The Bivariate Integer-Valued GARCH Model
A Bayesian Estimation Framework

Yuri Verges

DISSERTATION PRESENTED
TO
INSTITUTE OF MATHEMATICS AND STATISTICS
OF
UNIVERSITY OF SÃO PAULO
TO
OBTAINING THE TITLE
OF
MASTER IN SCIENCE

Program: Statistics
Advisor: Prof. Dr. Hedibert Freitas Lopes

São Paulo, Setembro de 2019

The Bivariate Integer-Valued GARCH Model

A Bayesian Estimation Framework

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 10/09/2019. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Hedibert Freitas Lopes - INSPER/IME-USP
- Prof. Dr. Ricardo Sandes Ehlers - ICMC - USP
- Profa. Dra. Verónica Andrea González-lópez - UNICAMP

Agradecimentos

Quero agradecer a minha esposa Andressa Scapim Verges, por estar ao meu lado nessa longa e demandante jornada. Eu te amo muito e sempre saiba da minha eterna gratidão pela compreensão, incentivo e apoio. Aos meus pais, agradeço pela gentileza, compreensão e exemplo que vocês sempre foram.

Ao meu orientador Hedibert Freitas Lopes, gostaria de expressar minha mais profunda gratidão. Obrigado pela disponibilidade, atenção, críticas construtivas, solidariedade e capacidade de sempre me fazer melhor. Tenho uma grande admiração pelo senhor. Aos meus professores Airlane Pereira Alencar, Florencia Graciela Leonardi, Luís Gustavo Esteves, Marcia D'Elia Branco, Miguel Natalio Abadi, Paulo Domingos Cordaro, Sergio Wechsler, Silvia Lopes de Paula Ferrari e Vanderlei da Costa Bueno, tão talentosos, capazes e atenciosos em minha jornada, deixo meus agradecimentos e carinho. Também aos professores Ricardo Sandes Ehlers e Verónica Andrea González-lópez, agradeço por terem participado de forma tão construtiva neste importante momento da minha vida, minha defesa de mestrado.

Em particular, gostaria de agradecer à nossa secretária do programa de pós-graduação, Regiane Fascina Prado Jacintho Guimarães, por sempre ser tão prestativa e cuidadosa com os alunos. Saiba da minha eterna gratidão por toda ajuda e cuidado durante meu curso.

Também agradeço à Tatiana Grecco por toda a amizade, confiança e espaço que sempre me foram depositados.

Aos meus amigos Lucas Cabral, Henrique Bolfarine e Helton Graziadei, gostaria de soubessem que sempre foi ótimo poder aprender com vocês. E, especialmente ao Frederic Brown, gostaria de agradecer pelo seu apoio, parceria e motivação constante.

Acknowledgements

I want to thank my wife Andressa Scapim Verges, for being by my side on this long and demanding journey. I love you so much and always know of my eternal gratitude for understanding, encouragement, and support. To my parents, I thank you for the kindness, understanding, and example you have always been.

To my advisor Hedibert Freitas Lopes, I would like to express my deepest gratitude. Thank you for your always availability, attention, constructive criticism, solidarity, and the ability to always make me better. I have great admiration for you. To my professors Airlane Pereira Alencar, Florencia Graciela Leonardi, Luís Gustavo Esteves, Marcia D'Elia Branco, Miguel Natalio Abadi, Paulo Domingos Cordaro, Sergio Wechsler, Silvia Lopes de Paula Ferrari, and Vanderlei da Costa Bueno, so talented, capable, and attentive in my journey, I leave my thanks and affection. Also, to the professors Ricardo Sandes Ehlers and Verónica Andrea González-lópez, I thank you for having participated in a very constructive way in this extremely important moment of my life, my master's defense.

In particular, I would like to thank our secretary of the graduation program Regiane Fascina Prado Jacintho Guimarães, for always being so helpful and careful with the students. Know of my eternal gratitude for all help and care during my course.

I also thank Tatiana Grecco for all the friendship, confidence, and space that has always been placed on me.

To my friends, Lucas Cabral, Henrique Bolfarine, Helton Graziadei, I would like to say that it has always been great to be able to learn from you. And, especially to Frederic Brown, I would like to thank you most sincerely for your always support, partnership, and motivation.

