Univariate and Bivariate ACD Models for High-Frequency Data Based on Birnbaum-Saunders and Related Distributions
UNIVARIATE AND BIVARIATE ACD MODELS FOR HIGH-FREQUENCY DATA BASED ON BIRNBAUM-SAUNDERS AND RELATED DISTRIBUTIONS

BY
TAO TAN

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree of
Doctor of Philosophy

McMaster University
© Copyright by Tao Tan, September 2018
Doctor of Philosophy (2018) McMaster University
(Mathematics) Hamilton, Ontario, Canada

TITLE: Univariate and Bivariate ACD Models for High-Frequency Data Based on Birnbaum-Saunders and Related Distributions

AUTHOR: Tao Tan

SUPERVISOR: Professor N. Balakrishnan

NUMBER OF PAGES: xv, 132
To my husband Dr. Hanjiang Dong

To my parents Zuwei Tan & Lunzhen Li

To my sister Rong Tan
Abstract

This thesis proposes a new class of bivariate autoregressive conditional median duration models for matched high-frequency data and develops some inferential methods for an existing univariate model as well as the bivariate models introduced here to facilitate model fitting and forecasting.

During the last two decades, the autoregressive conditional mean duration (ACD) model has been playing a dominant role in analyzing irregularly spaced high-frequency financial data. Univariate ACD models have been extensively discussed in the literature. However, some major challenges remain. The existing ACD models do not provide a good distributional fit to financial durations, which are right-skewed and often exhibit unimodal hazard rates. Birnbaum-Saunders (BS) distribution is capable of modeling a wide variety of positively skewed data. Median is not only a robust measure of central tendency, but also a natural scale parameter of the BS distribution. A class of conditional median duration models, the BS-ACD and the scale-mixture BS ACD models based on the BS, BS power-exponential and Student-$t$ BS (BSt) distributions, have been suggested in the literature to improve the quality of the model fit.

The BSt-ACD model is more flexible than the BS-ACD model in terms of kurtosis and skewness. In Chapter 2, we develop the maximum likelihood estimation method
for the BS-ACD model. The estimation is performed by utilizing a hybrid of optimization algorithms. The performance of the estimates is then examined through an extensive Monte Carlo simulation study. We also carry out model discrimination using both likelihood-based method and information-based criterion. Applications to real trade durations and comparison with existing alternatives are then made.

The bivariate version of the ACD model has not received attention due to non-synchronicity. Although some bivariate generalizations of the ACD model have been introduced, they do not possess enough flexibility in modeling durations since they are conditional mean-based and do not account for non-monotonic hazard rates.

Recently, the bivariate BS (BVBS) distribution has been developed with many desirable properties and characteristics. It allows for unimodal shapes of marginal hazard functions. In Chapter 3, upon using this bivariate BS distribution, we propose the BVBS-ACD model as a natural bivariate extension of the BS-ACD model. It enables us to jointly analyze matched duration series, and also capture the dependence between the two series. The maximum likelihood estimation of the model parameters and associated inferential methods have been developed. A Monte Carlo simulation study is then carried out to examine the performance of the proposed inferential methods. The goodness-of-fit and predictive performance of the model are also discussed. A real bivariate duration data analysis is provided to illustrate the developed methodology.

The bivariate Student-$t$ BS (BVBS$t$) distribution has been introduced in the literature as a robust extension of the BVBS distribution. It provides greater flexibility in terms of the kurtosis and skewness through the inclusion of an additional shape
parameter. In Chapter 4, we propose the BVBS\textit{t}-ACD model as a natural extension of the BST-ACD model to the bivariate case. We then discuss the maximum likelihood estimation of the model parameters. A simulation study is carried out to investigate the performance of these estimators. Model discrimination is then done by using information-based criterion. Methods for evaluating the goodness-of-fit and predictive ability of the model are also discussed. A simulated data example is used to illustrate the proposed model as compared to the BVBS-ACD model.

Finally, in Chapter 5, some concluding comments are made and also some problems for future research are mentioned.

**Key words**: High-frequency financial data; Autoregressive conditional duration model; Conditional quantile duration; Student-\textit{t} Birnbaum-Saunders distribution; Bivariate Birnbaum-Saunders distribution; Bivariate Student-\textit{t} Birnbaum-Saunders distribution; Maximum likelihood estimation; Nelder-Mead algorithm; BFGS method; Monte Carlo simulation; Density forecast; Goodness-of-fit; Model discrimination; Information-based criterion.
Acknowledgements

I would like to express my sincere appreciation to my supervisor, Prof. N. Balakrishnan, for accepting me as his student, for giving me many opportunities for success, for his guidance, support, encouragement and great patience throughout my Ph.D. journey, and also for careful reading of my thesis. Without him, this thesis would not have been completed and I would not be in the position I am today.

I am very grateful to my supervisory committee members, Prof. Shui Feng and Prof. Aaron Childs for their valuable advice which has definitely improved the content of this thesis, along with my external examiner, Prof. Nalini Ravishanker and defense chair Prof. Ben Bolker. I thank Prof. John Maheu for valuable discussions.

I would also like to thank all my friends and fellow students who inspired and helped me in various ways. I am thankful to all the faculty members and staff for all their help during my graduate studies in the Department of Mathematics and Statistics.

Finally, special thanks go to my husband, my parents and my sister for their endless love and support.
# Contents

Abstract iv

Acknowledgements vii

List of Figures xii

List of Tables xiii

1 Introduction 1

1.1 High-frequency financial data . . . . . . . . . . . . . . . . . . . . 1

1.1.1 Features of high-frequency financial data . . . . . . . . . . . . 1

1.1.2 Diurnal adjustment of durations . . . . . . . . . . . . . . . . . 2

1.1.3 Total Time on Test plot . . . . . . . . . . . . . . . . . . . . . . 3

1.2 Some ACD models . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.2.1 Standard ACD and Log-ACD . . . . . . . . . . . . . . . . . . 4

1.2.2 GG-ACD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.2.3 BS-ACD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

1.3 A brief literature review . . . . . . . . . . . . . . . . . . . . . . . . 7

1.4 Motivation and objectives . . . . . . . . . . . . . . . . . . . . . . . 9
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 Distributions for high-frequency trade duration</td>
<td>11</td>
</tr>
<tr>
<td>1.5.1 BS distribution</td>
<td>11</td>
</tr>
<tr>
<td>1.5.2 Generalized BS Distribution</td>
<td>12</td>
</tr>
<tr>
<td>1.5.3 BS\textsubscript{t} Distribution</td>
<td>14</td>
</tr>
<tr>
<td>1.5.4 Bivariate BS distribution</td>
<td>15</td>
</tr>
<tr>
<td>1.5.5 Bivariate Student-\textit{t} BS Distribution</td>
<td>17</td>
</tr>
<tr>
<td>1.6 Likelihood inference</td>
<td>18</td>
</tr>
<tr>
<td>1.6.1 Maximum likelihood estimation</td>
<td>18</td>
</tr>
<tr>
<td>1.6.2 The NM method</td>
<td>19</td>
</tr>
<tr>
<td>1.6.3 The BFGS method</td>
<td>21</td>
</tr>
<tr>
<td>1.6.4 Standard errors of MLEs</td>
<td>22</td>
</tr>
<tr>
<td>1.7 Simulation study</td>
<td>23</td>
</tr>
<tr>
<td>1.8 Model discrimination</td>
<td>23</td>
</tr>
<tr>
<td>1.8.1 Likelihood-based method</td>
<td>24</td>
</tr>
<tr>
<td>1.8.2 Information-based criterion</td>
<td>24</td>
</tr>
<tr>
<td>1.9 Goodness-of-fit test</td>
<td>25</td>
</tr>
<tr>
<td>1.10 Out-of-sample forecast evaluation</td>
<td>27</td>
</tr>
<tr>
<td>1.11 Scope of the thesis</td>
<td>27</td>
</tr>
<tr>
<td>2 BS\textit{t}-ACD model</td>
<td>29</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>29</td>
</tr>
<tr>
<td>2.2 BS\textit{t}-ACD Model and the log-likelihood function</td>
<td>30</td>
</tr>
<tr>
<td>2.3 The first derivatives of $\ln L(\theta)$</td>
<td>32</td>
</tr>
<tr>
<td>2.4 The second derivatives of $\ln L(\theta)$</td>
<td>34</td>
</tr>
<tr>
<td>2.5 Estimation and inference</td>
<td>39</td>
</tr>
</tbody>
</table>
2.6 Simulation study .............................................. 40
2.7 Model discrimination ........................................... 43
  2.7.1 Likelihood-based method .................................. 43
  2.7.2 Information-based criterion ................................. 45
2.8 Application to trade duration data ............................... 49
  2.8.1 Data description .......................................... 49
  2.8.2 Estimation results ........................................ 51
  2.8.3 Goodness-of-fit .......................................... 53
  2.8.4 Out-of-sample forecast evaluation .......................... 53

3 Bivariate BS-ACD model ........................................ 55
  3.1 Introduction ................................................ 55
  3.2 Bivariate BS-ACD model and the log-likelihood function .......... 56
  3.3 The first derivatives of \( \ln L(\theta) \) ................................ 58
  3.4 The second derivatives of \( \ln L(\theta) \) ............................... 59
  3.5 Estimation and inference ..................................... 66
  3.6 Simulation study ........................................... 67
  3.7 Application to trade duration data .............................. 68
    3.7.1 Data description ....................................... 68
    3.7.2 Estimation results ...................................... 75
    3.7.3 Model Comparison ....................................... 75
    3.7.4 Goodness-of-fit ....................................... 76
    3.7.5 Out-of-sample forecast evaluation ......................... 77

4 Bivariate BS\(_t\)-ACD model .......................... 82
4.1 Introduction .............................................................. 82
4.2 Bivariate BSt-ACD model and the log-likelihood function .... 83
4.3 The first derivatives of $\ln L(\theta)$ .................................. 85
4.4 The second derivatives of $\ln L(\theta)$ ............................... 87
4.5 Estimation and inference .............................................. 97
4.6 Simulation study ...................................................... 97
4.7 Model discrimination .................................................. 99
4.8 Illustrative example 1: Example revisited ......................... 109
4.9 Illustrative example 2: Simulated data ............................. 109
   4.9.1 Data description ................................................. 110
   4.9.2 Estimation results ............................................... 112
   4.9.3 Goodness-of-fit ............................................... 113
   4.9.4 Out-of-sample forecast evaluation ............................ 114

5 Concluding Remarks .................................................. 117
  5.1 Summary of Research .............................................. 117
  5.2 Future work .......................................................... 122

Bibliography .......................................................... 125
List of Figures

2.1 TTT plots for IBM (left) and JNJ (right) .............................................. 50
2.2 ACF plots for IBM (left) and JNJ (right) .............................................. 50
2.3 GCS plots for IBM (left) and JNJ (right) stocks .............................. 53
3.1 Histograms for matched duration series of JNJ (left) and PG (right) . 73
3.2 TTT plots for matched duration series of JNJ (left) and PG (right) . 74
3.3 ACF plots for matched duration series of JNJ (left) and PG (right) . 74
3.4 Histograms for $z_1$ (left) and $z_2$ (right) ........................................... 78
3.5 Histograms for $z_{1|2}$ (left) and $z_{2|1}$ (right) ................................. 78
3.6 ACF plots for $z_1$ (left) and $z_2$ (right) ........................................... 79
3.7 ACF plots for $z_{1|2}$ (left) and $z_{2|1}$ (right) ....................................... 79
3.8 QQ plots for GCS residuals of JNJ (left) and PG (right) stocks ....... 80
3.9 ACF plots for GCS residuals of JNJ (left) and PG (right) stocks ....... 80
4.1 Histograms for the simulated bivariate dataset of Series 1 (left) and
   Series 2 (right) ................................................................. 111
4.2 TTT plots for the simulated bivariate dataset of Series 1 (left) and
   Series 2 (right) ................................................................. 111
4.3 ACF plots for the simulated bivariate dataset of Series 1 (left) and
   Series 2 (right) ................................................................. 112
List of Tables

1.1 Hazard function shapes for the GG distribution . . . . . . . . . . . . 5
1.2 Moment of the Normal and Student-t distribution . . . . . . . . . . . 14
2.1 Simulation results for BS\textit{t}-ACD model when \( n = 3000 \) . . . 41
2.2 Simulation results for BS\textit{t}-ACD model when \( n = 5000 \) . . . 41
2.3 Simulation results for BS\textit{t}-ACD model when \( n = 10000 \) . . . 42
2.4 Simulation results for BS\textit{t}-ACD model when \( n = 20000 \) . . . 42
2.5 Model Discrimination by LR test . . . . . . . . . . . . . . . . . . . . 47
2.6 Model Discrimination by AIC and BIC . . . . . . . . . . . . . . . . 48
2.7 Summary statistics for the adjusted durations . . . . . . . . . . . . . 49
2.8 Estimation results for three ACD models . . . . . . . . . . . . . . . 52
2.9 \( P \)-values for out-of-sample tests . . . . . . . . . . . . . . . . . . 54
3.1 Simulation results for BVBS-ACD model when \( \rho = 0.95 \) . . . . 69
3.2 Simulation results for BVBS-ACD model when \( \rho = 0.5 \) . . . . 70
3.3 Simulation results for BVBS-ACD model when \( \rho = 0.25 \) . . . . 71
3.4 Simulation results for BVBS-ACD model when \( \rho = 0 \) . . . . . . 72
3.5 Summary statistics for matched data . . . . . . . . . . . . . . . . . 73
3.6 Estimation results for the BVBS-ACD model . . . . . . . . . . . . . 75
3.7 LR test and AICs. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
3.8 P-values for in-sample goodness of fit tests for the BVBS-ACD model based on density forecasts ........................................ 77
3.9 P-values for out-of-sample tests for the BVBS-ACD model based on density forecasts .................................................. 81
4.1 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.0 \) 100
4.2 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.2 \) 100
4.3 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.5 \) 101
4.4 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.9 \) 101
4.5 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.0 \) 102
4.6 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.2 \) 102
4.7 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.5 \) 103
4.8 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.9 \) 103
4.9 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 5000 \) and \( \rho = 0.0 \) 104
4.10 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 5000 \) and \( \rho = 0.2 \) 104
4.11 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 5000 \) and \( \rho = 0.5 \) 105
4.12 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 5000 \) and \( \rho = 0.9 \) 105
4.13 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 10000 \) and \( \rho = 0.0 \) 106
4.14 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 10000 \) and \( \rho = 0.2 \) 106
4.15 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 10000 \) and \( \rho = 0.5 \) 107
4.16 Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 10000 \) and \( \rho = 0.9 \) 107
4.17 Model Discrimination by AIC and BIC ........................................ 108
4.18 The maximized log-likelihood value and AIC versus degrees of freedom 
\( \nu \) ................................................................. 109
4.19 Summary statistics for the simulated bivariate dataset ................. 110
4.20 Estimation results for two bivariate ACD models . . . . . . . . . . . . 113
4.21 $P$-values for in-sample goodness-of-fit tests for the BVBS-$t$-ACD model
   based on density forecasts . . . . . . . . . . . . . . . . . . . . . . . . . . . 115
4.22 $P$-values for in-sample goodness-of-fit tests for the BVBS-ACD model
   based on density forecasts . . . . . . . . . . . . . . . . . . . . . . . . . . . 115
4.23 $P$-values for out-of-sample tests for the of BVBS-$t$-ACD model based
   on density forecasts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
4.24 $P$-values for out-of-sample tests for the of BVBS-ACD model based on
   density forecasts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
Chapter 1

Introduction

1.1 High-frequency financial data

1.1.1 Features of high-frequency financial data

Due to the rapid development of computing power, storage capacity and the adoption of electronic trading systems, a massive amount of financial data has become widely accessible on a transaction-by-transaction basis, so-called high-frequency financial data. High-frequency data are observations taken at a fine time scale (Tsay (2010)). Here, time is often measured in seconds. Typical examples include transaction times, transaction prices, bid and ask quotes and transaction volumes. High-frequency financial data play an important role in market microstructure theory, indicating the information with respect to fundamental asset prices and the behavior of market participants.

The distinctive feature of transaction data is the irregular spacing over time since the arrival times of market events (such as trades and quotes) are random. Durations
are the waiting times between market events such as trades, price and volume changes, which are quite informative and become crucial in understanding market activity and the underlying microstructure of the financial market.

Trade duration is the time elapsed between two consecutive trades or transactions of equity. This type of data have some special features. In addition to irregular spacing mentioned above, trade durations display a cluster effect. Short durations tend to be followed by short durations, and likewise for long durations, which leads to positive autocorrelations. Trade durations often exhibit diurnal patterns due to the intraday periodicity of daily trading activities. Trading activities are heavier in the beginning and closing of the trading day, and usually thinner in the middle of the day. Thus, we usually observe short durations during opening and closing hours and long durations during lunch hour. Another important feature of trade durations is that they often show unimodal hazard rates; see Section 1.1.3 for details on assessing the shape of the hazard.

The high-frequency financial data used in this thesis is from the New York Stock Exchange (NYSE) Trade and Quote (TAQ) database (note: the data sets were generously provided by Dr. Chad R. Bhatti, V.P., Citizens Bank). Transactions recorded before 9:30 am and after 4:00 pm were omitted, and simultaneous transactions were treated as a single transaction.

1.1.2 Diurnal adjustment of durations

Trade durations exhibit strong diurnal patterns or time-of-day effects, which indicates high trading activities during opening and closing hours and less activity around noon. In the analysis of duration data and ACD framework, it is common to remove
time-of-day effects first and then fit a ACD model to diurnally adjusted data.

The adjusted durations can be obtained by

\[ X_i = \tilde{X}_i / \hat{\phi} = \tilde{X}_i / \exp(\hat{s}) , \]

where \( \tilde{X}_i \) is the raw duration and \( \phi \) is the time-of-day effect. To estimate \( s \), we use quadratic splines and indicator variables on the thirty-minute intervals from 9:30 am to 4:00 pm; see Engle and Russell (1998), Engle (2000), Tsay (2010), Bhatti (2010), Hautsch (2012) and Leiva et al. (2014).

1.1.3 Total Time on Test plot

The Total Time on Test (TTT) plot can be used to identify the shape of the hazard function (see Aarset (1987) and Azevedo et al. (2012)). Its theoretical counterpart is the scaled TTT transform. If a life distribution \( F(y) \) is absolutely continuous and its survival function is \( S(y) = 1 - F(y) \), the scaled TTT transform is \( W(u) = H^{-1}(u)/H^{-1}(1) \) for \( 0 < u < 1 \), where \( H^{-1}(u) = \int_0^u S(y)dy \). If \( Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n} \) denote the order statistics with common pdf \( f(y) \) and cdf \( F(y) \), the empirical scaled TTT transform is given by \( W_n(r/n) = H^{-1}_n(r/n)/H^{-1}_n(1) = [\sum_{i=1}^r y_{i:n} + (n-r)y_{r:n}] / \sum_{i=1}^n y_{i:n} \), for \( r = 1, \ldots, n \). The TTT plot is obtained by plotting \((r/n, W_n(r/n))\), \( r = 1, \ldots, n \). For example, a linear TTT curve, which is randomly around the main diagonal of the unit square (the \( 45^\circ \)-line), implies a constant hazard function, a convex TTT curve indicates a decreasing hazard, and a concave TTT curve demonstrates an increasing hazard. If a TTT curve is concave (convex) and then convex (concave), the corresponding hazard is unimodal (bathtub).
1.2 Some ACD models

1.2.1 Standard ACD and Log-ACD

Engle and Russell (1998) introduced the following ACD\((p,q)\) model, which is one of the primary tools used in modeling durations:

\[
X_i = \psi_i \epsilon_i, \quad \psi_i = \omega + \sum_{j=1}^{p} \alpha_j X_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j},
\]  

(1.1)

where \(X_i\) denotes the duration, \(\psi_i = E(X_i|\mathcal{F}_{i-1})\) is the conditional expectation given \(\mathcal{F}_{i-1}\), the information set available at the \(i\)th duration, \(\epsilon_i\) is a series of i.i.d. positive random variables such that \(E(\epsilon_i) = 1\), \(\omega > 0\), \(\alpha_j \geq 0\) and \(\beta_j \geq 0\) are the unknown parameters, and \(\sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1\). To relax the positivity constraints on the model parameters, Bauwens and Giot (2000) suggested the Logarithmic ACD model of the form

\[
X_i = \psi_i \epsilon_i, \quad \ln \psi_i = \omega + \alpha \epsilon_{i-1} + \beta \ln \psi_{i-1}, \quad |\beta| < 1.
\]  

(1.2)

In these models, durations are assumed to follow either exponential or Weibull distributions. Therefore, the hazard functions are either constant or monotonic.

1.2.2 GG-ACD

It is necessary to briefly introduce the generalized gamma (GG) distribution before presenting details of the GG-ACD model. The PDF of the GG distribution with shape
parameters $v$ and $\eta$ and scale parameter $\omega$ is given by

$$f_X(x) = \frac{\eta}{\omega \Gamma(v)} \left(\frac{x}{\omega}\right)^{\eta - 1} \exp \left[-\left(\frac{x}{\omega}\right)^\eta\right], \quad x > 0, v, \omega, \eta > 0, \quad (1.3)$$

where $\Gamma(v) = \int_0^\infty u^{v-1} \exp(-u) du$ is the complete gamma function. The exponential ($v = \eta = 1$), gamma ($\eta = 1$), lognormal ($v \to \infty$) and Weibull ($v = 1$) distributions are all special cases of the GG distribution, denoted by $GG(v, \eta, \omega)$.

The CDF of the GG distribution can be expressed as $F_X(x) = \frac{\Gamma((x/\omega)^\eta, v)}{\Gamma(v)}$, where $\Gamma(y, v) = \int_y^\infty u^{v-1} \exp(-u) du$ is the lower incomplete gamma function.

If $X \sim GG(v, \eta, \omega)$, then $E(X^k) = \frac{\omega^k \Gamma(v+k/\eta)}{\Gamma(v)}$. In particular, $E(X) = \frac{\omega \Gamma(v+1/\eta)}{\Gamma(v)}$ and $V(X) = \omega^2 \left(\frac{\Gamma(v+2/\eta)}{\Gamma(v)} - \left(\frac{\Gamma(v+1/\eta)}{\Gamma(v)}\right)^2\right)$.

The GG distribution can produce all common types of hazard function (see Bhatti (2010)):

<table>
<thead>
<tr>
<th>$\eta$</th>
<th>$v\eta - 1 &lt; 0$</th>
<th>$v\eta - 1 &gt; 0$</th>
<th>$v\eta - 1 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta = 1$</td>
<td>Decreasing</td>
<td>Increasing</td>
<td>Constant</td>
</tr>
<tr>
<td>$\eta &gt; 1$</td>
<td>Bathub</td>
<td>Increasing</td>
<td>Increasing</td>
</tr>
<tr>
<td>$\eta &lt; 1$</td>
<td>Decreasing</td>
<td>Unimodal</td>
<td>Decreasing</td>
</tr>
</tbody>
</table>

To allow for non-monotonic hazards, Lunde (1999) proposed the GG-ACD model. It is one of the most flexible conditional mean duration models, which can be written as

$$X_i = \psi_i \epsilon_i, \quad \ln \psi_i = \alpha + \beta \ln \psi_{i-1} + \gamma \left[\frac{X_{i-1}}{\psi_{i-1}}\right], \quad |\beta| < 1, \quad (1.4)$$

where $X_i \sim GG(v, \eta, \psi_i)$, $\epsilon_i \overset{iid}{\sim} GG(v, \eta, 1)$ and $\psi_i = E(X_i \mid \mathcal{F}_{i-1}) = \frac{\omega \Gamma(v+1/\eta)}{\Gamma(v)}$. The pdf of GG-ACD model is given by
where \( \varphi(v, \eta) = \Gamma(v)/\Gamma(v + 1/\eta) \). The associated survival function is given by

\[
S_{GG}(x_i; v, \eta, \psi_i) = 1 - \Gamma(x_i^\eta|\psi_i\varphi(v, \eta)]^{-\eta}, v)/\Gamma(v), \ x_i, v, \eta, \psi_i > 0. \tag{1.6}
\]

### 1.2.3 BS-ACD

Bhatti (2010) introduced the autoregressive conditional median duration model based on the BS distribution. The primary motivation for the BS-ACD model is to provide an alternative to existing ACD models and enhance the quality of the model fit. The BS distribution fits duration data well in terms of the shapes of its density and hazard functions. The BS-ACD model is specified in terms of a time-varying conditional median duration \( \sigma_i = F_{X_i}^{-1}(0.5|\mathcal{F}_{i-1}) \), where \( F_{X_i}^{-1} \) is the inverse CDF or quantile function of the model, instead of the conditional mean duration in the usual ACD model, viz.,

\[
\ln \sigma_i = \alpha + \beta \ln \sigma_{i-1} + \gamma \left[ \frac{X_{i-1}}{\sigma_{i-1}} \right], \quad |\beta| < 1. \tag{1.7}
\]

Median is not only a robust measure of central tendency, but also a natural parameter of the BS distribution. The associated PDF with the BS-ACD model is given by

\[
f_{X_i}(x_i; \kappa, \sigma_i) = \frac{1}{2\sqrt{2\pi}\kappa \sigma_i} \left[ \left( \frac{\sigma_i^2}{x_i} \right)^{\frac{1}{2}} + \left( \frac{\sigma_i}{x_i} \right)^{\frac{3}{2}} \right] \exp \left[ -\frac{1}{2\kappa^2} \left( \frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2 \right) \right], \quad x_i, \kappa, \sigma_i > 0. \tag{1.8}
\]
The corresponding survival function is given by

$$S_{BS}(x_i, \kappa, \sigma_i) = 1 - \Phi \left[ \frac{1}{\kappa} \left( \sqrt{\frac{x_i}{\sigma_i}} - \sqrt{\frac{\sigma_i}{x_i}} \right) \right], \quad i = 1, 2, \ldots, n.$$  \hspace{1cm} (1.9)

\subsection*{1.3 A brief literature review}

The analysis of high-frequency financial data has received increasing attention recently in the literature. Recent research interest in duration models was largely spurred by the original work of Engle and Russell (1998), which introduced the autoregressive conditional duration (ACD) model to analyze the irregularly spaced trade durations of single assets. See Hautsch (2012), Pacurar (2008) and Saranjeet and Ramanathan (2018) for surveys on this topic.

Univariate ACD models have been well studied in the literature. An important extension is the logarithmic version of the ACD model, introduced by Bauwens and Giot (2000), which relaxes the positivity constraints on the model parameters. There exists different distributional assumptions on the conditional durations in the literature such as the exponential and Weibull distributions in Engle and Russell (1998), and the generalized gamma (GG) distribution in Lunde (1999). The first two distributions have either a flat or monotonic hazard rates. However, the last one is quite flexible and can produce all common types of hazard functions. Thus, the GG-ACD model is regarded as the status quo in the literature of ACD models. The parametric formulation of the ACD models determines the dynamics of the conditional mean duration.

The Birnbaum-Saunders (BS) distribution, introduced by Birnbaum and Saunders (1969), has been used quite successfully to model a wide variety of positively skewed
data; see also Johnson et al. (1995) for a concise review of the classical BS distribution and Leiva (2016) for an overview on recent theoretical developments and applications. All developments on BS and related distributions have recently been thoroughly reviewed in the discussion paper by Balakrishnan and Kundu (2018). Due to its good properties, Jin and Kawczak (2003) claimed that BS distribution deserves more attention in the study of high-frequency financial data, and these authors then utilized the BS kernel to estimate the probability density function of trade durations. As illustrated in Kundu et al. (2008), the hazard rate of BS distribution has a unimodal shape. Considering that the hazard function for duration data is usually unimodal, Bhatti (2010) suggested the BS-ACD model possessing a realistic distribution and an ACD specification in terms of conditional median instead of conditional mean in the original ACD model. Median is not only a robust measure of central tendency but also a natural parameter in the BS distribution. Thus, the model fit can be enhanced by modeling the conditional median duration instead of the conditional mean duration as done in the original ACD model. A recent extension of this conditional median-type ACD model is due to Leiva et al. (2014), which considered the BS, BS power-exponential and Student-t BS distributions, so-called scale-mixture BS distribution family (see Balakrishnan et al. (2009) and Díaz-García and Leiva (2005) for details on SBS distributions) and developed corresponding expectation-maximization (EM) algorithms for SBS-ACD models. Moreover, Saulo and Leão (2017) discussed the conditional median-type ACD model based on the log-symmetric distributions. Saulo et al. (2017a) compared the mean-based and median-based BS-ACD models regarding model-fitting, forecasting and influence analysis. In general, their study confirmed that the conditional median-type ACD models based on BS and associated
distributions are superior to the existing ACD model (the Generalized Gamma ACD model, see Lunde (1999)) in terms of model-fitting and forecasting.

1.4 Motivation and objectives

Despite the vast literature on duration series of single assets, a bivariate version of the ACD model has not received enough attention because of nonsynchronous trading/asynchronous observations. Only a few papers, with regard to the ACD model, have focused on jointly analyzing the duration series of trades and quotes of a single asset. Engle and Lunde (2003) proposed a censored bivariate ACD model using a semiparametric estimation approach. Mosconi and Olivetti (2005) introduced bivariate Exponential and Weibull ACD models (see Balakrishnan and Lai (2009) for detailed discussions on these bivariate distributions) in the same line.

However, both models share some common drawbacks. First, they only take into account the duration of one stock, and do not include information given by another stock. New available information to the stock market may affect not only a specific stock, but also the related stocks in the same industry sector or even the whole stock market (Simonsen (2007)). As a result, new information may lead to dependence between trade durations in different stocks. Therefore, modeling the dependence between two or multiple duration series remains a challenge. Second, durations often show unimodal hazard rates. The above bivariate ACD models fail to capture this main feature of duration data. A simulation study in Grammig and Maurer (2000) showed that the misspecification of the hazard rate can severely deteriorate the predictive ability of the ACD model. Third, the median is a better measure of central tendency than the mean in a skewed distribution. Hence, the model fit can
be improved by replacing the conditional mean in the original ACD model with the conditional median. Furthermore, there is no work till now on the topic of joint modeling of the the duration series for matched data.

Recently, Kundu et al. (2010) derived the bivariate Birnbaum-Saunders (BVBS) distribution through a transformation of the bivariate normal distribution (see Kundu et al. (2013) and Vilca et al. (2014a,b) for generalizations of the BVBS distribution). Its marginals are univariate BS distributions with unimodal hazard functions and its conditional distributions can be expressed in terms of normal distribution. Moreover, it has a correlation parameter which indirectly represents the dependence between the two BS random variables. The bivariate Student-t BS (BVBS\textit{t}) distribution was suggested by Vilca et al. (2014a) as a robust extension of the BVBS distribution, allowing more flexibility in kurtosis and skewness achieved with the inclusion of an additional shape parameter.

Our major objective is to jointly model the trade durations of two assets for matched data, pairs of durations with the same starting time (see Simonsen (2007)). In this thesis, we construct a bivariate ACD model based on BVBS and Student-t BVBS distributions, which facilitate the joint analysis and also to detect the dependence between two matched duration series.
1.5 Distributions for high-frequency trade duration

1.5.1 BS distribution

The CDF of a two-parameter BS random variable $X$ is given by

$$F_X(x; \kappa, \sigma) = \Phi \left[ \frac{1}{\kappa} \left( \sqrt{\frac{x}{\sigma}} - \sqrt{\frac{\sigma}{x}} \right) \right], \quad x > 0,$$

(1.10)

where $\Phi(\cdot)$ is the standard normal CDF, $\kappa > 0$ is the shape parameter and $\sigma > 0$ is the scale parameter which is also the median. The corresponding PDF is

$$f_X(x; \kappa, \sigma) = \frac{1}{2\sqrt{2\pi\kappa\sigma}} \left[ \left( \frac{\sigma}{\sqrt{x}} \right)^{\frac{3}{2}} + \left( \frac{x}{\sqrt{\sigma}} \right)^{\frac{3}{2}} \right] \exp \left[ -\frac{1}{2\kappa^2} \left( \frac{x}{\sigma} + \frac{\sigma}{x} - 2 \right) \right], \quad x > 0.$$  

(1.11)

This distribution is simply denoted by $BS(\kappa, \sigma)$.

