0$ for $j > v$. The difference among various volatility models relies on how σ_t^2 is related to σ_s^2's and y_s's with $s < t$. Further note that when $\alpha_i = 0$ for $i = 1, \cdots, m$, the model (1.5) - (1.6) becomes ARCH model which was introduced by Engle (1982). Due to the success of GARCH model and for alleviating its weakness, there appeared its revised versions such as the exponential GARCH (EGARCH) model of Nelson (1991), and the threshold GARCH (TGARCH) model of Glosten, Jagannathan and Runkle (1993) and Zakoian (1994). Even though the GARCH type models are popular in econometrics, there are still some weakness and disadvantages limiting their application in financial time series. For example, GARCH models require the volatility to be observable, but frequently in financial time series the volatility is not directly observable; moreover, when generalizing the univariate GARCH models to the multivariate cases, a proliferation of parameters can happen without a systematic and convincing approach to handle. For these reasons, even though some authors such as Harvey (2013, Chapter 4) considered t-error based GARCH model, we do not include this problem in the thesis.

Note that when GARCH models are compared to the SV models, the SV models assume from the beginning that volatility is unobservable, in agreement with the basic properties of many financial time series, and can be naturally generalized to multivariate cases. The SV models also capture the main empirical properties of many financial time series such as uncorrelated but dependent responses, more extreme values than normal case indicating higher kurtosis, and non-stationary variation over time. Furthermore, the SV models agree with and are the natural discretization of the modern continuous-time financial theory including the Black-Scholes theory and continuous-time Orstein-Uhlenbeck process, thus conceptually promising as a tool for applying modern financial theory to real data analysis. Therefore the SV models got popular soon in the area of financial time series, and also have many applications in econometrics.

In comparison with the GARCH model, the SV models involve an extra random variable η_t as shown in (1.2), increasing greatly the flexibility of the model in describing the evolution of σ_t^2, but also making the parameter estimation more difficult. For example, the responses $\{y_t\}$ are not conditionally Gaussian now, thus difficult to apply the maximum likelihood approach for parameter estimation. Nevertheless, there exist some widely used analytical estimation approaches, namely, the quasi-maximum likelihood (QML) of Nelson (1988) and Harvey et al. (1992), and the so-called generalized method of moments (GMM) of Anderson and Sorensen (1996). In QML, (1.1) is first written as

$$\log(y_t^2) \quad = \quad -1.27 + h_t + u_t, \tag{1.7}$$

where $u_t = \log \epsilon_t^2 + 1.27$ follows a $\log(\chi_1^2)$ distribution with $\mathrm{E}(u_t) = 0$ and $\mathrm{Var}(u_t) = \pi^2/2$. By treating $\{u_t\}$ as though it were $\mathrm{NID}(0, \pi^2/2)$, (1.7) and (1.2) form a linear state space model, with (1.7) being the measurement equation and (1.2) the transition equation. The Kalman filter can then be applied to obtain the prediction errors and their variances, which are used to construct the exact likelihood for conditionally Gaussian case, but only quasi-likelihood in this case where $\{\log(y_t^2)\}$ are not conditionally Gaussian. Then the parameters are estimated by numerically maximizing the resulting quasi-likelihood function. The approximations intrinsic to the QML estimators can make them biased and inefficient. As an illustration, Figure 1.1 gives a comparison between normal and $\log(\chi_1^2)$ densities, indicating that to approximate

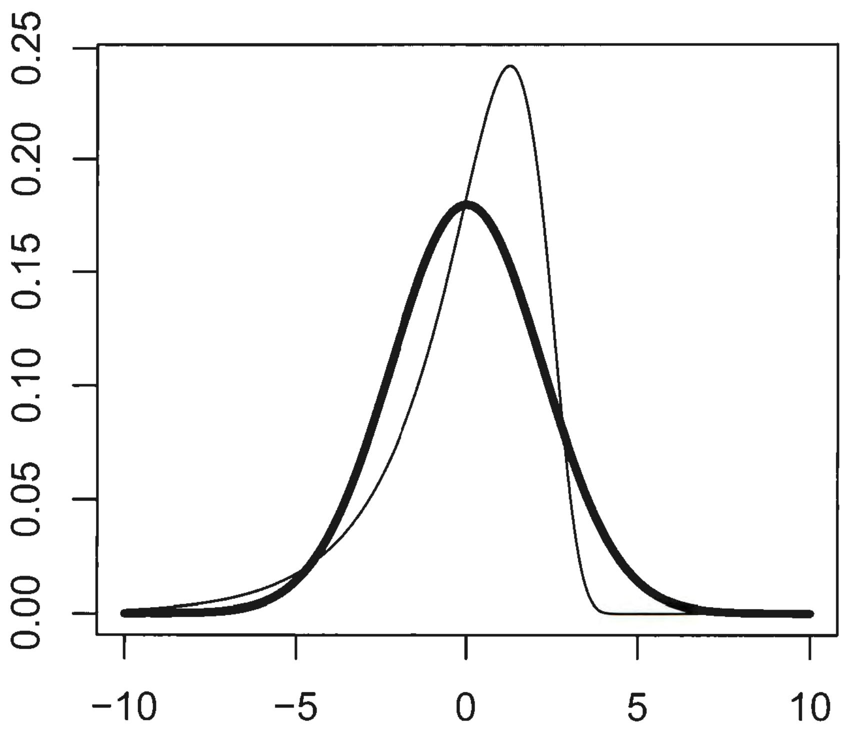

Figure 1.1: Comparison of the $\log(\chi_1^2)$ density (thin solid line) with the $N(0, \pi^2/2)$ density (thick solid line).

$\log(\chi_1^2)$ density with a normal density is rather inappropriate. In the GMM approach of Anderson and Sorensen (1996), with some arbitrariness, the authors first chose 24 unbiased moment functions to construct an 24×1 vector, then with some approximation and still some arbitrariness, the estimator of the covariance matrix of the vector was selected . The GMM estimating equation could then be written out and solved iteratively. This approach is too complicated and algebraically painstaking without showing any substantial efficiency gain in estimation [Anderson and Sorensen (1997), Ruiz (1997)] over other competing approaches such as the QML estimation approach, and the lack of guidelines for selecting the moment functions further makes this approach unconvincing and less favorable.

Note that to obtain consistent estimates in a finite sample set up (i.e, for a time series with moderate length), as opposed to the GMM and QML approaches, there exist several numerical approaches such as Bayesian approach by Jacquier et.al (1994) and the simulated ML (SML) approach which is considered to be an improvement over the so-called Markov chain Monte Carlo (MCMC) approach. For SML approach, we refer to Danielsson (1994), Shephard and Pitt (1997), Durbin and Koopman (1997), Liesenfel and Richard (2003) and Lee and Koopman (2004). It is, however, recognized that these numerical techniques are computationally intensive. For this reason, and also because in practice such as in financial or environmental analysis one may encounter a large time series, similar to Anderson and Sorensen (1996, 1997), Tagore (2010) revisited the basic SV model (1.1) - (1.2) and proposed simpler alternative inferences for the volatility parameters, as compared to the existing GMM [Anderson and Sorensen (1996)] and QML [Nelson (1988) and Harvey et al. (1992)] approaches. However, the author did not consider the estimation of the intercept parameter (γ_0) in the SV model, and also the model was confined to the normal based errors. In the thesis, we further pursue this inference problem to accommodate these issues.

1.2 Objective of the thesis

Because the intercept parameter γ_0 in the SV model (1.2) is important to understand the magnitude of time dependent variances, in this thesis, we further revisit the normality based SV model and inference procedures studied by Tagore (2010), and extend the estimation to include the intercept parameter γ_0. This inference for full family of parameters is then generalized to achieve our main goal where we will deal with t-distribution based SV model in order to accommodate much larger kurtosis. To be precise and clear, the specific objectives of the thesis are as follows.

In Chapter 2, we first modify the normality based MM inferences due to Tagore (2010) to accommodate the estimation of the additional parameter γ_0 (intercept).

The GQL approach considered by Tagore (2010) is also modified to include the estimation of γ_0. Note that even though QML approach was included in discussion, Tagore (2010) however did not study its relative finite sample performance with MM and GQL approaches. In this chapter, we propose a simplification to the existing QML estimating equations by approximating the associated covariance matrix with a simpler tri-diagonal matrix. All three approaches, i.e., MM, GQL and modified QML are compared through a simulation study for the estimation of all parameters including the intercept parameter γ_0. The applications of these three methods are also illustrated by analyzing a real life data set on exchange rates of some popular currencies.

In Chapter 3, we generalize the above normality based estimation approaches to the inferences for the t-distribution based SV model parameters. This generalization helps to make inferences for more volatile data as compared to normality based volatile data. Therefore, to understand the reflection of larger volatility, we study the kurtosis, and estimate them, using a t-distribution based SV model. Note however that this generalization is challenging because of the need of estimation of a further degrees of freedom parameter which reflects the heavy tails of the data as compared to standard normality based volatile data. The relative performances of the MM, GQL and modified QML approaches are compared through a new simulation study.

The thesis concludes in Chapter 4.

Chapter 2

Parameter Estimation for Gaussian Volatility Models

Under the assumption that ϵ_t $(t = 1, \cdots, T)$ in (1.1) follow the standard normal distribution, many authors have studied the inference for the parameters involved in the model (1.1) - (1.2). Recently, Tagore (2010) has dealt with the same model but proposed a simpler method of moments (MM) for the estimation of γ_1 and σ_η^2, as compared to the existing GMM and QML approaches. However, the author did not address the estimation of the intercept parameter γ_0 in (1.2). Thus, the inference remains incomplete. As pointed out in Chapter 1, in this chapter we revisit the same SV model and following Tagore (2010) we carry out the same simple MM estimation and discuss the estimation of all three parameters, namely γ_0, γ_1 and σ_η^2. For this purpose, we first briefly describe the estimation proposed by Tagore (2010) which we refer to as partial inference for SV model.

2.1 Partial Inference for SV Models

Unlike using 24 or 34 moments by GMM technique (Anderson and Sorensen (1996)), Tagore (2010) has used only two unbiased estimating equations for the estimation of two parameters γ_1 and σ_η^2.

2.1.1 Unbiased Moment Estimating Equation for σ_η^2

Tagore (2010) considers an intercept free volatility model, that is, chooses $\gamma_0 = 0$ in the volatility model (1.2). Furthermore, even though $\log\sigma_1^2$ is supposed to be random following $N\left(0, \frac{\sigma_\eta^2}{1-\gamma_1^2}\right)$ as shown in (1.3), for convenience, one may choose a small value close to mean zero for this initial variance. In order to develop an estimating equation for σ_η^2, it is clear from (1.2) that σ_t^2's maintain a non-stationary dynamic relationship, stationarity being a special case. Now if σ_t^2's were stationary, that is, $E[\sigma_t^2] = h^*(\sigma_\eta^2)$, a suitable constant function of σ_η^2, then one would have estimated

$h^*(\sigma_\eta^2)$ consistently by using the raw second order moment

$$S_1 = \frac{1}{T}\sum_{t=1}^{T} y_t^2, \tag{2.1}$$

$E[Y_t|\sigma_t^2]$ being zero. However, in the present case, σ_t^2's are unobservable and their log values satisfy a non stationary Gaussian AR(1) type relationship given by (1.2), with errors $\eta_t \overset{iid}{\sim} N(0, \sigma_\eta^2)$. This leads to the expected value of S_1 as

$$\begin{aligned} E[S_1] &= E_{\sigma_t^2} E[\frac{1}{T}\sum_{t=1}^{T} y_t^2 \mid \sigma_t^2] \\ &= \frac{1}{T}\left[\sigma_1^2 + \sum_{t=2}^{T} \exp\left(\gamma_1^{t-1}\log\sigma_1^2 + \frac{\sigma_\eta^2}{2}\sum_{r=0}^{t-2}\gamma_1^{2r}\right)\right] \\ &= g_1(\gamma_1, \sigma_\eta^2, \sigma_1^2), \quad say. \end{aligned} \tag{2.2}$$

Thus, for given γ_1 and σ_1^2, Tagore solved the unbiased estimating equation

$$S_1 - g_1(\gamma_1, \sigma_\eta^2, \sigma_1^2) = 0 \tag{2.3}$$

to obtain a consistent estimate for σ_η^2.

Note that the solution of (2.3) requires a good initial value for σ_η^2, which Tagore suggests to obtain by solving an asymptotic unbiased estimating equation. It is clear that for a suitable large T_0, for any $t > T_0$, $\gamma_1^{t-1} \to 0$ for $|\gamma_1| < 1$. Further, the expectation of y_t^2 for $t > T_0$ becomes stationary, producing

$$\begin{aligned} \lim_{t\to\infty} E[Y_t^2] &= \exp\left[\frac{\sigma_\eta^2}{2}\left(\frac{1}{1-\gamma_1^2}\right)\right] \\ &= g_{10}(\gamma_1, \sigma_\eta^2), \quad say. \end{aligned} \tag{2.4}$$

Thus, if the series under consideration is large, i.e, $T \to \infty$, replacing the exact $E[Y_t^2]$ for $t > T_0$ by $\lim_{t\to\infty} E[Y_t^2] = g_{10}(\gamma_1, \sigma_\eta^2)$, one can consistently estimate the stationary mean function, namely $g_{10}(.)$ by using $S_{10} = \frac{1}{T-T_0}\sum_{t=T_0+1}^{T} y_t^2$. Consequently, for known $\gamma_1 = \gamma_1(0)$, we may obtain a very reasonable initial value for σ_η^2 by solving

$$S_{10} - g_{10}(\gamma_1, \sigma_\eta^2) \;=\; 0\,. \tag{2.5}$$

We denote this initial value of σ_η^2 by $\sigma_\eta^2(0)$.

2.1.2 Unbiased Moment Estimating Equation for γ_1

Next, to construct an unbiased estimating equation for γ_1, one notices that γ_1 is the lag 1 dependence parameter in the Gaussian AR(1) model (1.2). Tagore, therefore, chooses a lag 1 based function given by

$$S_2 \;=\; \frac{1}{T-1}\sum_{t=2}^{T} y_{t-1}^2 y_t^2 \tag{2.6}$$

to construct the moment equation for γ_1. For this purpose, the expected value of S_2 is computed as

$$
\begin{aligned}
E[S_2] &= E_{\sigma_t^2} E[\frac{1}{T-1} \sum_{t=2}^{T} y_{t-1}^2 y_t^2] \\
&= \frac{1}{T-1} \Bigg[\sigma_1^2 \ \exp\left(\gamma_1 \log \sigma_1^2 + \frac{\sigma_\eta^2}{2} \right) \\
&\quad + \sum_{t=3}^{T} \exp\left(\gamma_1^{t-1} \log \sigma_1^2 + \gamma_1^{t-2} \log \sigma_1^2 + \frac{\sigma_\eta^2}{2} \{ (1+\gamma_1)^2 \sum_{l=0}^{t-3} \gamma_1^{2l} + 1 \} \right) \Bigg] \\
&= g_2(\gamma_1, \sigma_\eta^2, \sigma_1^2), \quad say,
\end{aligned} \tag{2.7}
$$

and the unbiased estimating equation

$$
S_2 - g_2(\gamma_1, \sigma_\eta^2, \sigma_1^2) = 0 \tag{2.8}
$$

is solved iteratively for γ_1.

Remark on Large Sample Estimation

When time series is quite long such as $T = 5000$ or more, one may obtain much simpler estimating equations than used in Sections 2.1.1 and 2.1.2.

Note that an initial value of σ_η^2 was obtained by (2.5) based on large sample. In fact, for large sample cases, one can always use this estimate. Thus, when large sample based estimate for γ_1 is available, we estimate σ_η^2 by using

$$
\tilde{\sigma}_\eta^2 = 2(1 - \tilde{\gamma}_1^2) \log(S_{10}), \tag{2.9}
$$

where $\tilde{\gamma}_1$ is the large sample based estimate for γ_1 to be obtained as follows.

To construct an asymptotic moment function based estimating equation for γ_1, we use $S_{20} = \dfrac{1}{T - T_0 - 1} \sum_{t=T_0+1}^{T} y_{t-1}^2 y_t^2$, and similar to (2.4), compute the asymptotic expectation of S_{20} by using

$$\begin{aligned} \lim_{t\to\infty} E[y_{t-1}^2 y_t^2] = \lim_{t\to\infty} g_2(\gamma_1, \sigma_\eta^2, \sigma_1^2) &= \exp\left[\frac{\sigma_\eta^2}{1-\gamma_1}\right] \\ &= g_{20}(\gamma_1, \sigma_\eta^2), \quad say, \end{aligned} \tag{2.10}$$

where the formula for $g_2(\gamma_1, \sigma_\eta^2, \sigma_1^2)$ is given in (2.7).

Hence the asymptotic moment based unbiased estimating equation for γ_1 is written as

$$S_{20} - g_{20}(\gamma_1, \sigma_\eta^2) = 0, \tag{2.11}$$

which has a closed-form solution

$$\tilde{\gamma}_1 = 1 - \frac{\sigma_\eta^2}{\ln(S_{20})}. \tag{2.12}$$

The improved value of γ_1 from (2.12) is then used in (2.9) for obtaining improved estimate for σ_η^2. This cycle of iterations continues until convergence.

2.2 Inferences for Complete SV models

Unlike in Section 2.1, we now include γ_0 in our estimation. Thus, we develop three estimating equations for three parameters, namely, γ_0, σ_η^2 and γ_1, as follows.

2.2.1 MM for All Parameters

In the presence of γ_0, we first generalize the MM estimating equations for γ_1 (2.7) and σ_η^2 (2.2), which now contain γ_0 as well, as given by

$$\begin{aligned} E[S_2] &= E_{\sigma_t^2} E[\frac{1}{T-1}\sum_{t=2}^{T} y_{t-1}^2 y_t^2] \\ &= \frac{e^{\gamma_0}}{T-1}\Bigg[\sigma_1^2 \ \exp\left(\gamma_1 \log\sigma_1^2 + \frac{\sigma_\eta^2}{2}\right) \\ &\quad + \sum_{t=3}^{T} \exp\Bigg(\gamma_0(1+\gamma_1)\sum_{l=0}^{t-3}\gamma_1^l + \gamma_1^{t-1}\log\sigma_1^2 + \gamma_1^{t-2}\log\sigma_1^2 \\ &\quad + \frac{\sigma_\eta^2}{2}\{(1+\gamma_1)^2 \sum_{l=0}^{t-3}\gamma_1^{2l} + 1\}\Bigg)\Bigg] \\ &= g_2(\gamma_0, \gamma_1, \sigma_\eta^2, \sigma_1^2), \quad say, \end{aligned} \tag{2.13}$$

and

$$\begin{aligned} E[S_1] &= E_{\sigma_t^2} E[\frac{1}{T}\sum_{t=1}^{T} y_t^2] \\ &= \frac{1}{T}\left[\sigma_1^2 + \sum_{t=2}^{T}\exp\left(\gamma_1^{t-1}\log\sigma_1^2 + \gamma_0\sum_{i=0}^{t-2}\gamma_1^i + \frac{\sigma_\eta^2}{2}\sum_{r=0}^{t-2}\gamma_1^{2r}\right)\right] \\ &= g_1(\gamma_0, \gamma_1, \sigma_\eta^2, \sigma_1^2), \quad say. \end{aligned} \tag{2.14}$$

Because γ_0 is unknown, it is important to estimate this parameter consistently. For this purpose, define

$$S_3 = \frac{1}{T}\sum_{t=1}^{T}\log\left(y_t^2\right), \tag{2.15}$$

and notice that

$$\mathrm{E}(S_3) = \left(\frac{\gamma_0}{1-\gamma_1}\right)\left(1-\frac{1}{T}\frac{1-\gamma_1^T}{1-\gamma_1}\right)+\frac{1}{T}h_1\left(\frac{1-\gamma_1^T}{1-\gamma_1}\right)+\kappa_1,$$

where $\kappa_1 = E\left(\log \epsilon_1^2\right) = -1.270363$. By equating $\mathrm{E}(S_3)$ to S_3 and solving for γ_0, we obtain the MM estimating equation for γ_0 as

$$\hat{\gamma}_0 = \frac{\left\{(S_3-\kappa_1)(1-\gamma_1)-\frac{1}{T}\log(\sigma_1^2)\left(1-\gamma_1^T\right)\right\}(1-\gamma_1)}{1-\gamma_1-\frac{1}{T}\left(1-\gamma_1^T\right)}. \tag{2.16}$$

For convenience of computation, we provide an estimation algorithm as follows.