Resumo

VERGES, Y. **O Modelo Bivariado GARCH Inteiro - Uma abordagem Bayesiana**. 2019. 120 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

Uma extensa literatura tem se desenvolvido sobre dados de contagem nos últimos anos, sendo ainda, pequena a contribuição que busca a abordagem multivariada desse problema. Este trabalho têm como objetivo analisar em maior profundidade e realizar a estimação bayesiana do modelo bivariado INGARCH proposto em Cui and Zhu [2017], onde a autoregressão estudada em Liu [2012] é estendida de forma a tratar eventos negativamente correlacionados. Uma vez que a função de probabilidade proposta em Lakshminarayana *et al.* [1999] demanda certa atenção para a não infração dos axiomas da probabilidade, uma análise minuciosa sobre esta nova distribuição foi executada. Para o procedimento de estimação Bayesiana, foi aplicado o algoritmo de Random Walk Metropolis-Hastings, sendo o tuning escolhido como em uma das abordagens em Bennett *et al.* [1996]. Uma análise exhaustiva em dados simulados foi desempenhada para o real entendimento de como se comporta o modelo proposto e, objetivando a aplicação em dados reais, um estudo sobre os dados de criminalidade de Pittsburgh e outro sobre os números de negócios realizados para os contratos futuros de Euro e Libra Esterlina negociados na CME (Chicago Merchandise Exchange) foram efetuados.

Palavras-chave: Inferência Bayesiana, Problema de Dados de Contagem/GARCH Inteiro, Modelagem de Correlação Negativa.

Abstract

VERGES, Y. **The Bivariate Integer-Valued GARCH Model - A Bayesian Estimation Framework**. 20189. 120 f. Dissertação - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

An extensive literature has been developed on counting data in recent years, and the contribution that seeks the multivariate approach to this problem is still small. This paper aims to analyze in greater depth and perform the Bayesian estimation of the bivariate INGARCH model proposed in Cui and Zhu [2017], where the autoregression studied in Liu [2012] is extended to treat negatively correlated events. Since the probability function proposed in Lakshminarayana *et al.* [1999] demands some attention to the non-infringement of the probability axioms, a thorough analysis of this new distribution has been performed. For the Bayesian estimation procedure, the Random Walk Metropolis-Hastings algorithm was applied, and tuning was chosen as in one of Bennett *et al.* [1996] approaches. An exhaustive analysis on simulated data was performed for the real understanding of how the proposed model behaves, and, aiming at the application in real data, a study on the Pittsburgh crime data and another on the number of trades for the futures contracts of Euro and British Pound at traded CME (Chicago Merchandise Exchange) were implemented.

Keywords: Bayesian Inference, Counting Data Problem/INGARCH, Negative Correlation Modeling.

Contents

Symbols List	vii
List of Figures	viii
List of Tables	xi
1 Introduction	1
1.1 Background	1
1.2 Contributions and Presentation of the Dissertation	4
2 The Bivariate Poisson distribution	5
2.1 Introduction	5
2.2 The Delta Domain	7
2.3 Sampling From the Bivariate Poisson Distribution	17
2.4 Examples	18
3 A Bivariate Poisson INGARCH model	24
3.1 The Model	24
3.2 Estimation	27
3.2.1 Prior	30
4 Results	31
4.1 Simulated Data Analysis	32
4.2 Real Data Application - Positive Cross-Correlation	45
4.2.1 Number of Trades Data	45
4.3 Real Data Application - Negative cross-correlation	48
4.3.1 Artificial Data	48
4.3.2 Crime Data	51
4.3.3 Forecasting	59
5 Conclusion	60
A Markov Chain Monte Carlo	62
A.0.1 Markov Chains	62
A.0.2 The Random Walk Metropolis-Hastings Algorithm	67

B	Some Results	69
B.0.1	Demonstration - Inequality (B.1)	69
B.0.2	Demonstration - Correlation Boundary - $\tau > 0$ and $\tilde{\delta} < 0$	69
B.0.3	Demonstration - Correlation Boundary - $\tau > 0$ and $\tilde{\delta} < 0$	70
B.0.4	Demonstration - Correlation Boundary - $\tau \leq 0$ and $\tilde{\delta} < 0$	70
C	Moments Calculation	72
C.1	Characteristic Function	72
C.2	First Order Moment	72
C.3	Second Order Moment	73
C.4	Log-Likelihood Derivatives	76
C.5	First Order Derivatives	76
C.6	Second Order Derivatives	78
C.7	Additional Tables	91
C.7.1	Simulated Data - Changing the number of data points/ Changing samples burned	91
D	Random Walk Metropolis Hasting - Trace Charts	100
E	Random Walk Metropolis Hasting - ACF Plots	114
E.1	Random Walk Metropolis Hasting - $\lambda_{1,t}$, $\lambda_{2,t}$, δ_t and ρ_t evolution	132
	Bibliography	140