The BS distribution possesses the scaling property, namely, for $c > 0$, $cX \sim BS(\kappa, c\sigma)$, and in particular, $\frac{X}{\sigma} \sim BS(\kappa, 1)$, and moreover it has the reciprocal property, namely, $\frac{1}{X} \sim BS(\kappa, \frac{1}{\sigma})$. There is a monotone transformation between the BS and normal distributions given by

$$Z = \frac{1}{\kappa} \left( \sqrt{\frac{X}{\sigma}} - \sqrt{\frac{\sigma}{X}} \right),$$  

(1.12)

or equivalently

$$X = \sigma \left[ \frac{1}{2}\kappa Z + \sqrt{\left( \frac{1}{2}\kappa Z \right)^2 + 1} \right]^2,$$  

(1.13)

where $Z \sim N(0, 1)$. Using the transformation in (1.13), the mean, variance and the
coefficients of skewness (CS) and kurtosis (CK) can be shown to be

\[
E(X) = \sigma\left(1 + \frac{1}{2}\kappa^2\right), \quad V(X) = (\kappa\sigma)^2\left(1 + \frac{5}{4}\kappa^2\right),
\]

\[
CS = \frac{4\kappa(11\kappa^2 + 6)}{(5\kappa^2 + 4)^2}, \quad CK = 3 + \frac{6\kappa^2(93\kappa^2 + 40)}{(5\kappa^2 + 4)^2}.
\]

Kundu et al. (2008) have shown that the hazard function of the BS distribution

\[
h_X(x; \kappa, \sigma) = \frac{f_X(x; \kappa, \sigma)}{1 - F_X(x; \kappa, \sigma)}
\]

has an unimodal shape.

### 1.5.2 Generalized BS Distribution

A random variable \(X\) is said to have a generalized Birnbaum-Saunders distribution if it allows the stochastic representation

\[
X = \sigma \left[\frac{1}{2}\kappa Z + \sqrt{\left(\frac{1}{2}\kappa Z\right)^2 + 1}\right]^2 \sim GBS(\kappa, \sigma; g)
\]

and

\[
Z = \frac{1}{\kappa} \left(\sqrt{\frac{X}{\sigma}} - \sqrt{\frac{\sigma}{X}}\right) \sim S(g),
\]

where \(\kappa > 0\) is the shape parameter, \(\sigma > 0\) is the scale parameter which is also the median and \(Z\) follows a standard symmetrical distribution, denoted by \(Z \sim S(g)\), with \(g\) being the kernel of the p.d.f. of \(Z\) (see Díaz-García and Leiva (2005) and Sanhueza et al. (2008)).

The GBS distribution has many good properties. For example, if \(X \sim GBS(\kappa, \sigma; g)\),
then

1. \(cX \sim GBS(\kappa, c\sigma; g)\) with \(c > 0\), and \(X^{-1} \sim GBS(\kappa, \sigma^{-1}; g)\);

2. \(U = Z^2 = \frac{1}{\kappa^2} \left( \frac{X}{\sigma} + \frac{\sigma}{X} - 2 \right) \sim G_{\chi^2}(g)\), a generalized chi-square distribution with one degree of freedom. The pdf of \(U\) is given by \(f_U(u) = cg(u)u^{-1/2}\), with \(u > 0\), where \(cg(\cdot)\) is the pdf of \(Z \sim S(g)\);

3. The quantile function of \(X\) is given by \(x(q) = F_X^{-1}(p) = \frac{\sigma}{4} (\kappa z_q + [\kappa^2 z_q^2 + 4]^{1/2})^2\), where \(z_q\) is the \(q\)th quantile of \(Z \sim S(g)\). If \(q = 0.5\), then \(x(0.5) = \sigma\) and so \(\sigma\) is the median;

4. If \(E(Z^k)\) exists, then

\[
E(X^k) = \sigma^k \sum_{j=0}^{k} \binom{2k}{2j} \sum_{i=0}^{j} \binom{j}{i} E(U^{k+i-j}) \left( \frac{\kappa}{2} \right)^{2(k+i-j)}
\]

(1.19)

(see Sanhueza et al. (2008) and Díaz-García and Leiva (2007)).

Thus, the mean, variance and the coefficients of skewness (CS) and kurtosis (CK) of \(X\) are given by

\[
E(X) = \frac{\sigma}{2} (2 + u_1 \kappa^2),
\]

(1.20)

\[
Var(X) = \frac{\sigma^2 \kappa^2}{4} [4u_1 + (2u_2 - u_1^2) \kappa^2],
\]

(1.21)

\[
CS = \frac{4\kappa [(3u_2 - 3u_1^2) + \frac{1}{2} (2u_3 - 3u_1 u_2 + u_1^3) \kappa^2]}{[4u_1 + (2u_2 - u_1^2) \kappa^2]^{3/2}},
\]

(1.22)

\[
CK = \frac{16u_2 + (32u_3 - 48u_1 u_2 + 24u_1^2) \kappa^2 + (8u_4 - 16u_1 u_3 + 12u_1^2 u_2 - 3u_1^4) \kappa^4}{[4u_1 + (2u_2 - u_1^2) \kappa^2]^2},
\]

(1.23)

where \(u_k = E(U^k)\) and \(U \sim G_{\chi^2}(g)\) (see Sanhueza et al. (2008)).
Table 1.2: Moment of the Normal and Student-t distribution

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$u_1(g)$</th>
<th>$u_2(g)$</th>
<th>$u_3(g)$</th>
<th>$u_4(g)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>1</td>
<td>3</td>
<td>15</td>
<td>105</td>
</tr>
<tr>
<td>Student-t</td>
<td>$\frac{\nu}{(\nu-2)}$</td>
<td>$\frac{3\nu^2}{(\nu-2)(\nu-4)}$</td>
<td>$\frac{15\nu^4}{(\nu-2)(\nu-4)(\nu-6)}$</td>
<td>$\frac{105\nu^4}{(\nu-2)(\nu-4)(\nu-6)(\nu-8)}$</td>
</tr>
<tr>
<td>$\nu &gt; 2$</td>
<td>$\nu &gt; 4$</td>
<td>$\nu &gt; 6$</td>
<td>$\nu &gt; 8$</td>
<td></td>
</tr>
</tbody>
</table>

(see Saulo et al. (2017b)).

1.5.3 BST Distribution

In (1.17), if $g(\cdot)$ is normal, then $X \sim BS(\kappa, \sigma)$. On the other hand, if $g(\cdot)$ is $t$ kernel, then $X \sim BSt(\kappa, \sigma, \nu)$. When $\nu$ goes to infinity, the BSt distribution tends to be the BS distribution.

If $X \sim BSt(\kappa, \sigma, \nu)$, the corresponding PDF of $X$ is given by

$$f_X(x; \kappa, \sigma, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{2\nu\pi\Gamma\left(\frac{\nu}{2}\right)\kappa\sigma} \left[\left(\frac{\sigma}{x}\right)^{\frac{1}{2}} + \left(\frac{\sigma}{x}\right)^{\frac{3}{2}}\right] \left[1 + \frac{(\frac{x}{\sigma} + \frac{\sigma}{x} - 2)}{\nu\kappa^2}\right]^{-\frac{\nu+1}{2}}, \quad (1.24)$$

where $\kappa > 0$ and $\nu > 0$ are shape parameters, $\nu$ is also known as the degree of freedom, and $\sigma > 0$ is the scale parameter.

The corresponding CDF can be written as

$$F(x; \kappa, \sigma, \nu) = \frac{1}{2} \left\{1 + \frac{I[\sigma(x)^2/][\sigma(x)^2+\nu]}{I[\sigma(x)^2/][\sigma(x)^2+\nu]} \left(1, \frac{\nu}{2}ight)\right\}, \quad (1.25)$$

where $a(x) = (\sqrt{x/\sigma} - \sqrt{\sigma/x})/\kappa$ and $I_x(a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt/ \int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function ratio (see Azevedo et al. (2012)).

Similarly, there is a monotone transformation between the BSt and Student-t distributions. Using the transformation in (1.17), the mean, variance and the coefficients
of skewness and kurtosis can be shown to be

\[ E(X) = \frac{\sigma^2}{2} \left( 2 + \frac{\nu}{\nu - 2} \kappa^2 \right), \quad \nu > 2, \quad (1.26) \]

\[ \text{Var}(X) = \frac{\sigma^2 \kappa^2}{4(\nu - 2)} \left[ 4 + \frac{\nu(5\nu - 8)\kappa^2}{(\nu - 2)(\nu - 4)} \right], \quad \nu > 4, \quad (1.27) \]

\[ CS = \frac{4\kappa(\nu - 1)\nu^2}{(\nu - 4)(\nu - 2)^2} \left[ \frac{6 + \frac{\nu(11\nu - 18)\kappa^2}{(\nu - 6)(\nu - 2)}}{4\nu} \right], \quad \nu > 6, \quad (1.28) \]

\[ CK = \frac{3\nu^2}{(\nu - 2)(\nu - 4)} \left[ 16 + \frac{8\nu(15\nu^2 - 42\nu + 32)\kappa^2}{(\nu - 6)(\nu - 2)^2} + \frac{\nu^2(211\nu^3 - 894\nu^2 + 1288\nu - 640)\kappa^4}{(\nu - 8)(\nu - 6)(\nu - 2)^3} \right], \quad \nu > 8. \quad (1.29) \]

The BS\textit{t} distribution also possesses unimodal hazard; see Azevedo \textit{et al.} (2012) for a thorough study of hazard functions of this family of distributions.

### 1.5.4 Bivariate BS distribution

Kundu \textit{et al.} (2010) introduced the bivariate Birnbaum-Saunders (BVBS) distribution. A bivariate random vector \((X_1, X_2)\) is said to follow a BVBS distribution with parameters \(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho\), if its joint CDF can be written as

\[ F_{X_1, X_2}(x_1, x_2) = \Phi_2 \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right); \rho \right], \quad x_1, x_2 > 0, \quad (1.30) \]

where \(\kappa_1, \sigma_1, \kappa_2, \sigma_2 > 0, -1 < \rho < 1\), and \(\Phi_2(z_1, z_2; \rho)\) is the joint CDF of a standard bivariate normal vector \((Z_1, Z_2)\) with correlation coefficient \(\rho\). We denote this distribution simply by \(BVBS(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho)\). The corresponding joint PDF of
\((X_1, X_2)\) is given by

\[
f_{X_1,X_2}(x_1, x_2) = \phi_2 \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right) ; \rho \right] \\
\times \frac{1}{2\kappa_1\sigma_1} \left[ \left( \frac{\sigma_1}{x_1} \right)^{\frac{1}{2}} + \left( \frac{\sigma_1}{x_1} \right)^{\frac{3}{2}} \right] \frac{1}{2\kappa_2\sigma_2} \left[ \left( \frac{\sigma_2}{x_2} \right)^{\frac{1}{2}} + \left( \frac{\sigma_2}{x_2} \right)^{\frac{3}{2}} \right],
\]

where \(\phi_2(z_1, z_2; \rho)\) is the joint PDF of \((Z_1, Z_2)\) defined by

\[
\phi_2(z_1, z_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} (z_1^2 + z_2^2 - 2\rho z_1 z_2) \right\}.
\]

If \((X_1, X_2) \sim BVBS(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho)\), then \(X_j \sim BS(\kappa_j, \sigma_j)\), \(j = 1, 2\). The conditional PDF of \(X_1\), given \(X_2 = x_2\), is given by

\[
f_{X_1|X_2}(x_1|x_2) = \frac{1}{2\kappa_1\sigma_1\sqrt{2\pi(1-\rho^2)}} \left[ \left( \frac{\sigma_1}{x_1} \right)^{\frac{1}{2}} + \left( \frac{\sigma_1}{x_1} \right)^{\frac{3}{2}} \right] \\
\times \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right) - \frac{\rho}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right) \right] ^2 \right\},
\]

The corresponding CDF of \(X_1\), given \(X_2 = x_2\), is

\[
F_{X_1|X_2}(x_1|x_2) = \Phi \left\{ \frac{\frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right) - \frac{\rho}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right)}{\sqrt{1-\rho^2}} \right\}.
\]

The relationship between the BVBS and bivariate normal distributions can be represented by \(Z_j = \frac{1}{\kappa_j} \left( \sqrt{\frac{X_j}{\sigma_j}} - \sqrt{\frac{\sigma_j}{X_j}} \right)\), or equivalently \(X_j = \sigma_j \left[ \frac{1}{2}\kappa_j Z_j + \sqrt{(\frac{1}{2}\kappa_j Z_j)^2 + 1} \right]^2\), for \(j = 1, 2\), where \((Z_1, Z_2) \sim N(0, \Sigma)\), with \(\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\). Thus, random numbers from the BVBS distribution can be easily generated using bivariate normal distribution (see Kundu et al. (2010)). Also, \(\rho_{X_1,X_2} \approx \rho\) (see Balakrishnan and Zhu (2015)).
1.5.5 Bivariate Student-t BS Distribution

A bivariate random vector \((X_1, X_2)\) is said to follow a bivariate Student-t BS distribution with parameters \(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho, \nu\), if its joint CDF can be written as

\[
F_{X_1, X_2}(x_1, x_2) = T_2 \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right); \rho, \nu \right], \quad x_1, x_2 > 0,
\]

where \(\kappa_1, \sigma_1, \kappa_2, \sigma_2 > 0\), \(-1 < \rho < 1\), \(\nu > 2\) and \(T_2(z_1, z_2; \rho, \nu)\) is the joint CDF of a bivariate \(t\) vector \((Z_1, Z_2)\) with correlation coefficient \(\rho\). We denote this distribution simply by \(\text{BVBS}t(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho, \nu)\).

The corresponding joint PDF of \((X_1, X_2)\) is given by

\[
f_{X_1, X_2}(x_1, x_2) = t_2 \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_1}{\sigma_1}} - \sqrt{\frac{\sigma_1}{x_1}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_2}{\sigma_2}} - \sqrt{\frac{\sigma_2}{x_2}} \right); \rho, \nu \right]
\times \frac{1}{2\kappa_1 \sigma_1} \left[ \left( \frac{\sigma_1}{x_1} \right)^{\frac{1}{2}} + \left( \frac{\sigma_1}{x_1} \right)^{\frac{3}{2}} \right] \frac{1}{2\kappa_2 \sigma_2} \left[ \left( \frac{\sigma_2}{x_2} \right)^{\frac{1}{2}} + \left( \frac{\sigma_2}{x_2} \right)^{\frac{3}{2}} \right],
\]

where \(t_2(z_1, z_2; \rho, \nu)\) is the joint PDF of \((Z_1, Z_2)\) defined by

\[
t_2(z_1, z_2; \rho, \nu) = \frac{1}{2\pi \sqrt{1 - \rho^2}} \left[ 1 + \frac{z_1^2 + z_2^2 - 2\rho z_1 z_2}{\nu(1 - \rho^2)} \right]^{-\frac{\nu + 2}{2}}.
\]

If \((X_1, X_2) \sim \text{BVBS}t(\kappa_1, \sigma_1, \kappa_2, \sigma_2, \rho, \nu)\), then \(X_j \sim \text{BST}(\kappa_j, \sigma_j, \nu)\), \(j = 1, 2\). The conditional PDF of \(X_1\), given \(X_2 = x_2\), is given by

\[
f_{X_1|X_2}(x_1|x_2) = f_{q(x_2)} \left( [a_{x_1}(\kappa_1, \sigma_1) - \mu_1(x_2)]^2 \right) A_{x_1}(\kappa_1, \sigma_1),
\]

where \(f_{q(x_2)}(u) = \frac{\Gamma(\frac{\nu + 2}{2})}{\sqrt{\pi} \Gamma(\frac{\nu}{2})} [\nu + q(x_2)]^{-\frac{\nu + 2}{2}} [\nu + q(x_2) + u]^{-\frac{\nu + 2}{2}}\), \(u \geq 0\), \(q(x_2) = a_{x_2}^2(\kappa_2, \sigma_2)\), \(a_{x_j}(\kappa_j, \sigma_j) = \frac{1}{\kappa_j} \left( \sqrt{\frac{x_j}{\sigma_j}} - \sqrt{\frac{\sigma_j}{x_j}} \right)\), \(j = 1, 2\), \(\kappa_1 = \sqrt{1 - \rho^2} \kappa_1\), \(\mu_1(x_2) = \rho a_{x_2}(\kappa_2, \sigma_2)\), and

\[
A_{x_j}(\kappa_j, \sigma_j) = \frac{\partial a_{x_j}(\kappa_j, \sigma_j)}{\partial x_j} = \frac{1}{\kappa_j \sigma_j} \left[ \left( \frac{\sigma_j}{x_j} \right)^{\frac{1}{2}} + \left( \frac{x_j}{\sigma_j} \right)^{\frac{3}{2}} \right].
\]
The corresponding joint CDF of \((X_1, X_2)\) can be written as
\[
P(X_1 \leq x_1, X_2 \leq x_2) = G_2(a_t(\kappa, \sigma); \rho, \nu) = G_2(a_{x_1}(\kappa_1, \sigma_1), a_{x_2}(\kappa_2, \sigma_2); \rho, \nu),
\]
where
\[
G_2(a_x(\kappa, \sigma); \rho, \nu) = G_2\left(\sqrt{\frac{\nu - 1}{\nu + 1}} a_x(\kappa, \sigma); \rho, \nu - 1\right)
+ C_\nu \left\{ \frac{a_{x_1}(\kappa_1, \sigma_1)}{1 + \nu + a_{x_1}^2(\kappa_1, \sigma_1)} G\left(\frac{\sqrt{\nu}(a_{x_2}(\kappa_2, \sigma_2) - \rho a_{x_1}(\kappa_1, \sigma_1))}{\sqrt{1 - \rho^2}} \sqrt{\frac{1}{1 + \nu + a_{x_2}^2(\kappa_2, \sigma_2)}}; \nu\right) \right. \\
+ \left. \frac{a_{x_2}(\kappa_2, \sigma_2)}{1 + \nu + a_{x_1}^2(\kappa_2, \sigma_2)} G\left(\frac{\sqrt{\nu}(a_{x_1}(\kappa_1, \sigma_1) - \rho a_{x_2}(\kappa_2, \sigma_2))}{\sqrt{1 - \rho^2}} \sqrt{\frac{1}{1 + \nu + a_{x_1}^2(\kappa_2, \sigma_2)}}; \nu\right) \right\}
\]
(1.39)  
where \(C_\nu = \frac{\Gamma\left(\frac{\nu}{2}\right)(\nu + 1)^{(\nu - 1)/2}}{2\sqrt{\pi} \Gamma\left(\frac{\nu + 1}{2}\right)}\) and \(G(.; \nu)\) denotes the cdf of the usual univariate Student-\(t\) distribution with \(\nu\) degrees of freedom.

1.6 Likelihood inference

1.6.1 Maximum likelihood estimation

Likelihood inference is a procedure of making a statement about a data generating process (DGP). The DGP can be represented by a statistical model we select. The goal of likelihood inference is to make a statement about the unknown parameters \(\theta\) that governs the model. The likelihood principle states that all the revelent information about the parameters is contained in the observed data, represented by the likelihood function. Therefore, the likelihood function is a function of \(\theta\) given the observed data and associated with the probability of the data given the parameters. The maximum likelihood estimator (MLE) is then obtained by finding the value of \(\theta\)
in the parameter space \( \Theta \) which is “most likely” to have generated the observed data, i.e.,

\[
\hat{\theta}_{MLE} = \arg\max_{\theta \in \Theta} L(\theta | x)
\]  

(1.41)

Setting the first derivative (gradient) of the log-likelihood function to zero

\[
\frac{\partial \ln L}{\partial \theta} = 0,
\]

(1.42)

the MLEs become the solution of the first-order conditions.

In many situations, however, there is no analytical solution or closed-form expressions. Hence, numerical optimization algorithms need to be used. Numerical methods for nonlinear optimization problems are iterative in nature. These algorithms begin by assuming starting values for the unknown parameters and then proceed iteratively. In our estimation method, we implemented a hybrid of numerical optimization algorithms, namely, NM followed by BFGS.

### 1.6.2 The NM method

Nelder-Mead (NM) simplex algorithm, proposed by Nelder and Mead (1965), is one of the most widely used direct search methods for nonlinear optimization in a multidimensional space

\[
x^* = \arg\min_x f(x)
\]

(1.43)

which targets finding the minimum of a nonlinear objective function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), or alternatively the maximum of \(-f\). Here, \( x^* \) is the solution. This algorithm is derivative-free and only requires the values of \( f \). It performs simplex search using heuristic ideas.
In geometry, a simplex is defined as the convex hull of a set of \((n + 1)\) affinely independent points, \(\{ \mathbf{x}_j \}^{n+1}_{j=1} (\mathbf{x}_j \in \mathbb{R}^n)\), called vertices of the simplex. For example, a two-vertex simplex with one dimension is a line segment and a three-vertex simplex is a triangle with two dimensions.

The Nelder-Mead algorithm starts with a nondegenerate simplex and then modifies the simplex at each iteration using the operations of reflection, contraction, expansion and shrinking respectively. The Nelder-Mead method iteratively generates a sequence of transformations of the working simplex to approximate an optimal point of \(f\). At each iteration, the vertices of the simplex are sorted according to the objective function values

\[
 f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \cdots \leq f(\mathbf{x}_{n+1}),
\]

(1.44)

where \(\mathbf{x}_1\) is called the best vertex and \(\mathbf{x}_{n+1}\) the worst.

The algorithm attempts to replace the worst vertex \(\mathbf{x}_{n+1}\) with a new point of the form

\[
 \mathbf{x}(\delta) = (1 + \delta) \overline{\mathbf{x}} - \delta \mathbf{x}_{n+1},
\]

(1.45)

where \(\overline{\mathbf{x}}\) is the centroid of the convex hull of \(\{ \mathbf{x}_j \}^n_{j=1}\) given by

\[
 \overline{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i.
\]

Here, \(\delta \in \mathbb{R}\) is a coefficient associated with a particular step, which can be the reflection, contraction and expansion, or shrinking step of the Nelder-Mead iteration. The algorithm terminates if either \(f(\mathbf{x}_{n+1}) - f(\mathbf{x}_1)\) is sufficiently small, or the number of function evaluations has exceeded an user-specified limit.
1.6.3 The BFGS method

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is one of the most popular quasi-Newton methods. BFGS not only possesses computational efficiency and good asymptotic convergence, but also very effective self-correcting properties (see Wright and Nocedal (1999)).

Suppose the objective function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) in (1.43) is convex and twice continuously differentiable. The well-known Newton method is derivative-based. It uses the quadratic Taylor expansion of \( f \) at the current point \( x_k \):

\[
f(x) \approx f(x_k) + (x - x_k)^T \nabla f(x_k) + \frac{1}{2} (x - x_k)^T H(x_k)(x - x_k),
\]

where \( H(x_k) = \nabla^2 f(x_k) \) is the Hessian of \( f(x) \) at \( x_k \). The First-Order Necessary Condition requires

\[
\nabla f(x_k) + H(x_k)(x - x_k) = 0
\]

which yields

\[
x_{k+1} = x_k - H(x_k)^{-1} \nabla f(x_k)
\]

Newton method computes the Hessian directly and converges quickly. However, the method assumes \( H(x_k) \) is nonsingular and positive-definite at each iteration and there is no guarantee that \( f(x_{k+1}) \leq f(x_k) \). For large \( n \), the evaluation of \( H \) can be computationally expensive.

Various modifications of the Newtons method, called quasi-Newton methods, mitigate some of these problems. Instead of the exact computation of \( H \), BFGS algorithm
approximates $H$ at each iteration by the update formula

$$H_{k+1} = H_k - \frac{(H_k s_k)(H_k s_k)^T}{s_k^T H_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$  

(1.49)

where $s_k = x_{k+1} - x_k$ and $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$. $H_{k+1}$ satisfies the secant equation $H_{k+1} s_k = y_k$ and converges to $H(x^*)$.

The BFGS quasi-Newton method with the analytic gradient will be faster, more stable and lead to more accurate estimates than a numerical gradient (see Bard (1974), Bolker (2008) and Mayorov (2011)).

1.6.4 Standard errors of MLEs

Under certain regularity conditions, the ML estimator $\hat{\theta}$ is $\sqrt{n}$-consistent and the asymptotic distribution is

$$\sqrt{n} \left( \hat{\theta} - \theta_0 \right) \xrightarrow{d} N \left( 0, nI(\theta_0)^{-1} \right),$$  

(1.50)

where $I(\theta_0) = -E \left[ \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} |_{\theta = \theta_0} \right] = -E \left[ H(\theta_0) \right]$ and $H = \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'}$ is the Hessian matrix. In order to estimate the asymptotic covariance matrix of $\hat{\theta}$, we consider the well-known consistent estimator $-H(\hat{\theta})^{-1}$, which simply evaluates the second derivative matrix of the log-likelihood function at the MLEs. The square roots of its diagonal elements provide the standard errors of the estimates of the parameters.
1.7 Simulation study

Monte Carlo simulations play an important role in research. Lots of statistical models and methods are introduced or validated based on evidence from Monte Carlo simulations. In this thesis, we conduct simulation studies to investigate the performance of inferential methods developed here. For the BST-ACD model, given a true parameter vector, various settings based on different sample sizes are considered. For the bivariate BS-ACD and BST-ACD models, we fix vectors of true parameters. A number of different scenarios concerning sample sizes and levels of correlation are then considered. We first generate 1000 simulated datasets from the true models. For each setting and each estimate, the performance of the MLEs of the models under consideration are then evaluated in terms of the mean, coefficients of skewness and kurtosis, bias and root mean squared error (RMSE) of the MLEs over 1000 replicates.

1.8 Model discrimination

In this thesis, we use model discrimination to evaluate the performance of the BST-ACD and BVBST-ACD models, respectively. We consider a set of candidate models, which are the special cases of the concerned general model. Based on the general model, we are then interested in assessing how often the true model gets selected in the set of candidate models. such a model discrimination can be done either by Likelihood-based test or by Information-based criterion.
1.8.1 Likelihood-based method

We study the performance of the likelihood-ratio test for testing the null hypothesis that the data generating process (DGP) can be described by one of the simple models versus an alternative hypothesis that the DGP can be described by a member of the general model family other than the one specified under the null hypothesis. The likelihood-ratio test statistic can be expressed as $\Lambda = 2(\hat{l} - \hat{l}_0)$, where $\hat{l}_0$ and $\hat{l}$ are the constrained and unconstrained maximized log-likelihood values. The large sample null distribution of $\Lambda$ is simple. According to Wilks’ theorem, under some regularity conditions, for testing $H_0 : \nu = \nu_0$ versus $H_1 : \nu = \nu_1$, $\Lambda \sim \chi^2_1$, where $\chi^2_1$ represents a chi-square distribution with one degree of freedom. However, such results are only valid in the interior of the parameter space. In the case of $H_0 : \nu \to \infty$, where the parameter is at the boundary, the asymptotic null distribution of $\Lambda$ can be approximated as $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$, i.e., a mixture of a point mass at zero and a $\chi^2_1$ distribution with equal weights and $P[\Lambda \leq \lambda] = \frac{1}{2} + \frac{1}{2}P[\chi^2_1 \leq \lambda]$ (See Self and Liang (1987) and Balakrishnan and Pal (2013)).

1.8.2 Information-based criterion

Two popular model selection criteria, AIC and BIC, are used here. They can be expressed as follows: $\text{AIC} = -2l + 2k$ and $\text{BIC} = -2l + k \ln(n)$, where $l$ stands for the maximized log-likelihood value, $p$ is the number of model parameters to be estimated, and $n$ is the sample size. The model with the minimum AIC or BIC is then chosen to be the working model.
1.9 Goodness-of-fit test

We can evaluate goodness-of-fit of the models under consideration by using in-sample density forecasts and the generalized Cox-Snell (GCS) residuals. Density forecasts have come to play an important role in both macroeconomics and finance (see Tay and Wallis (2000)). Diebold et al. (1998) introduced the density forecast technique, which could be employed to evaluate ACD models (see Bauwens et al. (2004), Hautsch (2012) and Leiva et al. (2014)). Let \(\{x_i\}_{i=1}^n\) denote a series generated from the series of true predictive densities \(\{f_i(x_i|\mathcal{F}_{i-1})\}_{i=1}^n\). Let \(\{p_i(x_i|\mathcal{F}_{i-1})\}_{i=1}^n\) be a series of one-step-ahead density forecasts. The null hypothesis is that the density forecasts should coincide with the true predictive densities, that is, \(\{f_i(x_i|\mathcal{F}_{i-1})\}_{i=1}^n = \{p_i(x_i|\mathcal{F}_{i-1})\}_{i=1}^n\). Next, a series of probability integral transforms (PIT) of \(\{x_i\}_{i=1}^n\) is obtained by \(z_i = \int_{-\infty}^{x_i} p_i(u)du\). Under the null hypothesis, the distribution of \(\{z_i\}_{i=1}^n\) should be i.i.d. \(U(0,1)\).

Diebold et al. (1999) extended the above idea to the multivariate case. Here, we are especially interested in the bivariate case. Let \(\{(x_{1i}, x_{2i})\}_{i=1}^n\) represent a bivariate series and \(\{p_i(x_{1i}, x_{2i}|\mathcal{F}_{i-1})\}_{i=1}^n\) be a series of joint density forecasts. The forecasts can be decomposed into \(p_i(x_{1i}, x_{2i}) = p_i(x_{1i})p_i(x_{2i}|x_{1i})\) and \(p_i(x_{1i}, x_{2i}) = p_i(x_{2i})p_i(x_{1i}|x_{2i})\), \(i = 1, \ldots, n\). By applying the PIT to \(p_i(x_{1i}), p_i(x_{2i}), p_i(x_{1i}|x_{2i})\) and \(p_i(x_{2i}|x_{1i})\), respectively, the corresponding series we obtain, \(z_1, z_2, z_{1|2}\) and \(z_{2|1}\), should be i.i.d \(U(0,1)\) individually under the null hypothesis. According to Diebold et al. (1999), for good bivariate forecasts, not only the \(z\) series but also the pooled series \(\{z_1, z_{2|1}\}\) and \(\{z_2, z_{1|2}\}\) should each be i.i.d. \(U(0,1)\).

We can use the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests to
assess the uniformity of the \( z \) series and the Ljung-Box test to examine the independence of the \( z \) series (see Mitchell and Wallis (2011), Rossi and Sekhposyan (2014) and Leiva et al. (2014)). The Ljung-Box test here is defined as

\[
H_0 : \{z_i\} \text{ is independently distributed.}
\]

\[
H_1 : \{z_i\} \text{ is not independently distributed.}
\]

The test statistic is

\[
Q_m = n(n + 2) \sum_{h=1}^{m} \frac{\hat{\rho}_h^2}{n - h}
\]

where \( \hat{\rho}_h = \sum_{i=h+1}^{n} (z_i - \bar{z})(z_{i-h} - \bar{z}) / \sum_{i=1}^{n} (z_i - \bar{z})^2 \) and \( \bar{z} = \sum_{i=1}^{n} z_i / n \). If \( \{z_i\} \) is an i.i.d. sequence satisfying \( E(z_i^2) < \infty \), then

\[
(\sqrt{n}\hat{\rho}_1, \ldots, \sqrt{n}\hat{\rho}_m) \xrightarrow{d} N_m(0, I)
\]

as \( n \to \infty \) (see Anderson (1971), Runde (1997), Fan and Yao (2003), Tsay (2010) and Fan and Yao (2017)), where \( I \) is the \( m \times m \) identity matrix. Obviously, under \( H_0 \), \( \sum_{h=1}^{m} n\hat{\rho}_h^2 \) is then approximately \( \chi_m^2 \) distribution. When \( n \) is large, \( (n+2)/(n-h) \approx 1 \). Therefore, under \( H_0 \), the Ljung-Box test statistics

\[
Q_m = \sum_{h=1}^{m} \frac{n+2}{n-h} n\hat{\rho}_h^2 \approx \sum_{h=1}^{m} n\hat{\rho}_h^2 \xrightarrow{d} \chi_m^2
\]

as \( n \to \infty \).