Step 1. For a small initial value $\gamma_1 = \gamma_1(0)$, we calculate $\gamma_0(0)$ by (2.16) and then choose the initial value of $\sigma_\eta^2 = \sigma_\eta^2(0)$ by solving the asymptotic unbiased estimating equation (2.5). To be specific,

$$\sigma_\eta^2(0) = 2\left(\ln(S_{10})-\frac{\gamma_0(0)}{1-\gamma_1(0)}\right)(1-\gamma_1^2(0)). \tag{2.17}$$

Step 2. Once the initial values are chosen/computed as in Step 1, we solve $S_2 - g_2(\gamma_0,\gamma_1,\sigma_\eta^2,\sigma_1^2) = 0$ iteratively to obtain an improved value for γ_1. Note that in large sample case, one may ignore σ_1^2 or put a small value. The iterative equation has the form

$$\hat{\gamma}_1(r+1) = \hat{\gamma}_1(r) + \left[\left(\frac{\partial g_2(\gamma_0,\gamma_1,\sigma_\eta^2,\sigma_1^2)}{\partial\gamma_1}\right)^{-1}\left(S_2 - g_2(\gamma_0,\gamma_1,\sigma_\eta^2,\sigma_1^2)\right)\right]_{[r]}, \tag{2.18}$$

where $\hat{\gamma}_1(r)$ is the value of γ_1 at the rth iteration, and $[.]_{[r]}$ denotes that the value of the expression in the square bracket is evaluated at $\gamma_1 = \hat{\gamma}_1(r)$.

Step 3. In this step, we estimate γ_0. For this, we use the estimate of γ_1 obtained from Step 2 in MM equation (2.16) for γ_0.

Step 4. The estimates of γ_0 and γ_1 obtained from Steps 2 and 3 are then used to solve $S_1 - g_1(\gamma_0, \gamma_1, \sigma_\eta^2, \sigma_1^2) = 0$ iteratively to obtain an improvement over $\sigma_\eta^2(0)$, computed in Step 1. The iterative equation has the form

$$\hat{\sigma}_\eta^2(r+1) = \hat{\sigma}_\eta^2(r) + \left[\left(\frac{\partial g_1(\gamma_0, \gamma_1, \sigma_\eta^2, \sigma_1^2)}{\partial \sigma_\eta^2} \right)^{-1} \left(S_1 - g_1(\gamma_0, \gamma_1, \sigma_\eta^2, \sigma_1^2) \right) \right]_{[r]}, \quad (2.19)$$

where $\hat{\sigma}_\eta^2(r)$ is the value of σ_η^2 at the rth iteration, and $[.]_{[r]}$ denotes that the value of the expression in the square bracket is evaluated at $\sigma_\eta^2 = \hat{\sigma}_\eta^2(r)$.

This 4 steps cycle of iterations continues until convergence. Let the final estimates for γ_0, γ_1 and σ_η^2 be denoted by $\hat{\gamma}_{0,MM}$, $\hat{\gamma}_{1,MM}$ and $\hat{\sigma}_{\eta,MM}^2$, respectively.

2.2.2 An Approximate GQL (AGQL) Approach for the Estimation of Parameters

Tagore has also used an approximation to construct a GQL approach following Sutradhar [Sutradhar (2003, Section 3.1)] for the estimation of two parameters γ_1 and σ_η^2. To save space we do not reproduce these equations. Rather, we now modify these equations by accommodating γ_0. Note that γ_0 will be estimated as before following the moment equation (2.16). For the purpose of approximation, we first write the

GQL estimating equations for γ_1 and σ_η^2 as functions of γ_0 as follows.

Consider two basic vectors of statistics as

$$\mathrm{u} = \left[y_1^2, \ldots, y_t^2, \ldots, y_T^2\right]', \quad \text{and} \quad \mathrm{v} = \left[y_1^2 y_2^2, \ldots, y_{t-1}^2 y_t^2, \ldots, y_{T-1}^2 y_T^2\right]', \tag{2.20}$$

respectively. Let

$$\begin{aligned} \lambda\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) &= E[U] = [\lambda_1, \ldots, \lambda_t, \ldots, \lambda_T]' \\ \psi\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) &= E[V] = [\psi_{12}, \ldots, \psi_{t-1,t}, \ldots, \psi_{T-1,T}]', \end{aligned} \tag{2.21}$$

and

$$\Sigma\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{Cov}(U), \qquad \Omega\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{Cov}(V). \tag{2.22}$$

One may then write the GQL estimating equations [Sutradhar (2004)] for σ_η^2 and γ_1 as

$$\frac{\partial \lambda'\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \sigma_\eta^2} \Sigma^{-1}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)\left(\mathrm{u} - \lambda\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)\right) = 0, \tag{2.23}$$

$$\frac{\partial \psi'\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \gamma_1} \Omega^{-1}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)\left(\mathrm{v} - \psi\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)\right) = 0 \tag{2.24}$$

respectively.

In (2.21),

$$\lambda_t\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = E[Y_t^2] = E_{\sigma_t^2} E[Y_t^2|\sigma_t^2]$$
$$= \begin{cases} \sigma_1^2 & \text{for } t{=}1 \\ \exp\left[\frac{\gamma_0}{1-\gamma_1} + \gamma_1^{t-1}(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}) + \dfrac{\sigma_\eta^2}{2}\displaystyle\sum_{r=0}^{t-2}\gamma_1^{2r}\right] & \text{for } t = 2, \ldots, T \end{cases} \tag{2.25}$$

[see also (2.2)], and

$$\psi_{t-1,t}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = E[Y_{t-1}^2 Y_t^2] = E_{\sigma_{t-1}^2, \sigma_t^2} E[Y_{t-1}^2 Y_t^2|\sigma_{t-1}^2, \sigma_t^2]$$
$$= \begin{cases} \sigma_1^2 \;\; \exp\left[\gamma_1 \log\sigma_1^2 + \dfrac{\sigma_\eta^2}{2}\right] & \text{for } t{=}2 \\ \exp\Bigg[\frac{2\gamma_0}{1-\gamma_1} + \gamma_1^{t-2}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) + \gamma_1^{t-1}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) & \\ \quad + \dfrac{\sigma_\eta^2}{2}\left((1+\gamma_1)^2 \displaystyle\sum_{l=0}^{t-3}\gamma_1^{2l} + 1\right)\Bigg] & \text{for } t = 3, \ldots, T. \end{cases} \tag{2.26}$$

Note that the diagonal elements of $\Sigma\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ matrix may be computed by using

$$\sigma_{tt}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{Var}[Y_t^2]$$
$$= \begin{cases} 3\sigma_1^4 - \lambda_1^2 & \text{for } t = 1 \\ 3\exp\left[\frac{2\gamma_0}{1-\gamma_1} + 2\gamma_1^{t-1}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) + 2\,\sigma_\eta^2[\sum_{r=0}^{t-2}\gamma_1^{2r}]\right] & \\ \quad -\lambda_t^2\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) & \text{for } t = 2, \ldots, T, \end{cases} \tag{2.27}$$

and for lag $k = 0,, \ldots, T-1$, the off-diagonal elements of the Σ matrix may be computed by using

$$\sigma_{t-k,t}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{Cov}[Y_t^2, Y_{t-k}^2]$$
$$= \begin{cases} \sigma_1^2\lambda_t\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) - \lambda_1\lambda_t\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) & \text{for } t = 1, \\ \exp\left[\frac{2\gamma_0}{1-\gamma_1} + \gamma_1^{t-1}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) + \gamma_1^{t-k-1}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right)\right. & \\ \left. +\frac{\sigma_\eta^2}{2}\left((1+\gamma_1^k)^2\sum_{l=0}^{t-k-2}\gamma_1^{2l} + \sum_{r=0}^{k-1}\gamma_1^{2r}\right)\right] - \lambda_{t-k}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)\lambda_t\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) & \\ & \text{for } t = 2, \ldots, T. \end{cases} \tag{2.28}$$

However, unlike the elements of $\Sigma\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$, the computation of the elements of $\Omega\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ is very complicated, resulting in further difficulties for $\Omega^{-1}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$.

As a remedy to the numerical difficulty, similar to Tagore (2010), we now provide an approximation to the GQL estimating equation (2.23) and (2.24), where we ignore the off-diagonal elements of both $\Sigma\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ and $\Omega\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ matrices.

Note that (2.25) and (2.27) give the kurtosis $\kappa_t(\gamma_1, \sigma_\eta^2)$ as

$$\kappa_t(\gamma_1, \sigma_\eta^2) = \frac{E\left(Y_t^4\right)}{\left[E\left(Y_t^2\right)\right]^2} = 3\exp\left[\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right] \tag{2.29}$$

Approximate GQL Equations:

When off-diagonal elements are ignored, we replace $\Sigma\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ by $\Sigma_d\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$, say, where

$$\Sigma_d\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{diag}[\mathrm{Var}(Y_1^2), \ldots, \mathrm{Var}(Y_t^2), \ldots, \mathrm{Var}(Y_T^2)], \tag{2.30}$$

with $\mathrm{Var}(Y_t^2) = \sigma_{tt}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ as in (2.27). Similarly, $\Omega\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ will be replaced by $\Omega_d\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$, where

$$\Omega_d\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = \mathrm{diag}[\mathrm{Var}(Y_1^2 Y_2^2), \ldots, \mathrm{Var}(Y_{t-1}^2 Y_t^2), \ldots, \mathrm{Var}(Y_{T-1}^2 Y_T^2)]. \tag{2.31}$$

In (2.31),

$$Var[Y_{t-1}^2 Y_t^2] =$$

$$\begin{cases} 9\,\sigma_1^4 \exp\left[\frac{2\gamma_0}{1-\gamma_1} + 2\gamma_1\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) + 2\,\sigma_\eta^2\right] - \psi_{12}^2\left(\gamma_0,\gamma_1,\sigma_\eta^2\right) & \text{for } t\text{=}2 \\ 9\ \exp\left[\frac{4\gamma_0}{1-\gamma_1} + 2\left(\gamma_1^{t-2}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right) + \gamma_1^{t-1}\left(\log\sigma_1^2 - \frac{\gamma_0}{1-\gamma_1}\right)\right)\right. & \\ \left. +2\,\sigma_\eta^2\left((1+\gamma_1)^2\ \sum_{l=0}^{t-3}\gamma_1^{2l} + 1\right)\right] - \psi_{t-1,t}^2\left(\gamma_0,\gamma_1,\sigma_\eta^2\right) & \text{for } t = 3,\ \ldots, T, \end{cases} \tag{2.32}$$

where $\psi_{t-1,t}\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)$ is given in (2.26).

The approximation based GQL (AGQL) estimating equations have the forms:

$$\frac{\partial\lambda'\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)}{\partial\sigma_\eta^2}\Sigma_d^{-1}\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)\left(\mathrm{u} - \lambda\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)\right) = 0, \tag{2.33}$$

and

$$\frac{\partial\psi'\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)}{\partial\gamma_1}\Omega_d^{-1}\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)\left(\mathrm{v} - \psi\left(\gamma_0,\gamma_1,\sigma_\eta^2\right)\right) = 0, \tag{2.34}$$

for σ_η^2 and γ_1, respectively.

Note that γ_0 is estimated by MM as in (2.16). Thus, after each estimation of γ_1 by (2.34), $\hat{\gamma}_0$ has to be updated by (2.16).

Let the final estimates obtained from (2.16), (2.34) and (2.33) be denoted by $\hat{\gamma}_{0,AGQL}$, $\hat{\gamma}_{1,AGQL}$ and $\hat{\sigma}^2_{\eta,AGQL}$ respectively.

The formulas for the first derivatives in (2.34) and (2.33) are given as follows: The first derivatives of $\psi_{t-1,t}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$ w.r.t γ_1 for $t = 2$ is given by

$$\frac{\partial \psi_{1,2}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \gamma_1} = \psi_{1,2}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) \log \sigma_1^2,$$

and for a general $t = 3, \ldots, T$, the derivative has the expression as

$$\begin{aligned}
\frac{\partial \psi_{t-1,t}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \gamma_1} &= \psi_{t-1,t}\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) \left[\frac{\gamma_0}{(1-\gamma_1)^2}\left[2-(1+\gamma_1)\gamma_1^{t-2}\right]\right. \\
&\quad -\frac{\gamma_0}{1-\gamma_1}\left[(t-2)(1+\gamma_1)\gamma_1^{t-3}+\gamma_1^{t-2}\right] \\
&\quad (t-2)\gamma_1^{t-3}\log\sigma_1^2+(t-1)\gamma_1^{t-2}\log\sigma_1^2 \\
&\quad \left.+\frac{\sigma_\eta^2}{2}\left(2\left(1+\gamma_1\right)\sum_{l=0}^{t-3}\gamma_1^{2l}+(1+\gamma_1)^2\sum_{l=0}^{t-3}(2l)\,\gamma_1^{(2l-1)}\right)\right].
\end{aligned}$$

The derivative of λ_t w.r.t σ_η^2 is

$$\frac{\partial \lambda_t\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \sigma_\eta^2} = \begin{cases} 0 & \text{for } t = 1 \\ \frac{1}{2}\lambda_t \sum_{r=0}^{t-2}\gamma_1^{2r} & \text{for } t = 2, \ldots, T. \end{cases}$$

We remark that the AGQL estimating equations for γ_1 (2.34) and σ_η^2 (2.33) are similar to the well known weighted least square (WLS) equations for the corresponding parameters.

2.2.3 A Modified QML (MQML) Approach for All Parameters

Note that so far we have made improvements over the MM and AGQL estimation done by Tagore (2010) by accommodating γ_0 parameter. We have also estimated this γ_0 parameter by the method of moment. As mentioned earlier, there also exists another widely used so-called QML (quasi Maximum likelihood) approach which estimates the three parameters, γ_0, γ_1 and σ_η^2, through a Kalman filtering approach, where predicated errors are used to form a conditional likelihood to obtain the likelihood equations. The existing QML approach is numerically not so cumbersome as compared to the GMM approach. However, because of normal approximation to the log chi-square distribution, it may not produce efficient estimation in all possible situations.

Further note that for an asymptotic comparison with the proposed simpler MM approach, Tagore (2010) has simplified QML equations for the estimation of γ_1 and σ_η^2. We however find that these estimating equations in Tagore (2010) can be improved for numerical approximation by modifying the so-called covariance matrix $\mathbf{\Phi}$ to be discussed below. This modification mainly aims to use a band form for this $\mathbf{\Phi}$ matrix. Furthermore, unlike Tagore (2010), we include γ_0 parameter in the proposed modified

QML approach. For the purpose, we first describe the existing QML approach as follows.

Existing QML Approach Including γ_0

We first rewrite (1.1) as

$$z_t = \log(y_t^2) = h_t + \log(\epsilon_t^2), \qquad t = 1, \cdots, T, \tag{2.35}$$

where $h_t = \log \sigma_t^2$, compute $\mathrm{E}(z_t) = m_t$ and $\mathrm{Cov}(z_u, z_t) = \Phi_{ut}$, and form

$$\mathbf{z} = (z_1, \cdots, z_t, \cdots, z_T)',$$

$$\mathbf{m} = (m_1, \cdots, m_t, \cdots, m_T)',$$

and

$$\mathbf{\Phi} = \mathrm{Cov}(\mathbf{z}) = (\Phi_{ut}).$$

It then follows that the quasi-maximum likelihood (QML) approach solves a normality based pseudo likelihood equation, where the log pseudo likelihood function has the form

$$\log L_Q^*\left(\gamma_0, \gamma_1, \sigma_\eta^2\right) = c_0 - \frac{1}{2}\log|\mathbf{\Phi}| - \frac{1}{2}[(\mathbf{z}-\mathbf{m})'\,\mathbf{\Phi}^{-1}\,(\mathbf{z}-\mathbf{m})], \tag{2.36}$$

[Shephard, 1996 eq:1.17]. In (2.36), c_0 is a normalization constant. The functions m_t and Φ_{ut} may be computed as in Tagore (2010) except that now m_t will involve γ_0.

The new formula for m_t and Φ_{ut} are as follows:

$$m_1 = E(z_1) = h_1 + E\left[\log(\epsilon_1^2)\right] = h_1 + \kappa_1,$$

and for $t \geq 2$

$$m_t = E(z_t) = \frac{\gamma_0\left(1-\gamma_1^{t-1}\right)}{1-\gamma_1} + \gamma_1^{t-1}h_1 + \kappa_1,$$

where $\kappa_1 = E\left[\log(\epsilon_1^2)\right] = -1.270363$.

$$\Phi_{11} = \mathrm{var}(z_1) = \mathrm{var}(h_1 + \log(\epsilon_1^2)) = \kappa_2,$$

where $\kappa_2 = \mathrm{var}(\log(\epsilon_1^2)) = \pi^2/2$, and for $t \geq 2$

$$\begin{aligned}
\Phi_{1t} &= \mathrm{cov}(z_1, z_t) = E\left[(z_1 - E(z_1))(z_t - E(z_t))\right] \\
&= E\left[\left(\log(\epsilon_1^2) - \kappa_1\right)\left(\eta_t + \eta_{t-1}\gamma_1 + \cdots + \eta_2\gamma_1^{t-2} + (\log(\epsilon_t^2) - \kappa_1)\right)\right] \\
&= 0, \\
\Phi_{tt} &= \mathrm{var}(z_t) = E\left[(z_t - E(z_t))^2\right] = E\left[\left(h_t - E(z_t) + \log(\epsilon_t^2)\right)^2\right] \\
&= E\left[\left\{\eta_t + \eta_{t-1}\gamma_1 + \cdots + \eta_2\gamma_1^{t-2} + (\log(\epsilon_t^2) - \kappa_1)\right\}^2\right] \\
&= \sigma_\eta^2\left[1 + \gamma_1^2 + \cdots + \gamma_1^{2(t-2)}\right] + \kappa_2 \\
&= \sigma_\eta^2\left[\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right] + \kappa_2,
\end{aligned}$$

and for $2 \leq u < t$

$$\begin{aligned}
\Phi_{ut} &= \operatorname{cov}(z_u, z_t) = E\left[(z_u - E(z_u))(z_t - E(z_t))\right] \\
&= E\left[\left\{\eta_u + \eta_{u-1}\gamma_1 + \cdots + \eta_2\gamma_1^{u-2} + (\log(\epsilon_u^2) - \kappa_1)\right\}\right. \\
&\qquad \left.\left\{\eta_t + \eta_{t-1}\gamma_1 + \cdots + \eta_2\gamma_1^{t-2} + (\log(\epsilon_t^2) - \kappa_1)\right\}\right] \\
&= \sigma_\eta^2\left[\gamma_1^{u+t-4} + \gamma_1^{u+t-6} + \cdots + \gamma_1^{t-u}\right] \\
&= \sigma_\eta^2\left[\frac{\gamma_1^{t-u} - \gamma_1^{u+t-2}}{1-\gamma_1^2}\right].
\end{aligned}$$

It then follows from (2.36) that the quasi maximum likelihood (QML) estimates for γ_1 and σ_η^2 can be obtained by solving

$$\frac{\partial \log L_Q^*\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \gamma_1} = 0 \qquad \text{and} \qquad \frac{\partial \log L_Q^*\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)}{\partial \sigma_\eta^2} = 0. \tag{2.37}$$

Defining $l = \log L_Q^*\left(\gamma_0, \gamma_1, \sigma_\eta^2\right)$, we have

$$\begin{aligned}
\frac{\partial l}{\partial \gamma_1} &= -\frac{1}{2}\frac{\partial \log|\Phi|}{\partial \gamma_1} - \frac{\partial (z-m)'}{\partial \gamma_1}\Phi^{-1}(z-m) - \frac{1}{2}(z-m)'\frac{\partial \Phi^{-1}}{\partial \gamma_1}(z-m) \\
\frac{\partial l}{\partial \sigma_\eta^2} &= -\frac{1}{2}\frac{\partial \log|\Phi|}{\partial \sigma_\eta^2} - \frac{1}{2}(z-m)'\frac{\partial \Phi^{-1}}{\partial \sigma_\eta^2}(z-m).
\end{aligned}$$