Symbols List

$\phi(t)_Y$	Characteristic Function of the Random Vector Y
c	Constant. $c = 1 - e^{-1}$
$\tilde{\delta}$	Constant linkage parameter
$\stackrel{\text{cid}}{=}$	Conditionally Independent
$\tilde{\delta}$	Constant Dependency Parameter
$\alpha,$	Constant that transmits the contribution of the previous Dynamic Dependency Parameter δ_{t-1} into δ_t
$\beta,$	Constant that transmits the contribution associated with the previous Data Vector Y_{t-1} into δ_t
ρ_{Y_1, Y_2}	Correlation Coefficient between Y_1 and Y_2
$\stackrel{\text{def}}{=}$	Definition
$D(\cdot)$	Dump Function
a_{ij}	Entry of the Matrix A, respective to the row i , column j
b_{ij}	Entry of the Matrix B, respective to the row i , column j
θ	Full Vector of Parameters
$F^-(\cdot)$	Generalized Inverse
$\stackrel{\text{ind}}{=}$	Independent
$B,$	Matrix that transmits the contribution of the previous data vector Y_{t-1} into λ_t
$A,$	Matrix that transmits the contribution of the previous vector of intensities λ_{t-1} into λ_t
Θ	Parameter Space
$\lambda_1, \lambda_2, \phi$	Poisson Intensities
$P_i(\cdot)$	Polynomial of Order i , $i \in \mathbb{N}$
$\mathbb{R}$	Real Line
Ω	Sample Space
$\mathbb{Z}$	Set of Integer Numbers
$\mathbb{N}$	Set of Natural Numbers
$\mathbb{Z}^+$	Set of Non-negative Integer Numbers
Δ_c	Region of validity of the linkage parameter
F	Cumulative Probability Function
F_{y_i}	Cumulative Probability Function Conditional on $Y_i = y_i$
θ_i	Vector of Parameters associated with the i 'th Poisson Rate's Autoregression
ψ	Vector of Parameters associated with the Linkage Parameter's Autoregression

List of Figures

2.1	Cui and Zhu [2017] third example of table 1 - Misspecified $\tilde{\delta}$ boundary resulting in not adequate $\tilde{\delta}$ selections	8
2.2	Δ_c proposed by Lakshminarayana <i>et al.</i> [1999]	11
2.3	Δ_c proposed by this work	11
2.4	Planes that limit where the correlation coefficient would lie between in accordance with Cui and Zhu [2017]	15
2.5	Planes that limit where the correlation coefficient would lie between in accordance with this work	15
2.6	Cui and Zhu [2017] - Table 1 - Probability density plots of the first and the second example	16
2.7	Trace Charts - Rates Without Time Dependence - M1 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	20
2.8	Trace Charts - Rates Without Time Dependence - M2 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	20
2.9	Trace Charts - Rates Without Time Dependence - M3 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	21
2.10	Trace Charts - Rates Without Time Dependence - M1 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	21
2.11	Trace Charts - Rates Without Time Dependence - Pittsburgh Crime Data - Data Points: 120. Burn In: 100000. Sample Size: 100000.	22
2.12	ACF Charts - Rates Without Time Dependence - Pittsburgh Crime Data - Data Points: 120. Burn In: 100000. Sample Size: 100000.	23
3.1	Model D - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$	25
3.2	Model D - 10000 Simulated Scenarios. $ \delta_{boundary,t} - \delta_t , \forall t \in \{1, \dots, 200\}$	26
3.3	Model DC1 - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$	26
3.4	Model DC1 - 10000 Simulated Scenarios. $ \delta_{boundary,t} - \delta_t , \forall t \in \{1, \dots, 200\}$	27
4.1	Model C - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$	34
4.2	Trace Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	38
4.3	ACF Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	39
4.4	Trace Charts - DC2 - Data Points: 500. Burn In: 5000000. Sample Size: 5000000.	40
4.5	ACF Charts - DC2 - Data Points: 500. Burn In: 5000000. Sample Size: 5000000.	41