Bhatti (2010) suggested to use the GCS residuals to assess the goodness-of-fit of the BS-ACD model. Suppose \( X \) is a random variable with survival function
Regardless of the form of \( S(x) \), the distribution of the random variable \( Y = -\ln(S(X)) \) is \( \exp(1) \) (see Collett (2003)). The GCS residual is then defined as \( r_{i}^{GCS} = -\ln(\hat{S}(x_{i}|F_{i-1})) \), for \( i = 1, \ldots, n \), where \( \hat{S}(x_{i}|F_{i-1}) \) is the fitted conditional survival function of the \( i \)th trade duration (see Engle and Lunde (2003), Bhatti (2010) and Leiva et al. (2014)). Under the assumption that the model is correctly specified, irrespective of the ACD specification, the GCS residuals should be i.i.d. standard exponential distribution.

### 1.10 Out-of-sample forecast evaluation

It is important to evaluate the model’s predictive performance on the data not used in the model estimation. So, we split the full data set into two subsets. We take the first part of the data set as the training set, or “in-sample” data, for the purpose of estimation and use the rest of the data as the test set or “out-of-sample” data, for prediction. The predictive performance of the model can then be assessed by using density forecasts detailed in Diebold et al. (1998) and Diebold et al. (1999), detailed above in Section 1.9.

### 1.11 Scope of the thesis

This thesis considers autoregressive conditional median duration models (ACD) in both univariate and bivariate settings. We first investigate the univariate BS\( t \)-ACD model and then propose natural bivariate extensions of the BS-ACD and BS\( t \)-ACD models for the analysis of matched durations.
The rest of this thesis proceeds as follows. In Chapter 2, we show how to implement the maximum likelihood estimation method to estimate the BS\textsubscript{t}-ACD model. Monte Carlo simulation study is conducted to evaluate the proposed method. The performance of the model is further examined by model discrimination. A real data example and comparison with two alternative models are finally provided.

In Chapters 3 and 4, we focus on the bivariate versions of ACD models. In Chapter 3, by making use of the BVBS distribution, we introduce the BVBS-ACD model to analyze two matched duration series jointly and measure the strength of dependence between them. The maximum likelihood estimation of model parameters and associated inferential methods are then discussed. The proposed model and its estimation method are then examined via a simulation study. The goodness-of-fit and predictive performance of the model are evaluated through in-sample and out-of-sample density forecasts. The methodology is illustrated finally by using a real high-frequency data.

In Chapter 4, we propose the BVBS\textsubscript{t}-ACD model and then discuss the maximum likelihood estimation of the model parameters based on a hybrid of optimization algorithms. We investigate the performance of the inferential methods developed here through a Monte Carlo simulation study. The goodness-of-fit and predictive performance of the model are also discussed. Model discrimination using information-based method is conducted. We give model fitting and forecasting results using the modeling procedure on a simulated bivariate dataset. The results show the superiority of the proposed approach in comparison with the BVBS-ACD model.

In Chapter 5, we give a summary of the thesis and suggest some possible directions for future research.
Chapter 2

BS\textsuperscript{t}-ACD model

2.1 Introduction

The generalized Birnbaum-Saunders (GBS) distribution was introduced by Díaz-García and Leiva (2005), which allows great flexibility in terms of the kurtosis and skewness and can therefore be used as an alternative to the classical BS distribution in robustness studies (see Barros et al. (2009), Balakrishnan et al. (2007), Leiva et al. (2008) and Barros et al. (2009)).

Leiva et al. (2014) proposed the BS\textsuperscript{t}-ACD model and developed associated EM algorithm for the estimation of model parameters. To estimate the BS-ACD model, Bhatti (2010) developed a maximum likelihood (ML) method based on a mixture of optimization algorithms, namely NM followed by BFGS. In this Chapter, as an extension of the ML estimation procedure developed by Bhatti (2010), we suggest a maximum likelihood method with an extra shape parameter to estimate the BS\textsuperscript{t}-ACD model.

The rest of this Chapter is organized as follows. The BS\textsuperscript{t}-ACD model and the
corresponding log-likelihood function is presented in Section 2.2. We derive the first and second derivatives of $\ln L(\theta)$ in Sections 2.3 and 2.4. In Section 2.5, we discuss the maximum likelihood based estimation method and associated inference for the model parameters. A simulation study is carried out to examine the performance of these estimates in Section 2.6. Model discrimination is carried out using the likelihood-based method and information-based criteria in Section 2.7. In Section 2.8, we illustrate the proposed methodology using high-frequency data on two stocks from the New York Stock Exchange. We present the applications of the BSt-ACD model to trade durations and then a comparison with existing alternatives, the BS-ACD and GG-ACD models.

### 2.2 BSt-ACD Model and the log-likelihood function

The BSt-ACD model allows more flexibility in kurtosis and skewness by including an extra parameter, the degrees of freedom of the Student-$t$ distribution. This may potentially facilitate robust parameter estimation.

By working with the density of $X_i$ directly, the BSt-ACD model can be expressed as

$$X_i = \sigma_i \epsilon_i, \quad \ln \sigma_i = \alpha + \beta \ln \sigma_{i-1} + \gamma \left( \frac{X_{i-1}}{\sigma_{i-1}} \right), \quad |\beta| < 1,$$

(2.1)

where $X_i \sim \text{BSt}(\kappa, \sigma_i, \nu)$ and $\epsilon_i \overset{\text{iid}}{\sim} \text{BSt}(\kappa, 1, \nu)$. 

The pdf associated with the BSt-ACD model is given by

\[
f_{X_i}(x_i; \theta) = \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{2\sqrt{\pi \Gamma\left(\frac{\nu}{2}\right)} \kappa x_i} \left[\left(\frac{\sigma_i}{x_i}\right)^{\frac{1}{2}} + \left(\frac{\sigma_i}{x_i}\right)^{\frac{3}{2}}\right]^2 \left[1 + \frac{\left(\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right)}{\nu \kappa^2}\right]^{-\frac{\nu + 1}{2}}\]

\[
= \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{2\sqrt{\pi \Gamma\left(\frac{\nu}{2}\right)} \kappa x_i} \left[\left(\frac{x_i}{\sigma_i}\right)^{\frac{1}{2}} + \left(\frac{\sigma_i}{x_i}\right)^{\frac{1}{2}}\right]^2 \left[1 + \frac{\left(\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right)}{\nu \kappa^2}\right]^{-\frac{\nu + 1}{2}},
\]

here, \( \theta = (\alpha, \beta, \gamma, \kappa, \nu)' \) is the model parameter vector. \( x_i > 0, \nu > 0, i = 1, \ldots, n. \)

The corresponding survival function can be expressed as

\[
S_{BS}(x_i; \kappa, \sigma_i, \nu) = 1 - \frac{1}{2} \left\{ 1 + I_{[a(x_i)^2]/[a(x_i)^2 + \nu]} \left( \frac{1}{2}, \frac{\nu}{2} \right) \right\},
\]

where \( a(x_i) = (\sqrt{x_i/\sigma_i} - \sqrt{\sigma_i/x_i})/\kappa \) and \( I_x(a, b) = \int_0^x t^{a-1}(1 - t)^{b-1} dt / \int_0^1 t^{a-1}(1 - t)^{b-1} dt \) is the incomplete beta ratio (see Leiva et al. (2014)).

For \( i = 1, 2, \ldots, n \), the individual log-likelihood function can be expressed as

\[
\ln l_i(\theta) = -\ln 2\sqrt{\pi} - \ln(\kappa) - \ln(x_i) + \ln \Gamma\left(\frac{\nu + 1}{2}\right) - \frac{1}{2} \ln(\nu) - \ln \Gamma\left(\frac{\nu}{2}\right)
\]

\[
+ \ln \left[\left(\frac{x_i}{\sigma_i}\right)^{\frac{1}{2}} + \left(\frac{\sigma_i}{x_i}\right)^{\frac{1}{2}}\right] - \frac{\nu + 1}{2} \ln \left[1 + \frac{\left(\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right)}{\nu \kappa^2}\right]
\]

The log-likelihood function is then given by

\[
\ln L(\theta) = \sum_{i=1}^{n} \left\{ -\ln 2\sqrt{\pi} - \ln(\kappa) - \ln(x_i) + \ln \Gamma\left(\frac{\nu + 1}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2} \ln(\nu)
\]

\[
+ \ln \left[\left(\frac{x_i}{\sigma_i}\right)^{\frac{1}{2}} + \left(\frac{\sigma_i}{x_i}\right)^{\frac{1}{2}}\right] - \frac{\nu + 1}{2} \ln \left[1 + \frac{\left(\frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2\right)}{\nu \kappa^2}\right] \right\}. \quad (2.4)
\]
2.3 The first derivatives of \( \ln L(\theta) \)

The first derivatives of \( \ln L(\theta) \) with respect to \( \alpha \), \( \beta \) and \( \gamma \) are given by

\[
\begin{align*}
\frac{\partial \ln L}{\partial \alpha} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \alpha}, \\
\frac{\partial \ln L}{\partial \beta} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \beta} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \beta}, \\
\frac{\partial \ln L}{\partial \gamma} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \gamma} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \gamma},
\end{align*}
\]

where

\[
\frac{\partial \ln l_i}{\partial \sigma_i} = \frac{x_i - \sigma_i}{2\sigma_i} \left( \frac{(\nu + 1)(x_i + \sigma_i)}{\kappa x_i \sigma_i \nu + (x_i - \sigma_i)^2} - \frac{1}{x_i + \sigma_i} \right),
\]

(2.5)

and by Lemma 2.3 in Mayorov (2011),

\[
\begin{align*}
\frac{\partial \sigma_i}{\partial \alpha} &= \sigma_i \left( 1 + \frac{1}{\sigma_i} \frac{\partial \sigma_i}{\partial \alpha} \left( \beta - \gamma \frac{x_i - 1}{\sigma_i} \right) \right), \\
\frac{\partial \sigma_i}{\partial \beta} &= \sigma_i \left( \ln \sigma_{i-1} + \frac{1}{\sigma_{i-1}} \frac{\partial \sigma_i}{\partial \beta} \left( \beta - \gamma \frac{x_i - 1}{\sigma_{i-1}} \right) \right), \\
\frac{\partial \sigma_i}{\partial \gamma} &= \sigma_i \left( \frac{x_i - 1}{\sigma_{i-1}} + \frac{1}{\sigma_{i-1}} \frac{\partial \sigma_i}{\partial \gamma} \left( \beta - \gamma \frac{x_i - 1}{\sigma_{i-1}} \right) \right); \tag{2.6, 2.7, 2.8}
\end{align*}
\]
also, from by Lemma 2.4 in Mayorov (2011),

\[
\frac{\partial \sigma_i}{\partial \alpha} = \sigma_i \left(1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-2} \left(\beta - \gamma \frac{x_l}{\sigma_l}\right)\right), \quad i \geq 1,
\]

(2.9)

\[
\frac{\partial \sigma_i}{\partial \beta} = \sigma_i \left(\ln \sigma_i - 1 + \sum_{k=0}^{i-2} \ln \sigma_k \prod_{l=k+1}^{i-1} \left(\beta - \gamma \frac{x_l}{\sigma_l}\right)\right), \quad i \geq 2,
\]

(2.10)

\[
\frac{\partial \sigma_i}{\partial \gamma} = \sigma_i \left(\frac{x_i}{\sigma_i} - 1 \frac{\sigma_i}{x_i} + \sum_{k=0}^{i-2} \frac{x_k}{\sigma_k} \prod_{l=k+1}^{i-1} \left(\beta - \gamma \frac{x_l}{\sigma_l}\right)\right), \quad i \geq 2,
\]

(2.11)

with \(\nabla \sigma_0 = (0, 0, 0)^t\) and \(\nabla \sigma_1 = (\sigma_1, \sigma_1 \ln \sigma_0, \sigma_1 x_0 / \sigma_0)^t\).

The first derivatives of \(\ln L(\theta)\) with respect to \(\kappa\) and \(\nu\) are given by

\[
\frac{\partial \ln L}{\partial \kappa} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa},
\]

\[
\frac{\partial \ln L}{\partial \nu} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \nu},
\]

where

\[
\frac{\partial \ln l_i}{\partial \kappa} = \frac{1}{\kappa} \left(\frac{(\nu + 1)(x_i - \sigma_i)^2}{\kappa^2 x_i \sigma_i \nu + (x_i - \sigma_i)^2} - 1\right),
\]

(2.12)

\[
\frac{\partial \ln l_i}{\partial \nu} = \frac{1}{2} \left\{ \Psi \left(\frac{\nu + 1}{2}\right) - \Psi \left(\frac{\nu}{2}\right) + \frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2 - \kappa^2 \right. \\
\left. + \frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2 + \nu \kappa^2 - \ln \left(1 + \frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2 + \nu \kappa^2\right)\right\}.
\]

(2.13)

with \(\Psi(x) = \frac{d}{dx} \ln (\Gamma(x)) = \frac{\Gamma'(x)}{\Gamma(x)}\) denoting the digamma function.
2.4 The second derivatives of $\ln(\theta)$

The second derivatives of $\ln(\theta)$ with respect to $\theta = (\alpha, \beta, \gamma, \kappa, \nu)'$ are as follows:

\[
\begin{align*}
\frac{\partial^2 \ln L}{\partial \alpha^2} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha} \frac{\partial^2 \ln l_i}{\partial \alpha} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \alpha} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \alpha^2} \right], \\
\frac{\partial^2 \ln L}{\partial \beta^2} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta} \frac{\partial^2 \ln l_i}{\partial \beta} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \beta} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \beta^2} \right], \\
\frac{\partial^2 \ln L}{\partial \gamma^2} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma} \frac{\partial^2 \ln l_i}{\partial \gamma} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \gamma} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \gamma^2} \right], \\
\frac{\partial^2 \ln L}{\partial \alpha \partial \beta} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha \partial \beta} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha} \frac{\partial^2 \ln l_i}{\partial \beta} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \alpha} \right) \left( \frac{\partial \sigma_i}{\partial \beta} \right) + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \alpha \partial \beta} \right], \\
\frac{\partial^2 \ln L}{\partial \alpha \partial \gamma} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha \partial \gamma} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha} \frac{\partial^2 \ln l_i}{\partial \gamma} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \alpha} \right) \left( \frac{\partial \sigma_i}{\partial \gamma} \right) + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \alpha \partial \gamma} \right], \\
\frac{\partial^2 \ln L}{\partial \beta \partial \gamma} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta \partial \gamma} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta} \frac{\partial^2 \ln l_i}{\partial \gamma} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_i^2} \left( \frac{\partial \sigma_i}{\partial \beta} \right) \left( \frac{\partial \sigma_i}{\partial \gamma} \right) + \frac{\partial \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \beta \partial \gamma} \right], \\
\frac{\partial^2 \ln L}{\partial \alpha \partial \kappa} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha \partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha} \frac{\partial^2 \ln l_i}{\partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \alpha \partial \kappa}, \\
\frac{\partial^2 \ln L}{\partial \beta \partial \kappa} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta \partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta} \frac{\partial^2 \ln l_i}{\partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \beta \partial \kappa}, \\
\frac{\partial^2 \ln L}{\partial \gamma \partial \kappa} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma \partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma} \frac{\partial^2 \ln l_i}{\partial \kappa} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \gamma \partial \kappa}, \\
\frac{\partial^2 \ln L}{\partial \alpha \partial \nu} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha} \frac{\partial^2 \ln l_i}{\partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \alpha \partial \nu}, \\
\frac{\partial^2 \ln L}{\partial \beta \partial \nu} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta} \frac{\partial^2 \ln l_i}{\partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \beta \partial \nu}, \\
\frac{\partial^2 \ln L}{\partial \gamma \partial \nu} &= \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma} \frac{\partial^2 \ln l_i}{\partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \gamma \partial \nu}.
\end{align*}
\]
\[
\frac{\partial^2 \ln L}{\partial \kappa^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa^2}, \\
\frac{\partial^2 \ln L}{\partial \kappa \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa \partial \nu}, \\
\frac{\partial^2 \ln L}{\partial \nu^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \nu^2},
\]

where

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_i^2} = \frac{1}{2\sigma_i} \left\{ \frac{(\nu + 1) (x_i - \sigma_i) (x_i + \sigma_i) \left( \kappa^2 \nu x_i - 2 (x_i - \sigma_i) \right)}{\left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2} \right. \\
+ \left. \frac{(\nu + 1) (x_i^2 + \sigma_i^2)}{\sigma_i \left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)} - \frac{x_i^2 + 2 \sigma_i x_i - \sigma_i^2}{\sigma_i (x_i + \sigma_i)^2} \right\},
\]

(2.14)

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_i \partial \kappa} = \frac{\kappa \nu (\nu + 1) x_i (\sigma_i^2 - x_i^2)}{\left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2}
\]

(2.15)
\[
\frac{\partial^2 \ln l_i}{\partial \sigma_i \partial \nu} = \frac{(\sigma_i^2 - x_i^2) \left( \kappa^2 \sigma_i \nu x_i - \nu (x_i - \sigma_i)^2 \right)}{2 \sigma_i \nu \left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2},
\]

(2.16)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa^2} = \frac{1}{\kappa^2} \left\{ 1 - \frac{(\nu + 1) (x_i - \sigma_i)^2 \left( (x_i - \sigma_i)^2 + 3 \kappa^2 \sigma_i \nu x_i \right)}{\left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2} \right\},
\]

(2.17)

\[
\frac{\partial^2 \ln l_i}{\partial \k \partial \nu} = \frac{(x_i - \sigma_i)^2 \left( 1 + \nu - (x_i - \sigma_i)^2 - \kappa^2 \sigma_i \nu x_i \right)}{\kappa \nu \left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2},
\]

(2.18)

\[
\frac{\partial^2 \ln l_i}{\partial \nu^2} = \frac{1}{2} \left\{ \frac{(x_i - \sigma_i)^2 \left( (\nu - 1) (x_i - \sigma_i)^2 - 2 \kappa^2 \sigma_i \nu x_i \right)}{\nu^2 \left( (x_i - \sigma_i)^2 + \kappa^2 \sigma_i \nu x_i \right)^2} + \frac{\Psi_1 \left( \frac{\nu + 1}{2} \right)}{2} - \frac{\Psi_1 \left( \frac{\nu}{2} \right)}{2} + \frac{1}{\nu^2} \right\},
\]

(2.19)

\(\Psi_1(x)\) is the second derivative of \(\ln(\Gamma(x))\) function.
\[
\frac{\partial^2 \sigma_i}{\partial \alpha^2} = \frac{\partial \sigma_i}{\partial \alpha} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_i \gamma \prod_{l=k}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{x_{i} \frac{\partial \sigma_{jl}}{\partial \alpha}}{\sigma_{jl}^2} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right), \\
i \geq 1, \quad (2.20)
\]

\[
\frac{\partial^2 \sigma_i}{\partial \beta^2} = \frac{\partial \sigma_i}{\partial \gamma} \left\{ \ln \sigma_{j,i-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right\} + \sigma_i \left\{ \sum_{k=0}^{i-2} \left( \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma x_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma}}{\sigma_{jl}} + 1 \right) \frac{\partial \sigma_{jk}}{\partial \gamma} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right\} + \frac{\partial \sigma_{j,i-1}}{\partial \gamma}, \\
i \geq 2, \quad (2.21)
\]

\[
\frac{\partial^2 \sigma_i}{\partial \gamma^2} = \frac{\partial \sigma_i}{\partial \gamma} \left\{ \frac{x_{j,i-1}}{\sigma_{j,i-1}} + \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right\} + \sigma_i \left\{ \sum_{k=0}^{i-2} \left( \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma x_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma}}{\sigma_{jl}} + 1 \right) \frac{\partial \sigma_{jk}}{\partial \gamma} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right\} + \frac{\partial \sigma_{j,i-1}}{\partial \gamma}, \\
i \geq 2, \quad (2.22)
\]
\[
\frac{\partial^2 \sigma_i}{\partial \alpha \partial \beta} = \frac{\partial \sigma_i}{\partial \beta} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_i \sum_{l=1}^{i-1} \prod_{k=l}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma x_{jl}}{\sigma_{jl}} \frac{\partial \sigma_{jl}}{\partial x_{jl}} + 1, \\
i \geq 1, \quad (2.23)
\]

\[
\frac{\partial^2 \sigma_i}{\partial \alpha \partial \gamma} = \frac{\partial \sigma_i}{\partial \gamma} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_i \sum_{l=1}^{i-1} \prod_{k=l}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma x_{jl}}{\sigma_{jl}} \frac{\partial \sigma_{jl}}{\partial x_{jl}} - \frac{x_{jl}}{\sigma_{jl}}, \\
i \geq 1, \quad (2.24)
\]

\[
\frac{\partial^2 \sigma_i}{\partial \beta \partial \gamma} = \frac{\partial \sigma_i}{\partial \gamma} \left\{ \ln \sigma_{j,i-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \right\} \\
+ \sigma_i \left\{ \sum_{k=0}^{i-2} \left( \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma x_{jl}}{\sigma_{jl}} \frac{\partial \sigma_{jl}}{\partial x_{jl}} - \frac{x_{jl}}{\sigma_{jl}} \right) + \frac{\partial \sigma_{j,i-1}}{\partial \gamma} \prod_{l=k+1}^{i-1} \left( \beta - \gamma \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\partial \sigma_{jl}}{\partial x_{jl}} + \frac{\partial \sigma_{j,i-1}}{\partial \gamma} \right\} \\
i \geq 2. \quad (2.25)
\]
2.5 Estimation and inference

We estimate the parameter \( \theta = (\alpha, \beta, \gamma, \kappa, \nu)' \) of the BS\( t\)-ACD model by maximizing the log-likelihood function in equation (2.4). We propose the following two-step procedure for this purpose of estimation. In the first step, we estimate the ACD parameters, \( \alpha, \beta, \) and \( \gamma, \) by using the NM algorithm with \( \kappa \) and \( \nu \) fixed at their initial values. Then in the second step, we estimate over the whole parameter space by employing the BFGS algorithm with analytical gradients. We have derived and implemented the analytical gradients (the first derivatives in Section 2.3) in the second step.

Under certain regularity conditions (smoothness and boundariness), the ML estimator \( \hat{\theta} \) is \( \sqrt{n} \)-consistent and the asymptotic distribution of it is

\[
\sqrt{n} \left( \hat{\theta} - \theta_0 \right) \xrightarrow{d} N_5 \left( 0, n \mathcal{I}(\theta_0)^{-1} \right)
\]

(2.26)

where \( \mathcal{I}(\theta_0) = -E \left[ \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \right]_{\theta = \theta_0} = -E \left[ H(\theta_0) \right] \) with \( H = \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \) being the Hessian matrix. In order to estimate the asymptotic covariance matrix of \( \hat{\theta} \), we consider the well-known consistent estimator \( -H(\hat{\theta})^{-1} \), which simply evaluates the second derivative matrix of the log-likelihood function at the MLEs. The square roots of its diagonal elements then provide the standard errors of the estimates.
2.6 Simulation study

In this Section, we present the results of a simulation study carried out for evaluating the performance of the maximum likelihood estimates determined by the procedure described in Section 2.5.

The simulation scenarios considered are as follows: sample sizes $n \in \{3000, 5000, 10000, 20000\}$ and the vector of parameters $\theta = (\alpha, \beta, \gamma, \kappa, \nu)' = (0.1, 0.9, 0.1, 1.0, 12)'$. To estimate the BST-ACD model, we set the starting values for the ACD parameters to be $(\alpha_0, \beta_0, \gamma_0)' = (0.01, 0.80, 0.01)'$. $\kappa_0$ is determined as $\kappa_0 = \sqrt{2 \left( \frac{\bar{x}}{\text{med}(X)} - 1 \right) \frac{\nu^{-2}}{\nu}}$. We set $\nu_0 = 3$.

For each sample size and each estimate, we computed the mean, coefficients of skewness and kurtosis, bias and root mean squared error (RMSE) of the MLEs. Tables 2.1-2.4 summarize the distribution of the MLEs over 1000 simulation trials. From the simulated values presented in Tables 2.1-2.4, we observe that the bias, MSE and RMSE are consistently small and tend to 0 as sample size increases from 3000 to 20,000. It reveals that the MLEs are asymptotically unbiased and consistent. Moreover, the standard errors of the estimates become smaller, and also the sample skewness and kurtosis tend to 0 and 3, respectively, with increasing sample size. This reveals that the asymptotic distribution of the MLEs are normal, as suggested by the asymptotic distribution of the MLEs stated above.
Table 2.1: Simulation results for BS$^t$-ACD model when $n = 3000$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\kappa}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>trueparam</td>
<td>0.1000</td>
<td>0.9000</td>
<td>0.1000</td>
<td>1.1000</td>
<td>12.0000</td>
</tr>
<tr>
<td>mean</td>
<td>0.0994</td>
<td>0.8972</td>
<td>0.1042</td>
<td>1.0909</td>
<td>10.9443</td>
</tr>
<tr>
<td>skew</td>
<td>0.7005</td>
<td>-0.6133</td>
<td>-0.0666</td>
<td>0.1661</td>
<td>3.4636</td>
</tr>
<tr>
<td>kurt</td>
<td>3.9595</td>
<td>3.9989</td>
<td>2.9727</td>
<td>3.2400</td>
<td>22.2730</td>
</tr>
<tr>
<td>bias</td>
<td>-0.0006</td>
<td>-0.0028</td>
<td>0.0042</td>
<td>-0.0091</td>
<td>-1.0557</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0031</td>
<td>0.0005</td>
<td>0.0002</td>
<td>0.0016</td>
<td>26.4100</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0558</td>
<td>0.0226</td>
<td>0.0126</td>
<td>0.0399</td>
<td>5.1391</td>
</tr>
<tr>
<td>SE</td>
<td>0.0558</td>
<td>0.0225</td>
<td>0.0119</td>
<td>0.0389</td>
<td>5.0320</td>
</tr>
</tbody>
</table>

Table 2.2: Simulation results for BS$^t$-ACD model when $n = 5000$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\kappa}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>trueparam</td>
<td>0.1000</td>
<td>0.9000</td>
<td>0.1000</td>
<td>1.1000</td>
<td>12.0000</td>
</tr>
<tr>
<td>mean</td>
<td>0.1010</td>
<td>0.8991</td>
<td>0.1007</td>
<td>1.0946</td>
<td>11.2733</td>
</tr>
<tr>
<td>skew</td>
<td>0.2384</td>
<td>-0.2098</td>
<td>-0.0053</td>
<td>0.1486</td>
<td>1.2490</td>
</tr>
<tr>
<td>kurt</td>
<td>2.9921</td>
<td>2.9386</td>
<td>2.9319</td>
<td>2.8445</td>
<td>5.2743</td>
</tr>
<tr>
<td>bias</td>
<td>0.0010</td>
<td>-0.0009</td>
<td>0.0007</td>
<td>-0.0054</td>
<td>-0.7267</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0004</td>
<td>4.0278</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0238</td>
<td>0.0096</td>
<td>0.0051</td>
<td>0.0193</td>
<td>2.0069</td>
</tr>
<tr>
<td>SE</td>
<td>0.0238</td>
<td>0.0096</td>
<td>0.0050</td>
<td>0.0185</td>
<td>1.8717</td>
</tr>
</tbody>
</table>
Table 2.3: Simulation results for BSt-ACD model when $n = 10000$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\kappa}$</th>
<th>$\hat{v}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>trueparam</td>
<td>0.1000</td>
<td>0.9000</td>
<td>0.1000</td>
<td>1.1000</td>
<td>12.0000</td>
</tr>
<tr>
<td>mean</td>
<td>0.1000</td>
<td>0.8997</td>
<td>0.1003</td>
<td>1.0953</td>
<td>11.3754</td>
</tr>
<tr>
<td>skew</td>
<td>0.1613</td>
<td>-0.1726</td>
<td>0.0819</td>
<td>0.1590</td>
<td>0.9950</td>
</tr>
<tr>
<td>kurt</td>
<td>3.1191</td>
<td>3.1408</td>
<td>2.8582</td>
<td>2.9853</td>
<td>4.4802</td>
</tr>
<tr>
<td>bias</td>
<td>0.0000</td>
<td>-0.0003</td>
<td>0.0003</td>
<td>-0.0047</td>
<td>-0.6246</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0003</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>2.3332</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0168</td>
<td>0.0068</td>
<td>0.0036</td>
<td>0.0141</td>
<td>1.5275</td>
</tr>
<tr>
<td>SE</td>
<td>0.0168</td>
<td>0.0068</td>
<td>0.0036</td>
<td>0.0133</td>
<td>1.3947</td>
</tr>
</tbody>
</table>

Table 2.4: Simulation results for BSt-ACD model when $n = 20000$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\kappa}$</th>
<th>$\hat{v}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>trueparam</td>
<td>0.1000</td>
<td>0.9000</td>
<td>0.1000</td>
<td>1.1000</td>
<td>12.0000</td>
</tr>
<tr>
<td>mean</td>
<td>0.0996</td>
<td>0.9001</td>
<td>0.1001</td>
<td>1.0967</td>
<td>11.5472</td>
</tr>
<tr>
<td>skew</td>
<td>0.1929</td>
<td>-0.1420</td>
<td>0.1562</td>
<td>-0.0386</td>
<td>0.2609</td>
</tr>
<tr>
<td>kurt</td>
<td>3.1248</td>
<td>3.3047</td>
<td>2.9425</td>
<td>2.9998</td>
<td>2.9496</td>
</tr>
<tr>
<td>bias</td>
<td>-0.0004</td>
<td>0.0001</td>
<td>0.0001</td>
<td>-0.0033</td>
<td>-0.4528</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
<td>1.4330</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0115</td>
<td>0.0046</td>
<td>0.0025</td>
<td>0.0106</td>
<td>1.1971</td>
</tr>
<tr>
<td>SE</td>
<td>0.0115</td>
<td>0.0046</td>
<td>0.0025</td>
<td>0.0100</td>
<td>1.1087</td>
</tr>
</tbody>
</table>
2.7 Model discrimination

The BS\textsuperscript{t}-ACD model provides more flexibility in terms of kurtosis and skewness by the inclusion of an additional parameter, namely, the degrees of freedom $\nu$ of the Student-$t$ distribution. The BS\textsubscript{7}, BS\textsubscript{9}, BS\textsubscript{12} and BS-ACD models are some special cases, for example. Using the BS\textsuperscript{t}-ACD model, it will then be of great interest to examine how often the true model gets selected in the set of candidate models and to choose a simple model that provides an adequate fit to the data. This model evaluation technique is called model discrimination. One may refer to the book by McLachlan and Peel (2000), and the recent papers by Balakrishnan and Peng (2006), Balakrishnan and Pal (2013), Balakrishnan and Pal (2016), and Balakrishnan et al. (2017) in this direction in different contexts. Two types of common selection criteria are used in the literature for this purpose, namely, likelihood ratio tests (LRTs) and information criterion.

2.7.1 Likelihood-based method

We study the performance of the likelihood ratio test (LRT) for testing the null hypothesis that the data generating process (DGP) can be described by one of the BS\textsubscript{7} ($H_0 : \nu = 7$), BS\textsubscript{9} ($H_0 : \nu = 9$), BS\textsubscript{12} ($H_0 : \nu = 12$) and BS ($H_0 : \nu \rightarrow \infty$) ACD models versus an alternative hypothesis that the DGP can be described by a member of the BS\textsuperscript{t}-ACD family other than the one specified in the null hypothesis.

First, we assume that the true data generating processes (DGP) is the BS\textsubscript{7}, BS\textsubscript{9}, BS\textsubscript{12} and BS-ACD models (true models), respectively. Next we simulate datasets from each true model and fit the BS\textsubscript{7}, BS\textsubscript{9}, BS\textsubscript{12}, BS and BS\textsuperscript{t}-ACD models to these datasets, respectively. We use 1000 replications of sample sizes of $n \in$
\{2000, 3000, 5000\} and the same parameter setting \((\alpha, \beta, \gamma, \kappa)' = (0.1, 0.9, 0.1, 1.0)'\) for all true models. For each simulated sample, we obtain the LRT statistics of the fitted \(\text{BS}_{t7}, \text{BS}_{t9}, \text{BS}_{t12}\) and \(\text{BS-ACD}\) models versus the fitted general \(\text{BS}_t\text{-ACD}\) model. We then evaluate the observed significance levels and powers of the LRT by calculating the proportion of the 1000 simulation trials that reside in the rejection of the null hypothesis at a nominal significance level of 0.05.

The results so obtained are reported in Table 2.5. The observed levels lie on the main diagonal, shown in bold, while other numbers are the observed powers. As the sample size increases, the observed levels decrease and the observed powers increase, which show an improvement in performance with sample size.