Because

$$\begin{aligned}
\frac{\partial \log|\Phi|}{\partial \gamma_1} &= \operatorname{trace}\left[\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\right] \\
\frac{\partial \Phi^{-1}}{\partial \gamma_1} &= -\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\Phi^{-1}
\end{aligned}$$

and so on, we write

$$\frac{\partial l}{\partial \gamma_1} = -\frac{1}{2}\text{trace}\left[\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\right] + \frac{\partial m'}{\partial \gamma_1}\Phi^{-1}(z-m) + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\Phi^{-1}(z-m) \tag{2.38}$$

$$\frac{\partial l}{\partial \sigma_\eta^2} = -\frac{1}{2}\text{trace}\left[\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\right] + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}(z-m). \tag{2.39}$$

We further define

$$\alpha = \begin{pmatrix} \gamma_1 \\ \sigma_\eta^2 \end{pmatrix},$$

and need to compute

$$\frac{\partial l}{\partial \alpha} = \begin{pmatrix} \frac{\partial l}{\partial \gamma_1} \\ \frac{\partial l}{\partial \sigma_\eta^2} \end{pmatrix}$$

and

$$\frac{\partial^2 l}{\partial \alpha \partial \alpha'} = \begin{pmatrix} \frac{\partial^2 l}{\partial {\gamma_1}^2} & \frac{\partial^2 l}{\partial \gamma_1 \partial \sigma_\eta^2} \\ \frac{\partial^2 l}{\partial \gamma_1 \partial \sigma_\eta^2} & \frac{\partial^2 l}{\partial {\sigma_\eta^2}^2} \end{pmatrix}$$

for the QML estimation of γ_1 and σ_η^2. To be specific, to solve the likelihood equations, namely,

$$\frac{\partial l}{\partial \alpha} = 0,$$

for α, we use the Taylor's expansion

$$\frac{\partial l}{\partial \alpha} + \frac{\partial^2 l}{\partial \alpha \partial \alpha'}(\alpha_{new} - \alpha) = 0,$$

leading to the Newton-Raphson iterative equation

$$\alpha_{r+1} = \alpha_r - \left[\left(\frac{\partial^2 l}{\partial\alpha\partial\alpha'}\right)^{-1}\left(\frac{\partial l}{\partial\alpha}\right)\right]_{[r]}. \tag{2.40}$$

The formulas for $\frac{\partial l}{\partial\gamma_1}$ and $\frac{\partial l}{\partial\sigma_\eta^2}$ are given in (2.38) and (2.39), respectively. The second order derivatives can be computed as follows.

$$\begin{aligned}
\frac{\partial^2 l}{\partial{\gamma_1}^2} &= -\frac{1}{2}\frac{\partial}{\partial\gamma_1}\left(\text{trace}\left[\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\right]\right) + \frac{\partial^2 m'}{\partial{\gamma_1}^2}\Phi^{-1}(z-m) + \frac{\partial m'}{\partial\gamma_1}\frac{\partial\Phi^{-1}}{\partial\gamma_1}(z-m) - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial m}{\partial\gamma_1} \\
&\quad - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) + (z-m)'\frac{\partial\Phi^{-1}}{\partial\gamma_1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) \\
&\quad + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial^2\Phi}{\partial{\gamma_1}^2}\Phi^{-1}(z-m) \\
&= -\frac{1}{2}\text{trace}\left[\frac{\partial\Phi^{-1}}{\partial\gamma_1}\frac{\partial\Phi}{\partial\gamma_1} + \Phi^{-1}\frac{\partial^2\Phi}{{\gamma_1}^2}\right] + \frac{\partial^2 m'}{\partial{\gamma_1}^2}\Phi^{-1}(z-m) + \frac{\partial m'}{\partial\gamma_1}\frac{\partial\Phi^{-1}}{\partial\gamma_1}(z-m) \\
&\quad - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial m}{\partial\gamma_1} - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) + (z-m)'\frac{\partial\Phi^{-1}}{\partial\gamma_1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) \\
&\quad + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial^2\Phi}{\partial{\gamma_1}^2}\Phi^{-1}(z-m)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial{\sigma_\eta^2}^2} &= -\frac{1}{2}\frac{\partial}{\partial\sigma_\eta^2}\left(\text{trace}\left[\Phi^{-1}\frac{\partial\Phi}{\partial\sigma_\eta^2}\right]\right) + (z-m)'\frac{\partial\Phi^{-1}}{\partial\sigma_\eta^2}\frac{\partial\Phi}{\partial\sigma_\eta^2}\Phi^{-1}(z-m) \\
&= -\frac{1}{2}\text{trace}\left[\frac{\partial\Phi^{-1}}{\partial\sigma_\eta^2}\frac{\partial\Phi}{\partial\sigma_\eta^2}\right] + (z-m)'\frac{\partial\Phi^{-1}}{\partial\sigma_\eta^2}\frac{\partial\Phi}{\partial\sigma_\eta^2}\Phi^{-1}(z-m),
\end{aligned}$$

where we applied $\frac{\partial m'}{\partial\sigma_\eta^2} = 0$ and $\frac{\partial^2\Phi}{\partial{\sigma_\eta^2}^2} = 0$.

$$\begin{aligned}
\frac{\partial^2 l}{\partial\gamma_1\partial\sigma_\eta^2} &= -\frac{1}{2}\text{trace}\left[\frac{\partial\Phi^{-1}}{\partial\gamma_1}\frac{\partial\Phi}{\partial\sigma_\eta^2} + \Phi^{-1}\frac{\partial^2\Phi}{\partial\gamma_1\partial\sigma_\eta^2}\right] - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\sigma_\eta^2}\Phi^{-1}(z-m) \\
&\quad + (z-m)'\frac{\partial\Phi^{-1}}{\partial\gamma_1}\frac{\partial\Phi}{\partial\sigma_\eta^2}\Phi^{-1}(z-m) + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial^2\Phi}{\partial\gamma_1\partial\sigma_\eta^2}\Phi^{-1}(z-m).
\end{aligned}$$

We now provide the formulas for all the first and second order derivatives that we need for (2.40).

$$\frac{\partial m_1}{\partial \gamma_1} = \frac{\partial m_1}{\partial \sigma_\eta^2} = 0.$$

For $t \geq 2$

$$\begin{aligned}
\frac{\partial m_t}{\partial \gamma_1} &= \gamma_0 \frac{1-\gamma_1^{t-1}-(t-1)(1-\gamma_1)\gamma_1^{t-2}}{(1-\gamma_1)^2} + (t-1)h_1\gamma_1^{t-2} \\
&= \gamma_0 \left[\frac{1-\gamma_1^{t-1}}{(1-\gamma_1)^2} - \frac{(t-1)\gamma_1^{t-2}}{1-\gamma_1}\right] + (t-1)h_1\gamma_1^{t-2} \\
\frac{\partial m_t}{\partial \sigma_\eta^2} &= 0 \\
\frac{\partial^2 m_t}{\partial \gamma_1^2} &= \gamma_0 \left\{\frac{2(1-\gamma_1)\left(1-\gamma_1^{t-1}\right)-(t-1)\gamma_1^{t-2}(1-\gamma_1)^2}{(1-\gamma_1)^4} - (t-1)\left[\frac{\gamma_1^{t-2}+(t-2)\gamma_1^{t-3}(1-\gamma_1)}{(1-\gamma_1)^2}\right]\right\} \\
&\quad +(t-1)(t-2)h_1\gamma_1^{t-3}.
\end{aligned}$$

For $t = 1, \cdots, T$,

$$\frac{\partial \Phi_{1t}}{\partial \gamma_1} = \frac{\partial \Phi_{1t}}{\partial \sigma_\eta^2} = 0.$$

For $t \geq 2$

$$\begin{aligned}
\frac{\partial \Phi_{tt}}{\partial \gamma_1} &= 2\sigma_\eta^2 \left[\frac{\gamma_1\left(1-\gamma_1^{2(t-1)}\right)-(t-1)\gamma_1^{2t-3}(1-\gamma_1^2)}{(1-\gamma_1^2)^2}\right] \\
&= 2\sigma_\eta^2 \left[\frac{\gamma_1\left(1-\gamma_1^{2(t-1)}\right)}{(1-\gamma_1^2)^2} - \frac{(t-1)\gamma_1^{2t-3}}{(1-\gamma_1^2)}\right] \\
\frac{\partial \Phi_{tt}}{\partial \sigma_\eta^2} &= \frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}.
\end{aligned}$$

For $2 \leq u < t$

$$\begin{aligned}
\frac{\partial \Phi_{ut}}{\partial \gamma_1} &= \sigma_\eta^2 \left[\frac{2\gamma_1 \left(\gamma_1^{t-u} - \gamma_1^{t+u-2}\right) + (1-\gamma_1^2)\left[(t-u)\gamma_1^{t-u-1} - (u+t-2)\gamma_1^{u+t-3}\right]}{(1-\gamma_1^2)^2} \right] \\
&= \sigma_\eta^2 \left[\frac{2\left(\gamma_1^{t-u+1} - \gamma_1^{t+u-1}\right)}{(1-\gamma_1^2)^2} + \frac{(t-u)\gamma_1^{t-u-1} - (u+t-2)\gamma_1^{u+t-3}}{(1-\gamma_1^2)} \right] \\
\frac{\partial \Phi_{ut}}{\partial \sigma_\eta^2} &= \frac{\gamma_1^{t-u} - \gamma_1^{u+t-2}}{1-\gamma_1^2}.
\end{aligned}$$

For $t \geq 2$

$$\begin{aligned}
\frac{\partial^2 \Phi_{tt}}{\partial \gamma_1 \partial \sigma_\eta^2} &= 2\frac{\gamma_1 \left(1-\gamma_1^{2(t-1)}\right) - (t-1)\gamma_1^{2t-3}(1-\gamma_1^2)}{\left(1-\gamma_1^2\right)^2} \\
\frac{\partial^2 \Phi_{tt}}{(\partial \sigma_\eta^2)^2} &= 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \Phi_{tt}}{(\partial \gamma_1)^2} &= 2\sigma_\eta^2 \left[\frac{4(1-\gamma_1^2)\gamma_1^2\left(1-\gamma_1^{2(t-1)}\right) + \left[1-(2t-1)\gamma_1^{2t-2}\right]\left(1-\gamma_1^2\right)^2}{\left(1-\gamma_1^2\right)^4} \right. \\
&\qquad \left. -(t-1)\frac{2\gamma_1^{2t-2} + (2t-3)\gamma_1^{2t-4}(1-\gamma_1^2)}{\left(1-\gamma_1^2\right)^2} \right] \\
&= 2\sigma_\eta^2 \left[\frac{4\gamma_1^2\left(1-\gamma_1^{2(t-1)}\right) + \left[1-(2t-1)\gamma_1^{2t-2}\right](1-\gamma_1^2)}{\left(1-\gamma_1^2\right)^3} \right. \\
&\qquad \left. -(t-1)\frac{2\gamma_1^{2t-2} + (2t-3)\gamma_1^{2t-4}(1-\gamma_1^2)}{\left(1-\gamma_1^2\right)^2} \right].
\end{aligned}$$

For $2 \leq u < t$

$$\frac{\partial^2 \Phi_{ut}}{\partial \gamma_1 \partial \sigma_\eta^2} = \frac{2\gamma_1 \left(\gamma_1^{t-u} - \gamma_1^{t+u-2}\right) + (1-\gamma_1^2)\left[(t-u)\gamma_1^{t-u-1} - (u+t-2)\gamma_1^{u+t-3}\right]}{(1-\gamma_1^2)^2}$$

$$\frac{\partial^2 \Phi_{ut}}{(\partial \sigma_\eta^2)^2} = 0$$

$$\frac{\partial^2 \Phi_{ut}}{(\partial \gamma_1)^2} =$$

$$\sigma_\eta^2 \left[2\frac{4(1-\gamma_1^2)\gamma_1\left(\gamma_1^{t-u+1}-\gamma_1^{u+t-1}\right)+\left[(t-u+1)\gamma_1^{t-u}-(u+t-1)\gamma_1^{u+t-2}\right](1-\gamma_1^2)^2}{(1-\gamma_1^2)^4} + \right.$$

$$\left. \frac{2\left[(t-u)\gamma_1^{t-u}-(u+t-2)\gamma_1^{u+t-2}\right]+(1-\gamma_1^2)\left[(t-u)(t-u-1)\gamma_1^{t-u-2}-(u+t-2)(u+t-3)\gamma_1^{u+t-4}\right]}{(1-\gamma_1^2)^2} \right]$$

$$= \sigma_\eta^2 \left[2\frac{4\left(\gamma_1^{t-u+2}-\gamma_1^{u+t}\right)+\left[(t-u+1)\gamma_1^{t-u}-(u+t-1)\gamma_1^{u+t-2}\right](1-\gamma_1^2)}{(1-\gamma_1^2)^3} + \right.$$

$$\left. \frac{2\left[(t-u)\gamma_1^{t-u}-(u+t-2)\gamma_1^{u+t-2}\right]+(1-\gamma_1^2)\left[(t-u)(t-u-1)\gamma_1^{t-u-2}-(u+t-2)(u+t-3)\gamma_1^{u+t-4}\right]}{(1-\gamma_1^2)^2} \right].$$

Note that in each iteration using (2.40), γ_0 is updated by the new γ_1 with (2.16).

Proposed Band Modification for Φ

Computation using full dimension of Φ is cumbersome and time consuming. This is especially true when inverting Φ. For this reason, we approximate Φ by a tridiagonal band matrix

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \Phi_{12} & \Phi_{22} & \Phi_{23} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \Phi_{23} & \Phi_{33} & \Phi_{34} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \Phi_{34} & \Phi_{44} & \Phi_{45} & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \Phi_{T-2,T-1} & \Phi_{T-1,T-1} & \Phi_{T-1,T} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \Phi_{T-1,T} & \Phi_{TT} \end{pmatrix}.$$

Then the determinant of Φ is given by the recursive formula

$$\det[\Phi]_{\{1,\cdots,n\}} = \Phi_{nn}\det[\Phi]_{\{1,\cdots,n-1\}} - \Phi_{n,n-1}\Phi_{n-1,n}\det[\Phi]_{\{1,\cdots,n-2\}}.$$

Because Φ is symmetric, it then follows that

$$\det[\Phi]_{\{1,\cdots,n\}} = \Phi_{nn}\det[\Phi]_{\{1,\cdots,n-1\}} - \Phi^2_{n-1,n}\det[\Phi]_{\{1,\cdots,n-2\}}. \tag{2.41}$$

Here $\det[\Phi]_{\{1,\cdots,k\}}$ denotes the kth principal minor, that is, $[\Phi]_{\{1,\cdots,k\}}$ is the submatrix formed by the first k rows and columns of Φ. If we also define $\det[\Phi]_{-1} = 0$ and $\det[\Phi]_0 = 1$, then the formula can start from $n = 1$. Note that the tridiagonalization makes the computation quite manageable as opposed to dealing with full Φ matrix.

In Usmani (1994), he gave an elegant and concise formula for the inverse of the tridiagonal matrix Φ as

$$\left(\Phi^{-1}\right)_{ij} = \begin{cases} (-1)^{i+j}\Phi_{i,i+1}\Phi_{i+1,i+2}\cdots\Phi_{j-1,j}\theta_{i-1}\phi_{j+1}/\theta_T, & i<j, \\ \theta_{i-1}\phi_{i+1}/\theta_T, & i=j, \\ (-1)^{i+j}\Phi_{j+1,j}\Phi_{j+2,j+1}\cdots\Phi_{i,i-1}\theta_{j-1}\phi_{i+1}/\theta_T, & i>j. \end{cases} \tag{2.42}$$

Here $\theta_i = \det[\Phi]_{\{1,\cdots,i\}}$, $i = 1,\cdots,T$, and the sequence $\{\phi_i\}$ is defined by the recurrence formula

$$\phi_i = \Phi_{ii}\phi_{i+1} - \Phi_{i,i+1}\Phi_{i+1,i}\phi_{i+2}, \qquad i = T, T-1, \cdots, 3, 2, 1$$

$$\phi_{T+1} = 1, \ \ \phi_{T+2} = 0\ . \tag{2.43}$$

If Φ^{-1} is approximated by a band matrix, then computing Φ^{-1} is relatively easy.

Because the log likelihood in the present setup has the formula given by

$$l = \log L_Q^* = c_0 - \frac{1}{2}\log|\Phi| - \frac{1}{2}[(z-m)'\Phi^{-1}(z-m)], \tag{2.44}$$

we may now apply (2.41) and (2.42) to (2.44) and compute the likelihood function in an easier way. Note however that computing $|\Phi|$, which appears in the formula for every element of Φ^{-1}, is still not so easy. Following Usmani (1994), we develop below a simpler formula to compute Φ^{-1}. This revision of the formula, i.e., alternative derivation of Φ^{-1}, is given as follows:

Eq. (2.41) can be written with θ_i as

$$\theta_{-1} = 0, \quad \theta_0 = 1$$
$$\theta_i = \Phi_{ii}\theta_{i-1} - \Phi_{i,i-1}\Phi_{i-1,i}\theta_{i-2}, \qquad i = 1, \cdots, T\,. \tag{2.45}$$

Divided by θ_{i-1}, it becomes

$$\frac{\theta_i}{\theta_{i-1}} \;=\; \Phi_{ii} - \Phi_{i,i-1}\Phi_{i-1,i}\left(\frac{\theta_{i-1}}{\theta_{i-2}}\right)^{-1}.$$

Defining

$$u_i \;=\; \frac{\theta_i}{\theta_{i-1}},$$

then we have

$$\begin{aligned} u_1 &= \frac{\theta_1}{\theta_0} = \theta_1 = \Phi_{11} \\ u_i &= \Phi_{ii} - \Phi_{i,i-1}\Phi_{i-1,i}u_{i-1}^{-1}, \qquad i = 2, \cdots, T\,. \end{aligned} \tag{2.46}$$

Dividing Eq. (2.43) by ϕ_{i+1}, we obtain

$$\frac{\phi_i}{\phi_{i+1}} = \Phi_{ii} - \Phi_{i,i+1}\Phi_{i+1,i}\frac{\phi_{i+2}}{\phi_{i+1}} \;\;\Rightarrow\;\; \frac{\phi_{i+1}}{\phi_i} = \frac{1}{\Phi_{ii} - \Phi_{i,i+1}\Phi_{i+1,i}\frac{\phi_{i+2}}{\phi_{i+1}}}.$$

Defining

$$v_i \;=\; \frac{\phi_{i+2}}{\phi_{i+1}},$$

we have

$$v_{i-1} = \frac{1}{\Phi_{ii} - \Phi_{i,i+1}\Phi_{i+1,i}v_i}, \quad v_{T-1} = \frac{1}{\Phi_{TT}}, \quad v_T = \frac{\phi_{T+2}}{\phi_{T+1}} = 0, \tag{2.47}$$

$$i = 2, \cdots, T-1.$$

Note that u_i's are the quotient of the neighbouring principal minors, and v_i's are the similar thing for the submatrices starting from the opposite conner of the matrix Φ, so they should be real numbers small enough to be processed by computers.

According to Usmani (1994), we have

$$\theta_T = \theta_j\phi_{j+1} - \Phi_{j+1,j}\Phi_{j,j+1}\theta_{j-1}\phi_{j+2}, \quad j = T, T-1, \cdots, 2, 1, \tag{2.48}$$

which can also be derived as follows:

$$\mathbf{\Phi} = \begin{pmatrix}
\Phi_{11} & \Phi_{12} & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 \\
\Phi_{12} & \Phi_{22} & \Phi_{23} & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 \\
0 & \Phi_{23} & \Phi_{33} & \Phi_{34} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 \\
0 & 0 & \Phi_{34} & \Phi_{44} & \Phi_{45} & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \Phi_{j-1,j} & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \Phi_{j,j-1} & \Phi_{j,j} & \Phi_{j,j+1} & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \Phi_{j+1,j} & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \Phi_{T-2,T-1} & \Phi_{T-1,T-1} & \Phi_{T-1,T} \\
0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \Phi_{T-1,T} & \Phi_{TT}
\end{pmatrix}.$$

Determinant of $\mathbf{\Phi}$ is the summation of a series of terms. Observing this matrix, we have the following statements that in these terms,

- $\Phi_{j-1,j}$ and $\Phi_{j,j-1}$ must appear together;
- $\Phi_{j,j+1}$ and $\Phi_{j+1,j}$ must appear together;
- $\Phi_{j,j}$, any of $\Phi_{j-1,j}$ and $\Phi_{j,j-1}$, or any of $\Phi_{j,j+1}$ and $\Phi_{j+1,j}$ cannot appear together.