It is easy to see that the \(\chi^2_1\) distribution provides only a reasonable approximation to the null distribution of the LRT when testing for the \(\text{BS}_{t12}\text{-ACD}\) model as the observed levels are not so close to the nominal level, varying from 8.2% to 9.5%. But, the \(\chi^2_1\) distribution provides a good approximation to the null distribution of the LRT when testing for the \(\text{BS}_{t7}\) and \(\text{BS}_{t9}\text{-ACD}\) models as the observed levels are indeed close to the nominal level. However, when testing for the \(\text{BS-ACD}\) model, the mixture chi-square distribution doesn’t provide a good approximation to the null distribution of the LRT as the observed levels are found to be considerably above the nominal level ranging from 37.5% to 42.4%. The asymptotic mixture chi-square form is not sufficient for this result. However, the observed levels decrease as the sample size increases. For substantially larger sample size, they come down to the nominal level, 5%.

The observed powers vary in different situations and increase as sample size increases. When the true model is \(\text{BS}_{t7}\text{-ACD}\), the test has high power to reject \(\text{BS}\) (the
observed power is 100% in all cases) and BST_{12}-ACD (the observed powers are from 89.7% to 99.8%) and moderate power to reject BST_9-ACD (the observed powers take values from 48.8% to 73.7%). In all cases, the test has high power to reject BST_7, BST_9 and BST_{12}-ACD models if the true DGP is BS-ACD. When the true model is the BST_9-ACD, the test has high power to reject BS-ACD (the observed power is 100% in all cases), moderate power to reject BST_{12}-ACD (the observed powers range from 45.7% to 71.0%) and low power to reject BST_7-ACD (the observed powers vary from 22.0% to 56.4%). When the true DGP is the BST_{12}-ACD, the test has high power to reject BS-ACD (the observed powers take values from 99.3% to 100%), good power to reject BST_7-ACD (the observed powers are from 64.8% to 99.0%) and low power to reject BST_9-ACD (the observed powers range from 14.3% to 42.4%).

Thus, the difference in model fitting between the BST_7, BST_9, BST_{12} or BS-ACD models is significant enough to be detected by the LRT under the BST-ACD setup.

### 2.7.2 Information-based criterion

We consider the same parameter settings as mentioned in Section 2.7.1 and generate data from the true models. For each dataset, we calculate the AIC and BIC values of the fitted the BST_7, BST_9, BST_{12} and BS-ACD models. Then, the performance of the AIC and BIC can be evaluated by the selection rates for each of the fitted models. Since all fitted models have the same number of parameters, the AIC and BIC choose the same model for the same dataset simulated from the true model, and in fact they correspond to the model with the largest maximized log-likelihood value.

These results are presented in Table 2.6. The selection rates for the correct models lie on the main diagonal, shown in bold, while other numbers are the selection rates for
the wrong models. The information criteria can distinguish between the $BSt_7$, $BSt_9$, $BSt_{12}$ and BS-ACD models with a relatively high selection rates for the true models. As the sample size increases, the selection rates for the correct models increase and the selection rates for the wrong models decrease, which show an improvement in performance with sample size.

The information criteria perform well when the true DGP is $BSt_7$-ACD and BS-ACD models. The selection rates of the correct models are from 84.7% to 91.5% if the true distribution is $BSt_7$ and from 71.3% to 92.2% if the true distribution is BS. The information criteria perform moderately when the true DGP is $BSt_9$-ACD and $BSt_{12}$-ACD models. The selection rates of the right model are from 49.3% to 72.9% if the true model is $BSt_9$-ACD and from 55.2% to 77.3% if the true DGP is $BSt_{12}$-ACD.

As expected, when the true models are $BSt_7$-ACD and $BSt_9$-ACD, the selection rates for $BSt_9$-ACD and $BSt_7$-ACD are higher than those of $BSt_{12}$-ACD, which implies that $BSt_7$-ACD is closer to $BSt_9$-ACD than $BSt_{12}$-ACD, which are consistent with those results obtained earlier by the LRT.

In general, the $BSt_7$-ACD and BS-ACD models are quite different in terms of skewness and kurtosis. For a small sample size, $n = 2000$, the information criteria can distinguish between the $BSt_7$ and BS-ACD models with a high selection rates for the true models. The $BSt_9$-ACD and $BSt_{12}$-ACD models are very close to each other. Thus, the corresponding selection rates of the correct models are low. But for a larger sample size, $n = 5000$, the information criteria can distinguish between them with a relatively high selection rates for the true models.
Table 2.5: Model Discrimination by LR test

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>True BS</th>
<th>ACD Model BS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n=2000$</td>
<td>$n=3000$</td>
</tr>
<tr>
<td>$BS_{t_7}$</td>
<td>0.071</td>
<td>0.052</td>
</tr>
<tr>
<td>$BS_{t_9}$</td>
<td>0.488</td>
<td>0.576</td>
</tr>
<tr>
<td>$BS_{t_{12}}$</td>
<td>0.897</td>
<td>0.967</td>
</tr>
<tr>
<td>BS</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>True BS</th>
<th>ACD Model BS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n=2000$</td>
<td>$n=3000$</td>
</tr>
<tr>
<td>$BS_{t_7}$</td>
<td>0.071</td>
<td>0.488</td>
</tr>
<tr>
<td>$BS_{t_9}$</td>
<td>0.202</td>
<td>0.143</td>
</tr>
<tr>
<td>$BS_{t_{12}}$</td>
<td>0.648</td>
<td>0.093</td>
</tr>
<tr>
<td>BS</td>
<td>1.000</td>
<td>0.995</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>True BS</th>
<th>ACD Model BS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n=2000$</td>
<td>$n=3000$</td>
</tr>
<tr>
<td>$BS_{t_7}$</td>
<td>0.071</td>
<td>0.064</td>
</tr>
<tr>
<td>$BS_{t_9}$</td>
<td>0.488</td>
<td>0.143</td>
</tr>
<tr>
<td>$BS_{t_{12}}$</td>
<td>0.897</td>
<td>0.883</td>
</tr>
<tr>
<td>BS</td>
<td>1.000</td>
<td>0.993</td>
</tr>
</tbody>
</table>
Table 2.6: Model Discrimination by AIC and BIC

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>BS(_t_7)</th>
<th>BS(_t_9)</th>
<th>BS(<em>t</em>{12})</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS(_t_7)</td>
<td>0.847</td>
<td>0.353</td>
<td>0.050</td>
<td>0.000</td>
</tr>
<tr>
<td>BS(_t_9)</td>
<td>0.142</td>
<td>0.493</td>
<td>0.389</td>
<td>0.001</td>
</tr>
<tr>
<td>BS(<em>t</em>{12})</td>
<td>0.010</td>
<td>0.154</td>
<td>0.552</td>
<td>0.286</td>
</tr>
<tr>
<td>BS</td>
<td>0.001</td>
<td>0.000</td>
<td>0.009</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>BS(_t_7)</th>
<th>BS(_t_9)</th>
<th>BS(<em>t</em>{12})</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS(_t_7)</td>
<td>0.866</td>
<td>0.263</td>
<td>0.012</td>
<td>0.000</td>
</tr>
<tr>
<td>BS(_t_9)</td>
<td>0.132</td>
<td>0.611</td>
<td>0.337</td>
<td>0.000</td>
</tr>
<tr>
<td>BS(<em>t</em>{12})</td>
<td>0.002</td>
<td>0.126</td>
<td>0.650</td>
<td>0.190</td>
</tr>
<tr>
<td>BS</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.810</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>BS(_t_7)</th>
<th>BS(_t_9)</th>
<th>BS(<em>t</em>{12})</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS(_t_7)</td>
<td>0.915</td>
<td>0.181</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>BS(_t_9)</td>
<td>0.085</td>
<td>0.729</td>
<td>0.225</td>
<td>0.000</td>
</tr>
<tr>
<td>BS(<em>t</em>{12})</td>
<td>0.000</td>
<td>0.090</td>
<td>0.773</td>
<td>0.078</td>
</tr>
<tr>
<td>BS</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.922</td>
</tr>
</tbody>
</table>


to trade duration data

In this section, we present the applications of the BS\(t\)-ACD model to some trade durations and a comparison with existing alternatives, namely, the BS-ACD and GG-ACD models.

2.8.1 Data description

We consider the trade durations of International Business Machines (IBM) and Johnson & Johnson (JNJ) stocks in the period of 28 consecutive trading days from January 2, 2002 to February 11, 2002. There are 91,819 (IBM) and 61,188 (JNJ) observations for each asset. The original data sets were studied in detail by Bhatti (2010) and Leiva et al. (2014).

After restricting the adjusted durations between 10:00 am and 4:00 pm, we obtain 89,171 (IBM) and 59,390 (JNJ) observations for each asset. The initial values \(\sigma_0\) for each day were set to be the median trade duration from 9:50 to 10:00 am for BS\(t\)-ACD model, and \(\psi_0\) to be the mean trade duration from 9:50 to 10:00 am for GG-ACD model.

As seen in Table 2.7, the two adjusted duration series are positive and right-skewed with high degree of kurtosis indicating heavy tails.

Table 2.7: Summary statistics for the adjusted durations

<table>
<thead>
<tr>
<th>Data</th>
<th>Min</th>
<th>Median</th>
<th>Mean</th>
<th>Max</th>
<th>SD</th>
<th>skew</th>
<th>kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>0.170</td>
<td>1.034</td>
<td>1.394</td>
<td>27.480</td>
<td>1.279</td>
<td>3.018</td>
<td>21.543</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.131</td>
<td>0.979</td>
<td>1.532</td>
<td>27.930</td>
<td>1.616</td>
<td>3.042</td>
<td>20.083</td>
</tr>
</tbody>
</table>
Figure 2.1: TTT plots for IBM (left) and JNJ (right)

Figure 2.2: ACF plots for IBM (left) and JNJ (right)
In Figure 2.1, the empirical scaled TTT transforms are plotted, from which they are seen to be first concave and then convex, revealing that both marginals may have unimodal shaped hazard rates. Therefore, the BST distribution could be a good fit to the duration data due to these shapes of its marginal density functions and hazard rates.

The Autocorrelation Function (ACF) plots in Figure 2.2 indicate that there are significant positive autocorrelations up to long lags in both series, which suggests that a ACD specification may be a reasonable choice (see Tsay (2010), Hautsch (2012) and Leiva et al. (2014)). Note: the plots are trimmed to be in the interval up to 0.1 even though the value should be 1 at 0. The ACF of the series represents the strength of linear dependence between $X_i$ and $X_{i-h}$ for $0 \leq h < n$. Here, the lag-$h$ autocorrelation of $X_i$ is defined as

$$\rho_h = \frac{Cov(X_i, X_{i-h})}{\sqrt{Var(X_i)Var(X_{i-h})}} = \frac{Cov(X_i, X_{i-h})}{Var(X_i)}. \tag{2.27}$$

Let $\bar{X} = \sum_{i=1}^{n} X_i / n$. Then the lag-$h$ sample autocorrelation of $X_i$ is

$$\hat{\rho}_h = \frac{\sum_{i=h+1}^{n} (X_i - \bar{X})(X_{i-h} - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}. \tag{2.28}$$

### 2.8.2 Estimation results

We employ the method described earlier in Section 2.5 to estimate the BST-ACD model for the IBM an JNJ duration series, respectively. For comparison purpose, we also consider two alternatives, namely, the BS-ACD and GG-ACD models, using the two-step estimation procedures as detailed in Bhatti (2010). The obtained results are
shown in Table 2.8. All estimates are statistically significant at 1% level. Overall, the BS-ACD and BSt-ACD models perform much better than the GG-ACD model in terms of AIC and BIC values. The BSt-ACD model outperforms all the considered models and the BS-ACD model provides close values according to AIC values for both series.

Table 2.8: Estimation results for three ACD models

<table>
<thead>
<tr>
<th>Data</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\gamma} )</th>
<th>( \hat{\kappa} )</th>
<th>( \hat{\nu} )</th>
<th>maxlnL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>-0.041</td>
<td>0.950</td>
<td>0.029</td>
<td>0.871</td>
<td>81.069</td>
<td>-108306.40</td>
<td>216622.80</td>
<td>216669.80</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0044)</td>
<td>(0.0014)</td>
<td>(0.0021)</td>
<td>(0.1664)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.016</td>
<td>0.979</td>
<td>0.011</td>
<td>1.012</td>
<td>142.443</td>
<td>-79467.52</td>
<td>158945.00</td>
<td>158990.00</td>
</tr>
<tr>
<td></td>
<td>(0.0015)</td>
<td>(0.0032)</td>
<td>(0.0010)</td>
<td>(0.0030)</td>
<td>(0.4033)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\gamma} )</th>
<th>( \hat{\kappa} )</th>
<th>maxlnL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>-0.040</td>
<td>0.950</td>
<td>0.029</td>
<td>0.882</td>
<td>-108314.50</td>
<td>216637.00</td>
<td>216674.60</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0044)</td>
<td>(0.0014)</td>
<td>(0.0021)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.016</td>
<td>0.978</td>
<td>0.011</td>
<td>1.019</td>
<td>-79468.83</td>
<td>158945.70</td>
<td>158981.60</td>
</tr>
<tr>
<td></td>
<td>(0.0015)</td>
<td>(0.0032)</td>
<td>(0.0010)</td>
<td>(0.0030)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data</th>
<th>( \hat{\alpha} )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\gamma} )</th>
<th>( \hat{\nu} )</th>
<th>( \hat{\eta} )</th>
<th>maxlnL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>-0.029</td>
<td>0.949</td>
<td>0.045</td>
<td>45.377</td>
<td>0.182</td>
<td>-108531.00</td>
<td>217072.00</td>
<td>217119.00</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0043)</td>
<td>(0.0021)</td>
<td>(0.1520)</td>
<td>(0.0005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.009</td>
<td>0.984</td>
<td>0.016</td>
<td>48.251</td>
<td>0.155</td>
<td>-79939.19</td>
<td>159888.40</td>
<td>159933.30</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
<td>(0.0025)</td>
<td>(0.0015)</td>
<td>(0.4665)</td>
<td>(0.0009)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The in-sample predictive model for the IBM series based on BSt-ACD model is given by

\[
\hat{\sigma}_i = \exp \left( -0.041 + 0.950 \ln \hat{\sigma}_{i-1} + 0.029 \left[ \frac{X_{i-1}}{\hat{\sigma}_{i-1}} \right] \right),
\]

and for the JNJ series it is given by

\[
\hat{\sigma}_i = \exp \left( -0.016 + 0.979 \ln \hat{\sigma}_{i-1} + 0.011 \left[ \frac{X_{i-1}}{\hat{\sigma}_{i-1}} \right] \right).
\]
2.8.3 Goodness-of-fit

We examine the distribution of GCS residuals for all the considered models through QQ plots. In general, QQ plots in Figure 2.3 indicate that GCS residual of BSs and BSt-ACD models seem to follow exp(1) distribution, considering that the 95th and 99th percentiles of this distribution are around 3.0 and 4.6, respectively. Furthermore, the BSt-ACD model seems to provide a slightly better fit to the data than the BS-ACD model. However, based on the GCS residuals, the GG-ACD model does not fit the JNJ series well although it provides a fairly good fit to the IBM series.

![Figure 2.3: GCS plots for IBM (left) and JNJ (right) stocks](image)

2.8.4 Out-of-sample forecast evaluation

We use the first nine tenth of the JNJ series, “in-sample” observations 1 to 53,451, for estimating the model parameters and then employ the resulting predictive model to the rest of the data, “out-of-sample” observations 53,452 to 59,390, to form density
forecasts. The corresponding evaluation results are shown in Table 2.9. For the BS-ACD and BS\textsubscript{t}-ACD models, at 1\% significance level, the null hypothesis that the \( z \) series is \textit{i.i.d.} \( U(0,1) \) can not be rejected according to the KS, AD and LB tests. The KS and AD tests show that the \( z \) series is from the uniform distribution, \( U(0,1) \). The LB tests indicate no serial correlation in the associated \( z \) and \( z^2 \) series.

Overall, the BS\textsubscript{t}-ACD model yields best forecasts, and the BS-ACD model provides close values, but the GG-ACD model is the worst one with regard to out-of-sample forecast evaluation.

Table 2.9: \( P \)-values for out-of-sample tests

<table>
<thead>
<tr>
<th>Model</th>
<th>AD</th>
<th>KS</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB(^2)(10)</th>
<th>LB(^2)(15)</th>
<th>LB(^2)(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS\textsubscript{t}-ACD</td>
<td>0.037</td>
<td>0.056</td>
<td>0.495</td>
<td>0.774</td>
<td>0.720</td>
<td>0.755</td>
<td>0.815</td>
<td>0.743</td>
</tr>
<tr>
<td>BS-ACD</td>
<td>0.035</td>
<td>0.044</td>
<td>0.490</td>
<td>0.771</td>
<td>0.717</td>
<td>0.753</td>
<td>0.814</td>
<td>0.741</td>
</tr>
<tr>
<td>GG-ACD</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.531</td>
<td>0.788</td>
<td>0.742</td>
<td>0.762</td>
<td>0.786</td>
<td>0.725</td>
</tr>
</tbody>
</table>

\textbf{Note:} KS represents Kolmogorov-Smirnov test, AD represents Anderson-Darling test, LB\((l)\) represents Ljung-Box test for \( z \) series over \( l \) lags. LB\(^2\)(\( l \)) represents Ljung-Box test for \( z^2 \) series over \( l \) lags.

For point-wise forecasts, one can also use measures like mean absolute error (MAE), mean absolute percentage error (MAPE) or mean squared error (MSE) to compare models based on out-of-sample forecast accuracy. We assess the predictive performance of the models by evaluating out-of-sample density forecasts of the models since density forecasts are popular in both macroeconomics and finance.
Chapter 3

Bivariate BS-ACD model

3.1 Introduction

Kundu et al. (2010) derived the bivariate Birnbaum-Saunders (BVBS) distribution through a transformation of the bivariate normal distribution (see Kundu et al. (2013) and Vilca et al. (2014a,b) for generalizations of the BVBS distribution). Its marginals are univariate BS distributions and its conditional distributions can be expressed in terms of normal distribution. Moreover, it has a correlation parameter which indirectly represents the dependence between the two BS random variables.

Our goal now is to construct a bivariate autoregressive conditional duration model based on the bivariate BS distribution, which would then facilitate us to jointly analyze and measure the strength of dependence between two matched duration series, pairs of durations with the same starting time (see Simonsen (2007)).

The rest of this Chapter is organized as follows. In Section 3.2, we present the BVBS-ACD model and the corresponding log-likelihood function. We derive the first and second derivatives of $\ln L(\theta)$ in Sections 3.3 and 3.4. In Section 3.5, we discuss the
maximum likelihood estimation and associated inference for the model parameters. A Monte Carlo simulation study is carried out in Section 3.6 to examine the properties of the MLEs. In Section 3.7, we provide an application of the BVBS-ACD model to a trade duration data set.

3.2 Bivariate BS-ACD model and the log-likelihood function

By working with the joint density of \((X_{1i}, X_{2i})\) directly, we propose the following BVBS-ACD model:

\[
X_{1i} = \sigma_{1i} \epsilon_{1i}, \quad \ln \sigma_{1i} = \alpha_1 + \beta_1 \ln \sigma_{1,i-1} + \gamma_1 \left[ \frac{X_{1,i-1}}{\sigma_{1,i-1}} \right], \quad |\beta_1| < 1, \tag{3.1}
\]

\[
X_{2i} = \sigma_{2i} \epsilon_{2i}, \quad \ln \sigma_{2i} = \alpha_2 + \beta_2 \ln \sigma_{2,i-1} + \gamma_2 \left[ \frac{X_{2,i-1}}{\sigma_{2,i-1}} \right], \quad |\beta_2| < 1, \tag{3.2}
\]

where \((X_{1i}, X_{2i}) \sim BVBS(\kappa_1, \sigma_{1i}, \kappa_2, \sigma_{2i}, \rho), \epsilon_{1i} \iid BS(\kappa_1, 1)\) and \(\epsilon_{2i} \iid BS(\kappa_2, 1)\).

The associated joint PDF of the BVBS-ACD model is given by

\[
f_{X_{1i},X_{2i}}(x_{1i}, x_{2i}; \theta) = \frac{1}{2\kappa_1 \sigma_{1i}} \left[ \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] \frac{1}{2\kappa_2 \sigma_{2i}} \left[ \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right]
\]

\[
\times \phi_2 \left[ \frac{1}{\kappa_1} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{\sigma_{1i}}{x_{1i}}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right); \rho \right]
\]

\[
= \frac{1}{2\kappa_1 x_{1i}} \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{3}{2}} + \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{3}{2}} \right] \frac{1}{2\kappa_2 x_{2i}} \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{3}{2}} \right] \tag{3.3}
\]
\begin{align*}
\times \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\kappa_1} \left( \frac{x_{1i}}{\sigma_{1i}} + \frac{\sigma_{1i}}{x_{1i}} - 2 \right) + \frac{1}{\kappa_1} \left( \frac{x_{2i}}{\sigma_{2i}} + \frac{\sigma_{2i}}{x_{2i}} - 2 \right) - \frac{2\rho}{\kappa_1\kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{\sigma_{1i}}{x_{1i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \right] \right\};
\end{align*}

here, \( \theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho)' \) is the model parameter. \( \phi_2 \) denotes the standard bivariate normal density function.

For \( i = 1, 2, \ldots, n \), the individual log-likelihood function can be expressed as

\begin{align*}
\ln l_i(\theta) &= -\ln(\kappa_1) - \ln(\kappa_2) - \frac{1}{2} \ln(1 - \rho^2) \\nonumber \\
&\quad + \ln \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{1}{2}} \right] - \frac{1}{2(1-\rho^2)\kappa_1} \left( \frac{x_{1i}}{\sigma_{1i}} + \frac{\sigma_{1i}}{x_{1i}} - 2 \right) \\nonumber \\
&\quad + \ln \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{1}{2}} \right] - \frac{1}{2(1-\rho^2)\kappa_2} \left( \frac{x_{2i}}{\sigma_{2i}} + \frac{\sigma_{2i}}{x_{2i}} - 2 \right) \\nonumber \\
&\quad + \frac{\rho}{(1-\rho^2)\kappa_1\kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{\sigma_{1i}}{x_{1i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right). \quad (3.4) \end{align*}

The log-likelihood function, without additive constant, is then given by

\begin{align*}
\ln L(\theta) &= \sum_{i=1}^{n} \ln l_i(\theta) = \sum_{i=1}^{n} \left\{ -\ln(\kappa_1) - \ln(\kappa_2) - \frac{1}{2} \ln(1 - \rho^2) \right. \\nonumber \\
&\quad + \ln \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{1}{2}} \right] - \frac{1}{2(1-\rho^2)\kappa_1} \left( \frac{x_{1i}}{\sigma_{1i}} + \frac{\sigma_{1i}}{x_{1i}} - 2 \right) \\nonumber \\
&\quad + \ln \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{1}{2}} \right] - \frac{1}{2(1-\rho^2)\kappa_2} \left( \frac{x_{2i}}{\sigma_{2i}} + \frac{\sigma_{2i}}{x_{2i}} - 2 \right) \\nonumber \\
&\quad + \left. \frac{\rho}{(1-\rho^2)\kappa_1\kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{\sigma_{1i}}{x_{1i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \right\}. \quad (3.5) \end{align*}
3.3 The first derivatives of $\ln L(\theta)$

The first derivatives of $\ln L(\theta)$ with respect to $\alpha_j$, $\beta_j$ and $\gamma_j$ are given by

\[
\begin{align*}
\frac{\partial \ln L}{\partial \alpha_j} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \alpha_j}, \\
\frac{\partial \ln L}{\partial \beta_j} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \beta_j}, \\
\frac{\partial \ln L}{\partial \gamma_j} &= \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \gamma_j},
\end{align*}
\]

where

\[
\begin{align*}
\frac{\partial \ln l_i}{\partial \sigma_{ji}} &= -\frac{x_{ji} - y_{ji}}{2\sigma_{ji}(x_{ji} + y_{ji})} + \frac{(x_{ji} - y_{ji})(x_{ji} + y_{ji})}{2x_{ji}y_{ji}^2(j - \rho^2)\kappa_j^2} \\
&\quad - \frac{\rho}{2\sigma_{ji}(j - \rho^2)\kappa_j\kappa_{ji'}} \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{y_{ji}}{y_{ji}}} \right) \left( \sqrt{\frac{x_{ji'}}{\sigma_{ji'}}} - \sqrt{\frac{y_{ji'}}{y_{ji'}}} \right),
\end{align*}
\]

here $j = 1, 2$, $j' = \{1, 2\} - \{j\}$ and

\[
\begin{align*}
\frac{\partial \sigma_{ji}}{\partial \alpha_j} &= \sigma_{ji} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \right), i \geq 1, \\
\frac{\partial \sigma_{ji}}{\partial \beta_j} &= \sigma_{ji} \left( \ln \sigma_{ji-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \right), i \geq 2, \\
\frac{\partial \sigma_{ji}}{\partial \gamma_j} &= \sigma_{ji} \left( \frac{x_{ji-1}}{\sigma_{ji-1}} + \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \right), i \geq 2,
\end{align*}
\]

with $\nabla \sigma_{j0} = (0, 0, 0)'$ and $\nabla \sigma_{j1} = (\sigma_{j1}, \sigma_{j1} \ln \sigma_{j0}, \sigma_{j1}x_{j0}/\sigma_{j0})'$.  

58
The first derivatives of $\ln L(\theta)$ with respect to $\kappa_j$, $j = 1, 2$ and $\rho$ are given by

$$\frac{\partial \ln L}{\partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_j},$$

$$\frac{\partial \ln L}{\partial \rho} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \rho},$$

where

$$\frac{\partial \ln l_i}{\partial \kappa_j} = -\frac{1}{\kappa_j} + \frac{1}{(1-\rho^2)\kappa_j^2} \left( \frac{x_{ji}}{\sigma_{ji}} + \frac{\sigma_{ji}}{x_{ji}} - 2 \right),$$

$$\frac{\partial \ln l_i}{\partial \rho} = \frac{\rho}{1-\rho^2} \left( \frac{x_{1i}}{\sigma_{1i}} - \frac{\sigma_{1i}}{x_{1i}} \right) \left( \frac{x_{2i}}{\sigma_{2i}} - \frac{\sigma_{2i}}{x_{2i}} \right),$$

(3.10)

$$\frac{\partial \ln l_i}{\partial \rho} = \frac{\rho}{1-\rho^2} \left( \frac{x_{1i}}{\sigma_{1i}} + \frac{\sigma_{1i}}{x_{1i}} - 2 \right) + \frac{1}{\kappa_1^2} \left( \frac{x_{2i}}{\sigma_{2i}} + \frac{\sigma_{2i}}{x_{2i}} - 2 \right)$$

$$+ \frac{1 + \rho^2}{(1-\rho^2)^2\kappa_1\kappa_2} \left( \frac{x_{1i}}{\sigma_{1i}} - \frac{\sigma_{1i}}{x_{1i}} \right) \left( \frac{x_{2i}}{\sigma_{2i}} - \frac{\sigma_{2i}}{x_{2i}} \right).$$

(3.11)

### 3.4 The second derivatives of $\ln L(\theta)$

The second derivatives of $\ln L(\theta)$ with respect to $\theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho)'$ are given by
\[
\frac{\partial^2 \ln L}{\partial \alpha_j^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \alpha_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_j} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_j^2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \beta_j^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \beta_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \beta_j} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \beta_j^2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \gamma_j^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \gamma_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \gamma_j} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \gamma_j^2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_j} \frac{\partial \sigma_{ji}}{\partial \beta_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \beta_j} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_j} \frac{\partial \sigma_{ji}}{\partial \gamma_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \beta_j} \frac{\partial \sigma_{ji}}{\partial \gamma_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \beta_j \partial \gamma_j} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_1 \partial \alpha_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_1 \partial \alpha_2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{1i}}{\partial \alpha_1 \partial \alpha_2} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{1i}^2} \left( \frac{\partial \sigma_{1i}}{\partial \alpha_1} \frac{\partial \sigma_{1i}}{\partial \alpha_2} \right) + \frac{\partial \ln l_i}{\partial \sigma_{1i}} \frac{\partial^2 \sigma_{1i}}{\partial \alpha_1 \partial \alpha_2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_1 \partial \beta_2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{1i}}{\partial \beta_1 \partial \beta_2} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{1i}^2} \left( \frac{\partial \sigma_{1i}}{\partial \beta_1} \frac{\partial \sigma_{1i}}{\partial \beta_2} \right) + \frac{\partial \ln l_i}{\partial \sigma_{1i}} \frac{\partial^2 \sigma_{1i}}{\partial \beta_1 \partial \beta_2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \gamma_1 \partial \gamma_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_1 \partial \gamma_2} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{1i}}{\partial \gamma_1 \partial \gamma_2} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{1i}^2} \left( \frac{\partial \sigma_{1i}}{\partial \gamma_1} \frac{\partial \sigma_{1i}}{\partial \gamma_2} \right) + \frac{\partial \ln l_i}{\partial \sigma_{1i}} \frac{\partial^2 \sigma_{1i}}{\partial \gamma_1 \partial \gamma_2} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \alpha_j \partial \beta_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_j} \frac{\partial \sigma_{ji}}{\partial \beta_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \beta_j} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_j} \frac{\partial \sigma_{ji}}{\partial \gamma_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} \right],
\]
\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i \partial \sigma_{ji}}{\partial \beta_j \partial \gamma_j} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \beta_j} \frac{\partial \sigma_{ji}}{\partial \gamma_j} \right) + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \beta_j \partial \gamma_j} \right],
\]
\[ \frac{\partial^2 \ln L}{\partial \alpha_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_j} \frac{\partial \sigma_j}{\partial \alpha_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \beta_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_j} \frac{\partial \sigma_j}{\partial \beta_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \gamma_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_j} \frac{\partial \sigma_j}{\partial \gamma_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \alpha_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_{j'}} \frac{\partial \sigma_j}{\partial \alpha_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \beta_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_{j'}} \frac{\partial \sigma_j}{\partial \beta_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \gamma_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_{j'}} \frac{\partial \sigma_j}{\partial \gamma_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \alpha_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \rho} \frac{\partial \sigma_j}{\partial \alpha_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \beta_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \rho} \frac{\partial \sigma_j}{\partial \beta_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \gamma_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \rho} \frac{\partial \sigma_j}{\partial \gamma_j}, \]

\[ \frac{\partial^2 \ln L}{\partial \kappa_j^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_j^2}, \]

\[ \frac{\partial^2 \ln L}{\partial \kappa_1 \partial \kappa_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_1 \partial \kappa_2}, \]

\[ \frac{\partial^2 \ln L}{\partial \kappa_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \rho}, \]

\[ \frac{\partial^2 \ln L}{\partial \rho^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \rho^2}, \]
where \( j = 1, 2 \), \( j' = \{1, 2\} - \{j\} \) and

\[
\frac{\partial^2 \ln l_i}{\partial \sigma^2_{ji}} = \rho \left( \frac{3 \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}}}{4 \kappa_1 \kappa_2 (1 - \rho^2)} \right) \left( \frac{x_{ji}}{\sigma_{ji}} - \sqrt{x_{ji} \sigma_{ji}^2} \right) + \frac{x_{ji}^2 + 2 \sigma_{ji} x_{ji} - \sigma_{ji}^2}{2 \sigma_{ji}^2 (x_{ji} + \sigma_{ji})^2} + \frac{x_{ji}}{\kappa_j^2 (\rho^2 - 1)} \sigma_{ji}^3, 
\]

(3.12)

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{1i} \partial \sigma_{2i}} = \rho \left( \frac{\sqrt{\frac{x_{11}}{\sigma_{11}}} + \sqrt{\frac{\sigma_{11}}{x_{11}}}}{4 \kappa_1 \kappa_2 (1 - \rho^2)} \right) \left( \frac{\sqrt{\frac{x_{21}}{\sigma_{21}}} + \sqrt{\frac{\sigma_{21}}{x_{21}}}}{\sigma_{11} \sigma_{2i}} \right),
\]

(3.13)

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_j} = \rho \left( \frac{\sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}}}{2 \kappa_j^2 \kappa_j' (1 - \rho^2)} \right) \left( \frac{x_{ji} \sigma_{ji}^2}{\sqrt{x_{ji} \sigma_{ji}^2}} \right) \frac{(x_{ji} - \sigma_{ji}) (x_{ji} + \sigma_{ji})}{\kappa_j^3 (1 - \rho^2) \sigma_{ji}^2 x_{ji}},
\]

(3.14)

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_{j'}} = \rho \left( \frac{\sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}}}{2 \kappa_j \kappa_{j'}^2 (1 - \rho^2) \sigma_{ji}} \right),
\]

(3.15)
\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} = \frac{\rho (x_{ji} - \sigma_{ji}) (x_{ji} + \sigma_{ji})}{\kappa_j^2 (1 - \rho^2)^2 \sigma_j^2 x_{ji}} - \frac{(\rho^2 + 1) \left( \sqrt{\frac{x_{j1}}{\sigma_{j1}}} + \sqrt{\frac{\sigma_{j1}}{x_{j1}}} \right) \left( \sqrt{\frac{x_{j2}}{\sigma_{j2}}} - \sqrt{\frac{\sigma_{j2}}{x_{j2}}} \right)}{2 \kappa_1 \kappa_2 (\rho - 1)^2 (\rho + 1)^2 \sigma_j^2},
\]

(3.16)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa_j^2} = \frac{2 \rho \left( \sqrt{\frac{x_{j1}}{\sigma_{j1}}} - \sqrt{\frac{\sigma_{j1}}{x_{j1}}} \right) \left( \sqrt{\frac{x_{j2}}{\sigma_{j2}}} - \sqrt{\frac{\sigma_{j2}}{x_{j2}}} \right)}{\kappa_j^3 \kappa_j' \left( 1 - \rho^2 \right)} - \frac{3 \left( \frac{x_{j1}}{\sigma_{j1}} + \frac{\sigma_{j1}}{x_{j1}} - 2 \right)}{\kappa_j^4 \left( 1 - \rho^2 \right)} + \frac{1}{\kappa_j^2},
\]

(3.17)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa_1 \partial \kappa_2} = \rho \left( \sqrt{\frac{x_{11}}{\sigma_{11}}} - \sqrt{\frac{\sigma_{11}}{x_{11}}} \right) \left( \sqrt{\frac{x_{21}}{\sigma_{21}}} - \sqrt{\frac{\sigma_{21}}{x_{21}}} \right),
\]

(3.18)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \rho} = \frac{2 \rho \left( \frac{x_{j1}}{\sigma_{j1}} + \frac{\sigma_{j1}}{x_{j1}} - 2 \right)}{\kappa_j^3 \left( 1 - \rho^2 \right)^2} - \frac{(\rho^2 + 1) \left( \sqrt{\frac{x_{j1}}{\sigma_{j1}}} - \sqrt{\frac{\sigma_{j1}}{x_{j1}}} \right) \left( \sqrt{\frac{x_{j2}}{\sigma_{j2}}} - \sqrt{\frac{\sigma_{j2}}{x_{j2}}} \right)}{\kappa_j^2 \kappa_j' (\rho - 1)^2 (\rho + 1)^2}.
\]

(3.19)
\[
\frac{\partial^2 \ln l_i}{\partial \rho^2} = \frac{2 \rho (\rho^2 + 3) \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)}{\kappa_1 \kappa_2 (1 - \rho^2)^3} (3 \rho^2 + 1) \left( \frac{x_{2i} x_{2j} - x_{2i} x_{2j}}{\kappa^2_2} + \frac{x_{1i} x_{1j} - x_{1i} x_{1j}}{\kappa^2_1} \right) + \frac{\rho^2 + 1}{(1 - \rho^2)^2},
\]

(3.20)

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha^2_j} = \frac{\partial \sigma_{ji}}{\partial \alpha_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j x_{jl} \sigma_{jl} \right) \right) + \sigma_{ji} \gamma_i \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j x_{jl} \sigma_{jl} \right) \sum_{l=k}^{i-1} \frac{x_{l} \partial \sigma_{jl}}{\partial \gamma_j} \frac{\beta_j - \gamma_j x_{jl} \sigma_{jl}}{\sigma_{jl}},
\]

\[i \geq 1, \quad (3.21)\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \beta^2_j} = \frac{\partial \sigma_{ji}}{\partial \beta_j} \left\{ \ln \sigma_{j, i-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} \sigma_{jl} \right) \right\}
+ \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \left( \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} \sigma_{jl} \right) \right) \sum_{l=k+1}^{i-1} \frac{\gamma_j x_{jl} \partial \sigma_{jl}}{\sigma_{jl}} + 1 \right\} + \frac{\partial \sigma_{jk}}{\partial \beta_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} \sigma_{jl} \right) \frac{\partial \sigma_{ji}}{\partial \beta_j} \frac{\sigma_{jk}}{\sigma_{j, i-1}},
\]

\[i \geq 2, \quad (3.22)\]
\[
\frac{\partial^2 \sigma_{ji}}{\partial \gamma_j^2} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left\{ \frac{x_{ji}}{\sigma_{ji}} + \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right\} \\
+ \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma_j}}{\sigma_{jl}} + \frac{\partial \sigma_{jk}}{\partial \gamma_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right\}, \\
i \geq 2, \quad (3.23)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \beta_j} = \frac{\partial \sigma_{ji}}{\partial \beta_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_{ji} \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma_j}}{\sigma_{jl}} + 1, \\
i \geq 1, \quad (3.24)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_{ji} \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma_j}}{\sigma_{jl}} - \frac{x_{jl}}{\sigma_{jl}}, \\
i \geq 1, \quad (3.25)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \beta_j \partial \gamma_j} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left\{ \ln \sigma_{ji,i-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right\} \\
+ \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma_j}}{\sigma_{jl}} + \frac{\partial \sigma_{jk}}{\partial \gamma_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \frac{x_{jl}}{\sigma_{jl}} \right) \right\}, \\
i \geq 2. \quad (3.26)
\]
3.5 Estimation and inference

We estimate the parameter \( \mathbf{\theta} = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho)' \) of the BVBS-ACD model by maximizing the log-likelihood function (without the additive constant) in equation (3.5). We do the estimation by using a hybrid of optimization algorithms, namely, Nelder-Meade (NM) followed by Broyden-Fletcher-Goldfarb-Shanno (BFGS) (see Bhatti (2010) and Leiva et al. (2014) for recent details). First, we apply the NM algorithm to estimate the ACD parameters \( \alpha_j, \beta_j, \) and \( \gamma_j, j = 1, 2 \) by fixing \( \kappa_1, \kappa_2 \) and \( \rho \) at their initial values. Next, we use BFGS algorithm to estimate over the entire parameter space. We have derived and implemented the analytical gradient (the first derivatives in Section 3.3) in the second step. The BFGS quasi-Newton method with the analytic gradients will be faster, more stable and lead to more accurate estimates than a numerical gradient method (see Bard (1974), Bolker (2008) and Mayorov (2011)).

The standard errors (SEs) of the MLEs of the model parameters can then be calculated as the square root of the diagonal elements of the negative of the inverse Hessian matrix, which is an estimator of the asymptotic covariance matrix.

Leiva et al. (2014) discussed the Wald test to check the statistical significance of individual model parameters. Let \( \theta \) denote a parameter of the BVBS-ACD model. The Wald statistic \( W = \frac{\hat{\theta} - \theta_0}{SE(\hat{\theta})} \) can be used to test the hypotheses \( H_0 : \theta = \theta_0 \) versus \( H_1 : \theta \neq \theta_0 \). Then, under \( H_0 \), \( W \) is asymptotically \( \mathcal{N}(0, 1) \), under the usual regularity conditions.
3.6 Simulation study

We examine the properties of the MLEs of the parameters of the BVBS-ACD model through a Monte Carlo simulation study. We use 1000 simulated samples of sample sizes $n = 500, 1000, 3000$ and 5000 from the BVBS-ACD model with the vector of true parameters

$$\theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho)' = (0.1, 0.9, 0.1, -0.1, 0.95, 0.1, 1.1, 1.2, \rho)' .$$

Here, we choose $\rho = 0, 0.25, 0.5, 0.95$. To estimate the BVBS-ACD model, we set the starting values for the ACD parameters

$$(\alpha_{10}, \beta_{10}, \gamma_{10}, \alpha_{20}, \beta_{20}, \gamma_{20})' = (0.01, 0.70, 0.01, -0.01, 0.8, 0.01)' .$$

$(\kappa_{10}, \kappa_{20}, \rho_0)$ then have well-defined initial values obtained by

$$\kappa_{j0} = \sqrt{2 \left( \frac{\bar{X}_j}{\text{med}(X_j)} - 1 \right)}, \quad j = 1, 2$$

(see (1.14)) and

$$\rho_0 = \text{Corr} \left( \frac{1}{\kappa_{10}} \left( \sqrt{\frac{X_1}{\hat{\sigma}_1}} - \sqrt{\frac{\hat{\sigma}_1}{X_1}} \right) , \frac{1}{\kappa_{20}} \left( \sqrt{\frac{X_2}{\hat{\sigma}_2}} - \sqrt{\frac{\hat{\sigma}_2}{X_2}} \right) \right) ,$$

where $\hat{\sigma}_j$ is estimated by

$$\hat{\sigma}_j = \exp \left( \alpha_{j0} + \beta_{j0} \ln \hat{\sigma}_{j,i-1} + \gamma_{j0} \left[ \frac{X_{j,i-1}}{\hat{\sigma}_{j,i-1}} \right] \right) ,$$
For each setting of the parameter values and each sample size, we calculate the mean, coefficients of skewness and kurtosis, bias and root mean squared error (RMSE) of the MLEs over 1000 replications. These results are presented in Tables 3.1-3.4. For different correlation coefficients, all the estimates possess low bias and RMSE which decrease with increased sample size, tending towards 0. This empirically displays the asymptotic unbiasedness, consistency and efficiency of the MLEs. Moreover, as $n$ increases, the empirical distributions of all the estimators become close to normal distribution, which may be seen from the values of skewness and kurtosis, for example.

### 3.7 Application to trade duration data

In this section, we present the application of BVBS-ACD model to a high-frequency trade duration data.

#### 3.7.1 Data description

We consider the trade durations of Johnson & Johnson (JNJ) and Procter & Gamble Company (PG) stocks in the period of 30 consecutive trading days from January 2, 2002 to February 13, 2002. There are 65,127 (JNJ) and 57,335 (PG) observations for each asset. The original data sets were studied in detail by Bhatti (2010) and Leiva et al. (2014).

After restricting two adjusted duration series for JNJ and PG stocks between 10:00 am and 4:00 pm, we obtain two matched series with 6,222 pairs of durations with the same starting time.
Table 3.1: Simulation results for BVBS-ACD model when $\rho = 0.95$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Statistics</th>
<th>$n$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>1000</td>
<td>3000</td>
</tr>
<tr>
<td>$\alpha_1 = 0.1$</td>
<td>Mean</td>
<td>0.0979</td>
<td>0.0988</td>
<td>0.0995</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.2000</td>
<td>0.3171</td>
<td>0.1825</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>6.4886</td>
<td>3.1753</td>
<td>2.9457</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0021</td>
<td>-0.0012</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0498</td>
<td>0.0329</td>
<td>0.0189</td>
</tr>
<tr>
<td>$\beta_1 = 0.9$</td>
<td>Mean</td>
<td>0.8923</td>
<td>0.8961</td>
<td>0.8986</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-0.3145</td>
<td>-0.3093</td>
<td>-0.2455</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.0404</td>
<td>3.0194</td>
<td>3.0778</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0077</td>
<td>-0.0039</td>
<td>-0.0014</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0206</td>
<td>0.0137</td>
<td>0.0077</td>
</tr>
<tr>
<td>$\gamma_1 = 0.1$</td>
<td>Mean</td>
<td>0.1112</td>
<td>0.1057</td>
<td>0.1020</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.0839</td>
<td>0.2243</td>
<td>0.0684</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.0643</td>
<td>3.0290</td>
<td>2.9507</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0112</td>
<td>0.0057</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0173</td>
<td>0.0107</td>
<td>0.0054</td>
</tr>
<tr>
<td>$\alpha_2 = -0.1$</td>
<td>Mean</td>
<td>-0.1065</td>
<td>-0.1036</td>
<td>-0.1011</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-0.0990</td>
<td>0.1592</td>
<td>0.0814</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>6.2791</td>
<td>3.112</td>
<td>3.1276</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0065</td>
<td>-0.0036</td>
<td>-0.0011</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0218</td>
<td>0.0139</td>
<td>0.0077</td>
</tr>
<tr>
<td>$\beta_2 = 0.95$</td>
<td>Mean</td>
<td>0.9456</td>
<td>0.9481</td>
<td>0.9492</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-0.5418</td>
<td>-0.5210</td>
<td>-0.2765</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.5363</td>
<td>3.6419</td>
<td>3.0362</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0044</td>
<td>-0.0019</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0117</td>
<td>0.0073</td>
<td>0.0040</td>
</tr>
<tr>
<td>$\gamma_2 = 0.1$</td>
<td>Mean</td>
<td>0.1058</td>
<td>0.1030</td>
<td>0.1009</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.1677</td>
<td>0.2066</td>
<td>0.1627</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.3693</td>
<td>3.2666</td>
<td>3.1211</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0058</td>
<td>0.0030</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0127</td>
<td>0.0082</td>
<td>0.0045</td>
</tr>
<tr>
<td>$\kappa_1 = 1.1$</td>
<td>Mean</td>
<td>1.1270</td>
<td>1.1151</td>
<td>1.1048</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.4710</td>
<td>0.2482</td>
<td>0.0683</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.0018</td>
<td>3.3397</td>
<td>3.0630</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0270</td>
<td>0.0151</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0514</td>
<td>0.0318</td>
<td>0.0154</td>
</tr>
<tr>
<td>$\kappa_2 = 1.2$</td>
<td>Mean</td>
<td>1.2059</td>
<td>1.2037</td>
<td>1.2007</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.2258</td>
<td>0.2272</td>
<td>0.1380</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.7508</td>
<td>3.3847</td>
<td>3.0840</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0059</td>
<td>0.0037</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0419</td>
<td>0.0283</td>
<td>0.0159</td>
</tr>
<tr>
<td>$\rho = 0.95$</td>
<td>Mean</td>
<td>0.9460</td>
<td>0.9479</td>
<td>0.9493</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-0.4434</td>
<td>-0.3525</td>
<td>-0.1592</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.3355</td>
<td>3.4547</td>
<td>2.9353</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0040</td>
<td>-0.0021</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0067</td>
<td>0.0041</td>
<td>0.0020</td>
</tr>
</tbody>
</table>
Table 3.2: Simulation results for BVBS-ACD model when $\rho = 0.5$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Statistics</th>
<th>$n$</th>
<th>500</th>
<th>1000</th>
<th>3000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1 = 0.1$</td>
<td>Mean</td>
<td></td>
<td>0.1138</td>
<td>0.1077</td>
<td>0.1025</td>
<td>0.1009</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>1.1298</td>
<td>0.6483</td>
<td>0.3646</td>
<td>0.2623</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>5.4679</td>
<td>3.3928</td>
<td>3.0443</td>
<td>3.0993</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0138</td>
<td>0.0077</td>
<td>0.0025</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0836</td>
<td>0.0552</td>
<td>0.0309</td>
<td>0.0232</td>
</tr>
<tr>
<td>$\beta_1 = 0.9$</td>
<td>Mean</td>
<td></td>
<td>0.8880</td>
<td>0.8936</td>
<td>0.8978</td>
<td>0.8989</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>-0.8633</td>
<td>-0.5516</td>
<td>-0.4102</td>
<td>-0.2501</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>4.5459</td>
<td>3.2770</td>
<td>3.1868</td>
<td>2.9495</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>-0.0120</td>
<td>-0.0064</td>
<td>-0.0022</td>
<td>-0.0011</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0359</td>
<td>0.0237</td>
<td>0.0131</td>
<td>0.0097</td>
</tr>
<tr>
<td>$\gamma_1 = 0.1$</td>
<td>Mean</td>
<td></td>
<td>0.1090</td>
<td>0.1047</td>
<td>0.1017</td>
<td>0.1011</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.1115</td>
<td>0.2722</td>
<td>0.0996</td>
<td>-0.0113</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>3.0447</td>
<td>3.1798</td>
<td>2.9707</td>
<td>2.8346</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0090</td>
<td>0.0047</td>
<td>0.0017</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0199</td>
<td>0.0130</td>
<td>0.0070</td>
<td>0.0053</td>
</tr>
<tr>
<td>$\alpha_2 = -0.1$</td>
<td>Mean</td>
<td></td>
<td>-0.0966</td>
<td>-0.0989</td>
<td>-0.0991</td>
<td>-0.0996</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.5843</td>
<td>0.3731</td>
<td>0.2382</td>
<td>0.1851</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>4.1691</td>
<td>3.6212</td>
<td>3.0768</td>
<td>3.2444</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0034</td>
<td>0.0011</td>
<td>0.0009</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0278</td>
<td>0.0189</td>
<td>0.0106</td>
<td>0.0078</td>
</tr>
<tr>
<td>$\beta_2 = 0.95$</td>
<td>Mean</td>
<td></td>
<td>0.9436</td>
<td>0.9472</td>
<td>0.9487</td>
<td>0.9493</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>-1.0474</td>
<td>-0.6883</td>
<td>-0.4069</td>
<td>-0.3821</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>5.6827</td>
<td>3.8742</td>
<td>3.1820</td>
<td>3.3638</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>-0.0064</td>
<td>-0.0028</td>
<td>-0.0013</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0195</td>
<td>0.0124</td>
<td>0.0068</td>
<td>0.0050</td>
</tr>
<tr>
<td>$\gamma_2 = 0.1$</td>
<td>Mean</td>
<td></td>
<td>0.1027</td>
<td>0.1014</td>
<td>0.1004</td>
<td>0.1003</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.2234</td>
<td>0.2921</td>
<td>0.2723</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>3.2393</td>
<td>3.4492</td>
<td>3.1149</td>
<td>2.9713</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0027</td>
<td>0.0014</td>
<td>0.0004</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0149</td>
<td>0.0099</td>
<td>0.0056</td>
<td>0.0041</td>
</tr>
<tr>
<td>$\kappa_1 = 1.1$</td>
<td>Mean</td>
<td></td>
<td>1.1257</td>
<td>1.1144</td>
<td>1.1049</td>
<td>1.1033</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.3089</td>
<td>0.1493</td>
<td>0.0360</td>
<td>-0.0686</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>3.3437</td>
<td>3.1565</td>
<td>2.9667</td>
<td>3.0989</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0257</td>
<td>0.0144</td>
<td>0.0049</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0503</td>
<td>0.0313</td>
<td>0.0153</td>
<td>0.0115</td>
</tr>
<tr>
<td>$\kappa_2 = 1.2$</td>
<td>Mean</td>
<td></td>
<td>1.2042</td>
<td>1.2027</td>
<td>1.2002</td>
<td>1.2003</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.0982</td>
<td>0.2395</td>
<td>0.0876</td>
<td>0.1018</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>3.3152</td>
<td>3.4159</td>
<td>3.0412</td>
<td>2.9583</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0042</td>
<td>0.0027</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0411</td>
<td>0.0284</td>
<td>0.0164</td>
<td>0.0121</td>
</tr>
<tr>
<td>$\rho = 0.5$</td>
<td>Mean</td>
<td></td>
<td>0.5089</td>
<td>0.5054</td>
<td>0.5014</td>
<td>0.5009</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td></td>
<td>0.0150</td>
<td>-0.0573</td>
<td>-0.0271</td>
<td>-0.1053</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td></td>
<td>2.9122</td>
<td>3.0392</td>
<td>2.7894</td>
<td>2.8607</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td></td>
<td>0.0089</td>
<td>0.0054</td>
<td>0.0014</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td></td>
<td>0.0357</td>
<td>0.0249</td>
<td>0.0133</td>
<td>0.0104</td>
</tr>
</tbody>
</table>

70
Table 3.3: Simulation results for BVBS-ACD model when $\rho = 0.25$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Statistics</th>
<th>$n$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>1000</td>
<td>3000</td>
<td>5000</td>
<td></td>
</tr>
<tr>
<td>$\alpha_1 = 0.1$</td>
<td>Mean</td>
<td>0.1128</td>
<td>0.1074</td>
<td>0.1025</td>
<td>0.1011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>1.3299</td>
<td>0.7418</td>
<td>0.4268</td>
<td>0.2879</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>6.4049</td>
<td>3.5942</td>
<td>3.1606</td>
<td>3.1271</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0128</td>
<td>0.0074</td>
<td>0.0025</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0928</td>
<td>0.0606</td>
<td>0.0343</td>
<td>0.0255</td>
<td></td>
</tr>
<tr>
<td>$\beta_1 = 0.9$</td>
<td>Mean</td>
<td>0.8877</td>
<td>0.8933</td>
<td>0.8977</td>
<td>0.8988</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-1.0251</td>
<td>-0.6212</td>
<td>-0.4544</td>
<td>-0.2828</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>5.1294</td>
<td>3.4056</td>
<td>3.2236</td>
<td>2.9648</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0123</td>
<td>-0.0067</td>
<td>-0.0023</td>
<td>-0.0012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0397</td>
<td>0.0259</td>
<td>0.0145</td>
<td>0.0106</td>
<td></td>
</tr>
<tr>
<td>$\gamma_1 = 0.1$</td>
<td>Mean</td>
<td>0.1102</td>
<td>0.1054</td>
<td>0.1019</td>
<td>0.1012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.1325</td>
<td>0.2836</td>
<td>0.1095</td>
<td>0.0027</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.0767</td>
<td>3.1226</td>
<td>2.9154</td>
<td>2.8908</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0102</td>
<td>0.0054</td>
<td>0.0019</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0220</td>
<td>0.0143</td>
<td>0.0077</td>
<td>0.0057</td>
<td></td>
</tr>
<tr>
<td>$\alpha_2 = -0.1$</td>
<td>Mean</td>
<td>-0.0966</td>
<td>-0.0990</td>
<td>-0.0991</td>
<td>-0.0997</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.6223</td>
<td>0.3911</td>
<td>0.2302</td>
<td>0.1596</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.1643</td>
<td>3.6524</td>
<td>3.0860</td>
<td>3.2282</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0034</td>
<td>0.0010</td>
<td>0.0009</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0310</td>
<td>0.0210</td>
<td>0.0117</td>
<td>0.0085</td>
<td></td>
</tr>
<tr>
<td>$\beta_2 = 0.95$</td>
<td>Mean</td>
<td>0.9428</td>
<td>0.9470</td>
<td>0.9486</td>
<td>0.9492</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-1.2030</td>
<td>-0.6811</td>
<td>-0.4479</td>
<td>-0.3517</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>6.3884</td>
<td>3.7265</td>
<td>3.2563</td>
<td>3.2519</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0072</td>
<td>-0.0030</td>
<td>-0.0014</td>
<td>-0.0008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0218</td>
<td>0.0136</td>
<td>0.0075</td>
<td>0.0054</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2 = 0.1$</td>
<td>Mean</td>
<td>0.1034</td>
<td>0.1017</td>
<td>0.1005</td>
<td>0.1004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.2714</td>
<td>0.2902</td>
<td>0.2840</td>
<td>-0.0007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.2682</td>
<td>3.3968</td>
<td>3.0376</td>
<td>2.8688</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0034</td>
<td>0.0017</td>
<td>0.0005</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0164</td>
<td>0.0108</td>
<td>0.0060</td>
<td>0.0045</td>
<td></td>
</tr>
<tr>
<td>$\kappa_1 = 1.1$</td>
<td>Mean</td>
<td>1.1253</td>
<td>1.1140</td>
<td>1.1049</td>
<td>1.1034</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.3513</td>
<td>0.1474</td>
<td>0.0373</td>
<td>-0.0764</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.3813</td>
<td>3.1881</td>
<td>2.9441</td>
<td>3.1389</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0253</td>
<td>0.0140</td>
<td>0.0049</td>
<td>0.0034</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0505</td>
<td>0.0313</td>
<td>0.0153</td>
<td>0.0114</td>
<td></td>
</tr>
<tr>
<td>$\kappa_2 = 1.2$</td>
<td>Mean</td>
<td>1.2040</td>
<td>1.2022</td>
<td>1.2001</td>
<td>1.2002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.2705</td>
<td>0.1860</td>
<td>0.0394</td>
<td>0.0589</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.3483</td>
<td>3.3121</td>
<td>2.9944</td>
<td>2.9714</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0040</td>
<td>0.0022</td>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0416</td>
<td>0.0282</td>
<td>0.0164</td>
<td>0.0121</td>
<td></td>
</tr>
<tr>
<td>$\rho = 0.25$</td>
<td>Mean</td>
<td>0.2656</td>
<td>0.2592</td>
<td>0.2526</td>
<td>0.2516</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.0799</td>
<td>-0.0045</td>
<td>-0.0007</td>
<td>-0.0876</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>2.9182</td>
<td>3.0210</td>
<td>2.7867</td>
<td>2.8551</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0156</td>
<td>0.0092</td>
<td>0.0026</td>
<td>0.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0462</td>
<td>0.0319</td>
<td>0.0167</td>
<td>0.0130</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.4: Simulation results for BVBS-ACD model when $\rho = 0$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Statistics</th>
<th>$n$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>500</td>
<td>1000</td>
<td>3000</td>
<td>5000</td>
</tr>
<tr>
<td>$\alpha_1 = 0.1$</td>
<td>Mean</td>
<td>0.1189</td>
<td>0.1072</td>
<td>0.1015</td>
<td>0.1005</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>1.9843</td>
<td>1.0356</td>
<td>0.5140</td>
<td>0.3007</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>12.4607</td>
<td>5.4310</td>
<td>3.3849</td>
<td>2.8156</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0189</td>
<td>0.0072</td>
<td>0.0015</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.1074</td>
<td>0.0651</td>
<td>0.0345</td>
<td>0.0254</td>
</tr>
<tr>
<td>$\beta_1 = 0.9$</td>
<td>Mean</td>
<td>0.8841</td>
<td>0.8930</td>
<td>0.8980</td>
<td>0.8991</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-1.7870</td>
<td>-0.8688</td>
<td>-0.4545</td>
<td>-0.1977</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>10.5058</td>
<td>4.9659</td>
<td>3.3610</td>
<td>2.7907</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0159</td>
<td>-0.0070</td>
<td>-0.0020</td>
<td>-0.0009</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0465</td>
<td>0.0279</td>
<td>0.0148</td>
<td>0.0108</td>
</tr>
<tr>
<td>$\gamma_1 = 0.1$</td>
<td>Mean</td>
<td>0.1123</td>
<td>0.1060</td>
<td>0.1021</td>
<td>0.1011</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.4415</td>
<td>0.1910</td>
<td>0.2933</td>
<td>0.1809</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.8432</td>
<td>3.5541</td>
<td>3.3634</td>
<td>3.2788</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0123</td>
<td>0.0060</td>
<td>0.0021</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0244</td>
<td>0.0152</td>
<td>0.0079</td>
<td>0.0060</td>
</tr>
<tr>
<td>$\alpha_2 = -0.1$</td>
<td>Mean</td>
<td>-0.0953</td>
<td>-0.0979</td>
<td>-0.0993</td>
<td>-0.0995</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.6719</td>
<td>0.3909</td>
<td>0.3022</td>
<td>0.0918</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>5.2592</td>
<td>3.3247</td>
<td>3.2052</td>
<td>2.8505</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0047</td>
<td>0.0021</td>
<td>0.0007</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0345</td>
<td>0.0218</td>
<td>0.0119</td>
<td>0.0088</td>
</tr>
<tr>
<td>$\beta_2 = 0.95$</td>
<td>Mean</td>
<td>0.9411</td>
<td>0.9456</td>
<td>0.9484</td>
<td>0.9492</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>-1.3671</td>
<td>-0.6772</td>
<td>-0.4254</td>
<td>-0.2484</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>6.8392</td>
<td>3.7528</td>
<td>3.4520</td>
<td>3.0809</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>-0.0089</td>
<td>-0.0044</td>
<td>-0.0016</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0241</td>
<td>0.0142</td>
<td>0.0077</td>
<td>0.0057</td>
</tr>
<tr>
<td>$\gamma_2 = 0.1$</td>
<td>Mean</td>
<td>0.1038</td>
<td>0.1022</td>
<td>0.1009</td>
<td>0.1004</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.1192</td>
<td>0.1136</td>
<td>-0.0729</td>
<td>-0.0438</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.3123</td>
<td>2.7931</td>
<td>2.9934</td>
<td>2.8415</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0038</td>
<td>0.0022</td>
<td>0.0009</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0169</td>
<td>0.0108</td>
<td>0.0060</td>
<td>0.0047</td>
</tr>
<tr>
<td>$\kappa_1 = 1.1$</td>
<td>Mean</td>
<td>1.1260</td>
<td>1.1123</td>
<td>1.1046</td>
<td>1.1030</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.2113</td>
<td>-0.0531</td>
<td>0.0977</td>
<td>0.2128</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.4990</td>
<td>3.2449</td>
<td>2.9873</td>
<td>2.8949</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0260</td>
<td>0.0123</td>
<td>0.0046</td>
<td>0.0030</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0509</td>
<td>0.0297</td>
<td>0.0154</td>
<td>0.0119</td>
</tr>
<tr>
<td>$\kappa_2 = 1.2$</td>
<td>Mean</td>
<td>1.2058</td>
<td>1.2043</td>
<td>1.2017</td>
<td>1.2010</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.3055</td>
<td>0.0370</td>
<td>0.0812</td>
<td>0.0353</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>4.3661</td>
<td>2.8992</td>
<td>2.8667</td>
<td>2.8285</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0058</td>
<td>0.0043</td>
<td>0.0017</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0434</td>
<td>0.0283</td>
<td>0.0156</td>
<td>0.0120</td>
</tr>
<tr>
<td>$\rho = 0$</td>
<td>Mean</td>
<td>0.0186</td>
<td>0.0102</td>
<td>0.0048</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.0696</td>
<td>0.0890</td>
<td>0.1807</td>
<td>-0.0745</td>
</tr>
<tr>
<td></td>
<td>Kurt</td>
<td>3.0213</td>
<td>3.0437</td>
<td>3.0644</td>
<td>3.1268</td>
</tr>
<tr>
<td></td>
<td>Bias</td>
<td>0.0186</td>
<td>0.0102</td>
<td>0.0048</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0493</td>
<td>0.0335</td>
<td>0.0187</td>
<td>0.0144</td>
</tr>
</tbody>
</table>
As seen in Figure 3.1 and Table 3.5, the two matched duration series are positive and right-skewed with high degree of kurtosis indicating heavy tails.

![Histograms](image)

**Figure 3.1: Histograms for matched duration series of JNJ (left) and PG (right)**

<table>
<thead>
<tr>
<th>Data</th>
<th>Min</th>
<th>Median</th>
<th>Mean</th>
<th>Max</th>
<th>SD</th>
<th>skew</th>
<th>kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>JNJ</td>
<td>0.131</td>
<td>0.973</td>
<td>1.499</td>
<td>18.248</td>
<td>1.594</td>
<td>2.886</td>
<td>15.699</td>
</tr>
<tr>
<td>PG</td>
<td>0.113</td>
<td>0.963</td>
<td>1.522</td>
<td>14.967</td>
<td>1.624</td>
<td>2.603</td>
<td>12.840</td>
</tr>
</tbody>
</table>

In Figure 3.2, the empirical scaled TTT transforms are presented first concave and then convex, revealing that both marginals may have unimodal hazard rates.

Hence, the BVBS distribution could be a good fit for the duration data due to the shape of its marginal density functions and hazard rates.