So we have

$$\begin{aligned} |\mathbf{\Phi}| &= \text{summation of terms without } \Phi_{j,j+1} \text{ and } \Phi_{j+1,j} \\ &\quad + \text{summation of terms with } \Phi_{j,j+1} \text{ and } \Phi_{j+1,j} \\ &= \theta_j \phi_{j+1} - \Phi_{j+1,j}\Phi_{j,j+1}\theta_{j-1}\phi_{j+2} \, . \end{aligned}$$

According to Formula (2.42), for $i < T$,

$$\begin{aligned} \left(\Phi^{-1}\right)_{ii} &= \frac{\theta_{i-1}\phi_{i+1}}{\theta_i\phi_{i+1} - \Phi_{i+1,i}\Phi_{i,i+1}\theta_{i-1}\phi_{i+2}} = \frac{1}{\frac{\theta_i}{\theta_{i-1}} - \Phi_{i+1,i}\Phi_{i,i+1}\frac{\phi_{i+2}}{\phi_{i+1}}} \\ &= \frac{1}{u_i - \Phi_{i+1,i}\Phi_{i,i+1}v_i} \, , \end{aligned}$$

and

$$\left(\Phi^{-1}\right)_{TT} = \frac{\theta_{T-1}\phi_{T+1}}{\theta_T} = \frac{1}{u_T} \, .$$

For $i < j$, according to Formula (2.42), for $j < T$,

$$
\begin{aligned}
\left(\Phi^{-1}\right)_{ij} &= \frac{(-1)^{i+j}\Phi_{i,i+1}\Phi_{i+1,i+2}\cdots\Phi_{j-1,j}\theta_{i-1}\phi_{j+1}}{\theta_j\phi_{j+1} - \Phi_{j+1,j}\Phi_{j,j+1}\theta_{j-1}\phi_{j+2}} = \frac{(-1)^{i+j}\Phi_{i,i+1}\Phi_{i+1,i+2}\cdots\Phi_{j-1,j}}{\frac{\theta_j}{\theta_{i-1}} - \Phi_{j+1,j}\Phi_{j,j+1}\frac{\theta_{j-1}}{\theta_{i-1}}\frac{\phi_{j+2}}{\phi_{j+1}}} \\
&= \frac{(-1)^{i+j}\Phi_{i,i+1}\Phi_{i+1,i+2}\cdots\Phi_{j-1,j}}{u_i\cdots u_j - \Phi_{j+1,j}\Phi_{j,j+1}u_i\cdots u_{j-1}v_j},
\end{aligned}
$$

and

$$
\left(\Phi^{-1}\right)_{iT} = \frac{(-1)^{i+T}\Phi_{i,i+1}\Phi_{i+1,i+2}\cdots\Phi_{T-1,T}}{u_i\cdots u_T}.
$$

Since Φ and Φ^{-1} are symmetric matrices, the above formula define the whole Φ^{-1}. We can see that these formulas contain only u_i's and v_i's without θ_i's and ϕ_i's, so they can be processed by computers. We numerically tested the program to invert Φ with these formula. The order of the maximum element differences between $\Phi\Phi^{-1}$ and the identity matrix are given in the following table for different T.

T	10	100	1000	3000
order of difference	10^{-16}	10^{-14}	10^{-13}	10^{-12}

The formula for calculating the second derivatives of log-likelihood l are revised for this approach as

$$
\begin{aligned}
\frac{\partial^2 l}{\partial{\gamma_1}^2} =& \frac{1}{2}\text{trace}\left[\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1} - \Phi^{-1}\frac{\partial^2\Phi}{{\gamma_1}^2}\right] + \frac{\partial^2 m'}{\partial{\gamma_1}^2}\Phi^{-1}(z-m) - \frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial m}{\partial\gamma_1} \\
& - 2\frac{\partial m'}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) - (z-m)'\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}\frac{\partial\Phi}{\partial\gamma_1}\Phi^{-1}(z-m) \\
& + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial^2\Phi}{\partial{\gamma_1}^2}\Phi^{-1}(z-m) \qquad (2.49)
\end{aligned}
$$

$$\frac{\partial^2 l}{\partial {\sigma_\eta^2}^2} = \frac{1}{2}\text{trace}\left[\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\right] - (z-m)'\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}(z-m) \qquad (2.50)$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \gamma_1 \partial \sigma_\eta^2} =& \frac{1}{2}\text{trace}\left[\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2} - \Phi^{-1}\frac{\partial^2 \Phi}{\partial \gamma_1 \partial \sigma_\eta^2}\right] - \frac{\partial m'}{\partial \gamma_1}\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}(z-m) \\ & - (z-m)'\Phi^{-1}\frac{\partial \Phi}{\partial \gamma_1}\Phi^{-1}\frac{\partial \Phi}{\partial \sigma_\eta^2}\Phi^{-1}(z-m) + \frac{1}{2}(z-m)'\Phi^{-1}\frac{\partial^2 \Phi}{\partial \gamma_1 \partial \sigma_\eta^2}\Phi^{-1}(z-m).\end{aligned} \qquad (2.51)$$

Since Φ is a tridiagonal matrix, the elements of Φ^{-1} goes to zero with the increase of the distance of the elements from the diagonal line, which we call off-diagonal distance. So we can approximate Φ^{-1} with a band matrix of half-band-width p, i.e., $(\Phi^{-1})_{ij} = 0$ if $|j-i| > p$. The half-band-width of the tridiagonal matrix Φ is 1.

Note that under this proposed MQML approach, we still estimate γ_0 by using MM estimating equation (2.16), and γ_1 and σ_η^2 are obtained by solving the MQML estimating equations (2.40), with second order derivatives calculated by (2.49)-(2.51). For convenience, we denote all these estimates under the MQML approach as $\hat{\alpha}_{MQML} = \left(\hat{\gamma}_{1,MQML}, \hat{\sigma}^2_{\eta,MQML}\right)$ and $\hat{\gamma}_{0,MQML}$ respectively.

2.3 Illustration of the Estimation Approaches: A Simulation Study

Recall that the existing GMM approach is cumbersome and can be less efficient as compared to the proposed MM approach. Note that the proposed approaches, the MM approach in particular, are much more simpler than the QML and GMM approaches. Similar to Tagore (2010), in this section, we examine both small and large sample estimation performances of the proposed MM and AGQL approaches through a simulation study. The difference between Tagore (2010) and the present simulation lies in the fact that we are now also estimating γ_0, whereas Tagore (2010) evaluated the performance of the estimating methods only when $\gamma_0 = 0$. For the purpose, we choose T as small as 200, and several finite but large values such as T = 1000, 2000, and 3000. Note that these values of T are chosen to indicate that unlike the existing GMM approach (where length of time series requires to be infinitely large such as T =10,000 or 15,000, ..., and so on), the proposed approaches produce good estimates based on a practically reasonable length of the time series.

In the small sample case the initial variance σ_1^2 will have some effects on the estimation of the main volatility parameters, as expected. We assume that it can be estiamated effectively from previous data, then since $\log \sigma_1^2$ is assumed to have a normal distribution with mean $\gamma_0/(1-\gamma_1)$ as shown in (1.3), we choose the value of $\gamma_0/(1-\gamma_1)$ for $\log(\sigma_1^2)$ in our simulation study. Now, to examine the small sample estimation performance for γ_0, γ_1 and σ_η^2 by the MM approach, we solve the MM estimating equation (2.8) for γ_1, and (2.3) for σ_η^2, iteratively, γ_0 being updated by (2.16) after obtaining new γ_1. For the large sample case, we solve the asymptotic estimating equations (2.11) and (2.17) for γ_1 and σ_η^2, respectively, and still update γ_0 by (2.16) after obtaining new γ_1. The simulated means (SM) along with simulated standard errors (SSE) for the MM estimates based on 500 simulations are reported in Tables 2.1 - 2.3. For the estimation of γ_0, γ_1 and σ_η^2 by the AGQL approach, we solve the AGQL estimating equations (2.34) for γ_1, and (2.33) for σ_η^2, iteratively, and update γ_0 by (2.16) after obtaining new γ_1. Note that these equations are available for both small and large T. The simulated means and their standard errors for the AGQL estimates are given in Tables 2.4 - 2.6.

The columns 6 to 9 in Tables 2.1 - 2.3 show that for a reasonably large time series with length between T= 2000 and 10,000, the proposed MM approach performs very well in estimating γ_0, γ_1 and σ_η^2. This is a big improvement over the existing GMM and MQML approaches mainly because of the fact that proposed MM approach is simpler and computationally quite efficient. Also, unlike the existing GMM and MQML approaches, the MM approach does not encounter any convergence problems.

Table 2.1: Simulated mean (SM) and simulated standard error (SSE) of MM estimates for selected parameter values based on 500 simulations.

		Time Series Length (T)						
Parameters	Quantity	200	500	1000	2000	3000	5000	10,000
γ_0 =0.05	SM	0.0318	0.0395	0.0409	0.0487	0.0478	0.0488	0.0492
	SSE	0.0967	0.0779	0.0606	0.0428	0.0341	0.0253	0.0180
γ_1 =0.25	SM	0.2500	0.2151	0.2190	0.2267	0.2510	0.2550	0.2555
	SSE	0.3706	0.3405	0.3014	0.2543	0.2224	0.1774	0.1231
σ_η^2 =0.25	SM	0.3227	0.2693	0.2537	0.2399	0.2378	0.2421	0.2461
	SSE	0.2097	0.1474	0.1213	0.0866	0.0810	0.0639	0.0435
γ_0 =0.05	SM	0.0361	0.0434	0.0518	0.0486	0.0508	0.0503	0.0498
	SSE	0.1201	0.0835	0.0606	0.0434	0.0351	0.0241	0.0188
γ_1 =0.25	SM	0.2180	0.2415	0.2282	0.2419	0.2505	0.2462	0.2485
	SSE	0.3501	0.3133	0.2463	0.1789	0.1478	0.1288	0.0921
σ_η^2 =0.5	SM	0.5195	0.4706	0.4644	0.4900	0.4802	0.4880	0.4934
	SSE	0.2937	0.2209	0.1552	0.1092	0.0873	0.0708	0.0493
γ_0 =0.05	SM	0.0425	0.0479	0.0486	0.0477	0.0508	0.0500	0.0511
	SSE	0.1527	0.0989	0.0658	0.0484	0.0351	0.0278	0.0198
γ_1 =0.25	SM	0.1664	0.1978	0.2087	0.2190	0.2279	0.2376	0.2502
	SSE	0.3000	0.2491	0.2067	0.1532	0.1236	0.1133	0.0996
σ_η^2 =1.0	SM	0.9458	0.9485	0.9676	0.9910	0.9947	0.9970	0.9841
	SSE	0.3770	0.2889	0.2066	0.1577	0.1186	0.1111	0.1008
γ_0 =0.05	SM	0.0536	0.0430	0.0517	0.0484	0.0495	0.0489	0.0492
	SSE	0.1073	0.0694	0.0524	0.0335	0.0281	0.0237	0.0149
γ_1 =0.5	SM	0.3010	0.3945	0.4409	0.4790	0.4826	0.4974	0.5058
	SSE	0.3597	0.3493	0.2909	0.2363	0.2055	0.1768	0.1208
σ_η^2 =0.25	SM	0.3561	0.2733	0.2597	0.2469	0.2456	0.2435	0.2457
	SSE	0.2287	0.1777	0.1353	0.1127	0.0926	0.0822	0.0586
γ_0 =0.05	SM	0.0576	0.0510	0.0502	0.0516	0.0508	0.0498	0.0498
	SSE	0.1239	0.0822	0.0503	0.0331	0.0275	0.0233	0.0166
γ_1 =0.5	SM	0.3780	0.4222	0.4502	0.4677	0.4761	0.4980	0.4965
	SSE	0.3412	0.2882	0.2336	0.1785	0.1617	0.1379	0.1001
σ_η^2 =0.5	SM	0.5309	0.5038	0.5012	0.4992	0.4967	0.4875	0.4948
	SSE	0.3015	0.2531	0.2023	0.1528	0.1270	0.1190	0.0816
γ_0 =0.05	SM	0.0631	0.0571	0.0530	0.0570	0.0513	0.0538	0.0518
	SSE	0.1547	0.0899	0.0631	0.0388	0.0332	0.0265	0.0182
γ_1 =0.5	SM	0.3181	0.3853	0.4187	0.4446	0.4560	0.4705	0.4859
	SSE	0.2837	0.2315	0.2007	0.1522	0.1387	0.1371	0.1120
σ_η^2 =1.0	SM	1.0633	1.0220	1.0280	1.0316	1.0264	1.0088	0.9962
	SSE	0.4482	0.3382	0.3110	0.2334	0.2063	0.2176	0.1738

Table 2.2: Table 2.1 Contd....

		Time Series Length (T)						
Parameters	Quantity	200	500	1000	2000	3000	5000	10,000
γ_0 =0.1	SM	0.0757	0.0856	0.0958	0.1012	0.0965	0.0987	0.0996
	SSE	0.1241	0.0967	0.0637	0.0535	0.0405	0.0321	0.0225
γ_1 =0.25	SM	0.2199	0.2104	0.2172	0.2247	0.2533	0.2600	0.2474
	SSE	0.3547	0.3443	0.3005	0.2455	0.2183	0.1708	0.1259
σ_η^2 =0.25	SM	0.3329	0.2770	0.2507	0.2435	0.2405	0.2425	0.2468
	SSE	0.2152	0.1604	0.1149	0.0861	0.0792	0.0628	0.0419
γ_0 =0.1	SM	0.0911	0.0932	0.0969	0.0978	0.0995	0.1006	0.1000
	SSE	0.1354	0.0979	0.0664	0.0450	0.0386	0.0293	0.0201
γ_1 =0.25	SM	0.2289	0.2117	0.2438	0.2308	0.2473	0.2438	0.2474
	SSE	0.3400	0.3053	0.2455	0.1833	0.1524	0.1247	0.0859
σ_η^2 =0.5	SM	0.4937	0.4755	0.4726	0.4932	0.4877	0.4921	0.4964
	SSE	0.2903	0.2044	0.1551	0.1145	0.0903	0.0756	0.0487
γ_0 =0.1	SM	0.0978	0.1219	0.1067	0.1012	0.1031	0.1007	0.1012
	SSE	0.1551	0.1042	0.0729	0.0493	0.0401	0.0321	0.0232
γ_1 =0.25	SM	0.1731	0.1768	0.2081	0.2267	0.2221	0.2426	0.2428
	SSE	0.3031	0.2420	0.1966	0.1500	0.1207	0.1251	0.0881
σ_η^2 =1.0	SM	0.9223	0.9364	0.9710	0.9814	0.9946	0.9865	0.9955
	SSE	0.3892	0.2716	0.2058	0.1489	0.1130	0.1221	0.0737
γ_0 =0.1	SM	0.1113	0.1182	0.1071	0.1061	0.0963	0.0993	0.0985
	SSE	0.1339	0.0969	0.0759	0.0521	0.0388	0.0358	0.0255
γ_1 =0.5	SM	0.2930	0.3529	0.4347	0.4542	0.4991	0.4914	0.5046
	SSE	0.3689	0.3367	0.3025	0.2213	0.1986	0.1743	0.1293
σ_η^2 =0.25	SM	0.3380	0.2901	0.2522	0.2570	0.2447	0.2473	0.2447
	SSE	0.2152	0.1688	0.1385	0.1076	0.0968	0.0838	0.0620
γ_0 =0.1	SM	0.1056	0.0999	0.1082	0.1021	0.1006	0.1017	0.0996
	SSE	0.1378	0.0853	0.0659	0.0449	0.0362	0.0316	0.0231
γ_1 =0.5	SM	0.3530	0.4366	0.4406	0.4715	0.4879	0.4939	0.5034
	SSE	0.3511	0.2872	0.2381	0.1796	0.1514	0.1297	0.1029
σ_η^2 =0.5	SM	0.5462	0.4943	0.5010	0.4993	0.4959	0.4889	0.4882
	SSE	0.3319	0.2501	0.1943	0.1553	0.1320	0.1104	0.0848
γ_0 =0.1	SM	0.1471	0.1245	0.1148	0.1100	0.1055	0.1041	0.1021
	SSE	0.1613	0.1012	0.0699	0.0474	0.0392	0.0315	0.0270
γ_1 =0.5	SM	0.3281	0.3763	0.4153	0.4554	0.4644	0.4816	0.4877
	SSE	0.2928	0.2422	0.1888	0.1574	0.1417	0.1275	0.1136
σ_η^2 =1.0	SM	1.0251	1.0492	1.0439	1.0145	1.0107	0.9980	0.9971
	SSE	0.4599	0.3811	0.2911	0.2373	0.2079	0.1982	0.1804

Table 2.3: Table 2.1 Contd....

		Time Series Length (T)						
Parameters	Quantity	200	500	1000	2000	3000	5000	10,000
γ_0 =0.2	SM	0.1597	0.1901	0.1962	0.1985	0.1981	0.1975	0.2006
	SSE	0.1505	0.1253	0.1002	0.0786	0.0631	0.0474	0.0330
γ_1 =0.25	SM	0.2192	0.2089	0.2334	0.2446	0.2435	0.2531	0.2471
	SSE	0.3700	0.3454	0.3141	0.2636	0.2147	0.1672	0.1180
σ_η^2 =0.25	SM	0.3422	0.2732	0.2481	0.2377	0.2460	0.2414	0.2483
	SSE	0.2201	0.1614	0.1217	0.0913	0.0757	0.0592	0.0419
γ_0 =0.2	SM	0.1858	0.1988	0.2061	0.2016	0.2027	0.1979	0.1997
	SSE	0.1562	0.1210	0.0879	0.0601	0.0511	0.0392	0.0289
γ_1 =0.25	SM	0.2016	0.2151	0.2184	0.2397	0.2318	0.2506	0.2513
	SSE	0.3433	0.3210	0.2421	0.1845	0.1489	0.1348	0.0921
σ_η^2 =0.5	SM	0.5092	0.4685	0.4827	0.4829	0.4953	0.4918	0.4981
	SSE	0.2779	0.2138	0.1625	0.1028	0.0932	0.0791	0.0518
γ_0 =0.2	SM	0.2144	0.2159	0.2072	0.2033	0.2039	0.2016	0.2037
	SSE	0.1882	0.1144	0.0815	0.0570	0.0476	0.0422	0.0285
γ_1 =0.25	SM	0.1635	0.1927	0.1957	0.2197	0.2354	0.2489	0.2410
	SSE	0.3036	0.2322	0.1854	0.1495	0.1315	0.1189	0.0795
σ_η^2 =1.0	SM	0.9455	0.9599	0.9900	0.9935	0.9873	0.9746	0.9947
	SSE	0.4156	0.2741	0.2000	0.1542	0.1309	0.1061	0.0708
γ_0 =0.2	SM	0.2248	0.2502	0.2224	0.2061	0.2036	0.1978	0.1970
	SSE	0.1837	0.1520	0.1155	0.0964	0.0764	0.0681	0.0471
γ_1 =0.5	SM	0.3122	0.3450	0.4250	0.4737	0.4830	0.4994	0.5053
	SSE	0.3802	0.3364	0.2853	0.2409	0.1945	0.1760	0.1223
σ_η^2 =0.25	SM	0.3621	0.2878	0.2631	0.2484	0.2484	0.2452	0.2444
	SSE	0.2420	0.1648	0.1360	0.1127	0.0946	0.0830	0.0590
γ_0 =0.2	SM	0.2513	0.2317	0.2143	0.2082	0.2090	0.2084	0.1974
	SSE	0.1887	0.1385	0.0964	0.0784	0.0663	0.0496	0.0440
γ_1 =0.5	SM	0.3370	0.4010	0.4469	0.4824	0.4747	0.4733	0.5059
	SSE	0.3654	0.3129	0.2267	0.1827	0.1595	0.1206	0.1090
σ_η^2 =0.5	SM	0.5451	0.5033	0.4992	0.4850	0.5012	0.5077	0.4867
	SSE	0.3245	0.2499	0.1933	0.1546	0.1339	0.0992	0.0894
γ_0 =0.2	SM	0.2562	0.2495	0.2241	0.2235	0.2071	0.2063	0.2044
	SSE	0.1885	0.1263	0.0942	0.0671	0.0606	0.0566	0.0458
γ_1 =0.5	SM	0.3564	0.3772	0.4253	0.4355	0.4768	0.4848	0.4907
	SSE	0.2971	0.2315	0.1997	0.1411	0.1418	0.1361	0.1102
σ_η^2 =1.0	SM	1.0179	1.0547	1.0334	1.0385	1.0004	0.9897	0.9945
	SSE	0.4658	0.3673	0.3054	0.2066	0.2200	0.2113	0.1779

Table 2.4: Simulated mean (SM) and simulated standard error (SSE) of AGQL estimates for selected parameter values based on 500 simulations.