The ACF plots in Figure 3.3 indicate that there is a positive autocorrelation in both series, which suggests that a ACD specification is a reasonable choice, as well.
Figure 3.2: TTT plots for matched duration series of JNJ (left) and PG (right)

Figure 3.3: ACF plots for matched duration series of JNJ (left) and PG (right)
3.7.2 Estimation results

Table 3.6: Estimation results for the BVBS-ACD model.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\gamma}_2$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\kappa}_2$</th>
<th>$\hat{\rho}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.007</td>
<td>0.991</td>
<td>0.004</td>
<td>-0.010</td>
<td>0.991</td>
<td>0.006</td>
<td>1.024</td>
<td>1.077</td>
<td>0.041</td>
</tr>
<tr>
<td>SE</td>
<td>0.002</td>
<td>0.005</td>
<td>0.001</td>
<td>0.002</td>
<td>0.004</td>
<td>0.001</td>
<td>0.009</td>
<td>0.010</td>
<td>0.013</td>
</tr>
</tbody>
</table>

We employ the method described earlier in Section 3.5 to estimate the BVBS-ACD model with the matched duration data. The obtained results are shown in Table 3.6. All estimates are statistically significant at 1% level according to the Wald test. The estimated correlation coefficient is 0.041, which implies a weak positive correlation between the two matched duration series. The in-sample predictive model is given by

$$\hat{\sigma}_{1i} = \exp \left( -0.007 + 0.991 \ln \hat{\sigma}_{1,i-1} + 0.004 \left[ \frac{X_{1,i-1}}{\hat{\sigma}_{1,i-1}} \right] \right),$$

$$\hat{\sigma}_{2i} = \exp \left( -0.010 + 0.991 \ln \hat{\sigma}_{2,i-1} + 0.006 \left[ \frac{X_{2,i-1}}{\hat{\sigma}_{2,i-1}} \right] \right).$$

3.7.3 Model Comparison

To verify whether it is necessary to include the correlation parameter $\rho$ in the model, we conduct model comparison by using likelihood ratio test and Akaike information criterion (AIC).

We re-estimate the BVBS-ACD model by taking $\rho = 0$. Under $H_0: \rho = 0$, the likelihood ratio test statistic is defined by $2(\ln L_1 - \ln L_0)$. AIC is given by $-2\ln L + 2k$, where $\ln L$ is the maximized log-likelihood value of the model of interest and $k$ is the corresponding number of estimated parameters. As seen in Table 3.7, the $p$-value of
the likelihood ratio test is 0.001, which does not provide sufficient evidence towards the restricted model. The difference in AIC values between the two models is 8.502, which indicates that the AIC value of the restricted model is substantially larger than that of the unrestricted model. Both these findings suggest that the proposed BVBS-ACD model is a better choice than the restricted model in terms of overall model fit.

Table 3.7: LR test and AICs.

<table>
<thead>
<tr>
<th></th>
<th>Maximized log-likelihood</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restricted model</td>
<td>3098.554</td>
<td>-6181.108</td>
</tr>
<tr>
<td>Unrestricted model</td>
<td>3103.805</td>
<td>-6189.610</td>
</tr>
<tr>
<td>Difference</td>
<td>5.251</td>
<td>-8.502</td>
</tr>
<tr>
<td>p-value of LR test</td>
<td>0.001</td>
<td>—</td>
</tr>
</tbody>
</table>

3.7.4 Goodness-of-fit

To evaluate the goodness-of-fit of the BVBS-ACD model, we investigate the in-sample one-step-ahead density forecasts implied by the predictive model. We examine the uniformity of the $z$ series by Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests along with histogram plots, and also check the independence of the $z$ series by Ljung-Box (LB) test along with ACF plots. Results for these tests are shown in Table 3.8. At 1% significance level, the null hypothesis that the series $z_1$, $z_2$, $z_{1|2}$ and $z_{2|1}$ are all i.i.d. $U(0, 1)$ can not be rejected. The KS and AD tests, along with the histograms in Figures 3.4 and 3.5 seem to coincide with uniformity of the corresponding $z$ series. The LB tests, along with the ACF plots in Figures 3.6 and 3.7, show the absence of serial correlation in the associated $z$ and $z^2$ series.
Table 3.8: $P$-values for in-sample goodness of fit tests for the BVBS-ACD model based on density forecasts

<table>
<thead>
<tr>
<th>$z$ series</th>
<th>KS</th>
<th>AD</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB$^2$(10)</th>
<th>LB$^2$(15)</th>
<th>LB$^2$(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>0.012</td>
<td>0.017</td>
<td>0.119</td>
<td>0.177</td>
<td>0.195</td>
<td>0.435</td>
<td>0.616</td>
<td>0.730</td>
</tr>
<tr>
<td>$z_2$</td>
<td>0.032</td>
<td>0.037</td>
<td>0.682</td>
<td>0.747</td>
<td>0.480</td>
<td>0.421</td>
<td>0.561</td>
<td>0.422</td>
</tr>
<tr>
<td>$z_{1</td>
<td>2}$</td>
<td>0.011</td>
<td>0.019</td>
<td>0.111</td>
<td>0.174</td>
<td>0.185</td>
<td>0.404</td>
<td>0.597</td>
</tr>
<tr>
<td>$z_{2</td>
<td>1}$</td>
<td>0.025</td>
<td>0.033</td>
<td>0.701</td>
<td>0.782</td>
<td>0.481</td>
<td>0.431</td>
<td>0.590</td>
</tr>
</tbody>
</table>

Note: LB$(l)$ represents LB test for $z$ series over $l$ lags, where $l = 10, 15$ and $20$. LB$^2$(l) represents LB test for $z^2$ series over $l$ lags.

We also investigate the estimated GCS residuals obtained by (1.9). We examine the distribution and independence of the GCS residuals by QQ and ACF plots, respectively. The QQ plots in Figure 3.8 show that the GCS residuals seem to follow the exp(1) distribution considering that the 95th and 99th percentiles of this distribution are around 3.0 and 4.6, respectively. The ACF plots in Figure 3.9 demonstrate a lack of serial correlation in the GCS residuals. We can, therefore, conclude that the BVBS-ACD model may provide a good fit for these data.

3.7.5 Out-of-sample forecast evaluation

We use the first six seventh of the matched data, i.e., “in-sample” observations 1 to 5333, to estimate the model parameters, and then employ the resulting predictive model to the rest of the data, i.e., “out-of-sample” observations 5334 to 6222, to form density forecasts. These evaluation results are presented in Table 3.9. At 10% significance level, the null hypothesis that the series $z_1$, $z_2$, $z_{1|2}$ and $z_{2|1}$ are all i.i.d. $U(0, 1)$ cannot be rejected. The KS and AD tests both seem to support the uniformity of the corresponding $z$ series. The LB tests indicate no serial correlation is presented.
Figure 3.4: Histograms for $z_1$ (left) and $z_2$ (right)

Figure 3.5: Histograms for $z_{1|2}$ (left) and $z_{2|1}$ (right)
Figure 3.6: ACF plots for $z_1$ (left) and $z_2$ (right)

Figure 3.7: ACF plots for $z_{1|2}$ (left) and $z_{2|1}$ (right)
Figure 3.8: QQ plots for GCS residuals of JNJ (left) and PG (right) stocks

Figure 3.9: ACF plots for GCS residuals of JNJ (left) and PG (right) stocks
in the associated $z$ and $z^2$ series.

Furthermore, the $p$-values from the KS, AD, LB(15) and LB$^2$(15) tests for $\{z_1, z_{2|1}\}$ are 0.400, 0.230, 0.229 and 0.425. The corresponding values for $\{z_2, z_{1|2}\}$ are 0.452, 0.222, 0.215 and 0.418. Consequently, the null hypothesis that the pooled series $\{z_1, z_{2|1}\}$ and $\{z_2, z_{1|2}\}$ are i.i.d. $U(0, 1)$ can not be rejected at level 10%. These results all reveal that the BVBS-ACD model may provide good bivariate forecasts.

Table 3.9: $P$-values for out-of-sample tests for the BVBS-ACD model based on density forecasts

| $z$ series $|$ $z^2$ series | KS  | AD  | LB(10) | LB(15) | LB(20) | LB$^2$(10) | LB$^2$(15) | LB$^2$(20) |
|-----------------|-----|-----|--------|--------|--------|-----------|-----------|-----------|
| $z_1$           | 0.179 | 0.106 | 0.529 | 0.300 | 0.258 | 0.816    | 0.433    | 0.561    |
| $z_2$           | 0.478 | 0.278 | 0.400 | 0.361 | 0.303 | 0.152    | 0.170    | 0.160    |
| $z_{1|2}$       | 0.185 | 0.102 | 0.567 | 0.316 | 0.238 | 0.839    | 0.453    | 0.559    |
| $z_{2|1}$       | 0.508 | 0.267 | 0.401 | 0.390 | 0.297 | 0.139    | 0.174    | 0.149    |

**Note:** LB$(l)$ represents LB test for $z$ series over $l$ lags, where $l = 10, 15$ and 20. LB$^2$(l) represents LB test for $z^2$ series over $l$ lags.
Chapter 4

Bivariate BS\textsuperscript{t}-ACD model

4.1 Introduction

Kundu et al. (2010) extended the BS distribution to the bivariate case and proposed the bivariate Birnbaum-Saunders (BVBS) distribution. The bivariate Student-	extit{t} BS (BVBS\textit{t}) distribution was suggested by Vilca et al. (2014a) as a robust extension of the BVBS distribution; see Kundu et al. (2013) for generalized multivariate Birnbaum-Saunders distributions. Saulo et al. (2017b) studied the moment estimation of the parameters of the BVBS\textit{t} distribution within the generalized bivariate Birnbaum-Saunders family.

Our goal here is to jointly model the trade durations of two assets based on matched data, pairs of durations with the same starting time (see Simonsen (2007)). In the last Chapter, we constructed a bivariate ACD model based on BVBS distribution which allowed us to jointly analyze and measure the strength of the dependence of two matched duration series. In the present Chapter, we propose a bivariate ACD model based on BVBS\textit{t} distribution which is quite flexible and accounts for higher
kurtosis with the inclusion of an additional parameter, namely, the degrees of freedom of the Student-\(t\) distribution. This can potentially facilitate robust parameter estimation of the model parameters.

The rest of this Chapter is organized as follows. In Section 4.2, we propose the BVBS\(t\)-ACD model and the corresponding log-likelihood function. We derive the first and second derivatives of \(\ln L(\theta)\) in Sections 4.3 and 4.4. In Section 4.5, we discuss the maximum likelihood estimation and associated inference for the model parameters. A Monte Carlo simulation study is carried out in Section 4.6 to examine the properties of the MLEs. Model discrimination is discussed in Section 4.7. In Section 4.8 and 4.9, we provide an application of the BVBS\(t\)-ACD model to bivariate datasets, and then compare the results with those of the BVBS-ACD model.

4.2 Bivariate BS\(t\)-ACD model and the log-likelihood function

By working with the joint density of \((X_{1i}, X_{2i})\) directly, we propose the following BVBS\(t\)-ACD model:

\[
X_{1i} = \sigma_{1i} \epsilon_{1i}, \quad \ln \sigma_{1i} = \alpha_1 + \beta_1 \ln \sigma_{1,i-1} + \gamma_1 \left[ \frac{X_{1,i-1}}{\sigma_{1,i-1}} \right], \quad |\beta_1| < 1, \quad (4.1)
\]

\[
X_{2i} = \sigma_{2i} \epsilon_{2i}, \quad \ln \sigma_{2i} = \alpha_2 + \beta_2 \ln \sigma_{2,i-1} + \gamma_2 \left[ \frac{X_{2,i-1}}{\sigma_{2,i-1}} \right], \quad |\beta_2| < 1, \quad (4.2)
\]

where \((X_{1i}, X_{2i}) \sim \text{BVBS}\(t\)(\(\kappa_1, \sigma_{1i}, \kappa_2, \sigma_{2i}, \rho, \nu\)), \quad \epsilon_{1i} \overset{\text{iid}}{\sim} \text{BS}\(t\)(\(\kappa_1, 1, \nu\)) and \(\epsilon_{2i} \overset{\text{iid}}{\sim} \text{BS}\(t\)(\(\kappa_2, 1, \nu\)).
The associated joint PDF of the BVBS-t-ACD model is given by

$$f_{X_{1i}, X_{2i}}(x_{1i}, x_{2i}; \theta)$$

$$= \frac{1}{2 \kappa_1 \sigma_{1i}} \left[ \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] \frac{1}{2 \kappa_2 \sigma_{2i}} \left[ \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right]$$

$$\times t_2 \left[ \frac{1}{\kappa_1} \left( \frac{x_{1i}}{\sigma_{1i}} - \sqrt{\frac{\sigma_{1i}}{x_{1i}}} \right), \frac{1}{\kappa_2} \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right); \rho, \nu \right]$$

$$= \frac{1}{2 \kappa_1 x_{1i}} \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] \frac{1}{2 \kappa_2 x_{2i}} \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right]$$

$$\times \frac{1}{2 \pi \sqrt{1 - \rho^2}} \left[ 1 + \frac{1}{\pi^2} \left( \frac{x_{1i}^2 + x_{2i}^2}{x_{1i}} - 2 \right) + \frac{1}{\pi^2} \left( \frac{\sigma_{1i}^2 + \sigma_{2i}^2}{x_{2i}} - 2 \right) - \frac{2 \rho}{\pi^2 \sigma_{1i}^2} \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \right]^{-\frac{\nu + 2}{2}},$$

where, $\theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)'$ is the model parameters; in the above, $t_2$ denotes the bivariate $t$ density function.

For $i = 1, 2, \ldots, n$, the individual log-likelihood function can be expressed as

$$\ln l_i(\theta) = -\ln(\kappa_1) - \ln(\kappa_2) - \frac{1}{2} \ln(1 - \rho^2) + \ln \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] + \ln \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right]$$

$$- \frac{\nu + 2}{2} \ln \left[ 1 + \frac{1}{\sigma_{1i}^2} \left( \frac{x_{1i}^2}{x_{1i}} + \frac{\sigma_{1i}^2}{x_{1i}} - 2 \right) + \frac{1}{\sigma_{2i}^2} \left( \frac{x_{2i}^2}{x_{2i}} + \frac{\sigma_{2i}^2}{x_{2i}} - 2 \right) - \frac{2 \rho}{\sigma_{1i}^2 \sigma_{2i}^2} \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \right].$$

The log-likelihood function, without the additive constant, is then given by

$$\ln L(\theta) = \sum_{i=1}^{n} \left\{ -\ln(\kappa_1) - \ln(\kappa_2) - \frac{1}{2} \ln(1 - \rho^2) + \ln \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] + \ln \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right] \right.$$

$$- \frac{\nu + 2}{2} \ln \left[ 1 + \frac{1}{\sigma_{1i}^2} \left( \frac{x_{1i}^2}{x_{1i}} + \frac{\sigma_{1i}^2}{x_{1i}} - 2 \right) + \frac{1}{\sigma_{2i}^2} \left( \frac{x_{2i}^2}{x_{2i}} + \frac{\sigma_{2i}^2}{x_{2i}} - 2 \right) - \frac{2 \rho}{\sigma_{1i}^2 \sigma_{2i}^2} \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \left( \sqrt{\frac{\sigma_{1i}}{x_{1i}}} - \sqrt{\frac{\sigma_{2i}}{x_{2i}}} \right) \right].$$

$$= \left\{ -\ln(\kappa_1) - \ln(\kappa_2) - \frac{1}{2} \ln(1 - \rho^2) + \ln \left[ \left( \frac{x_{1i}}{\sigma_{1i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{1i}}{x_{1i}} \right)^{\frac{3}{2}} \right] + \ln \left[ \left( \frac{x_{2i}}{\sigma_{2i}} \right)^{\frac{1}{2}} + \left( \frac{\sigma_{2i}}{x_{2i}} \right)^{\frac{3}{2}} \right] \right.$$
4.3 The first derivatives of $\ln L(\theta)$

The first derivatives of $\ln L(\theta)$ with respect to $\alpha_j$, $\beta_j$ and $\gamma_j$ are given by

$$
\frac{\partial \ln L}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \alpha_j},
$$

$$
\frac{\partial \ln L}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \beta_j},
$$

$$
\frac{\partial \ln L}{\partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial \sigma_{ji}}{\partial \gamma_j},
$$

where

$$
\frac{\partial \ln l_i}{\partial \sigma_{ji}} = -\frac{\nu+2}{2} \left( \frac{\rho \left( \sqrt{\frac{x_{1i}}{\sigma_{ji}}} + \sqrt{\frac{x_{ji}}{\sigma_{ji}}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{1i}}{\sigma_{ji}}} \right)}{\kappa_1 \kappa_2 \sigma_{ji}} + \frac{1}{\kappa_j} \frac{x_{ji}}{\sigma_{ji}^2} \right)
+ \frac{\sigma_{ji} - x_{ji}}{2 \sigma_{ji} (x_{ji} + \sigma_{ji})};
$$

(4.5)

here $j = 1, 2, j' = \{1, 2\} \setminus \{j\}$ and

$$
Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)
$$

$$
= -\frac{2 \rho \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{1i}}} \right)}{\kappa_1 \kappa_2}
+ \frac{x_{2i}}{\kappa_{2i}} + \frac{\sigma_{2i}}{x_{2i}} - 2 \frac{x_{1i}}{\kappa_{1i}} + \frac{\sigma_{1i}}{x_{1i}} - 2 \frac{1}{\kappa_1^2} + \frac{1}{\kappa_2^2} + \nu \left( 1 - \rho^2 \right);
$$

(4.6)
also,

\[
\frac{\partial \sigma_{ji}}{\partial \alpha_j} = \sigma_{ji} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right), \quad i \geq 1, \quad (4.7)
\]

\[
\frac{\partial \sigma_{ji}}{\partial \beta_j} = \sigma_{ji} \left( \ln \sigma_{j,i} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right), \quad i \geq 2, \quad (4.8)
\]

\[
\frac{\partial \sigma_{ji}}{\partial \gamma_j} = \sigma_{ji} \left( \frac{x_{j,i-1}}{\sigma_{j,i-1}} + \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk}} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right), \quad i \geq 2, \quad (4.9)
\]

with \( \nabla \sigma_{j0} = (0, 0, 0)' \) and \( \nabla \sigma_{j1} = (\sigma_{j1}, \sigma_{j1} \ln \sigma_{j0}, \sigma_{j1} x_{j0}/\sigma_{j0})' \).

The first derivatives of \( \ln L(\boldsymbol{\theta}) \) with respect to \( \kappa_j, j = 1, 2, \rho \) and \( \nu \) are given by

\[
\frac{\partial \ln L}{\partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \kappa_j},
\]

\[
\frac{\partial \ln L}{\partial \rho} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \rho},
\]

\[
\frac{\partial \ln L}{\partial \nu} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \nu},
\]

where

\[
\frac{\partial \ln l_i}{\partial \kappa_j} = - \frac{(\nu + 2) \left( \left( \frac{\sqrt{x_{1i}}}{\sigma_{1i}} - \frac{\sqrt{x_{2i}}}{\sigma_{2i}} \right) \left( \frac{\sqrt{x_{1i}}}{\sigma_{1i}} - \frac{\sqrt{x_{2i}}}{\sigma_{2i}} \right) - \frac{x_{1i} \sigma_{1i} + x_{2i} \sigma_{2i} - 2}{\kappa_j} \right)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} - \frac{1}{\kappa_j},
\]

\[
\frac{\partial \ln l_i}{\partial \rho} = - \frac{\rho (\nu + 2) \left( Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu \right)}{(1 - \rho^2) Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} + \frac{(\nu + 2) \left( \frac{\sqrt{x_{1i} \sigma_{1i}} \sigma_{1i} + \sqrt{x_{2i} \sigma_{2i}} \sigma_{2i}}{x_{1i} \kappa_1 + x_{2i} \kappa_2} \right)}{\kappa_1 \kappa_2 Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} + \frac{\rho}{1 - \rho^2};
\]

\[
\frac{\partial \ln l_i}{\partial \nu} = \frac{(\nu + 2) \left( Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu \right)}{2 \nu Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} - \frac{1}{2} \log \left( \frac{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)}{(1 - \rho^2) \nu} \right); \quad (4.10)
\]
in the above, \(Q(x_1, x_2; \kappa_1, \kappa_2, \rho, \nu)\) is defined as (4.6).

### 4.4 The second derivatives of \(\ln L(\theta)\)

The second derivatives of \(\ln L(\theta)\) with respect to \(\theta^T = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)^T\) are as follows:

\[
\frac{\partial^2 \ln L}{\partial \alpha_1^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_1^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_1} \frac{\partial \ln l_i}{\partial \alpha_1} = n \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \alpha_1} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_1^2} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_1^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_1^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \beta_1} \frac{\partial \ln l_i}{\partial \beta_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \beta_1} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \beta_1^2} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_1^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_1^2} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \gamma_1} \frac{\partial \ln l_i}{\partial \gamma_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \left( \frac{\partial \sigma_{ji}}{\partial \gamma_1} \right)^2 + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \gamma_1^2} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_1 \partial \beta_1} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_1 \partial \beta_1} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_1} \frac{\partial \ln l_i}{\partial \beta_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \alpha_1} \frac{\partial \sigma_{ji}}{\partial \beta_1} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_1 \partial \beta_1} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_1 \partial \gamma_1} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_1 \partial \gamma_1} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_1} \frac{\partial \ln l_i}{\partial \gamma_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \alpha_1} \frac{\partial \sigma_{ji}}{\partial \gamma_1} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_1 \partial \gamma_1} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_2 \partial \beta_1} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_2 \partial \beta_1} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_2} \frac{\partial \ln l_i}{\partial \beta_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \alpha_2} \frac{\partial \sigma_{ji}}{\partial \beta_1} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_2 \partial \beta_1} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_2 \partial \gamma_1} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_2 \partial \gamma_1} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \alpha_2} \frac{\partial \ln l_i}{\partial \gamma_1} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \alpha_2} \frac{\partial \sigma_{ji}}{\partial \gamma_1} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \alpha_2 \partial \gamma_1} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_1 \partial \beta_2} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \beta_1} \frac{\partial \ln l_i}{\partial \beta_2} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \beta_1} \frac{\partial \sigma_{ji}}{\partial \beta_2} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \beta_1 \partial \beta_2} \right],
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_1 \partial \gamma_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_1 \partial \gamma_2} = \sum_{i=1}^{n} \frac{\partial \ln l_i}{\partial \gamma_1} \frac{\partial \ln l_i}{\partial \gamma_2} = \sum_{i=1}^{n} \left[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} \frac{\partial \sigma_{ji}}{\partial \gamma_1} \frac{\partial \sigma_{ji}}{\partial \gamma_2} + \frac{\partial \ln l_i}{\partial \sigma_{ji}} \frac{\partial^2 \sigma_{ji}}{\partial \gamma_1 \partial \gamma_2} \right].
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \beta_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \beta_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \sigma_{ji}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \gamma_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{j'i} \partial \sigma_{j'i}},
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}},
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_{j'}} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \sigma_{ji}}.
\]
\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} \frac{\partial \sigma_{ji}}{\partial \alpha_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} \frac{\partial \sigma_{ji}}{\partial \alpha_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} \frac{\partial \sigma_{ji}}{\partial \beta_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} \frac{\partial \sigma_{ji}}{\partial \gamma_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \alpha_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \alpha_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \nu} \frac{\partial \sigma_{ji}}{\partial \alpha_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \beta_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \beta_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \beta_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \nu} \frac{\partial \sigma_{ji}}{\partial \beta_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \gamma_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \gamma_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \gamma_j} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \nu} \frac{\partial \sigma_{ji}}{\partial \gamma_j},
\]

\[
\frac{\partial^2 \ln L}{\partial \kappa_j^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_j^2},
\]

\[
\frac{\partial^2 \ln L}{\partial \kappa_1 \partial \kappa_2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_1 \partial \kappa_2},
\]

\[
\frac{\partial^2 \ln L}{\partial \kappa_j \partial \rho} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \rho},
\]

\[
\frac{\partial^2 \ln L}{\partial \kappa_j \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \nu},
\]

\[
\frac{\partial^2 \ln L}{\partial \rho \partial \nu} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \rho \partial \nu},
\]

\[
\frac{\partial^2 \ln L}{\partial \rho^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \rho^2},
\]

\[
\frac{\partial^2 \ln L}{\partial \nu^2} = \sum_{i=1}^{n} \frac{\partial^2 \ln l_i}{\partial \nu^2}.
\]
where \(Q(x_1, x_2; \kappa_1, \kappa_2, \rho, \nu)\) is defined as (4.6); also,

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji}^2} = \frac{(\nu + 2)}{2Q(x_1, x_2; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{1}{\kappa_j^2} \left( \frac{x_{ji}}{\kappa_j} \right)^2 + \frac{\rho}{\kappa_1 \kappa_2 \sigma_{ji}} \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{x_{ji}}{\sigma_{ji}}} \right)^2 \right\}
\]

\[
= \frac{\nu + 2}{2 \kappa_1 \kappa_2 \sigma_{ji}^2} \left\{ \frac{3}{\kappa_j^2} \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}}} \right)^2 + \frac{3}{\kappa_j^2} \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}}} \right)^2 \right\}
\]

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{1i} \partial \sigma_{2i}} = \frac{(\nu + 2)}{2Q(x_1, x_2; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{\rho}{\kappa_1 \kappa_2 \sigma_{1i} \sigma_{2i}} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} + \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)^2 + \frac{1}{\kappa_1 \kappa_2 \sigma_{1i}^2} \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} + \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)^2 \right\}
\]

\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{1i} \partial \sigma_{2i}} = \frac{(\nu + 2)}{2Q(x_1, x_2; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{\rho}{\kappa_1 \kappa_2 \sigma_{1i} \sigma_{2i}} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} + \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)^2 + \frac{1}{\kappa_1 \kappa_2 \sigma_{1i}^2} \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} + \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)^2 \right\}
\]

(4.11)

(4.12)
\[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_j} = \frac{(\nu + 2)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) + \frac{1}{x_{ji}} - \frac{x_{ji}}{\sigma_{ji}^3} \right\} \right. \\
+ \left. \left( \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}'}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) \frac{(x_{ji}' + \frac{\sigma_{ji}'}{x_{ji}'}) - 2}{\kappa_j^2 \kappa_j'} \right) \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}'}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) + \frac{1}{x_{ji}} - \frac{x_{ji}}{\sigma_{ji}^3} \right) \right\} \\
+ \left. Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) \right\}, \\
(4.13) \\
\]

\[ \frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \kappa_{j'}} = \frac{(\nu + 2)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}'}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) + \frac{1}{x_{ji}} - \frac{x_{ji}}{\sigma_{ji}^3} \right\} \right. \\
+ \left. \left( \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}'}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) \frac{(x_{ji}' + \frac{\sigma_{ji}'}{x_{ji}'}) - 2}{\kappa_j \kappa_{j'}} \right) \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{\sigma_{ji}}{x_{ji}}} \right) \left( \sqrt{\frac{x_{ji}'}{\sigma_{ji}'}} - \sqrt{\frac{\sigma_{ji}'}{x_{ji}'}} \right) + \frac{1}{x_{ji}} - \frac{x_{ji}}{\sigma_{ji}^3} \right) \right\} \\
+ \left. Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) \right\}, \\
(4.14) \\
\]
\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \rho} = \frac{(\nu + 2)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) \right. \\
\left. \frac{2 \kappa_1 \kappa_2}{\sigma_{ji}} \right\} \\
- \frac{\kappa_1 \kappa_2 Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)}{\rho (Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}}} \right) + \frac{1}{\kappa_j^2} \right) \\
+ \frac{(1 - \rho^2) Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)}{\rho \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} + \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) + \frac{1}{\kappa_j^2} \right) \left( 1 - \rho^2 \right) \right\},
\]
\]
\[
(4.15)
\]
\[
\frac{\partial^2 \ln l_i}{\partial \sigma_{ji} \partial \nu} = \frac{\nu Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)}{(\nu + 2) (Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) \left( \sqrt{\frac{x_{ji}}{\sigma_{ji}}} - \sqrt{\frac{x_{ji}}{\sigma_{ji}^*}} \right) + \frac{1}{\kappa_j^2} \right) \left( 1 - \rho^2 \right)} \left\{ \frac{1}{2 Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \right\},
\]
\[
(4.16)
\]
\[
\frac{\partial^2 \ln l_i}{\partial \kappa_j^2} = \frac{\nu + 2}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{2 \left( \rho \left( \frac{\sqrt{\frac{\sigma_{11}}{x_{1i}}} - \sqrt{\frac{\sigma_{11}}{x_{1i}}}}{\kappa_j} \right) - \frac{x_{ji} + \sigma_{ji} x_{ji} - 2}{\kappa_j^2} \right) - 1}{\kappa_j^2} \right\} + \frac{1}{\kappa_j^2},
\]

(4.17)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa_1 \partial \kappa_2} = \frac{\nu + 2}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{\rho \left( \frac{\sqrt{\frac{\sigma_{11}}{x_{1i}}} - \sqrt{\frac{\sigma_{11}}{x_{1i}}}}{\kappa_1 \kappa_2} \right) - \frac{\sqrt{\frac{\sigma_{11}}{x_{1i}}} - \sqrt{\frac{\sigma_{11}}{x_{1i}}}}{\kappa_2^2} \right\} + \frac{1}{\kappa_j^2},
\]

(4.18)

\[
\frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \rho} = \frac{\nu + 2}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{\rho \left( \frac{\sqrt{\frac{\sigma_{11}}{x_{1i}}} - \sqrt{\frac{\sigma_{11}}{x_{1i}}}}{\kappa_1 \kappa_2} \right) - \frac{\sqrt{\frac{\sigma_{11}}{x_{1i}}} - \sqrt{\frac{\sigma_{11}}{x_{1i}}}}{\kappa_2^2} \right\} + \frac{1}{\kappa_j^2},
\]

(4.19)
\[
\frac{\partial^2 \ln l_i}{\partial \kappa_j \partial \nu} = \frac{1}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{\rho}{\kappa_j^2 \kappa_j'} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}}, \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right) - \frac{x_{ji} + \sigma_{ji} - 2}{\kappa_j^3} \right\},
\]

\[
(4.20)
\]

\[
\frac{\partial^2 \ln l_i}{\partial \rho \partial \nu} = \frac{1}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ (\nu + 2) \left( \frac{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} - (1 - \rho^2) \nu \right) - 2 \right\}
\]

\[
\times \left\{ \frac{\left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)}{\kappa_1 \kappa_2} - \frac{\rho}{(1 - \rho^2)} \left( Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu \right) \right\},
\]

\[
(4.21)
\]

\[
\frac{\partial^2 \ln l_i}{\partial \rho^2} = \frac{(\nu + 2)}{Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{2 \rho}{\kappa_1 \kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}}, \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right) \right\}
\]

\[
- \frac{\left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right)}{\kappa_1 \kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right) - 2 \rho \nu \right),
\]

\[
- \frac{2 \rho}{\kappa_1 \kappa_2} \left( \sqrt{\frac{x_{1i}}{\sigma_{1i}}} - \sqrt{\frac{x_{1i}}{\sigma_{2i}}} \right) \left( \sqrt{\frac{x_{2i}}{\sigma_{2i}}} - \sqrt{\frac{x_{2i}}{\sigma_{2i}}} \right) + \rho \nu \right) \left( Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu \right)
\]

\[
- \frac{(1 + \rho^2) \left( Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2) \nu \right)}{(1 - \rho^2)^2} \left( 1 - \frac{2 \rho^2}{(1 - \rho^2)^2} \right),
\]

\[
(4.22)
\]
\[
\frac{\partial^2 \ln l_i}{\partial \nu^2} = \frac{(Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu) - (1 - \rho^2)\nu)}{\nu Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \left\{ \frac{-(\nu + 2)}{2\nu} + 1 - \frac{(\nu + 2)(1 - \rho^2)}{2Q(x_{1i}, x_{2i}; \kappa_1, \kappa_2, \rho, \nu)} \right\},
\]

(4.23)

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha_j^2} = \frac{\partial \sigma_{ji}}{\partial \alpha_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \right) + \sigma_{ji} \gamma \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \sum_{l=k}^{i-1} \frac{x_l \frac{\partial \sigma_{jl}}{\partial \alpha_j}}{\sigma_{jl}^2 \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right)},
\]

\[
\quad \text{if } i \geq 1, \quad (4.24)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \beta_j^2} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left\{ \ln \sigma_{j,i-1} + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \right\}
\]

\[
\quad + \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \left( \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \sum_{l=k+1}^{i-1} \frac{\gamma_j x_{jl} \frac{\partial \sigma_{jl}}{\partial \gamma_j}}{\sigma_{jl}^2 \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right)} + \frac{\partial \sigma_{jk}}{\partial \gamma_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j x_{jl} / \sigma_{jl} \right) \frac{\partial \sigma_{j,i-1}}{\partial \gamma_j} \right) \right\},
\]

\[
\quad \text{if } i \geq 2, \quad (4.25)
\]
\[
\frac{\partial^2 \sigma_{ji}}{\partial \gamma_j^2} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left\{ \frac{x_{ji} - 1}{\sigma_{ji} - 1} + \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk} l=0} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right\} \\
+ \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \frac{x_{jk}}{\sigma_{jk} l=0} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \prod_{l=k+1}^{i-1} \frac{\gamma_j x_{jl} \partial \sigma_{jl}}{\sigma_{jl}} 1 + \frac{\partial \sigma_{jk}}{\partial \gamma_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right\}, \\
i \geq 2, \quad (4.26)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \beta_j} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_{ji} \sum_{l=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma_j x_{jl} \partial \sigma_{jl}}{\sigma_{jl}} - \frac{x_{jl}}{\sigma_{jl}}, \\
i \geq 1, \quad (4.27)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \alpha_j \partial \gamma_j} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left( 1 + \sum_{k=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right) + \sigma_{ji} \sum_{l=1}^{i-1} \prod_{l=k}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k}^{i-1} \frac{\gamma_j x_{jl} \partial \sigma_{jl}}{\sigma_{jl}} - \frac{x_{jl}}{\sigma_{jl}}, \\
i \geq 1, \quad (4.28)
\]

\[
\frac{\partial^2 \sigma_{ji}}{\partial \beta_j \partial \gamma_j} = \frac{\partial \sigma_{ji}}{\partial \gamma_j} \left\{ \ln \sigma_{ji} - 1 + \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right\} \\
+ \sigma_{ji} \left\{ \sum_{k=0}^{i-2} \ln \sigma_{jk} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \sum_{l=k+1}^{i-1} \frac{\gamma_j x_{jl} \partial \sigma_{jl}}{\sigma_{jl}} - \frac{x_{jl}}{\sigma_{jl}} + \frac{\partial \sigma_{jk}}{\partial \gamma_j} \prod_{l=k+1}^{i-1} \left( \beta_j - \gamma_j \frac{x_{jl}}{\sigma_{jl}} \right) \right\}, \\
i \geq 2. \quad (4.29)
\]
4.5 Estimation and inference

We estimate the parameter \( \theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)' \) of the BVBS\( t \)-ACD model by maximizing the log-likelihood function (without the additive constant) in equation (4.4). We maximize the above log-likelihood function by a two-step procedure. In the first step, we fix \( \kappa_1, \kappa_2, \rho \) and \( \nu \) at their initial values and maximize \( \ln L(\theta) \) with respect to the ACD parameters \( \alpha_j, \beta_j, \) and \( \gamma_j, j = 1, 2 \), by utilizing the NM algorithm. Then, in the second step, we estimate over the full parameter space by employing the BFGS algorithm. We have derived and implemented the analytical gradients (the first derivatives in Section 4.3) in the second step. The BFGS quasi-Newton method with the analytic gradient will be faster, more stable and lead to more accurate estimates than a numerical gradient (see Bard (1974), Bolker (2008) and Mayorov (2011)).

Under the regularity conditions, the ML estimator \( \hat{\theta} \) is consistent and \( \hat{\theta} \overset{d}{\sim} N_{10}(\theta_0, I(\theta_0)^{-1}) \), where \( I(\theta_0) = -E \left[ \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta}' \right] \bigg|_{\theta = \theta_0} = -E [H(\theta_0)]; H = \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta}' \) is the Hessian matrix. In order to estimate the asymptotic covariance matrix of \( \hat{\theta} \), we consider the consistent estimator \( -H(\hat{\theta})^{-1} \). The standard errors can then be approximated by the square roots of the diagonal elements of this matrix.

4.6 Simulation study

In order to assess the performance of the maximum likelihood estimators of the parameters of the BVBS\( t \)-ACD model, we carry out a simulation study for different sample sizes and correlation parameter \( \rho \), based on the method detailed in Section 4.5. For sample sizes \( n=1000, 3000, 5000 \) and \( 10000 \), we simulate 1000 samples from
the BVBS\textsubscript{t}-ACD model with the vector of true parameters

\[ \boldsymbol{\theta} = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)' = (0.1, 0.9, 0.1, -0.1, 0.95, 0.1, 1.0, 1.05, \rho, 12)' \]

Here, the value of \( \rho \) are chosen as 0.0, 0.2, 0.5 and 0.9. We choose the starting value for ACD parameters to be

\[ (\alpha_{10}, \beta_{10}, \gamma_{10}, \alpha_{20}, \beta_{20}, \gamma_{20})' = (0.01, 0.80, 0.01, -0.01, 0.80, 0.01)' \]

The initial values of shape parameters \( \kappa_{j0}, j = 1, 2 \), and the correlation parameter \( \rho_0 \) are given by

\[ \kappa_{j0} = \sqrt{2 \left( \frac{X_j}{\text{med}(X_j)} - 1 \right) \frac{\nu - 2}{\nu}}, \quad j = 1, 2, \]

and

\[ \rho_0 = \text{Corr} \left( \frac{1}{\kappa_{10}} \left( \sqrt{\frac{X_1}{\hat{\sigma}_1}} - \sqrt{\frac{X_1}{\hat{\sigma}_1}} \right), \frac{1}{\kappa_{20}} \left( \sqrt{\frac{X_2}{\hat{\sigma}_2}} - \sqrt{\frac{X_2}{\hat{\sigma}_2}} \right) \right), \]

where \( \hat{\sigma}_j \) is obtained as

\[ \hat{\sigma}_{ji} = \exp \left( \alpha_{j0} + \beta_{j0} \ln \hat{\sigma}_{j,i-1} + \gamma_{j0} \left[ \frac{X_{ji-1}}{\hat{\sigma}_{j,i-1}} \right] \right), \]

\( j = 1, 2, i = 1, 2, \ldots, n \). We set \( \nu_0 = 3 \).

The performance of the MLEs are examined in terms of the mean, coefficients of skewness and kurtosis, bias and root mean squared error (RMSE) of the MLEs we computed over 1000 replications for each sample size and for each level of correlation.

These results are presented in Tables 4.1-4.16. From these tables, we see that all estimators perform well across different levels of correlation. The bias, MSE and
RMSE decrease toward 0 quickly as sample size increases, which demonstrates the asymptotic unbiasedness and consistency properties of the MLEs. An increase in the sample size shrinks the standard errors of the mean. The skewness and kurtosis obtained from the estimators are close to their limiting values of 0 and 3 for a reasonable sample size. Clearly, the asymptotic normal approximation for the estimators improves with increasing sample size.

4.7 Model discrimination

As mentioned earlier, the BVBS\(t\)-ACD model allows more flexibility in terms of the kurtosis through an extra parameter, namely, the degrees of freedom of the Student-\(t\) distribution. For the purpose of model discrimination, we consider a collection of its special cases, BVBS\(t_7\), BVBS\(t_9\), BVBS\(t_{12}\) and BVBS-ACD models. From the general BVBS\(t\)-ACD model, we are interested in investigating how often the true model gets selected in the set of candidate models and select a parsimonious model that fits the data adequately. This model evaluation technique is called model discrimination. One may refer to the book by McLachlan and Peel (2000), and the recent papers by Balakrishnan and Peng (2006), Balakrishnan and Pal (2013), Balakrishnan and Pal (2016), and Balakrishnan \textit{et al.} (2017).

Two well-known model selection criteria, AIC and BIC, are utilized here. They are given by 
\[\text{AIC} = -2l + 2k\] and 
\[\text{BIC} = -2l + k\log(n),\] where \(l\) stands for the maximized log-likelihood value, \(p\) is the number of model parameters to be estimated, and \(n\) is the sample size. The model with the lowest AIC or BIC will be selected. Given that the true data generating process (DGP) are the BVBS\(t_7\), BVBS\(t_9\), BVBS\(t_{12}\) and BV-
Table 4.1: Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.0 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1047</td>
<td>0.8952</td>
<td>0.1044</td>
<td>-0.0998</td>
<td>0.9468</td>
<td>0.1021</td>
<td>1.0020</td>
<td>1.0437</td>
<td>0.0093</td>
<td>11.4213</td>
</tr>
<tr>
<td>Skew</td>
<td>1.1843</td>
<td>-1.0812</td>
<td>0.1918</td>
<td>0.3639</td>
<td>-0.6107</td>
<td>0.0669</td>
<td>0.0534</td>
<td>-0.0130</td>
<td>-0.1307</td>
<td>1.6467</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0047</td>
<td>-0.0048</td>
<td>0.0044</td>
<td>0.0002</td>
<td>-0.0032</td>
<td>0.0021</td>
<td>0.0020</td>
<td>-0.0063</td>
<td>0.0093</td>
<td>-0.5787</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0034</td>
<td>0.0006</td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0012</td>
<td>8.8907</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0586</td>
<td>0.0249</td>
<td>0.0132</td>
<td>0.0196</td>
<td>0.0130</td>
<td>0.0116</td>
<td>0.0307</td>
<td>0.0320</td>
<td>0.0345</td>
<td>2.9817</td>
</tr>
<tr>
<td>SE</td>
<td>0.0584</td>
<td>0.0244</td>
<td>0.0124</td>
<td>0.0196</td>
<td>0.0126</td>
<td>0.0114</td>
<td>0.0306</td>
<td>0.0314</td>
<td>0.0332</td>
<td>2.9265</td>
</tr>
</tbody>
</table>

Table 4.2: Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 1000 \) and \( \rho = 0.2 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1014</td>
<td>0.8964</td>
<td>0.1044</td>
<td>-0.0974</td>
<td>0.9468</td>
<td>0.1007</td>
<td>1.0029</td>
<td>1.0452</td>
<td>0.2052</td>
<td>11.6583</td>
</tr>
<tr>
<td>Skew</td>
<td>0.6768</td>
<td>-0.5418</td>
<td>0.1008</td>
<td>0.7212</td>
<td>-1.0215</td>
<td>0.0819</td>
<td>0.0092</td>
<td>0.0513</td>
<td>-0.0403</td>
<td>1.6153</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0014</td>
<td>-0.0036</td>
<td>0.0044</td>
<td>0.0026</td>
<td>-0.0032</td>
<td>0.0007</td>
<td>0.0029</td>
<td>-0.0048</td>
<td>0.0052</td>
<td>-0.3417</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0027</td>
<td>0.0005</td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0011</td>
<td>8.8651</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0521</td>
<td>0.0221</td>
<td>0.0131</td>
<td>0.0198</td>
<td>0.0138</td>
<td>0.0106</td>
<td>0.0297</td>
<td>0.0314</td>
<td>0.0332</td>
<td>2.9774</td>
</tr>
<tr>
<td>SE</td>
<td>0.0521</td>
<td>0.0218</td>
<td>0.0124</td>
<td>0.0196</td>
<td>0.0134</td>
<td>0.0106</td>
<td>0.0295</td>
<td>0.0310</td>
<td>0.0328</td>
<td>2.9592</td>
</tr>
</tbody>
</table>
Table 4.3: Simulation results for BVBS\(_t\)-ACD model when \(n = 1000\) and \(\rho = 0.5\).

<table>
<thead>
<tr>
<th>(\hat{\alpha}_1)</th>
<th>(\hat{\beta}_1)</th>
<th>(\hat{\gamma}_1)</th>
<th>(\hat{\alpha}_2)</th>
<th>(\hat{\beta}_2)</th>
<th>(\hat{\gamma}_2)</th>
<th>(\hat{\kappa}_1)</th>
<th>(\hat{\kappa}_2)</th>
<th>(\hat{\rho})</th>
<th>(\hat{\nu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1052</td>
<td>0.8956</td>
<td>0.1034</td>
<td>-0.0991</td>
<td>0.9474</td>
<td>0.1012</td>
<td>1.0034</td>
<td>1.0446</td>
<td>0.5038</td>
</tr>
<tr>
<td>Skew</td>
<td>0.6547</td>
<td>-0.5637</td>
<td>0.1934</td>
<td>0.3029</td>
<td>-0.5333</td>
<td>0.0845</td>
<td>0.0651</td>
<td>0.0091</td>
<td>-0.2359</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0052</td>
<td>-0.0044</td>
<td>0.0034</td>
<td>0.0009</td>
<td>-0.0026</td>
<td>0.0012</td>
<td>0.0034</td>
<td>-0.0054</td>
<td>0.0038</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0026</td>
<td>0.0005</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0006</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0514</td>
<td>0.0219</td>
<td>0.0116</td>
<td>0.0171</td>
<td>0.0113</td>
<td>0.0100</td>
<td>0.0304</td>
<td>0.0320</td>
<td>0.0253</td>
</tr>
<tr>
<td>SE</td>
<td>0.0511</td>
<td>0.0214</td>
<td>0.0111</td>
<td>0.0171</td>
<td>0.0110</td>
<td>0.0100</td>
<td>0.0302</td>
<td>0.0315</td>
<td>0.0250</td>
</tr>
</tbody>
</table>

Table 4.4: Simulation results for BVBS\(_t\)-ACD model when \(n = 1000\) and \(\rho = 0.9\).

<table>
<thead>
<tr>
<th>(\hat{\alpha}_1)</th>
<th>(\hat{\beta}_1)</th>
<th>(\hat{\gamma}_1)</th>
<th>(\hat{\alpha}_2)</th>
<th>(\hat{\beta}_2)</th>
<th>(\hat{\gamma}_2)</th>
<th>(\hat{\kappa}_1)</th>
<th>(\hat{\kappa}_2)</th>
<th>(\hat{\rho})</th>
<th>(\hat{\nu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1004</td>
<td>0.8974</td>
<td>0.1032</td>
<td>-0.1010</td>
<td>0.9484</td>
<td>0.1014</td>
<td>0.9941</td>
<td>1.0372</td>
<td>0.8989</td>
</tr>
<tr>
<td>Skew</td>
<td>0.0467</td>
<td>0.1223</td>
<td>-0.0223</td>
<td>0.0741</td>
<td>-0.0413</td>
<td>0.0024</td>
<td>0.0872</td>
<td>0.3339</td>
<td>-0.3515</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0004</td>
<td>-0.0026</td>
<td>0.0032</td>
<td>-0.0010</td>
<td>-0.0016</td>
<td>0.0014</td>
<td>-0.0059</td>
<td>-0.0128</td>
<td>-0.0011</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0013</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0012</td>
<td>0.0000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0361</td>
<td>0.0150</td>
<td>0.0093</td>
<td>0.0131</td>
<td>0.0079</td>
<td>0.0079</td>
<td>0.0306</td>
<td>0.0349</td>
<td>0.0066</td>
</tr>
<tr>
<td>SE</td>
<td>0.0361</td>
<td>0.0148</td>
<td>0.0087</td>
<td>0.0130</td>
<td>0.0078</td>
<td>0.0078</td>
<td>0.0300</td>
<td>0.0324</td>
<td>0.0065</td>
</tr>
</tbody>
</table>
Table 4.5: Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.0 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1002</td>
<td>0.8988</td>
<td>0.1015</td>
<td>-0.0990</td>
<td>0.9487</td>
<td>0.1004</td>
<td>0.9998</td>
<td>1.0474</td>
<td>0.0030</td>
<td>11.6888</td>
</tr>
<tr>
<td>Skew</td>
<td>0.4678</td>
<td>-0.3376</td>
<td>0.1518</td>
<td>0.2680</td>
<td>-0.4674</td>
<td>0.2388</td>
<td>0.0792</td>
<td>0.1310</td>
<td>0.1439</td>
<td>0.8277</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0002</td>
<td>-0.0012</td>
<td>0.0015</td>
<td>0.0010</td>
<td>-0.0013</td>
<td>0.0004</td>
<td>-0.0002</td>
<td>-0.0026</td>
<td>0.0030</td>
<td>-0.3112</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0010</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0003</td>
<td>0.0004</td>
<td>0.0004</td>
<td>2.4541</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0316</td>
<td>0.0136</td>
<td>0.0073</td>
<td>0.0111</td>
<td>0.0075</td>
<td>0.0061</td>
<td>0.0171</td>
<td>0.0188</td>
<td>0.0192</td>
<td>1.5666</td>
</tr>
<tr>
<td>SE</td>
<td>0.0316</td>
<td>0.0136</td>
<td>0.0071</td>
<td>0.0111</td>
<td>0.0074</td>
<td>0.0061</td>
<td>0.0171</td>
<td>0.0186</td>
<td>0.0190</td>
<td>1.5361</td>
</tr>
</tbody>
</table>

Table 4.6: Simulation results for BVBS\textsuperscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.2 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1015</td>
<td>0.8985</td>
<td>0.1014</td>
<td>-0.0989</td>
<td>0.9488</td>
<td>0.1002</td>
<td>0.9999</td>
<td>1.0469</td>
<td>0.2026</td>
<td>11.7179</td>
</tr>
<tr>
<td>Skew</td>
<td>0.3924</td>
<td>-0.3209</td>
<td>-0.0016</td>
<td>0.2728</td>
<td>-0.4342</td>
<td>-0.0300</td>
<td>-0.0064</td>
<td>-0.0297</td>
<td>0.1409</td>
<td>0.7954</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0015</td>
<td>-0.0015</td>
<td>0.0014</td>
<td>0.0011</td>
<td>-0.0012</td>
<td>0.0002</td>
<td>-0.0001</td>
<td>-0.0031</td>
<td>0.0026</td>
<td>-0.2861</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0009</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0004</td>
<td>2.5165</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0306</td>
<td>0.0129</td>
<td>0.0076</td>
<td>0.0105</td>
<td>0.0068</td>
<td>0.0059</td>
<td>0.0177</td>
<td>0.0185</td>
<td>0.0195</td>
<td>1.5864</td>
</tr>
<tr>
<td>SE</td>
<td>0.0305</td>
<td>0.0128</td>
<td>0.0074</td>
<td>0.0104</td>
<td>0.0067</td>
<td>0.0059</td>
<td>0.0177</td>
<td>0.0183</td>
<td>0.0193</td>
<td>1.5611</td>
</tr>
</tbody>
</table>
Table 4.7: Simulation results for BVBS\textsubscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.5 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1017</td>
<td>0.8985</td>
<td>0.1012</td>
<td>-0.0989</td>
<td>0.9489</td>
<td>0.1001</td>
<td>1.0001</td>
<td>1.0471</td>
<td>0.5014</td>
<td>11.7283</td>
</tr>
<tr>
<td>Skew</td>
<td>0.3472</td>
<td>-0.3114</td>
<td>0.0211</td>
<td>0.2399</td>
<td>-0.3930</td>
<td>-0.0215</td>
<td>-0.0215</td>
<td>-0.0344</td>
<td>0.1039</td>
<td>0.8159</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0017</td>
<td>-0.0015</td>
<td>0.0012</td>
<td>0.0011</td>
<td>-0.0011</td>
<td>0.0001</td>
<td>0.0001</td>
<td>-0.0029</td>
<td>0.0014</td>
<td>-0.2717</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0008</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0002</td>
<td>2.5618</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0277</td>
<td>0.0117</td>
<td>0.0068</td>
<td>0.0095</td>
<td>0.0062</td>
<td>0.0054</td>
<td>0.0176</td>
<td>0.0186</td>
<td>0.0151</td>
<td>1.6006</td>
</tr>
<tr>
<td>SE</td>
<td>0.0276</td>
<td>0.0116</td>
<td>0.0067</td>
<td>0.0094</td>
<td>0.0061</td>
<td>0.0054</td>
<td>0.0176</td>
<td>0.0184</td>
<td>0.0151</td>
<td>1.5781</td>
</tr>
</tbody>
</table>

Table 4.8: Simulation results for BVBS\textsubscript{t}-ACD model when \( n = 3000 \) and \( \rho = 0.9 \)

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1006</td>
<td>0.8990</td>
<td>0.1011</td>
<td>-0.0999</td>
<td>0.9494</td>
<td>0.1003</td>
<td>0.9964</td>
<td>1.0439</td>
<td>0.8996</td>
<td>11.2408</td>
</tr>
<tr>
<td>Skew</td>
<td>0.2912</td>
<td>-0.2795</td>
<td>0.0928</td>
<td>0.1622</td>
<td>-0.3255</td>
<td>0.0311</td>
<td>-0.0189</td>
<td>0.0220</td>
<td>0.0609</td>
<td>0.7550</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0006</td>
<td>-0.0010</td>
<td>0.0011</td>
<td>0.0001</td>
<td>-0.0006</td>
<td>0.0003</td>
<td>-0.0036</td>
<td>-0.0061</td>
<td>-0.0004</td>
<td>-0.7592</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0004</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>2.7348</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0201</td>
<td>0.0082</td>
<td>0.0052</td>
<td>0.0074</td>
<td>0.0043</td>
<td>0.0044</td>
<td>0.0179</td>
<td>0.0194</td>
<td>0.0039</td>
<td>1.6537</td>
</tr>
<tr>
<td>SE</td>
<td>0.0201</td>
<td>0.0082</td>
<td>0.0051</td>
<td>0.0074</td>
<td>0.0043</td>
<td>0.0044</td>
<td>0.0175</td>
<td>0.0184</td>
<td>0.0038</td>
<td>1.4699</td>
</tr>
</tbody>
</table>
Table 4.9: Simulation results for BVBS\textit{-ACD} model when $n = 5000$ and $\rho = 0.0$.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\gamma}_2$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\kappa}_2$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1004</td>
<td>0.8994</td>
<td>0.1007</td>
<td>-0.0995</td>
<td>0.9492</td>
<td>0.1002</td>
<td>0.9995</td>
<td>1.0477</td>
<td>0.0019</td>
<td>11.7867</td>
</tr>
<tr>
<td>Skew</td>
<td>0.3106</td>
<td>-0.3392</td>
<td>0.0134</td>
<td>0.1315</td>
<td>-0.2489</td>
<td>0.1334</td>
<td>0.0488</td>
<td>-0.0035</td>
<td>0.0406</td>
<td>0.5259</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0004</td>
<td>-0.0006</td>
<td>0.0007</td>
<td>0.0005</td>
<td>-0.0008</td>
<td>0.0002</td>
<td>-0.0005</td>
<td>-0.0023</td>
<td>0.0019</td>
<td>-0.2133</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0002</td>
<td>1.3949</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0242</td>
<td>0.0101</td>
<td>0.0055</td>
<td>0.0082</td>
<td>0.0054</td>
<td>0.0046</td>
<td>0.0137</td>
<td>0.0136</td>
<td>0.0156</td>
<td>1.1811</td>
</tr>
<tr>
<td>SE</td>
<td>0.0242</td>
<td>0.0101</td>
<td>0.0055</td>
<td>0.0082</td>
<td>0.0054</td>
<td>0.0046</td>
<td>0.0137</td>
<td>0.0134</td>
<td>0.0155</td>
<td>1.1622</td>
</tr>
</tbody>
</table>

Table 4.10: Simulation results for BVBS\textit{-ACD} model when $n = 5000$ and $\rho = 0.2$.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\gamma}_2$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\kappa}_2$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1005</td>
<td>0.8994</td>
<td>0.1006</td>
<td>-0.0995</td>
<td>0.9492</td>
<td>0.1002</td>
<td>0.9998</td>
<td>1.0479</td>
<td>0.2015</td>
<td>11.8201</td>
</tr>
<tr>
<td>Skew</td>
<td>0.2878</td>
<td>-0.3073</td>
<td>0.0163</td>
<td>0.1145</td>
<td>-0.2483</td>
<td>0.1261</td>
<td>0.0415</td>
<td>-0.0242</td>
<td>0.0179</td>
<td>0.5267</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0005</td>
<td>-0.0006</td>
<td>0.0006</td>
<td>0.0005</td>
<td>-0.0008</td>
<td>0.0002</td>
<td>-0.0002</td>
<td>-0.0021</td>
<td>0.0015</td>
<td>-0.1799</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0002</td>
<td>1.3950</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0235</td>
<td>0.0098</td>
<td>0.0054</td>
<td>0.0080</td>
<td>0.0053</td>
<td>0.0045</td>
<td>0.0137</td>
<td>0.0136</td>
<td>0.0150</td>
<td>1.1811</td>
</tr>
<tr>
<td>SE</td>
<td>0.0235</td>
<td>0.0097</td>
<td>0.0053</td>
<td>0.0080</td>
<td>0.0053</td>
<td>0.0045</td>
<td>0.0137</td>
<td>0.0135</td>
<td>0.0149</td>
<td>1.1679</td>
</tr>
</tbody>
</table>
Table 4.11: Simulation results for BVBS\(^t\)-ACD model when \( n = 5000 \) and \( \rho = 0.5 \)

<table>
<thead>
<tr>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1006</td>
<td>0.9994</td>
<td>0.1005</td>
<td>-0.0986</td>
<td>0.9494</td>
<td>0.1001</td>
<td>0.9994</td>
<td>1.0475</td>
<td>0.5007</td>
</tr>
<tr>
<td>Skew</td>
<td>0.2462</td>
<td>-0.2606</td>
<td>0.0434</td>
<td>0.1069</td>
<td>-0.2151</td>
<td>0.0791</td>
<td>0.0179</td>
<td>-0.0691</td>
<td>-0.0148</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0006</td>
<td>-0.0006</td>
<td>0.0005</td>
<td>0.0004</td>
<td>-0.0006</td>
<td>0.0001</td>
<td>-0.0006</td>
<td>-0.0025</td>
<td>0.0007</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0004</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0209</td>
<td>0.0087</td>
<td>0.0048</td>
<td>0.0072</td>
<td>0.0048</td>
<td>0.0040</td>
<td>0.0135</td>
<td>0.0137</td>
<td>0.0117</td>
</tr>
<tr>
<td>SE</td>
<td>0.0209</td>
<td>0.0087</td>
<td>0.0047</td>
<td>0.0071</td>
<td>0.0047</td>
<td>0.0040</td>
<td>0.0135</td>
<td>0.0135</td>
<td>0.0116</td>
</tr>
</tbody>
</table>

Table 4.12: Simulation results for BVBS\(^t\)-ACD model when \( n = 5000 \) and \( \rho = 0.9 \)

<table>
<thead>
<tr>
<th>( \hat{\delta}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\delta}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0999</td>
<td>0.8996</td>
<td>0.1005</td>
<td>-0.1001</td>
<td>0.9497</td>
<td>0.1002</td>
<td>0.9970</td>
<td>1.0454</td>
<td>0.8997</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1434</td>
<td>-0.1861</td>
<td>0.1117</td>
<td>0.0959</td>
<td>-0.1760</td>
<td>0.0386</td>
<td>-0.0659</td>
<td>-0.1003</td>
<td>-0.0624</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.0001</td>
<td>-0.0004</td>
<td>0.0005</td>
<td>-0.0001</td>
<td>-0.0003</td>
<td>0.0002</td>
<td>-0.0030</td>
<td>-0.0046</td>
<td>-0.0003</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0149</td>
<td>0.0060</td>
<td>0.0036</td>
<td>0.0054</td>
<td>0.0033</td>
<td>0.0032</td>
<td>0.0136</td>
<td>0.0144</td>
<td>0.0030</td>
</tr>
<tr>
<td>SE</td>
<td>0.0149</td>
<td>0.0060</td>
<td>0.0036</td>
<td>0.0054</td>
<td>0.0033</td>
<td>0.0032</td>
<td>0.0133</td>
<td>0.0136</td>
<td>0.0030</td>
</tr>
</tbody>
</table>
Table 4.13: Simulation results for BVBS\textsubscript{t}-ACD model when $n = 10000$ and $\rho = 0.0$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\gamma}_2$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\kappa}_2$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1004</td>
<td>0.8996</td>
<td>0.1003</td>
<td>-0.1000</td>
<td>0.9498</td>
<td>0.1002</td>
<td>0.9993</td>
<td>1.0486</td>
<td>0.0008</td>
<td>11.8615</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1088</td>
<td>-0.1241</td>
<td>-0.0315</td>
<td>0.1266</td>
<td>-0.1636</td>
<td>0.0570</td>
<td>-0.0760</td>
<td>-0.0029</td>
<td>0.1040</td>
<td>0.3358</td>
</tr>
<tr>
<td>Kurt</td>
<td>2.7984</td>
<td>2.8015</td>
<td>3.1167</td>
<td>2.9711</td>
<td>3.0264</td>
<td>2.9084</td>
<td>3.0068</td>
<td>2.7009</td>
<td>3.0058</td>
<td>3.2362</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0004</td>
<td>-0.0004</td>
<td>0.0003</td>
<td>0.0000</td>
<td>-0.0002</td>
<td>0.0002</td>
<td>-0.0007</td>
<td>-0.0014</td>
<td>0.0008</td>
<td>-0.1385</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.6708</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0170</td>
<td>0.0072</td>
<td>0.0038</td>
<td>0.0055</td>
<td>0.0036</td>
<td>0.0033</td>
<td>0.0096</td>
<td>0.0099</td>
<td>0.0107</td>
<td>0.8190</td>
</tr>
<tr>
<td>SE</td>
<td>0.0170</td>
<td>0.0072</td>
<td>0.0037</td>
<td>0.0055</td>
<td>0.0036</td>
<td>0.0033</td>
<td>0.0095</td>
<td>0.0098</td>
<td>0.0107</td>
<td>0.8076</td>
</tr>
</tbody>
</table>

Table 4.14: Simulation results for BVBS\textsubscript{t}-ACD model when $n = 10000$ and $\rho = 0.2$

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\alpha}_1$</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\gamma}_1$</th>
<th>$\hat{\alpha}_2$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\gamma}_2$</th>
<th>$\hat{\kappa}_1$</th>
<th>$\hat{\kappa}_2$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1006</td>
<td>0.8995</td>
<td>0.1003</td>
<td>-0.1000</td>
<td>0.9498</td>
<td>0.1002</td>
<td>0.9994</td>
<td>1.0487</td>
<td>0.2006</td>
<td>11.8732</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1258</td>
<td>-0.1388</td>
<td>-0.0411</td>
<td>0.1137</td>
<td>-0.1852</td>
<td>0.0540</td>
<td>-0.0768</td>
<td>0.0053</td>
<td>0.0893</td>
<td>0.3892</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0006</td>
<td>-0.0005</td>
<td>0.0003</td>
<td>0.0000</td>
<td>-0.0002</td>
<td>0.0002</td>
<td>-0.0006</td>
<td>-0.0013</td>
<td>0.0006</td>
<td>-0.1268</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.6929</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0168</td>
<td>0.0071</td>
<td>0.0037</td>
<td>0.0054</td>
<td>0.0035</td>
<td>0.0032</td>
<td>0.0096</td>
<td>0.0099</td>
<td>0.0103</td>
<td>0.8324</td>
</tr>
<tr>
<td>SE</td>
<td>0.0168</td>
<td>0.0071</td>
<td>0.0037</td>
<td>0.0054</td>
<td>0.0035</td>
<td>0.0032</td>
<td>0.0096</td>
<td>0.0098</td>
<td>0.0103</td>
<td>0.8231</td>
</tr>
</tbody>
</table>

106
Table 4.15: Simulation results for BVBS\textit{t}-ACD model when \(n = 10000\) and \(\rho = 0.5\)

<table>
<thead>
<tr>
<th>(\hat{\alpha}_1)</th>
<th>(\hat{\beta}_1)</th>
<th>(\hat{\gamma}_1)</th>
<th>(\hat{\alpha}_2)</th>
<th>(\hat{\beta}_2)</th>
<th>(\hat{\gamma}_2)</th>
<th>(\hat{k}_1)</th>
<th>(\hat{k}_2)</th>
<th>(\hat{\rho})</th>
<th>(\hat{\nu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1007</td>
<td>0.8995</td>
<td>0.1002</td>
<td>-0.1000</td>
<td>0.9499</td>
<td>0.1001</td>
<td>0.9996</td>
<td>1.0488</td>
<td>0.5002</td>
</tr>
<tr>
<td>Skew</td>
<td>0.1474</td>
<td>-0.1495</td>
<td>-0.0162</td>
<td>0.0890</td>
<td>-0.1908</td>
<td>0.0513</td>
<td>-0.0747</td>
<td>0.0340</td>
<td>0.0718</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0007</td>
<td>-0.0005</td>
<td>0.0002</td>
<td>0.0000</td>
<td>-0.0001</td>
<td>0.0001</td>
<td>-0.0004</td>
<td>-0.0012</td>
<td>0.0002</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0152</td>
<td>0.0064</td>
<td>0.0034</td>
<td>0.0048</td>
<td>0.0032</td>
<td>0.0029</td>
<td>0.0095</td>
<td>0.0099</td>
<td>0.0080</td>
</tr>
<tr>
<td>SE</td>
<td>0.0152</td>
<td>0.0064</td>
<td>0.0034</td>
<td>0.0049</td>
<td>0.0032</td>
<td>0.0029</td>
<td>0.0095</td>
<td>0.0099</td>
<td>0.0080</td>
</tr>
</tbody>
</table>

Table 4.16: Simulation results for BVBS\textit{t}-ACD model when \(n = 10000\) and \(\rho = 0.9\)

<table>
<thead>
<tr>
<th>(\hat{\alpha}_1)</th>
<th>(\hat{\beta}_1)</th>
<th>(\hat{\gamma}_1)</th>
<th>(\hat{\alpha}_2)</th>
<th>(\hat{\beta}_2)</th>
<th>(\hat{\gamma}_2)</th>
<th>(\hat{k}_1)</th>
<th>(\hat{k}_2)</th>
<th>(\hat{\rho})</th>
<th>(\hat{\nu})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1001</td>
<td>0.8998</td>
<td>0.1003</td>
<td>-0.1001</td>
<td>0.9499</td>
<td>0.1002</td>
<td>0.9982</td>
<td>1.0475</td>
<td>0.8998</td>
</tr>
<tr>
<td>Skew</td>
<td>0.0573</td>
<td>-0.0204</td>
<td>0.0619</td>
<td>-0.0404</td>
<td>-0.0449</td>
<td>0.0759</td>
<td>-0.0337</td>
<td>0.0265</td>
<td>0.0353</td>
</tr>
<tr>
<td>Bias</td>
<td>0.0001</td>
<td>-0.0002</td>
<td>0.0003</td>
<td>-0.0001</td>
<td>-0.0001</td>
<td>0.0002</td>
<td>-0.0018</td>
<td>-0.0025</td>
<td>-0.0002</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0109</td>
<td>0.0045</td>
<td>0.0026</td>
<td>0.0038</td>
<td>0.0023</td>
<td>0.0023</td>
<td>0.0096</td>
<td>0.0102</td>
<td>0.0020</td>
</tr>
<tr>
<td>SE</td>
<td>0.0109</td>
<td>0.0045</td>
<td>0.0026</td>
<td>0.0038</td>
<td>0.0023</td>
<td>0.0023</td>
<td>0.0094</td>
<td>0.0099</td>
<td>0.0020</td>
</tr>
</tbody>
</table>
BS-ACD models (true models), respectively, we fit the BVBS\textsubscript{t7}, BVBS\textsubscript{t9}, BVBS\textsubscript{t12} and BVBS-ACD models to the simulated datasets generated from each true model, respectively.