		Time Series Length (T)				
Parameters	Quantity	200	500	1000	2000	3000
γ_0 =0.05	SM	0.0256	0.0379	0.0422	0.0485	0.0483
	SSE	0.0947	0.0763	0.0613	0.0426	0.0339
γ_1 =0.25	SM	0.2718	0.2219	0.1997	0.2188	0.2410
	SSE	0.3633	0.3097	0.2695	0.2396	0.2083
σ^2_η =0.25	SM	0.3051	0.2741	0.2618	0.2449	0.2410
	SSE	0.2140	0.1446	0.1154	0.0844	0.0772
γ_0 =0.05	SM	0.0309	0.0454	0.0514	0.0488	0.0507
	SSE	0.1154	0.0846	0.0604	0.0432	0.0352
γ_1 =0.25	SM	0.2273	0.2176	0.2240	0.2432	0.2493
	SSE	0.3205	0.2762	0.2347	0.1810	0.1485
σ^2_η =0.5	SM	0.5285	0.4846	0.4680	0.4873	0.4813
	SSE	0.2849	0.2046	0.1485	0.1087	0.0883
γ_0 =0.05	SM	0.0343	0.0506	0.0495	0.0475	0.0514
	SSE	0.1390	0.0972	0.0660	0.0481	0.0359
γ_1 =0.25	SM	0.1601	0.1904	0.2060	0.2174	0.2233
	SSE	0.2878	0.2386	0.2032	0.1532	0.1245
σ^2_η =1.0	SM	0.9621	0.9607	0.9718	0.9941	1.0010
	SSE	0.3747	0.2769	0.2014	0.1575	0.1198
γ_0 =0.05	SM	0.0440	0.0448	0.0543	0.0509	0.0515
	SSE	0.1049	0.0709	0.0537	0.0342	0.0275
γ_1 =0.5	SM	0.3048	0.3488	0.4048	0.4459	0.4607
	SSE	0.3485	0.3016	0.2620	0.2104	0.1853
σ^2_η =0.25	SM	0.3598	0.3050	0.2796	0.2613	0.2553
	SSE	0.2317	0.1675	0.1240	0.1015	0.0812
γ_0 =0.05	SM	0.0562	0.0493	0.0527	0.0536	0.0516
	SSE	0.1158	0.0800	0.0508	0.0320	0.0271
γ_1 =0.5	SM	0.3547	0.4002	0.4217	0.4598	0.4660
	SSE	0.3189	0.2603	0.2110	0.1659	0.1503
σ^2_η =0.5	SM	0.5449	0.5186	0.5182	0.5083	0.5059
	SSE	0.2750	0.2261	0.1781	0.1403	0.1164
γ_0 =0.05	SM	0.0666	0.0625	0.0526	0.0566	0.0516
	SSE	0.1481	0.0873	0.0617	0.0381	0.0323
γ_1 =0.5	SM	0.3179	0.3728	0.3992	0.4372	0.4547
	SSE	0.2773	0.2217	0.1800	0.1431	0.1400
σ^2_η =1.0	SM	1.0451	1.0472	1.0669	1.0449	1.0293
	SSE	0.4367	0.3260	0.2720	0.2092	0.2050

Table 2.5: Table 2.4 Contd....

Parameters	Quantity	Time Series Length (T)				
		200	500	1000	2000	3000
γ_0 =0.1	SM	0.0586	0.0838	0.0993	0.1009	0.0976
	SSE	0.1071	0.0872	0.0606	0.0521	0.0406
γ_1 =0.25	SM	0.2555	0.2269	0.2029	0.2261	0.2443
	SSE	0.3536	0.3135	0.2835	0.2416	0.2067
σ_η^2 =0.25	SM	0.3324	0.2818	0.2502	0.2435	0.2437
	SSE	0.2122	0.1550	0.1104	0.0855	0.0761
γ_0 =0.1	SM	0.0767	0.0921	0.1006	0.0973	0.1006
	SSE	0.1215	0.0953	0.0693	0.0448	0.0382
γ_1 =0.25	SM	0.2345	0.1933	0.2226	0.2334	0.2395
	SSE	0.3205	0.2721	0.2329	0.1825	0.1461
σ_η^2 =0.5	SM	0.5086	0.4919	0.4796	0.4924	0.4905
	SSE	0.2868	0.1997	0.1418	0.1138	0.0870
γ_0 =0.1	SM	0.0883	0.1193	0.1047	0.1028	0.1031
	SSE	0.1442	0.0965	0.0731	0.0495	0.0398
γ_1 =0.25	SM	0.1929	0.1726	0.2156	0.2206	0.2309
	SSE	0.2797	0.2306	0.1947	0.1465	0.1190
σ_η^2 =1.0	SM	0.9338	0.9465	0.9635	0.9838	0.9926
	SSE	0.3840	0.2529	0.2081	0.1424	0.1138
γ_0 =0.1	SM	0.0988	0.1107	0.1149	0.1096	0.1037
	SSE	0.1138	0.0866	0.0728	0.0508	0.0366
γ_1 =0.5	SM	0.3027	0.3393	0.3922	0.4351	0.4611
	SSE	0.3442	0.3056	0.2783	0.2018	0.1781
σ_η^2 =0.25	SM	0.3569	0.3071	0.2698	0.2662	0.2590
	SSE	0.2134	0.1649	0.1254	0.0990	0.0845
γ_0 =0.1	SM	0.1008	0.1149	0.1112	0.1033	0.1020
	SSE	0.1289	0.0879	0.0639	0.0415	0.0345
γ_1 =0.5	SM	0.3272	0.3985	0.4101	0.4650	0.4821
	SSE	0.3121	0.2684	0.2127	0.1656	0.1417
σ_η^2 =0.5	SM	0.5861	0.5199	0.5314	0.5083	0.5029
	SSE	0.3127	0.2330	0.1764	0.1419	0.1226
γ_0 =0.1	SM	0.1380	0.1150	0.1145	0.1114	0.1065
	SSE	0.1448	0.0962	0.0670	0.0466	0.0376
γ_1 =0.5	SM	0.3155	0.3714	0.4116	0.4448	0.4644
	SSE	0.2639	0.2170	0.1800	0.1488	0.1315
σ_η^2 =1.0	SM	1.0757	1.0917	1.0624	1.0313	1.0189
	SSE	0.4326	0.3555	0.2803	0.2189	0.1941

Table 2.6: Table 2.4 Contd....

		Time Series Length (T)				
Parameters	Quantity	200	500	1000	2000	3000
γ_0 =0.2	SM	0.1389	0.1890	0.2027	0.1997	0.2017
	SSE	0.1290	0.1082	0.0983	0.0774	0.0620
γ_1 =0.25	SM	0.2510	0.1909	0.2240	0.2454	0.2302
	SSE	0.3469	0.3059	0.2947	0.2569	0.2054
σ^2_η =0.25	SM	0.3582	0.2934	0.2464	0.2371	0.2479
	SSE	0.2231	0.1509	0.1170	0.0892	0.0718
γ_0 =0.2	SM	0.1656	0.1881	0.2070	0.2017	0.1983
	SSE	0.1344	0.0968	0.0860	0.0604	0.0491
γ_1 =0.25	SM	0.2097	0.2221	0.2053	0.2307	0.2402
	SSE	0.3135	0.2614	0.2286	0.1850	0.1478
σ^2_η =0.5	SM	0.5377	0.4971	0.4907	0.4893	0.4972
	SSE	0.2735	0.1885	0.1536	0.1055	0.0905
γ_0 =0.2	SM	0.1815	0.2108	0.2108	0.2030	0.2054
	SSE	0.1558	0.1090	0.0802	0.0575	0.0483
γ_1 =0.25	SM	0.1830	0.1868	0.1908	0.2244	0.2294
	SSE	0.2695	0.2190	0.1796	0.1493	0.1253
σ^2_η =1.0	SM	0.9768	0.9747	0.9882	0.9881	0.9921
	SSE	0.3823	0.2711	0.1914	0.1503	0.1273
γ_0 =0.2	SM	0.2045	0.2271	0.2339	0.2219	0.2155
	SSE	0.1454	0.1089	0.1054	0.0844	0.0669
γ_1 =0.5	SM	0.3037	0.3534	0.3858	0.4328	0.4521
	SSE	0.3312	0.2882	0.2621	0.2059	0.1677
σ^2_η =0.25	SM	0.4074	0.3248	0.279	0.2690	0.2643
	SSE	0.2329	0.1543	0.1227	0.0965	0.0831
γ_0 =0.2	SM	0.2034	0.2244	0.2230	0.2176	0.2144
	SSE	0.1308	0.1132	0.0927	0.0719	0.0583
γ_1 =0.5	SM	0.3724	0.4095	0.4255	0.4605	0.4617
	SSE	0.3022	0.2567	0.2074	0.1637	0.1412
σ^2_η =0.5	SM	0.5763	0.5304	0.5242	0.5009	0.5112
	SSE	0.2987	0.2298	0.1767	0.1353	0.1211
γ_0 =0.2	SM	0.2425	0.2367	0.2365	0.2218	0.2152
	SSE	0.1566	0.1194	0.0876	0.0679	0.0640
γ_1 =0.5	SM	0.3219	0.3961	0.4056	0.4380	0.4605
	SSE	0.2603	0.2185	0.1766	0.1439	0.1416
σ^2_η =1.0	SM	1.0991	1.0329	1.0602	1.0384	1.0240
	SSE	0.4497	0.3419	0.2553	0.2133	0.2045

Table 2.7: Simulated mean (SM) and simulated standard error (SSE) of MQML estimates for selected parameter values based on 500 simulations.

		Time Series Length (T)	
Parameters	Quantity	1000	3000
γ_0 =0.05	SM	0.0481	0.0433
	SSE	0.0569	0.0368
γ_1 =0.5	SM	0.4329	0.5171
	SSE	0.4059	0.3368
σ_η^2 =0.25	SM	0.3239	0.2616
	SSE	0.2855	0.2193
γ_0 =0.05	SM	0.0451	0.0508
	SSE	0.0504	0.0282
γ_1 =0.5	SM	0.4929	0.4689
	SSE	0.2018	0.1125
σ_η^2 =1.0	SM	0.9960	1.0318
	SSE	0.4845	0.2974
γ_0 =0.2	SM	0.2340	0.1979
	SSE	0.1629	0.1285
γ_1 =0.5	SM	0.3794	0.4911
	SSE	0.4094	0.3299
σ_η^2 =0.25	SM	0.3697	0.2756
	SSE	0.3046	0.2263
γ_0 =0.2	SM	0.1962	0.2102
	SSE	0.0857	0.0455
γ_1 =0.5	SM	0.4937	0.4689
	SSE	0.2043	0.1123
σ_η^2 =1.0	SM	0.9960	1.0318
	SSE	0.4900	0.2972

The results in columns 6 and 7 in Tables 2.4 - 2.6 show that the proposed AGQL approach performs similarly to the MM approach. Note however that to save time and space we have considered T = 2000 and 3000 in this case. As the length of the series increases, both MM and AGQL approach appears to perform better as expected. To be specific, when T = 3000, for example, the MM approach provides estimates for $\gamma_0 = 0.05$, $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 1.0$ as $\hat{\gamma}_{0,MM} = 0.0513$ with its simulated standard error 0.0332, $\hat{\gamma}_{1,MM} = 0.456$ with its simulated standard error 0.139 and $\hat{\sigma}^2_{\eta,MM} = 1.026$ with its standard error 0.206. For the same parameter values, when T = 10,000, the MM approach produces $\hat{\gamma}_{0,MM} = 0.0518$ with its simulated standard error 0.0182, $\hat{\gamma}_{1,MM} = 0.486$ with its simulated standard error 0.112 and $\hat{\sigma}^2_{\eta,MM} = 0.996$ with its standard error 0.174. Thus, it is clear that the MM approach works very well even if the length of the series is as small as T = 3000. However, as expected, the standard errors of the estimates improves considerably when T increased from 3000 to 10,000. As mentioned earlier, the AGQL approach behaves similarly to the MM approach. For example, for the same parameter values, when T = 3000 the AGQL estimates give $\hat{\gamma}_{0,AGQL} = 0.0516$ with simulated standard error 0.0323, $\hat{\gamma}_{1,AGQL} = 0.455$ with its simulated standard error 0.140 and $\hat{\sigma}^2_{\eta,AGQL}= 1.029$ with its simulated standard error 0.205. Thus AGQL approach appears to produce similar estimates for γ_0, γ_1 and σ_η^2 with similar standard errors. So we can use either of them.

As far as the small sample performance is concerned, both MM and AGQL approaches provides somewhat reasonable, but not so satisfactory estimates. For example, when T = 500, the MM approach provides estimates for $\gamma_0 = 0.05$, $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ as $\hat{\gamma}_{0,MM} = 0.0510$, $\hat{\gamma}_{1,MM} = 0.422$ and $\hat{\sigma}^2_{\eta,MM} = 0.504$, respectively, with corresponding simulated standard errors 0.0822, 0.288 and 0.253. For the same parameter values, the AGQL provides $\hat{\gamma}_{0,AGQL} = 0.0493$ with its simulated standard error 0.0800, $\hat{\gamma}_{1,AGQL} = 0.400$ with its simulated standard error 0.260 and $\hat{\sigma}^2_{\eta,AGQL} = 0.519$ with its standard error 0.226. These and other results in Tables 2.1 - 2.6 indicate that the estimates of σ_η^2 appears to be close to the true values, whereas the estimates of γ_0 and especially γ_1 are not so satisfactory. But, the estimates of γ_0 and γ_1 get closer to the true values when the length of the series is increased.

The simulated means and standard errors of MQML method for selected parameter values based on 500 simulations are given in Table 2.7. The estimation results are not as good as those from the MM and AGQL approaches given in Tables 2.1 and 2.4. For example, for $\gamma_0 = 0.05$, $\gamma_1 = 0.5$, $\sigma_\eta^2 = 0.25$, and T = 1000 case, the MQML estimates give $\hat{\gamma}_{0,MQML} = 0.0481$, $\hat{\gamma}_{1,MQML} = 0.433$ and $\hat{\sigma}^2_{\eta,MQML} = 0.324$, which are all farther away from the true parameter values than the MM estimates of $\hat{\gamma}_{0,MM} = 0.0517$, $\hat{\gamma}_{1,MM} = 0.441$ and $\hat{\sigma}^2_{\eta,MM} = 0.260$, while the simulated standard errors for $\hat{\gamma}_{0,MQML}$, $\hat{\gamma}_{1,MQML}$ and $\hat{\sigma}^2_{\eta,MQML}$ in MQML approach are respectively 0.0569, 0.406 and 0.285, which are all greater than the corresponding simulated standard errors in MM approach, which are 0.0524, 0.291 and 0.135 respectively. In Tagore (2010), it has been established that the MM and AGQL approaches are asymptotically more efficient than the MQML method. Our above simulation results agree with this asymptotic conclusion.

Note that in the simulation study we have considered small values for γ_0 (0.05, 0.1 and 0.2) and moderately large values for γ_1 (0.25 and 0.5) along with a wider range for σ_η^2 ranging from 0.25 to 1.0. These small values for γ_0 and γ_1 are expected mostly in practice, because of the fact that σ_t^2 is an exponential function in these parameters. However, the proposed method works well for larger γ_1 values. For example, the real life data that we discuss below provides $\hat{\gamma}_1 \simeq 0.7$ which was obtained without any convergence problem.

2.4 Illustration of the Estimation Approaches: A Real Life Data Analysis

In this section, the stochastic volatility model is fitted to US-Dollar/Swiss-Franc exchange rate. The data consist of daily observations of weekdays close exchange rates from July 24, 2007 to July 24, 2012, which are denoted as P_t.

Univariate models were fitted to the log return of the exchange rates with mean subtracted, that is,

$$Y_t = \triangle \log P_t - \left(\sum_{t=1}^{T-1} \triangle \log P_t \right) / (T-1), \qquad t = 1, \cdots, T-1,$$

where $\triangle \log P_t = \log P_{t+1} - \log P_t$. In Ruiz (1994), the QML estimation method was applied to this log return of Yen/Dollar exchange rate from 1/10/81 to 28/6/85 to fit the stochastic volatility model. Note that the QML estimation method was also applied by some researchers in multivariate setup. For example, Harvey et.al (1994) illustrated the fitting of multivariate stochastic volatility model with the log return of Pound/Dollar, Deutschmark/Dollar, Yen/Dollar and Swiss-Franc/Dollar exchange rates from 1/10/81 to 28/6/85.

Turning back to the US-Dollar/Swiss-Franc exchange rate data set, we use the variance of the first 100 y_t's to estimate σ_1^2, and then fit the models (1.1) - (1.2) to the next 1000 y_t's. That is, we estimate the parameters of the model, namely, γ_0, γ_1 and σ_η^2, by using the proposed MM, AGQL and MQML approaches. The estimates of the parameters for these three approaches are given in Table 2.8.

Now to examine the performance of these three estimation approaches, we choose to compare the fitted mean and variance of y_t^2 with the observed mean and variance of y_t^2, respectively. For the calculation of fitted mean and variance of y_t^2, we simply use the estimates of the parameters from Table 2.8 to the formulas for $\mathrm{E}[Y_t^2]$ in (2.25)

Table 2.8: The estimated parameter values for fitting the stochastic volatility model to the 1000 observations of the log return of the daily US-Dollar/Swiss-Franc exchange rates from July 24, 2007 to July 24, 2012.

Method	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\sigma}_\eta^2$
MM	-3.585664	0.6897393	1.191857
AGQL	-3.880882	0.6641959	1.269635
MQML	-10.87192	0.05930448	5.834532

and $\mathrm{Var}[Y_t^2]$ in (2.27). Next for the computation of the observed mean and variance of y_t^2, we consider a group of 50 observations represented by y_{t+l}^2 $(l = 0, \cdots, 49)$, take the mean and variance of these 50 observations, and report them at time point $t = 1, 51, 101, \cdots$, and so on. The observed means are displayed in Figure 2.1 and variances are displayed in Figure 2.2. The expected values $\mathrm{E}[Y_t^2]$ obtained by MM, AGQL and MQML approaches are also given in Figure 2.1, whereas the estimated $\mathrm{Var}[Y_t^2]$ under MM and AGQL approaches are given in Figure 2.2.