Table 4.17: Model Discrimination by AIC and BIC

<table>
<thead>
<tr>
<th>Fitted ACD Model</th>
<th>True BVBS\textsubscript{t}-ACD Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BVBS\textsubscript{t7} BVBS\textsubscript{t9} BVBS\textsubscript{t12} BVBS</td>
</tr>
<tr>
<td>BVBS\textsubscript{t7}</td>
<td>0.963 0.084 0.000 0.000</td>
</tr>
<tr>
<td>BVBS\textsubscript{t9}</td>
<td>0.037 0.878 0.120 0.001</td>
</tr>
<tr>
<td>BVBS\textsubscript{t12}</td>
<td>0.000 0.038 0.880 0.032</td>
</tr>
<tr>
<td>BVBS</td>
<td>0.000 0.000 0.000 0.967</td>
</tr>
</tbody>
</table>

For sample sizes \( n = 5000 \), we simulated 1000 samples from the BVBS\textsubscript{t}-ACD model with the vector of true parameters

\[
\theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)' = (0.1, 0.9, 0.1, -0.1, 0.95, 0.1, 1.0, 1.05, 0.5, \nu)'.
\]

Here, we chose the degree of freedom \( \nu \) to be 7, 9, 12 and +\( \infty \). We then assess the performance of the AIC and BIC by the selection rates for each of the fitted model.

As seen in Table 4.17, AIC/BIC can distinguish between the BVBS\textsubscript{t7}, BVBS\textsubscript{t9}, BVBS\textsubscript{t12} and BVBS-ACD models with a relatively high selection rates for the true models. AIC/BIC perform best when the true models are BVBS\textsubscript{t7}-ACD and BVBS-ACD. The selection rate of the correct model is 96.3\% if the true DGP is BVBS\textsubscript{t7}-ACD and 96.7\% if the true DGP is BVBS-ACD. AIC/BIC perform well when the true models are BVBS\textsubscript{t9}-ACD and BVBS\textsubscript{t12}-ACD. The selection rate of the right model is 87.8\% if the true DGP is BVBS\textsubscript{t9}-ACD and 88.0\% if the true DGP is BVBS\textsubscript{t12}-ACD.
As expected, when the true models are BVBS\(_t^7\)-ACD and BVBS\(_t^9\)-ACD, the selection rates for BVBS\(_t^7\)-ACD are higher than those of BVBS\(_t^{12}\)-ACD, which implies that BVBS\(_t^7\)-ACD is closer to BVBS\(_t^9\)-ACD than BVBS\(_t^{12}\)-ACD.

4.8 Illustrative example 1: Example revisited

Does the BVBS-ACD model provide the best fit for the real data in Chapter 3 or can we improve the fit using the BVBS\(_t\)-ACD model? In this section, we fit the BVBS\(_t\)-ACD model with different degree of freedom to the matched data. We choose \( \nu = 7, 9, 12 \) and \( +\infty \). The corresponding maximized log-likelihood and AIC values, for different choices of \( \nu \), are presented in Table 4.18. As \( \nu \) increases, the maximized log-likelihood value monotonically increases and the AIC value monotonically decreases. Thus, the BVBS-ACD model gives the best fit.

Table 4.18: The maximized log-likelihood value and AIC versus degrees of freedom \( \nu \).

<table>
<thead>
<tr>
<th>( \nu )</th>
<th>Maximized log-likelihood</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>2962.930</td>
<td>-5907.859</td>
</tr>
<tr>
<td>9</td>
<td>3009.919</td>
<td>-6001.838</td>
</tr>
<tr>
<td>12</td>
<td>3045.486</td>
<td>-6072.972</td>
</tr>
<tr>
<td>+( \infty )</td>
<td><strong>3103.805</strong></td>
<td>-6189.610</td>
</tr>
</tbody>
</table>

4.9 Illustrative example 2: Simulated data

In this section, we provide an example based on a simulated dataset to illustrate the BVBS\(_t\)-ACD model developed here. We compare the results of the BVBS\(_t\)-ACD
model with those of the BVBS-ACD model.

### 4.9.1 Data description

For sample size \( n = 5000 \), we generated a bivariate dataset from a BVBS\( t_9 \)-ACD model with the vector of true parameters

\[
\theta = (\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \kappa_1, \kappa_2, \rho, \nu)'
\]

\[
= (-0.007, 0.991, 0.004, -0.010, 0.991, 0.008, 1.024, 1.077, 0.2, 9)'
\]

(4.30)

As seen in Figure 4.1 and Table 4.19, the two simulated series are positive and right-skewed with higher degree of kurtosis indicating heavy tails.

Table 4.19: Summary statistics for the simulated bivariate dataset

<table>
<thead>
<tr>
<th>Data</th>
<th>Min</th>
<th>Median</th>
<th>Mean</th>
<th>Max</th>
<th>SD</th>
<th>skew</th>
<th>kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series 1</td>
<td>0.025</td>
<td>1.019</td>
<td>1.708</td>
<td>28.140</td>
<td>2.202</td>
<td>4.212</td>
<td>31.134</td>
</tr>
<tr>
<td>Series 2</td>
<td>0.018</td>
<td>1.505</td>
<td>2.692</td>
<td>57.880</td>
<td>3.590</td>
<td>4.206</td>
<td>33.626</td>
</tr>
</tbody>
</table>

In Figure 4.2, the empirical scaled TTT transforms are first concave and then convex, revealing that both marginals may possess unimodal shaped hazard rates.

So, the BVBS\( t \) distribution could be a good fit to the duration data due to the shape of its marginal density functions and hazard rates.

The ACF plots in Figure 4.3 indicate that there is a positive autocorrelation in both series, which suggests that a ACD specification may be a reasonable choice.
Figure 4.1: Histograms for the simulated bivariate dataset of Series 1 (left) and Series 2 (right)

Figure 4.2: TTT plots for the simulated bivariate dataset of Series 1 (left) and Series 2 (right)
Figure 4.3: ACF plots for the simulated bivariate dataset of Series 1 (left) and Series 2 (right)

### 4.9.2 Estimation results

We employ the method described earlier in Section 4.5 to estimate the BVBS$_t$-ACD model with the simulated bivariate dataset. The estimation results of both BVBS-ACD and BVBS$_t$-ACD models are presented in Table 4.20. All estimates are statistically significant at 1% level. The BVBS$_t$-ACD model is seen to provide much better fit to the data. The difference in AIC and BIC values between the two models are $-265.57$ and $-250.535$, respectively, which indicates that the AIC and BIC values of the BVBS$_t$-ACD model are substantially smaller than those of the BVBS-ACD model.
Table 4.20: Estimation results for two bivariate ACD models

<table>
<thead>
<tr>
<th>BVBS-ACD model</th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
<th>maxlnL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.004</td>
<td>0.997</td>
<td>0.003</td>
<td>-0.009</td>
<td>0.990</td>
<td>0.007</td>
<td>1.036</td>
<td>1.097</td>
<td>0.213</td>
<td>9.107</td>
<td>1560.22</td>
<td>-3100.44</td>
<td>-2950.096</td>
</tr>
<tr>
<td>SE</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.004</td>
<td>0.002</td>
<td>0.014</td>
<td>0.015</td>
<td>0.015</td>
<td>0.720</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BVBS-ACD model</th>
<th>( \hat{\alpha}_1 )</th>
<th>( \hat{\beta}_1 )</th>
<th>( \hat{\gamma}_1 )</th>
<th>( \hat{\alpha}_2 )</th>
<th>( \hat{\beta}_2 )</th>
<th>( \hat{\gamma}_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
<th>( \hat{\rho} )</th>
<th>maxlnL</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-0.005</td>
<td>0.998</td>
<td>0.003</td>
<td>-0.010</td>
<td>0.991</td>
<td>0.008</td>
<td>1.173</td>
<td>1.240</td>
<td>0.215</td>
<td>1426.435</td>
<td>-2834.87</td>
<td>-2699.561</td>
</tr>
<tr>
<td>SE</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td>0.003</td>
<td>0.001</td>
<td>0.012</td>
<td>0.012</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Then, the in-sample predictive model of the BVBS-ACD model is given by

\[
\hat{\sigma}_{1i} = \exp \left( -0.004 + 0.997 \ln \hat{\sigma}_{1,i-1} + 0.003 \left[ \frac{X_{1,i-1}}{\hat{\sigma}_{1,i-1}} \right] \right),
\]

\[
\hat{\sigma}_{2i} = \exp \left( -0.009 + 0.990 \ln \hat{\sigma}_{2,i-1} + 0.007 \left[ \frac{X_{2,i-1}}{\hat{\sigma}_{2,i-1}} \right] \right)
\]

while that of the BVBS-ACD model is given by

\[
\hat{\sigma}_{1i} = \exp \left( -0.005 + 0.998 \ln \hat{\sigma}_{1,i-1} + 0.003 \left[ \frac{X_{1,i-1}}{\hat{\sigma}_{1,i-1}} \right] \right),
\]

\[
\hat{\sigma}_{2i} = \exp \left( -0.010 + 0.991 \ln \hat{\sigma}_{2,i-1} + 0.008 \left[ \frac{X_{2,i-1}}{\hat{\sigma}_{2,i-1}} \right] \right).
\]

### 4.9.3 Goodness-of-fit

To evaluate the goodness-of-fit of the BVBS\(t\) and BVBS-ACD models, we investigate the in-sample one-step-ahead density forecasts implied by the predictive models. We examine the uniformity of the \(z\) series by Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests and also check the independence of the \(z\) series by Ljung-Box (LB) tests. Results of these tests are presented in Tables 4.21 and 4.22.
As seen in Table 4.21, at 10% significance level, the null hypotheses that the series \(z_1, z_2, z_{1|2}, z_{2|1}\) and the pooled series \(\{z_1, z_{2|1}\}\) and \(\{z_2, z_{1|2}\}\) are i.i.d. \(U(0,1)\) can not be rejected. The LB tests show the absence of serial correlation in the associated \(z\) and \(z^2\) series. Hence the BVBSt-ACD model provides a good fit to the data. As expected, the BVBS-ACD model doesn’t fit data well. the null hypothesis that the \(z\) series should each be i.i.d. \(U(0,1)\) are rejected.

### 4.9.4 Out-of-sample forecast evaluation

We use the first three fifth of the simulated data, “in-sample” observations 1 to 3000, to estimate the model parameters and then employ the resulting predictive model to the rest of the data, “out-of-sample” observations 3001 to 5000, to form density forecasts. These evaluation results are presented in Tables 4.23 and 4.24. A seen in Table 4.23, at 10% significance level, the null hypothesis that the series \(z_1, z_2, z_{1|2}, z_{2|1}\) and the pooled series \(\{z_1, z_{2|1}\}\) and \(\{z_2, z_{1|2}\}\) are i.i.d. \(U(0,1)\) can not be rejected. The KS and AD tests both seem to support the uniformity of the corresponding \(z\) series. The LB tests indicate no serial correlation in the associated \(z\) and \(z^2\) series. As shown in Table 4.24, at roughly 1% significance level, the null hypothesis that the series \(z_1, z_2, z_{1|2}, z_{2|1}\) and the pooled series \(\{z_1, z_{2|1}\}\) and \(\{z_2, z_{1|2}\}\) are i.i.d. \(U(0,1)\) can not be rejected. Therefore, these results all reveal that the BVBSt-ACD model yields better bivariate forecasts than the BVBS-ACD model.
Table 4.21: $P$-values for in-sample goodness-of-fit tests for the BVBS$t$-ACD model based on density forecasts

<table>
<thead>
<tr>
<th>$z$ series</th>
<th>KS</th>
<th>AD</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB$^2$(10)</th>
<th>LB$^2$(15)</th>
<th>LB$^2$(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>0.938</td>
<td>0.993</td>
<td>0.798</td>
<td>0.688</td>
<td>0.787</td>
<td>0.808</td>
<td>0.756</td>
<td>0.768</td>
</tr>
<tr>
<td>$z_2$</td>
<td>0.914</td>
<td>0.894</td>
<td>0.522</td>
<td>0.602</td>
<td>0.546</td>
<td>0.610</td>
<td>0.781</td>
<td>0.680</td>
</tr>
<tr>
<td>$z_{1</td>
<td>2}$</td>
<td>0.957</td>
<td>0.987</td>
<td>0.683</td>
<td>0.561</td>
<td>0.698</td>
<td>0.709</td>
<td>0.653</td>
</tr>
<tr>
<td>$z_{2</td>
<td>1}$</td>
<td>0.824</td>
<td>0.962</td>
<td>0.741</td>
<td>0.718</td>
<td>0.657</td>
<td>0.782</td>
<td>0.858</td>
</tr>
<tr>
<td>$z_{1</td>
<td>2}$</td>
<td>0.913</td>
<td>0.975</td>
<td>0.918</td>
<td>0.560</td>
<td>0.693</td>
<td>0.918</td>
<td>0.560</td>
</tr>
<tr>
<td>$z_{2</td>
<td>1}$</td>
<td>0.935</td>
<td>0.960</td>
<td>0.778</td>
<td>0.477</td>
<td>0.637</td>
<td>0.760</td>
<td>0.610</td>
</tr>
</tbody>
</table>

**Note:** LB($l$) represents LB test for $z$ series over $l$ lags, where $l = 10, 15$ and 20. LB$^2$($l$) represents LB test for $z^2$ series over $l$ lags.

Table 4.22: $P$-values for in-sample goodness-of-fit tests for the BVBS-ACD model based on density forecasts

<table>
<thead>
<tr>
<th>$z$ series</th>
<th>KS</th>
<th>AD</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB$^2$(10)</th>
<th>LB$^2$(15)</th>
<th>LB$^2$(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>0.007</td>
<td>0.001</td>
<td>0.819</td>
<td>0.715</td>
<td>0.791</td>
<td>0.828</td>
<td>0.782</td>
<td>0.772</td>
</tr>
<tr>
<td>$z_2$</td>
<td>0.052</td>
<td>0.004</td>
<td>0.524</td>
<td>0.597</td>
<td>0.517</td>
<td>0.590</td>
<td>0.763</td>
<td>0.647</td>
</tr>
<tr>
<td>$z_{1</td>
<td>2}$</td>
<td>0.006</td>
<td>0.001</td>
<td>0.725</td>
<td>0.601</td>
<td>0.737</td>
<td>0.739</td>
<td>0.667</td>
</tr>
<tr>
<td>$z_{2</td>
<td>1}$</td>
<td>0.022</td>
<td>0.005</td>
<td>0.749</td>
<td>0.724</td>
<td>0.655</td>
<td>0.758</td>
<td>0.858</td>
</tr>
<tr>
<td>$z_{1</td>
<td>2}$</td>
<td>0.001</td>
<td>&lt; 0.001</td>
<td>0.925</td>
<td>0.573</td>
<td>0.699</td>
<td>0.936</td>
<td>0.766</td>
</tr>
<tr>
<td>$z_{2</td>
<td>1}$</td>
<td>0.002</td>
<td>&lt; 0.001</td>
<td>0.813</td>
<td>0.492</td>
<td>0.633</td>
<td>0.809</td>
<td>0.631</td>
</tr>
</tbody>
</table>

**Note:** LB($l$) represents LB test for $z$ series over $l$ lags, where $l = 10, 15$ and 20. LB$^2$($l$) represents LB test for $z^2$ series over $l$ lags.
Table 4.23: *P*-values for out-of-sample tests for the *BVBS*-ACD model based on density forecasts

<table>
<thead>
<tr>
<th></th>
<th>KS</th>
<th>AD</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB^2(10)</th>
<th>LB^2(15)</th>
<th>LB^2(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>z_1</td>
<td>0.560</td>
<td>0.435</td>
<td>0.400</td>
<td>0.562</td>
<td>0.315</td>
<td>0.411</td>
<td>0.556</td>
<td>0.289</td>
</tr>
<tr>
<td>z_2</td>
<td>0.164</td>
<td>0.145</td>
<td>0.773</td>
<td>0.757</td>
<td>0.535</td>
<td>0.938</td>
<td>0.857</td>
<td>0.690</td>
</tr>
<tr>
<td>z_{1</td>
<td>2}</td>
<td>0.609</td>
<td>0.357</td>
<td>0.260</td>
<td>0.468</td>
<td>0.281</td>
<td>0.273</td>
<td>0.470</td>
</tr>
<tr>
<td>z_{2</td>
<td>1}</td>
<td>0.055</td>
<td>0.096</td>
<td>0.759</td>
<td>0.853</td>
<td>0.670</td>
<td>0.889</td>
<td>0.942</td>
</tr>
<tr>
<td>z_1, z_{2</td>
<td>1}</td>
<td>0.444</td>
<td>0.572</td>
<td>0.562</td>
<td>0.698</td>
<td>0.313</td>
<td>0.560</td>
<td>0.669</td>
</tr>
<tr>
<td>z_2, z_{1</td>
<td>2}</td>
<td>0.971</td>
<td>0.985</td>
<td>0.540</td>
<td>0.662</td>
<td>0.364</td>
<td>0.632</td>
<td>0.630</td>
</tr>
</tbody>
</table>

**Note:** LB(l) represents LB test for \(z\) series over \(l\) lags, where \(l = 10, 15\) and 20. LB^2(l) represents LB test for \(z^2\) series over \(l\) lags.

Table 4.24: *P*-values for out-of-sample tests for the *BVBS*-ACD model based on density forecasts

<table>
<thead>
<tr>
<th></th>
<th>KS</th>
<th>AD</th>
<th>LB(10)</th>
<th>LB(15)</th>
<th>LB(20)</th>
<th>LB^2(10)</th>
<th>LB^2(15)</th>
<th>LB^2(20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>z_1</td>
<td>0.064</td>
<td>0.028</td>
<td>0.633</td>
<td>0.746</td>
<td>0.632</td>
<td>0.750</td>
<td>0.835</td>
<td>0.706</td>
</tr>
<tr>
<td>z_2</td>
<td>0.195</td>
<td>0.062</td>
<td>0.767</td>
<td>0.763</td>
<td>0.547</td>
<td>0.944</td>
<td>0.865</td>
<td>0.706</td>
</tr>
<tr>
<td>z_{1</td>
<td>2}</td>
<td>0.035</td>
<td>0.037</td>
<td>0.445</td>
<td>0.699</td>
<td>0.675</td>
<td>0.568</td>
<td>0.833</td>
</tr>
<tr>
<td>z_{2</td>
<td>1}</td>
<td>0.082</td>
<td>0.035</td>
<td>0.700</td>
<td>0.825</td>
<td>0.661</td>
<td>0.857</td>
<td>0.936</td>
</tr>
<tr>
<td>z_1, z_{2</td>
<td>1}</td>
<td>0.013</td>
<td>0.002</td>
<td>0.594</td>
<td>0.729</td>
<td>0.407</td>
<td>0.725</td>
<td>0.796</td>
</tr>
<tr>
<td>z_2, z_{1</td>
<td>2}</td>
<td>0.016</td>
<td>0.007</td>
<td>0.573</td>
<td>0.721</td>
<td>0.466</td>
<td>0.573</td>
<td>0.721</td>
</tr>
</tbody>
</table>

**Note:** LB(l) represents LB test for \(z\) series over \(l\) lags, where \(l = 10, 15\) and 20. LB^2(l) represents LB test for \(z^2\) series over \(l\) lags.
Chapter 5

Concluding Remarks

In this chapter, we first review the research contributions of this dissertation, and then point out some directions for future research.

5.1 Summary of Research

Replacing the conditional mean in the original ACD model with the conditional median, Bhatti (2010) developed the BS-ACD model which allows an unimodal hazard rate. A recent extension of this conditional median-type ACD model has been given by Leiva et al. (2014), which considered the BS, BS power-exponential and Student-\( t \) BS distributions, so called the scale-mixture BS distribution family; see Balakrishnan et al. (2009) and Díaz-García and Leiva (2005) for details on SBS distributions. Saulo and Leão (2017) discussed the conditional median-type ACD model based on log-symmetric distributions. Saulo et al. (2017a) compared the mean-based and median-based BS-ACD models for model-fitting, forecasting and influence analysis. Moreover, their study confirmed that the conditional median-type ACD models based on BS
and associated distributions are superior to the existing ACD model (the Generalized Gamma ACD model, see Lunde (1999)) in terms of model-fitting and forecasting.

In this thesis, we have first considered the univariate BS\( t \)-ACD model and then introduced bivariate ACD models based on BVBS and BVBST distributions. The main contributions include three parts. First, we have developed the maximum likelihood estimation method for the BS\( t \)-ACD model. Next, in spite of the vast literature on duration series of single assets, the modeling of bivariate duration series has not yet received much attention due to non-synchronosity. With the univariate ACD model, we only can analyze one stock at a time and the other stock at another time. We cannot do the analysis jointly. So, the natural question is how to model them together. For this purpose, we proposed the BVBS-ACD and BVBST-ACD models based on matched trade durations. The maximum likelihood estimation of model parameters based on a two-step procedure, NM followed by BFGS, and associated inferential methods have been developed. The goodness-of-fit and predictive performance of the models have been discussed.

In Chapter 2, we have developed the maximum likelihood estimation method for the BSt autoregressive conditional duration model. We have proposed the following two-step approach for the parameter estimation. In the first step, we estimate the ACD parameters by using the NM Algorithm with \( \kappa \) and \( \nu \) fixed at their initial values. In the second step, we estimate over the whole parameter space by the BFGS algorithm. We have derived and implemented analytical gradients in the second step. The standard errors of the MLEs of the model parameters have been calculated by inverting the observed information matrix evaluated at the MLEs. We have examined the properties of the MLEs of the BSt-ACD model parameters through a Monte
Carlo simulation study. The bias, MSE and RMSE are found to be consistently small and tend to 0 as sample size increases, showing that the MLEs are asymptotically unbiased and consistent. The standard errors of the means become smaller, and the sample skewness and kurtosis are found to close to 0 and 3, respectively, with increasing sample size. This suggests the asymptotic normal distribution-sample theory of MLEs. We have also evaluated the performance of the model by model discrimination using both likelihood-based and information-based method. The results of the likelihood-ratio test show that the chi-square distribution provides only a reasonable approximation to the null distribution of the likelihood-ratio test when testing for the BS\(_{12}\)-ACD model. But, the chi-square distribution provides a good approximation to the null distribution of the likelihood ratio test when testing for the BS\(_{7}\)-ACD and BS\(_{9}\)-ACD models. However, when testing for the BS-ACD model, the mixture chi-square distribution didn’t provide a good approximation to the null distribution of the likelihood ratio test. Furthermore, we have seen that when the true model is BS\(_{7}\)-ACD and BS-ACD, the test has reasonable power to reject the other candidate models. However, when the true model is BS\(_{9}\)-ACD and BS\(_{12}\)-ACD, the test have low powers to reject BS\(_{7}\)-ACD and BS\(_{9}\)-ACD, respectively. When investigating the performance of AIC and BIC, we first note that both of them result in the same selection rates. Hence, in this case, inference can be based on either AIC or BIC. The results show that the information-based criteria perform well in discriminating between the BS\(_{7}\), BS\(_{9}\), BS\(_{12}\) and BS-ACD models and show an improvement in performance with sample size. Finally, we have illustrated the proposed methodology using real high frequency data on two stocks from the New York Stock Exchange. In general, the BS\(_{7}\)-ACD model outperform all the considered models and the BS-ACD
model provides close values, but the GG-ACD model turns out to be worse in terms of AIC values, goodness-of-fit and out-of-sample density forecasts.

In Chapter 3, we have proposed a bivariate autoregressive conditional duration model based on the bivariate BS distribution, allowing us to jointly analyze the dependence of two matched duration series. The maximum likelihood estimation of model parameters and associated inferential methods have been developed. We have done the estimation by using a hybrid of optimization algorithms, NM followed by BFGS. First, we have applied the NM Algorithm to estimate the ACD parameters by fixing $\kappa_1$, $\kappa_2$ and $\rho$ at their initial values. We have then used the BFGS algorithm to estimate over the entire parameter space. We have derived and implemented analytical gradients in the second step. The standard errors have been estimated from the negative of the inverse Hessian matrix evaluated at the MLEs. We have conducted a simulation study to evaluate the properties of the model as well as the performance of the inferential methods developed here. For different correlation coefficients, all the estimators exhibit lower Bias and RMSEs which decrease with increased sample size, tending towards 0. This empirically supports the asymptotic unbiasedness, consistency and efficiency of the MLEs. Moreover, as $n$ increases, the empirical distributions of all the estimators become close to normal distribution in terms of skewness and kurtosis. Finally, we have illustrated the proposed methodology using real high frequency data on two stocks from the New York Stock Exchange. The estimation results reveal a weak positive correlation between the two matched duration series. To justify that it is necessary to include the correlation parameter in the model, we conduct model comparison between the restricted model ($\rho = 0$) and the full model by using likelihood ratio test and AIC. The results suggest that the proposed
BVBS-ACD model is a much better choice than the restricted model. In-sample and out-of-sample density forecasts have been employed to assess the goodness-of-fit and predictive ability of the model, which supports the performance of the BVBS-ACD model developed here.

In Chapter 4, we have proposed a bivariate autoregressive conditional duration model based on the bivariate Student-\(t\) BS distribution, which facilitates the joint modeling of trade durations and detect the dependence of two matched duration series. A further advantage of the proposed model is that it allows more flexibility in terms of the kurtosis and skewness through the inclusion of an additional shape parameter. The maximum likelihood estimation of model parameters and associated inferential methods have been developed. We have proposed a two-step estimation procedure. First, we fixed the distribution parameters, \(\kappa_1, \kappa_2, \rho\) and \(\nu\), at their initial values and maximized the log-likelihood function with respect to the ACD parameters through the NM Algorithm. Next, we estimated over the full parameter space by employing the BFGS algorithm. We have derived and implemented the analytical gradients in the second step. The standard errors have been obtained by inverting the negative Hessian evaluated at the MLEs. We have conducted a simulation study to evaluate the properties of the model as well as the performance of the inferential methods developed here. A number of different scenarios have been taken into account concerning the values of correlation coefficient and sample size. All estimators perform well across different levels of correlation, demonstrating the asymptotic unbiasedness and consistency properties of the MLEs. The closeness normal approximation for the estimates is seen through skewness and kurtosis when the sample size increases. Furthermore, we have evaluated the performance of the model by model
discrimination based on information-based method. AIC/BIC can distinguish between the BVBS\(t_7\), BVBS\(t_9\), BVBS\(t_{12}\) and BVBS-ACD models with a relatively high selection rates for the true models. Finally, we have illustrated the proposed methodology by a simulated bivariate dataset and compared the results of the BVBS\(t\)-ACD model with those of the BVBS-ACD model. The suggested model provides better fit to the data in terms of AIC/BIC and in-sample density forecasts, and yields better bivariate forecast according to out-of-sample density forecasts. These show the main advantages of the BVBS\(t\)-ACD model developed here.

5.2 Future work

We now describe some research topics, that would be of interest to investigate further. First, as continuation of this work, one may develop Bayesian inference of univariate and bivariate ACD models based on BS and associated distributions, such as BS-ACD, BS\(t\)-ACD, BVBS-ACD and BVBS\(t\)-ACD models, discussed in this thesis. In Chapter 4, we proposed the BVBS\(t\)-ACD model and a two-step procedure to obtain the MLEs of model parameters. One may consider developing an EM algorithm, for this maximum likelihood estimation of model parameters.

Rahul et al. (2018) have proposed an univariate autoregressive moving average (ARMA) model based on BS distribution. One may further consider constructing a bivariate ARMA model based on BVBS and BVBS\(t\) distributions. It will then be of great interest to investigate model misspecification between ACD and ARMA models. Instead of modeling conditional expected duration in the ACD model, for bivariate or multivariate duration data, we may consider constructing the autoregressive conditional intensity Model based on BVBS, BVBS\(t\) and even multivariate BS
Another direction is to construct copula-based ACD Models based on BVBS, BVBS$t$ and multivariate BS distributions.

In Chapter 3 and 4, we have proposed bivariate ACD models with a constant covariance structure. One may introduce a time-varying covariance/correlation into the current models and develop corresponding inferential methods.

Instead of modeling matched duration data, one may construct bivariate ACD model for the two full duration series by treating one of the duration series as censored.

In this thesis, our research focuses on ACD (1,1) instead of ACD $(p, q)$ allowing $p > 1$ and $q > 1$. The general model can be defined and the structure of the model can be explained. However, the numbers of model parameters will increase. The model estimation and fitting to a given dataset will be more complicated. For example, if we take $p = 2$ and $q = 2$, i.e., BVBS-ACD (2,2), we will have 4 more ACD parameters due to the second order, in addition to 9 parameters in the BVBS-ACD (1,1) model. That is 13 model parameters in total. We need to consider how to come up with initial values and how stable the estimation procedure is, etc. Besides, we tried to apply the BVBS-ACD (2,2) model to real data. The model fitting did not improve substantially. That is why we concentrated on the BVBS-ACD (1,1) model.

Finally, the multivariate extension of the work in this thesis would pose a challenging task and would be of great interest as it would be of great practical value. When we move from a BVBS-ACD to MVBS-ACD (see Kundu, Balakrishnan, and Jamalizadeh (2013)), the model definition is not a problem. Suppose we take $p$ variate. We have to take $x_{1i}, x_{2i}, ..., x_{pi}$. But now the number of model parameters would have exponentially increased. What would be the trivariate case? We will
have 3 more ACD parameters from $\sigma_3$, $\alpha_3$, $\beta_3$ and $\gamma_3$. We will have 3 more distribution parameters, 1 shape parameter $\kappa_3$ and 2 correlation parameters $\rho_{13}$ and $\rho_{23}$. 6 more parameters have been introduced simply because we went one dimension up. That is 15 model parameters in total. One can imagine what will happen if we go to higher dimensions. We need to come up with good initial values, a stable estimation procedure and appropriate tools to validate the model.
Bibliography


Bauwens, L. and Giot, P. (2000). The logarithmic ACD model: an application to the