In Table 2.8, the estimation results for γ_0, γ_1 and σ_η^2 from MM and AGQL approaches are quite close to each other, which is consistent with Figures 2.1 and 2.2, where the curves for means and variances under the MM and AGQL approaches almost overlap each other. All these agree with our observation in Section 2.3 that the MM and AGQL methods give similar estimation results. In contrast, the MQML estimation results are quite different from those by MM and AGQL approaches. Figures 2.1 and 2.2 indicate that the MM and AGQL estimates are considerably better

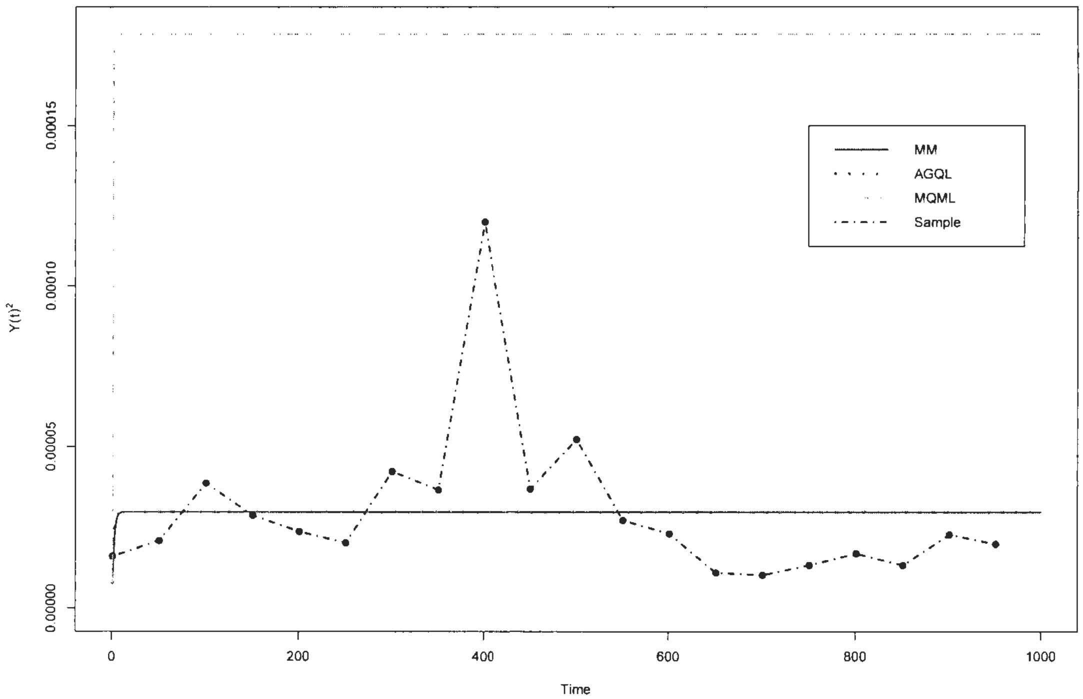

Figure 2.1: The estimated means of Y_t^2 for the 1000 data from the log return with mean subtracted of the US-Dollar/Swiss-Franc daily exchange rates from July 24, 2007 to July 24, 2012.

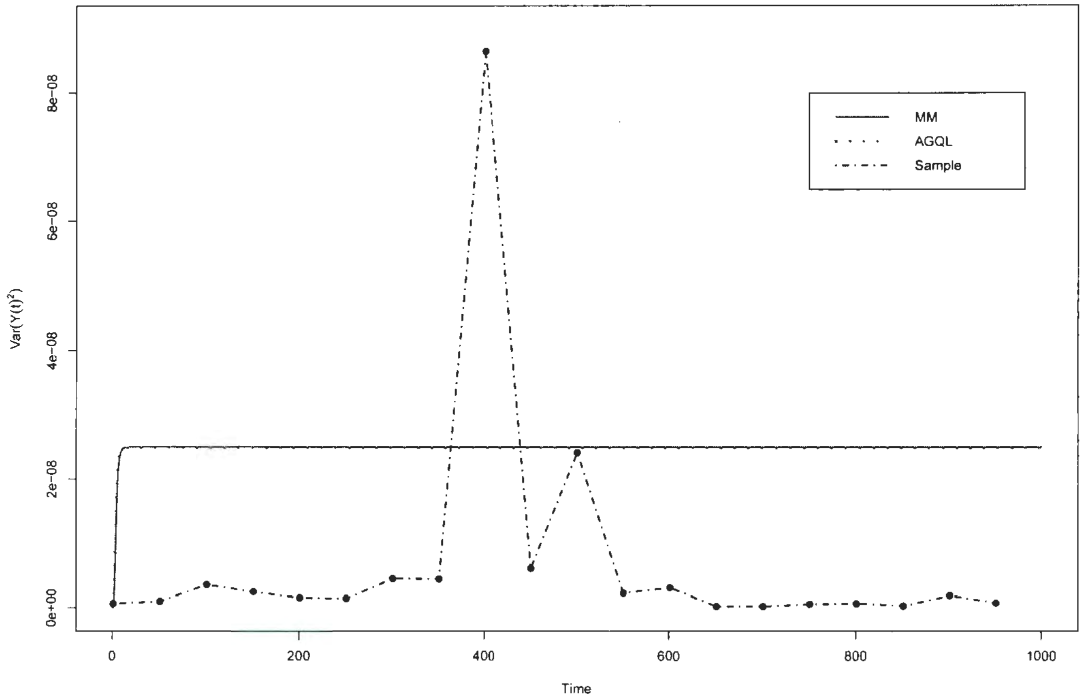

Figure 2.2: The estimated variances of Y_t^2 for the 1000 data from the log return with mean subtracted of the US-Dollar/Swiss-Franc daily exchange rates from July 24, 2007 to July 24, 2012.

than the MQML estimates.

In Figure 2.1, in general, the sample means of $\{y_t^2\}$ appear to lie either above or below the $\mathrm{E}\left[Y_t^2\right]$ level in a sequence, which shows positive correlations among y_t^2. This positive correlation can also be understood in theory by studying the patterns for lag covariances. To be specific, because

$$\mathrm{Cov}\left(Y_t^2, Y_{t+k}^2\right) = \lambda_t \lambda_{t+k} \left\{ \exp\left[2\gamma_1^k \left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2} \right) \right] - 1 \right\}, \quad (2.52)$$

it is obvious that for $0 < \gamma_1 < 1$, $\mathrm{Cov}\left(Y_t^2, Y_{t+k}^2\right) > 0$ and $\rightarrow 0$ as $\gamma_1 \rightarrow 0$ for all positive integer k. For negative γ_1, obviously correlations will be negative for odd k, but for the present data set γ_1 estimate is found to be positive.

To explain this issue further, the MM and AGQL approaches give γ_1 estimates as $\hat{\gamma}_{1,MM} = 0.690$ and $\hat{\gamma}_{1,AGQL} = 0.664$, which are all positive and can account for the positive correlation in Figure 2.1, while the MQML estimate of $\hat{\gamma}_{1,MQML} = 0.0593$ can be too small to explain this positive correlation. In addition, in Figure 2.1, the values for sample means appear to be close to the curves for the estimated mean by MM and AGQL approaches, indicating a reasonable fitting, while these sample values appear to be far away from the mean curve under the MQML method, implying that the MQML approach can not be applied to fit this data set, which may be due to the large standard errors of the MQML estimators.

Similar to the results for means shown in Figure 2.1, the curves for variances under

MM and AGQL approaches almost overlap each other, indicating the similarity of the estimates by the two approaches. Except the first several time points, the estimated variance of y_t^2 by MQML approach is on the order of 10^{-5}, which is considerably larger than the sample variances and the estimated variance values by MM and AGQL approaches, causing a scaling problem to accommodate the MQML results in the same figure. In comparison, the MM and AGQL curves have good agreement with the sample estimation points in Figure 2.2, indicating a better estimation than the MQML approach.

Chapter 3

Parameter Estimation for t-Distribution Based Volatility Models

In Chapter 2, we have discussed an improved estimation technique for the existing SV model with Gaussian error, as compared to the existing competitive approaches. Note that there also exist some studies dealing with SV models under the assumption that ϵ_t $(t = 1, \cdots, T)$ follow a heavy tailed t-distribution instead of Gaussian distribution. This produces much larger kurtosis than using normal distribution based SV model. See, for example, Harvey et al. (1994, Section 6) and Lee and Koopman (2004, Section 6). However, in these studies, the estimation of the parameters have been done mainly following the GMM and/or QML approaches.

As far as the models for time dependent variances are concerned, in Chapter 2, similar to the existing studies, we have considered lognormal distribution based AR(1) type model. The t-distribution based SV models are also developed based on lognormal AR(1) type relationship for the variances. There also exist some SV models where certain positive orthant distributions such as exponential and gamma distributions are used to model the time dependent variances (Abraham et al., 2006). However, they still use normal distribution for $\{\epsilon_t\}$. One could use t-distribution for those $\{\epsilon_t\}$ on top of using the gamma distributions based model for the variances. This is however beyond the scope of the present thesis.

In this chapter, we study the existing t-errors based SV model with variances satisfying the lognormal AR(1) model, but provide an improved estimation technique for all parameters including the degrees of freedom of the t-error distribution. To be more specific, in this chapter, we deal with a generalization of the model considered in Chapter 2, but use simpler MM, AGQL and MQML approaches for the estimation of the parameters. The degrees of freedom parameter in all cases is estimated by using the MM.

3.1 t-Distribution Based SV Models

We now turn back to the t-distribution based SV model studied by Harvey et al. (1994) [see also Lee and Koopman (2004)], but use the improved estimation technique introduced in Chapter 2 [see also Tagore (2010)]. To be specific, we recall the model form (1.1) - (1.2) and rewrite it here with a change in distribution of ϵ_t,. That is,

$$\begin{aligned} y_t|\sigma_t &= \sigma_t\epsilon_t \qquad t = 1, \cdots, T & (3.1)\\ \ln(\sigma_t^2) \equiv h_t &= \gamma_0 + \gamma_1 h_{t-1} + \eta_t \qquad t = 2, \cdots, T, & (3.2)\end{aligned}$$

where $\epsilon_t \overset{iid}{\sim} t(0, 1, \nu)$. That is, we use the distribution of ϵ_t as

$$f(\epsilon_t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(1 + \left(\frac{\epsilon_t^2}{\nu}\right)\right)^{(\nu+1)/2}}, \tag{3.3}$$

where ν is referred to as the shape or degrees of freedom parameter. Note that the t-distribution in (3.3) has the basic properties:

$$\begin{aligned} \mathrm{E}\left(\epsilon_t\right) &= 0 & (3.4)\\ \mathrm{Var}\left(\epsilon_t\right) &= \frac{\nu}{\nu - 2} & (3.5)\\ \mathrm{E}\left(\epsilon_t^4\right) &= \frac{3\nu^2}{(\nu-2)(\nu-4)}. & (3.6)\end{aligned}$$

Further note that the t-distribution (3.3) for ϵ_t, consequently, produces larger kurtosis for the data $\{y_t\}$ as compared to the Gaussian distribution based kurtosis. To be specific, to compute kurtosis under the t-model (3.1) - (3.2), we evaluate $E\left(Y_t^2\right)$ and $E\left(Y_t^4\right)$ as

$$\begin{aligned}
E(Y_t^2) &= E(\epsilon_t^2)E(\sigma_t^2) \\
&= \left[\frac{\nu}{\nu-2}\right]\exp\left\{\frac{\gamma_0}{1-\gamma_1}+\gamma_1^{t-1}\left(h_1-\frac{\gamma_0}{1-\gamma_1}\right)+\frac{\sigma_\eta^2}{2}\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right\}
\end{aligned} \tag{3.7}$$

$$\begin{aligned}
E(Y_t^4) &= E\left(\sigma_t^4\right)E\left(\epsilon_t^4\right) = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]E\left[\sigma_t^4\right] = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]E\left[e^{2h_t}\right] \\
&= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]\exp\left\{\frac{2\gamma_0}{1-\gamma_1}+2\gamma_1^{t-1}\left(h_1-\frac{\gamma_0}{1-\gamma_1}\right)+2\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right\}.
\end{aligned} \tag{3.8}$$

The above moments in (3.7) - (3.8) yield the kurtosis as

$$\begin{aligned}
\kappa_t^*(\nu,\gamma_1,\sigma_\eta^2) &= \frac{E\left(Y_t^4\right)}{\left[E\left(Y_t^2\right)\right]^2} = \frac{3\nu^2}{(\nu-2)(\nu-4)}\frac{(\nu-2)^2}{\nu^2}\exp\left[\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right] \\
&= \frac{3(\nu-2)}{(\nu-4)}\exp\left[\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right],
\end{aligned} \tag{3.9}$$

which is a function of ν parameter as well. When (3.9) is compared to the kurtosis (2.29) under the normal error, it is clear that the kurtosis in (3.9) under t-error distribution is $\frac{\nu-2}{\nu-4}$ times larger than the Gaussian based kurtosis. Thus, the kurtosis (3.9) models the heavy tails of the data through ν, σ_η^2 and γ_1.

3.2 Estimation of Parameters including Degrees of Freedom Parameters

As compared to the last chapter, it is clear that we now have one more (additional) parameter ν to estimate. Thus, all together we estimate $\gamma_0, \gamma_1, \sigma_\eta^2$ and ν.

Because ν is a new parameter as compared to the Gaussian model, we need an extra estimating equation for ν, whereas we can use the same estimating equations for other parameters as in Chapter 2, however, with slight adjustment due to the involvement of ν.

3.2.1 MM Approach for All Parameters

For the estimation of σ_η^2 and γ_1, we recall their equations from (2.3) to (2.8), and now make the adjustment for replacing Gaussian distribution of $\{\epsilon_t\}$ with t-distribution.

Thus, the moment equation for estimating σ_η^2 is given by

$$E[S_1] = E_{\sigma_t^2} E[\frac{1}{T}\sum_{t=1}^{T} y_t^2] = \frac{\nu}{\nu-2} E[\frac{1}{T}\sum_{t=1}^{T} \sigma_t^2] = \frac{\nu}{\nu-2} g_1 = S_1 \qquad (3.10)$$

where S_1 is defined in (2.1), and g_1 is given by (2.14). Comparing with the moment equation (2.14) for the Gaussian case, the only change is the insertion of multiplying factor $\frac{\nu}{\nu-2}$ in the equation.

Similarly, the new moment equation for estimating γ_1 is given by

$$\begin{aligned} E[S_2] &= E_{\sigma_t^2} E[\frac{1}{T-1}\sum_{t=2}^{T} y_{t-1}^2 y_t^2] = \left(\frac{\nu}{\nu-2}\right)^2 E[\frac{1}{T-1}\sum_{t=2}^{T}\sigma_{t-1}^2\sigma_t^2] \\ &= \left(\frac{\nu}{\nu-2}\right)^2 g_2 = S_2, \end{aligned} \tag{3.11}$$

where S_2 is defined in (2.6), and g_2 is given by (2.13). Comparing with the moment equation (2.13) for the Gaussian case, the only change is the insertion of multiplying factor $\left(\frac{\nu}{\nu-2}\right)^2$ in the equation.

Next for γ_0 estimation, we modify the Gaussian based equation as follows.

$$\hat{\gamma}_0 = \left.\frac{\{[S_3 - E(\log\epsilon_1^2)](1-\gamma_1) - \frac{1}{T}h_1(1-\gamma_1^T)\}(1-\gamma_1)}{1-\gamma_1-\frac{1}{T}(1-\gamma_1^T)}\right|_{\gamma_1=\hat{\gamma}_1} \tag{3.12}$$

where $S_3 = \frac{1}{T}\sum_{t=1}^{T}\log(y_t^2)$ as in (2.15), and since $\epsilon_t = \frac{u}{\sqrt{v/\nu}}$ for any t, with u being a standard normal random variable and v an independent chi-square random variable of degrees of freedom ν,

$$\begin{aligned} E\log\epsilon_t^2 &= E\log u^2 - E\log v + \log\nu \\ &= \psi\left(\frac{1}{2}\right) + \log 2 - \psi\left(\frac{\nu}{2}\right) - \log 2 + \log\nu \\ &= \left[\psi\left(\frac{1}{2}\right) - \log\left(\frac{1}{2}\right)\right] - \left[\psi\left(\frac{\nu}{2}\right) - \log\left(\frac{\nu}{2}\right)\right]. \end{aligned} \tag{3.13}$$

Note that $E\log v$ is general as compared to finding $E\log u^2$, as u^2 is simply χ_1^2. For $E\log v$, we use the formula available from Chan (1993). In (3.13), $\psi(x)$ is diagamma function, which is calculated by the following formula [Beal, M. J. (2003)]

$$\begin{aligned}
\psi(x) &= \ln(x) - \frac{1}{2x} - \frac{1}{12x^2} + \frac{1}{120x^4} - \frac{1}{252x^6} + \frac{1}{240x^8} - \frac{5}{660x^{10}} \\
&\quad + \frac{691}{32760x^{12}} - \frac{7}{84x^{14}} + \frac{3617}{8160x^{16}} - \frac{43867}{14364x^{18}} + O\left(\frac{1}{x^{20}}\right) \\
\psi(x+1) &= \frac{1}{x} + \psi(x) .
\end{aligned}$$

We followed Beal's suggestion [Beal, M. J. (2003)] of using the above recurrence to shift x to a value greater than 6 and then applying the above expansion with terms above x^{14} cut off, which yields "more than enough precision".

As far as the estimation of ν parameter is concerned, we write an additional equation by using sample statistics

$$S_4 = \frac{1}{T}\sum_{t=1}^{T} Y_t^4,$$

for which

$$\begin{aligned}
E(S_4) &= \frac{1}{T}\sum_{t=1}^{T} E(Y_t^4) = \frac{1}{T}\sum_{t=1}^{T} E(\epsilon_t^4)E(\sigma_t^4) \\
&= \frac{\nu^2}{(\nu-2)(\nu-4)}\, g_4,
\end{aligned} \tag{3.14}$$

with

$$g_4 = \frac{3}{T}\sum_{t=1}^{T} \exp\left\{\frac{2\gamma_0}{1-\gamma_1} + 2\gamma_1^{t-1}\left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right) + 2\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right\}, \tag{3.15}$$

where we used the result for $E(\sigma_t^4)$ in (2.27). Equating (3.14) with S_4 and combining it with Eqs. (3.10), we have

$$\frac{S_1^2}{S_4} = \frac{g_1^2}{g_4}\left(\frac{\nu-4}{\nu-2}\right), \tag{3.16}$$

yielding the estimating equation for ν as

$$\nu = 2 + \frac{2}{1 - \frac{S_1^2 g_4}{S_4 g_1^2}}. \tag{3.17}$$

For the purpose of kurtosis estimation, we assume that $\nu > 4$.

For $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma_\eta^2 = 0.25, \nu = 10$ and 500 simulations, the estimation results are given in Table 3.1, which shows that the estimates are reasonably good. This is especially true for γ_1 and ν, for which the estimates are reasonably quite close to the true parameter values, for even $T = 1000$. Note that under the normal SV model, we have chosen γ_0 values $\gamma_0 = 0.05, 0.1$ and 0.2. The estimates were found to be reasonable. Here we have taken $\gamma_0 = 0.05$ case only to see how much this and other parameters are affected by ν. Estimates were found to be good here as well. Other values of γ_0 were not chosen to save space.

		$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\sigma}^2_\eta$	$\hat{\nu}$
$T = 1000$	Simulation Mean	0.03530586	0.4202329	0.2280030	10.55707
	Simulation SE	0.05059033	0.3799676	0.1658846	8.846058
$T = 3000$	Simulation Mean	0.03990508	0.4929931	0.2164728	11.31905
	Simulation SE	0.03811798	0.3000582	0.1426227	14.73339

Table 3.1: Estimation results for $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma^2_\eta = 0.25, \nu = 10$ and 500 simulations with method of moments.

3.2.2 AGQL for All Parameters

Recall that from (2.34) and (2.33), the AGQL estimating equations for γ_1 and σ^2_η have the form

$$\frac{\partial \psi'}{\partial \gamma_1}\Omega_d^{-1}(\mathrm{v} - \psi) \;=\; 0$$

and

$$\frac{\partial \lambda'}{\partial \sigma^2_\eta}\Sigma_d^{-1}(\mathrm{u} - \lambda) \;=\; 0\,,$$

respectively, where

$$\mathrm{u} \;=\; [y_1^2, y_2^2, \cdots, y_t^2, \cdots, y_{T-1}^2, y_T^2]'$$

and

$$\mathrm{v} \;=\; [y_1^2 y_2^2, \cdots, y_{t-1}^2 y_t^2, \cdots, y_{T-1}^2 y_T^2]'.$$

However, their expectations and covariances will have slightly different formula because of the use of t-distribution for ϵ_t. Thus, we need to compute

$$\begin{aligned}\lambda &= E[\mathrm{U}]\\ \Sigma &= \mathrm{cov}(\mathrm{U})\\ \psi &= E[\mathrm{V}]\\ \Omega &= \mathrm{cov}(\mathrm{V})\end{aligned}$$

under t-distribution for ϵ_t. The new formulas are as follows.

$$\lambda_1 = E(Y_1^2) = \left[\frac{\nu}{\nu-2}\right] E(\sigma_1^2) = \left[\frac{\nu}{\nu-2}\right] e^{h_1}, \tag{3.18}$$

and for $t \geq 2$

$$\begin{aligned}&\lambda_t = E(Y_t^2) = E(\epsilon_t^2)E(\sigma_t^2)\\ &= \left[\frac{\nu}{\nu-2}\right] \exp\left\{\frac{\gamma_0}{1-\gamma_1} + \gamma_1^{t-1}\left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right) + \frac{\sigma_\eta^2}{2}\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right\}.\end{aligned} \tag{3.19}$$

$$E\left(Y_1^2 Y_1^2\right) = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] \sigma_1^4 = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] e^{2h_1} \tag{3.20}$$

$$Var\left(Y_1^2\right) = E\left(Y_1^2 Y_1^2\right) - \lambda_1^2 = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] e^{2h_1} - e^{2h_1}.$$

For $t \geq 2$

$$E\left(Y_t^4\right) = E\left(\sigma_t^4\right) E\left(\epsilon_t^4\right) = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] E\left[\sigma_t^4\right] = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] E\left[e^{2h_t}\right]$$
$$= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right] \exp\left\{\frac{2\gamma_0}{1-\gamma_1} + 2\gamma_1^{t-1}\left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right) + 2\sigma_\eta^2\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1^2}\right)\right\} \tag{3.21}$$

$$Var\left(Y_t^2\right) = E\left(Y_t^4\right) - \lambda_t^2.$$

$$\psi_1 = E[Y_1^2 Y_2^2] = E[Y_1^2]E[Y_2^2] = \lambda_1\lambda_2,$$

and for $t \geq 2$

$$\psi_t = E[Y_t^2 Y_{t+1}^2] = E_{\sigma_t^2,\sigma_{t+1}^2}[Y_t^2 Y_{t+1}^2 | \sigma_t^2, \sigma_{t+1}^2]$$
$$= \left(\tfrac{\nu}{\nu-2}\right)^2 \exp\left\{\tfrac{2\gamma_0}{1-\gamma_1} + \left(h_1 - \tfrac{\gamma_0}{1-\gamma_1}\right)\gamma_1^{t-1}(1+\gamma_1) + \tfrac{\sigma_\eta^2}{2}\left[(1+\gamma_1)\left(\tfrac{1-\gamma_1^{2(t-1)}}{1-\gamma_1}\right) + 1\right]\right\}.$$

$$\frac{\partial\psi_1}{\partial\gamma_1} = \lambda_1\frac{\partial\lambda_2}{\partial\gamma_1} = \psi_1 h_1.$$

For $t \geq 2$

$$\frac{\partial\psi_t}{\partial\gamma_1} = \psi_t\left\{\frac{2\gamma_0}{(1-\gamma_1)^2} + \left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right)\left[(t-1)\gamma_1^{t-2}(1+\gamma_1) + \gamma_1^{t-1}\right] - \frac{\gamma_0}{(1-\gamma_1)^2}\gamma_1^{t-1}(1+\gamma_1)\right.$$
$$\left. + \frac{\sigma_\eta^2}{2}\left[\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1} + (1+\gamma_1)\frac{\left[1-\gamma_1^{2(t-1)} - 2(t-1)\gamma_1^{2t-3}(1-\gamma_1)\right]}{(1-\gamma_1)^2}\right]\right\}.$$

$$\begin{aligned} E\left[Y_1^4 Y_2^4\right] &= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 E\left[\sigma_1^4\sigma_2^4\right] = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 \sigma_1^4 E\left[e^{2h_2}\right] \\ &= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 \sigma_1^4 \exp\left\{\frac{2\gamma_0}{1-\gamma_1} + 2\gamma_1\left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right) + 2\sigma_\eta^2\right\}. \end{aligned}$$

For $t \geq 2$

$$\begin{aligned} E\left[Y_t^4 Y_{t+1}^4\right] &= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 \left[\sigma_t^4\sigma_{t+1}^4\right] = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 \left[e^{2(h_t+h_{t+1})}\right] \\ &= \left[\frac{3\nu^2}{(\nu-2)(\nu-4)}\right]^2 \exp\left\{\frac{4\gamma_0}{1-\gamma_1} + 2\left(h_1 - \frac{\gamma_0}{1-\gamma_1}\right)\gamma_1^{t-1}(1+\gamma_1)\right. \\ &\quad \left. +2\sigma_\eta^2\left[(1+\gamma_1)\left(\frac{1-\gamma_1^{2(t-1)}}{1-\gamma_1}\right)+1\right]\right\}. \end{aligned}$$

Then

$$\mathrm{Var}(Y_t^2 Y_{t+1}^2) = E[Y_t^4 Y_{t+1}^4] - \psi_t^2.$$

For convenience, we now write an algorithm for the desired estimation, with following two steps.

Step 1: Estimating γ_1 iteratively with

$$\hat{\gamma}_1(r+1) = \hat{\gamma}_1(r) + \left[\left(\frac{\partial\psi'}{\partial\gamma_1}\Omega_d^{-1}\frac{\partial\psi}{\partial\gamma_1}\right)^{-1}\left(\frac{\partial\psi'}{\partial\gamma_1}\Omega_d^{-1}(\mathrm{v}-\psi)\right)\right]_{[r]}.$$

In each iteration, γ_0 should be updated by the new γ_1 with

$$\hat{\gamma}_0 = \left.\frac{\left\{\left[S_3 - E\left(\log\epsilon_1^2\right)\right](1-\gamma_1) - \frac{1}{T}h_1\left(1-\gamma_1^T\right)\right\}(1-\gamma_1)}{1-\gamma_1-\frac{1}{T}\left(1-\gamma_1^T\right)}\right|_{\gamma_1=\hat{\gamma}_1},$$

where

$$S_3 = \frac{1}{T}\sum_{t=1}^{T}\log\left(y_t^2\right),$$

and ν should be updated by the new ν with

$$\nu = 2 + \frac{2}{1 - \frac{S_1^2 g_4}{S_4 g_1^2}} .$$

Step 2: Estimating σ_η^2 iteratively with

$$\hat{\sigma}_\eta^2(r+1) = \hat{\sigma}_\eta^2(r) + \left[\left(\frac{\partial \lambda'}{\partial \sigma_\eta^2} \Sigma_d^{-1} \frac{\partial \lambda}{\partial \sigma_\eta^2} \right)^{-1} \left(\frac{\partial \lambda'}{\partial \sigma_\eta^2} \Sigma_d^{-1} (\mathrm{u} - \lambda) \right) \right]_{[r]} .$$

This two step circle of iterations continues until convergence.

For $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma_\eta^2 = 0.25, \nu = 10$ and 500 simulations, the estimation results are given in Table 3.2. It appears from this table that similar to the MM approach, the AGQL approach also estimates the parameters well. When standard errors are compared, they appear to be similar under both MM and AGQL approaches for the estimation of γ_0, γ_1 and σ_η^2. However, the AGQL approach appears to estimate ν with much smaller standard error as compared to the MM approach.

		$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\sigma}_\eta^2$	$\hat{\nu}$
$T = 1000$	Simulation Mean	0.03645553	0.4153435	0.1887665	9.469515
	Simulation SE	0.05856634	0.4302302	0.1732873	5.950574
$T = 3000$	Simulation Mean	0.03828042	0.5063921	0.2072618	10.50157
	Simulation SE	0.03791588	0.3157161	0.1489828	9.251301

Table 3.2: Estimation results for $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma_\eta^2 = 0.25, \nu = 10$ and 500 simulations with AGQL.

3.2.3 Modified QML (MQML) for All Parameters

We have discussed the quasi-maximum likelihood approach in Section 2.2.3 for Gaussian model where, for z_t defined in (2.35), $z = (z_1, \cdots, z_t, \cdots, z_T)'$ was approximated by a multivariate normal distribution, even though truly $\log \epsilon_t^2$ follows a $\log \chi_1^2$ distribution under normal $\{\epsilon_t\}$. However, in the present case, we assume $\{\epsilon_t\}$ follow t-distribution as in (3.3). This change in distribution of ϵ_t influences the MQML estimating equations only through $\kappa_1 = E[\log \epsilon_t^2]$ and $\kappa_2 = \text{var}[\log \epsilon_t^2]$. So for t-distribution based SV model, the only change we need to make is to calculate $E[\log \epsilon_t^2]$ and $\text{var}[\log \epsilon_t^2]$ for $\epsilon_t \sim t(0, 1, \nu)$, for replacing the κ_1 and κ_2 in Section 2.2.3 by the new $E[\log \epsilon_t^2]$ and $\text{var}[\log \epsilon_t^2]$ respectively, and estimate ν by

$$\nu = 2 + \frac{2}{1 - \frac{S_1^2 g_4}{S_4 g_1^2}},$$

then all the other formulas remain the same as in Section 2.2.3.

$E[\log \epsilon_t^2]$ was computed in (3.13). We now compute $\text{var}[\log \epsilon_t^2]$ as follows. Let $x \sim \text{Gamma}(\alpha, \beta)$ and

$$y = \log(x) \Longrightarrow x = e^y.$$

The pdf of y is

$$\begin{aligned} f(y) &= \frac{\beta^\alpha}{\Gamma(\alpha)} e^{(\alpha-1)y} e^{-\beta e^y} e^y = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{\alpha y - e^{(y+\log\beta)}} = \frac{1}{\Gamma(\alpha)} e^{\alpha(y+\log\beta) - e^{(y+\log\beta)}} \\ &= \frac{1}{\Gamma(\alpha)} e^{\alpha w - e^w}, \end{aligned}$$

where $w = y + \log\beta$ being a log-gamma random variable, and according to Chan (1993)

$$\begin{aligned} E[y] &= E[w] - \log\beta = \psi(\alpha) - \log(\beta) \\ \text{var}[y] &= \text{var}[w] = \psi'(\alpha). \end{aligned} \tag{3.22}$$

Applying (3.22) we now compute the variance of $\log\epsilon_t^2 = \log z^2 - \log v + \log\nu$ as

$$\text{var}[\log\epsilon_t^2] = \text{var}[\log z^2] + \text{var}[\log v] = \psi'\left(\frac{1}{2}\right) + \psi'\left(\frac{\nu}{2}\right). \tag{3.23}$$

For $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma_\eta^2 = 0.25, \nu = 10$ and 500 simulations, the estimation results are shown in Table 3.3. Comparing with Tables 3.1 and 3.2 for the proposed MM and AGQL approaches respectively, the estimation of ν by MQML is as good as those from MM and AGQL approaches. However, the estimate of γ_1 by MQML is relatively much farther away from the true parameter value than the other two approaches, and the estimates of γ_0 and σ_η^2 by MQML are considerably worse than those from the other two approaches, showing that the proposed MM and AGQL approaches are improved estimating methods for the SV models.

		$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\sigma}_\eta^2$	$\hat{\nu}$
$T = 1000$	Simulation Mean	0.01815647	0.6104038	0.1500680	9.080930
	Simulation SE	0.04933629	0.4295255	0.1948482	6.594675
$T = 3000$	Simulation Mean	0.02552030	0.6542555	0.1455136	9.653351
	Simulation SE	0.03669045	0.3459653	0.1671992	8.623688

Table 3.3: Estimation results for $\gamma_0 = 0.05, \gamma_1 = 0.5, \sigma_\eta^2 = 0.25, \nu = 10$ and 500 simulations with MQML.

3.2.4 Kurtosis for t-distribution Case

To understand the volatility, that is, to understand the changes in variance pattern in the time series, it is recommended to examine the kurtosis of the data over time. See, for example, Jacquier et.al (1994, p.387) Shephard (1996, p.23), Mills (1999, p.129), Ruiz (2004, p.615) and Tsay (2005, p.134)). The kurtosis for Gaussian SV model and t-distribution based SV model can be calculated with (2.29) and (3.9) respectively, whereas it is known that Gaussian distribution based kurtosis is: $\kappa = 3$, and the t-distribution based kurtosis is: $\kappa = 3(\nu - 2)/(\nu - 4)$. The plots of kurtosis for t-distribution based SV model with different ν values, and for normal distribution based SV model are given in Figures 3.1 - 3.5. Also the standard kurtosis for t_ν (t-distribution with ν degrees of freedom) and normal distributions are shown in the same figures. For comparison, they are plotted with the same kurtosis range. We have also plotted in Figure 3.4 the estimated kurtosis with MM both for normal and t-distribution based SV models.

Note that it is expected that the kurtosis under the Gaussian SV model (1.1) - (1.2) will be larger than the Gaussian based kurtosis (=3). Similarly, the kurtosis under the t-distribution based SV model (3.1) - (3.2) will also be larger than the t-distribution based kurtosis ($= 3(\nu-2)/(\nu-4)$). Now because t-distribution has heavier tails than the Gaussian distribution, it is expected that the kurtosis for t-distribution based SV model will be much larger than the simpler Gaussian distribution based kurtosis. Further note that these figures 3.1 - 3.5 provide a clear feeling on the changes in the magnitude of kurtosis for various t-distributions. To be specific, when ν gets smaller, the kurtosis gets larger. Figure 3.4 shows that the estimated kurtosis are quite close to the true kurtosis as expected. This is because the MM approach produces satisfying estimates for all the parameters.

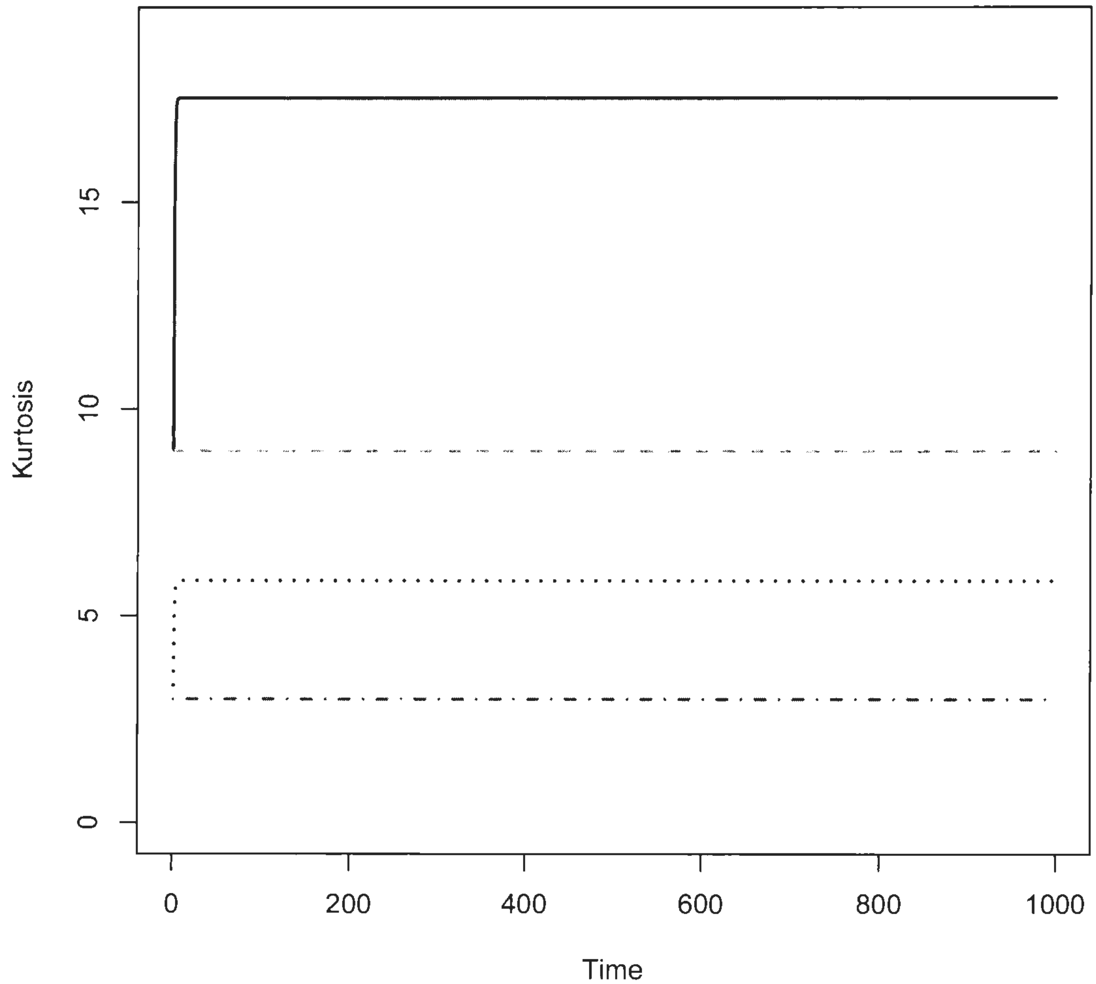

Figure 3.1: The kurtosis for the t-distribution-based SV model with parameters $\nu = 5, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (—), for t-distribution with parameter $\nu = 5$ (- - -), for the normal distribution-based SV model with parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ ($\cdots\cdots$), and for normal distribution ($\cdot - \cdot -$).

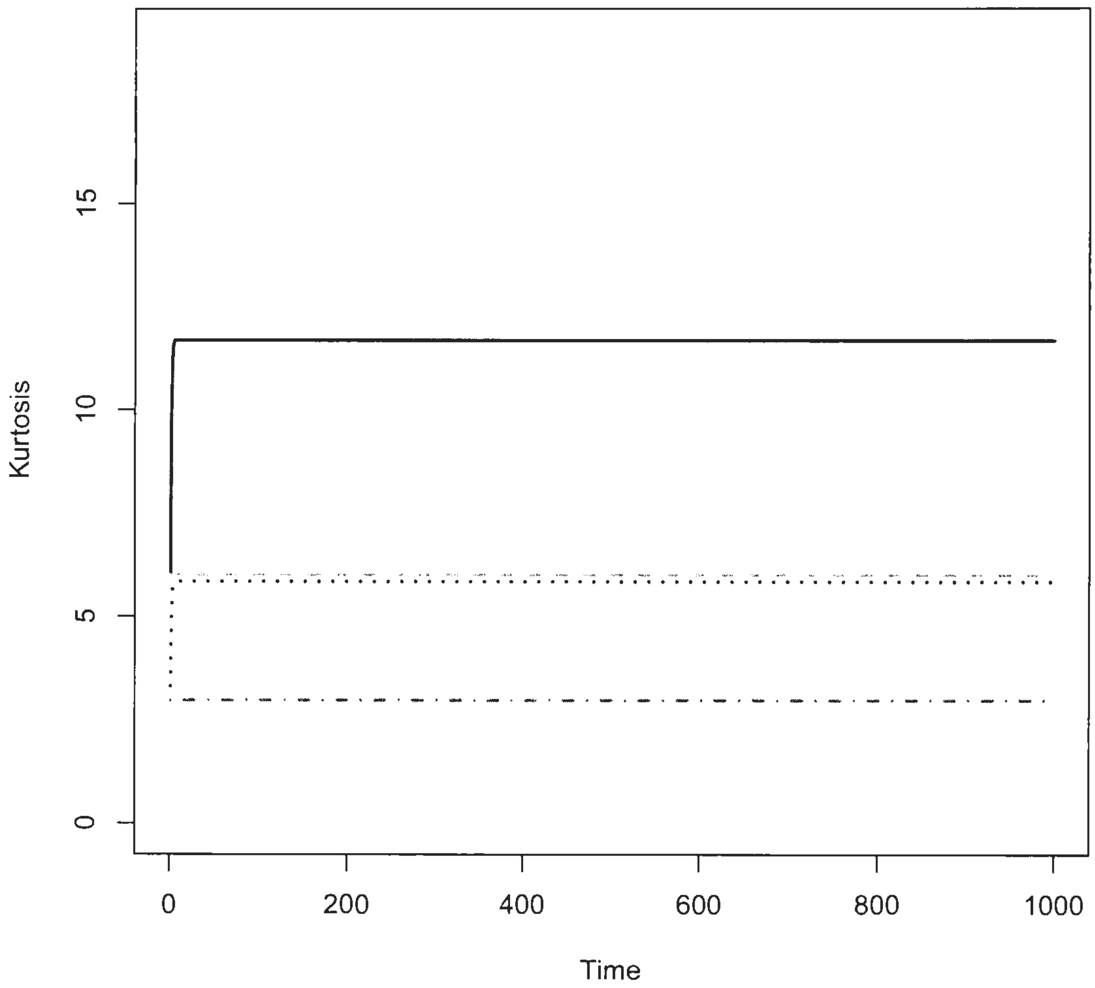

Figure 3.2: The kurtosis for the t-distribution-based SV model with parameters $\nu = 6, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (—), for t-distribution with parameter $\nu = 6$ (- - -), for the normal distribution-based SV model with parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ ($\cdots\cdots$), and for normal distribution ($\cdot - \cdot -$).

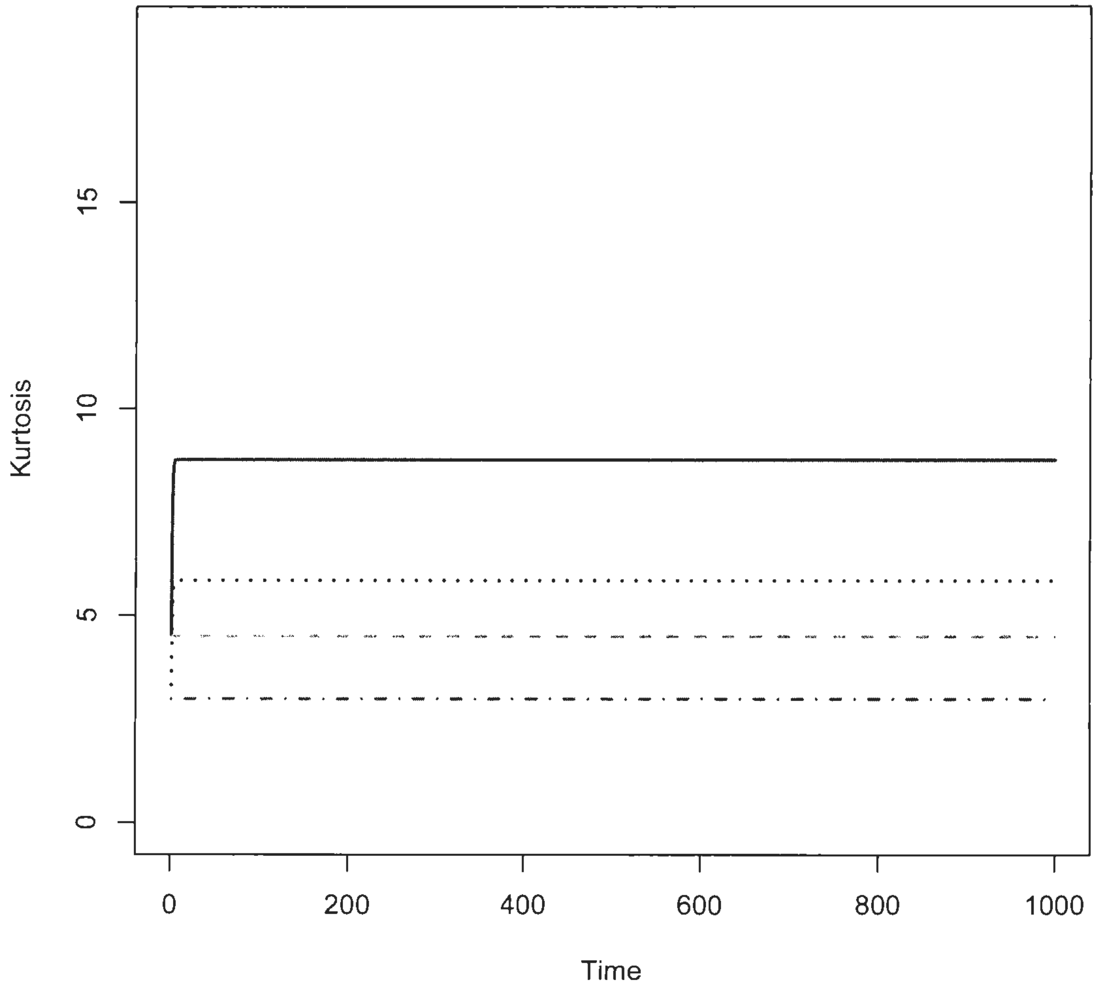

Figure 3.3: The kurtosis for the t-distribution-based SV model with parameters $\nu = 8, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (—), for t-distribution with parameter $\nu = 8$ (- - -), for the normal distribution-based SV model with parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ $(\cdots\cdots)$, and for normal distribution $(\cdot - \cdot -)$.

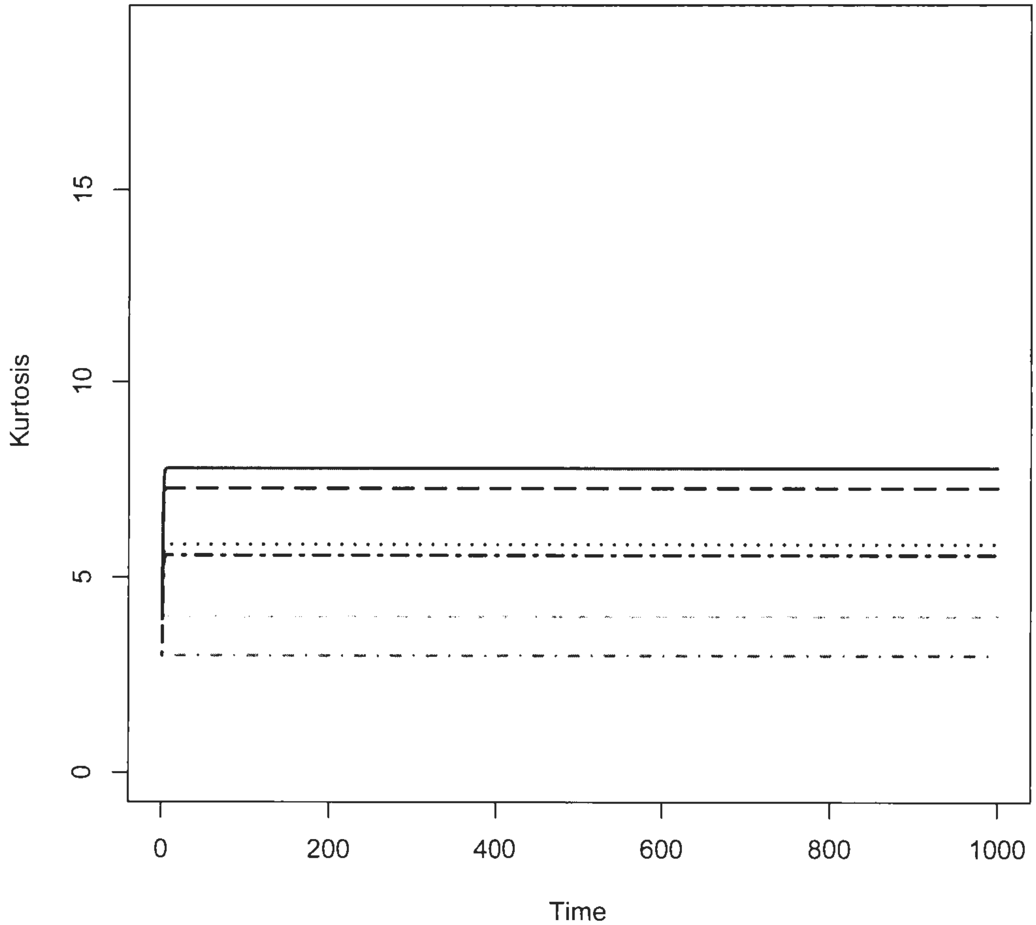

Figure 3.4: The kurtosis for the t-distribution-based SV model with parameters $\nu = 10, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (—), for t-distribution with parameter $\nu = 10$ (- - -), for the normal distribution-based SV model with parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (······), and for normal distribution (· — · —), and the estimated kurtosis for the t-distribution-based SV model with $T = 3000$ and parameters $\nu = 10, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (– – –), and for the normal distribution-based SV model with $T = 3000$ and parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (– - - -).

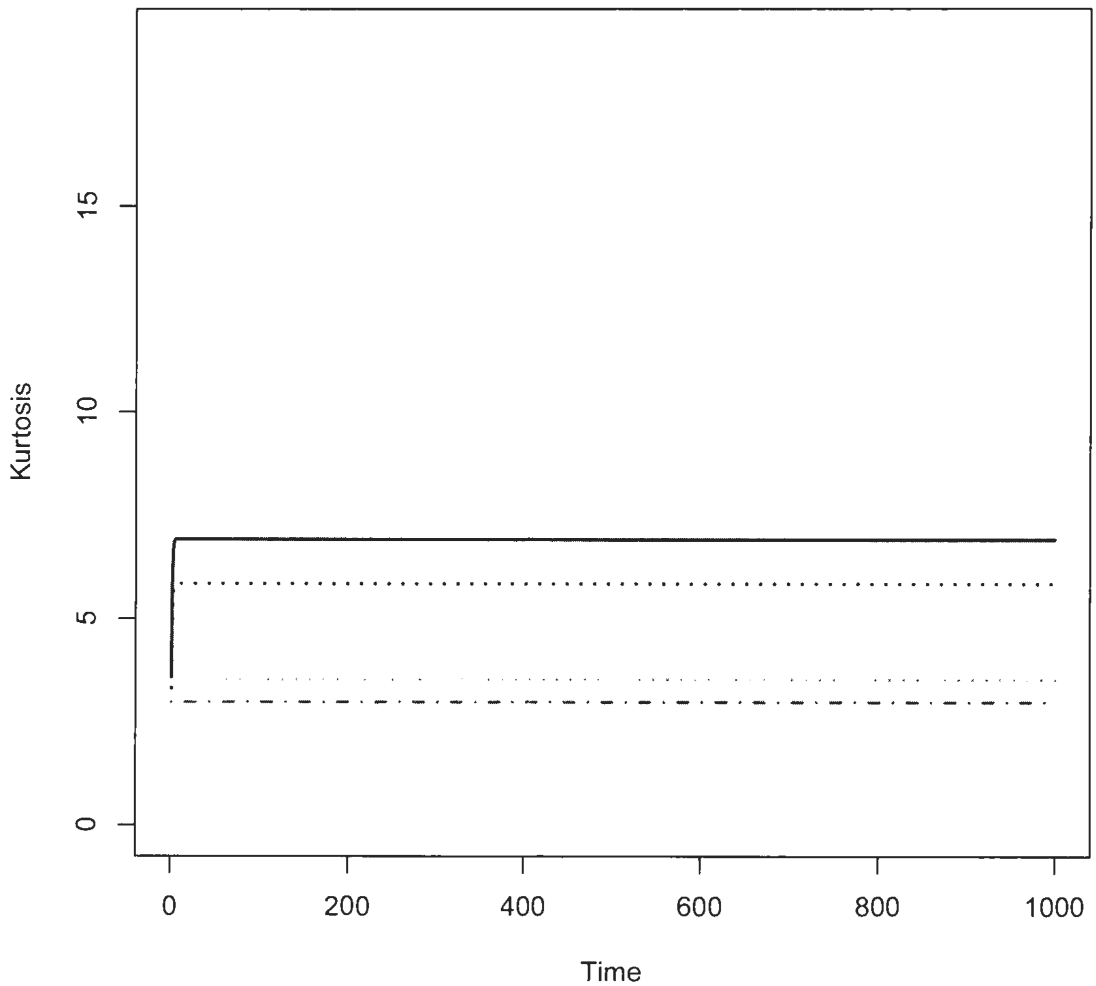

Figure 3.5: The kurtosis for the t-distribution-based SV model with parameters $\nu = 15, \gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ (—), for t-distribution with parameter $\nu = 15$ (- - -), for the normal distribution-based SV model with parameters $\gamma_1 = 0.5$ and $\sigma_\eta^2 = 0.5$ $(\cdots\cdots)$, and for normal distribution $(\cdot - \cdot -)$.

Chapter 4

Concluding Remarks

To fit the volatility model, the existing GMM and QML approaches are widely used. However, the GMM approach uses a large number of moments [Anderson and Sorensen (1996, p. 350-351), Anderson and Sorensen (1997, sections 3, p. 399-400)] to construct the GMM estimating equations for the consistent estimation of the volatility parameters, and the QML approach uses a normal approximation to a log chi-square distribution that arises in the construction of the so-called likelihood estimating equations. In this thesis, it is demonstrated that unlike the GMM approach, the moment estimating equations for three volatility parameters can be constructed by using only three unbiased moment functions selected carefully following the nature or definition of the parameters. This simpler approach has been referred to as the

MM (method of moments) approach. As the GMM approach is complex, it was not included for any comparison in this thesis. We have also provided a modification to the QML approach by using a modified (simpler) covariance matrix involved in the QML estimating equation. The finite sample behaviour of the proposed MM estimation approach was studied intensively, and it is found that the MM approach works very well in estimating all volatility parameters for time series size as small as 1000. The drawbacks of the QML approach is discussed. An AGQL approach was also considered. This AGQL approach performs similarly to the MM approach, however, it is computationally more involved than the MM approach. All three methods, MM, AGQL and QML were applied to fit the SV model (1.1) - (1.2) to a real life financial time series with length T = 1000, and it is found that the MM and AGQL approaches provide relatively much better fitting than the QML approach.

We also applied the MM, AGQL and MQML approaches to the heavy tailed t-distribution based SV model (3.1) - (3.2), and proposed the moment estimation for the degrees of freedom of the t-distribution. Simulation study shows that the three approaches give reasonably good estimates of the model parameters including the degrees of freedom parameter. Finally we compared the kurtosis for the SV models under Gaussian and t distributions. Our results show that the estimated kurtosis are quite close to the true kurtosis because the proposed MM approach produces satisfactory estimates for all the parameters. It was further demonstrated that the kurtosis under the Gaussian SV model (1.1) - (1.2) are larger, as expected, than the Gaussian based kurtosis (=3); and similarly the kurtosis under the t-distribution based SV model (3.1) - (3.2) are larger than the t-distribution based kurtosis ($= 3(\nu-2)/(\nu-4)$). Consequently, because the t-distribution has heavier tails than the Gaussian distribution, the kurtosis for t-distribution based SV model are much larger than the simpler Gaussian distribution based kurtosis. Consequently, this heavy tails based SV model becomes practically useful as some financial data appear to show high volatility.

Bibliography

Abraham, B., Balakrishna, N. and Sivakumar, R. (2006). Gamma stochastic volatility models. *Journal of Forecasting* **25**, 153-171.

Abromovitz, M. and Stegun, N. C. (1970). *Handbook of Mathematical Functions.* New York: Dover Publications Inc.

Ahn, S. and Schmidt, P. (1995). Efficient Estimation of Models for Dynamic Panel Data. *Journal of Econometrics* **68**, 5-28.

Anderson, J. (2001). On the normal inverse Gaussian stochastic volatility model. *Journal of Business and Economic Statistics* **19**, 44-54.

Andersen, T. G. and Sorensen, B. E. (1996). GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *Journal of Business and Economic Statistics* **14**, 328-352.

Andersen, T. G. and Sorensen, B. E. (1997). GMM and QML asymptotic standard deviations in stochastic volatility models: Comments on Ruiz (1994). *Journal of Econometrics* **76**, 397-403.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data : Monte Carlo Evidence and an Application to Employment Equations. *Review of Economics and Statistics* **58**, 277-298.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* **24**, 1-13.

Beal, M. J. (2003). Variational Algorithms for Approximate Bayesian Inference (PhD thesis). The Gatsby Computational Neuroscience Unit, University College London. pp. 265-266.

Black, R. (1986). Noise. *Journal of Finance* **3**, 529-543.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

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

Broto, C. and Ruiz, E. (2004). Estimation Methods For Stochastic Volatility Models: A Survey. *Journal of Economic Surveys* **18**, 613-649.

Chamberlain, G. (1992). Sequential Moment Restrictions in Panel Data-Comment. *Journal of Business and Economic Statistics* **10**, 20-26.

Chan, P. S. (1993). A Statistical Study of Log-Gamma Distribution (PhD thesis). McMaster University.

Danielsson, J. (1994). Stochastic Volatility in asset prices: estimation with simulated maximum likelihood. *Journal of Econometrics* **64**, 375-400.

Durbin, J. and Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika* **84**, 669-684.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometica* **50**, 987-1007.

Gallant, A. R., Hsieh, D. and Tauchen, G. (1997). Estimation of stochastic volatility model with diagnostics. *Journal of Econometrics* **81**, 159-192.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). On the relation between the expected value and the volatility of nominal excess return on stocks. *Journal of Finance* **48**, 1779-1801.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029-1054.

Harvey, A. C. (1998). Long memory in stochastic volatility. In J. Knight and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, pp., 307-320. Oxford: Butterworth Heinemann.

Harvey, A. C., Ruiz, E. and Shephard, N. (1994). Multivariate shochastic variance models. *Review of Econometrica* **61**, 247-264.

Harvey, A. C. and Shephard, N. G. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics* **14**, 429-434.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: with applications to financial and economic time series.* Cambridge University Press.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics* **12**, 371-417.

Jacquier, E., Polson, N. G. and Rossi, P. E., (1995). Models and prior for multivariate stochastic volatility. *CIRANO*, Working Paper, 95s-18.

Keane, M. and Runkle, D. (1992). On the Estimation of Panel Data Models With Serial Correlation When Instruments Are not Strictly Exogenous. *Journal of Business and Economic Statistics* **10**, 1-9.

Koopman, S. J., Harvey, A. C., Doornik, J. A. and Shephard, N. (1995). *Stamp 5.0: Structural Time Series Analyser, Modeller, and Predictor,* London: Chapman and Hall.

Lee, K. M. and Koopman, S. J. (2004). Estimating stochastic volatility models: A comparison of two importance samplers. *Studies in Nonlinear Dymaics and Economics* **8**, 1-15.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: conditional normality versus heavy tailed distributions. *Journal of Applied Econometrics* **15**, 137-160.

Liesenfeld, R. and Richard, J. F. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of Empirical Finance* **207**, 1-27.

Mahieu, R. J. and Schotman, P. C. (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics* **13**, 333-360.

Melino, A. and Turnbull, S. M. (1990). Pricing foregin currency options with stochastic volatility. *Journal of Econometrics* **45**, 239-265.

Mills, T. C. (1999). *The Econometric Modelling of Finance Time Series.* Cambridge University Press.

Nelson, D. B. (1988). The time series behavior of stock market volatility and returns. (Unpublished PhD Dissertation, Massachusetts Institute of Technology, Cambridge, MA).

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **59**, 347-370.

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

Ruiz, E. (1997). QML and GMM estimators of stochastic volatility models: Response to Anderson and Sorensen. *Journal of Econometrics* **76**, 405.

Sandmann, G. and Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics* **87**, 271-301.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. *In Time Series Models in Econometrics,Finance and Other Fields*, Cox, D.R, Hinkley, D.V,Barndorff-Nielsen,O.E (eds). *Chapman & Hall: London;* 1-67.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-Gaussian parameter-driven models. *Biometrica* **84**, 653-667.

So, M. K. P., Li, W. K. and Lam, K. (2002). A threshold stochastic volatility model. *Journal of Forecasting* **21**, 473-500.

Steel, M. F. J. (1998). Bayesian analysis of stochastic volatility models with flexible tails. *Econometric Reviews* **17(2)**, 109-143.

Sutradhar, B. C. (2003). An review on regression models for discrete longitudinal responses. *Statistical Science* **18**, 377-393.

Sutradhar, B. C. (2004). On exact quasilikelihood inference in generalized linear mixed models. *The Indian Journal of Statistics* **66**, 261-289.

Tagore, V. (2010). A simple and efficient estimation approach for stochastic volatility models. (Unpublished PhD Thesis, Memorial University of Newfoundland, St. John's, NL, Canada).

Tsay, R. S. (2005). Analysis of Finanical Time Series. *Wiley Interscience; New York.*

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes - A Study of Daily Sugar Prices, 1961-79. *Time Series Analysis: Theory and Practice 1, Edited by O.D Anderson* . North-Holland Publishing Company.

Taylor, S. J. (1986). *Modelling Financial Time Series.* John Wiley, Chichester.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* **4**, 183-204.

Usmani, R. (1994). Inversion of a tridiagonal Jacobi matrix, *Linear Algebra Appl* **212/213**, 413-414.

Watanabe, T. and Asai, M. (2001). Stochastic volatility models with heavy tailed distributions: A Bayesian approach. *IMES Discussion Paper Series 2001-E-17.*

Zakoian, J. M. (1994). Threshold heteroscedastic models. *Journal of Economic Dynamics and Control* **18**, 931-955.