4.6	CME Futures Number of Trades Data - Descriptive Analysis - EC1, Euro Future Contract; BP1, British Pound Future Contract	46
4.7	Model Trade - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ histograms $\forall t \in \{1, \dots, 361\}$	47
4.8	Pittsburgh Crime Data - Descriptive Analysis - ROB, Robbery; CMIS Criminal mischief	53
4.9	Trace Charts - Crime Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.	54
4.10	ACF Charts - Crime Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.	55
4.11	Trace Charts - Crime Non-Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.	56
4.12	ACF Charts - Crime Non-Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.	57
4.13	Cui and Zhu [2017] Pittsburgh Poisson Intensities given estimates. Diagonal case. $\lambda_{1,t}$ and $\lambda_{2,t}$ stands for the Robbery(ROB) and Criminal mischiefs(CMIS) respectively	58
A.1	Example: Graph diagram of a transient chain.	65
A.2	Example: Graph diagram of a irreducible chain.	65
A.3	Example: Graph diagram of an aperiodic chain.	66
D.1	Trace Charts - A1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	100
D.2	Trace Charts - A2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	101
D.3	Trace Charts - A3 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	102
D.4	Trace Charts - A4 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	103
D.5	Trace Charts - B - Data Points: 500. Burn In: 500000. Sample Size: 100000.	104
D.6	Trace Charts - C - Data Points: 500. Burn In: 500000. Sample Size: 100000.	105
D.7	Trace Charts - D - Data Points: 500. Burn In: 500000. Sample Size: 100000.	106
D.8	Trace Charts - DC1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	107
D.9	Trace Charts - E1 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	108
D.10	Trace Charts - E1 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	109
D.11	Trace Charts - E2 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	110
D.12	Trace Charts - E2 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	111
D.13	Trace Charts - Trades Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.	112
D.14	Trace Charts - Trades Non-Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.	113
E.1	ACF Charts - Rates Without Time Dependence - M2 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	115
E.2	ACF Charts - Rates Without Time Dependence - M3 - Data Points: 100. Burn In: 100000. Sample Size: 100000.	116
E.3	ACF Charts - A1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	117

E.4	ACF Charts - A2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	118
E.5	ACF Charts - A3 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	119
E.6	ACF Charts - A4 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	120
E.7	ACF Charts - B - Data Points: 500. Burn In: 500000. Sample Size: 100000.	121
E.8	ACF Charts - C - Data Points: 500. Burn In: 500000. Sample Size: 100000.	122
E.9	ACF Charts - D - Data Points: 500. Burn In: 500000. Sample Size: 100000.	123
E.10	ACF Charts - DC1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	124
E.11	ACF Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.	125
E.12	ACF Charts - E1 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	126
E.13	ACF Charts - E1 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	127
E.14	ACF Charts - E2 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	128
E.15	ACF Charts - E2 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.	129
E.16	ACF Charts - Trades Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.	130
E.17	ACF Charts - Trades Non-Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.	131

List of Tables

2.1	Bivariate Poisson Simulation - Rates Without Time Dependence - MCMC sampling size: 100000. Burn in: 100000. Data Points:100	19
2.2	Bivariate Poisson Crime Data - Rates Without Time Dependence - MCMC sampling size: 100000. Burn in: 100000. Data Points:120	23
4.1	Simulation - Parameters Configuration	32
4.2	Simulated Data - MCMC Output Statistics - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500	33
4.3	Simulated Data - MCMC Output Statistics (Cont.) - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500	35
4.4	Simulated Data - MCMC Output Statistics (Cont.) - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500	36
4.5	Simulated Data - MCMC Output Statistics - DC1 Large BurnIn - Large Sample - Data Points: 500	37
4.6	Simulated Data - MCMC Output Statistics - DC2 Large BurnIn - Large Sample - Data Points: 500	37
4.7	Simulation results for the proposed models A1, A2, A3 and A4. 200 Replications. MCMC sampling size: 100000. BurnIn: 500000.	43
4.8	Simulation results for the proposed model B. 200 Replications. BurnIn: 500000. . . .	44
4.9	Simulation results for the proposed model C. 200 Replications. BurnIn: 500000. . . .	44
4.10	Simulation results for the proposed model D. 200 Replications. BurnIn: 500000. . . .	44
4.11	Simulation results for the proposed model DC1. 200 Replications. BurnIn: 500000. . .	45
4.12	Simulation results for the proposed model DC2. 200 Replications. BurnIn: 500000. . .	45
4.13	Number Trades data - Descriptive Analysis	46
4.14	Estimates - Number of Trades Data	48
4.15	Artificial Data E1 - Descriptive Analysis	49
4.16	Artificial Data E2 - Descriptive Analysis	49
4.17	Estimates for Liu's, Cui's and proposed model - Artificial data E1	50
4.18	Estimates for Liu's, Cui's and proposed model - Artificial data E2	51
4.19	Crime data - Descriptive Analysis	53
4.20	Estimates for Liu's, Cui's and proposed model - Crime Data	54
4.21	MADE and MSE of Predicted Poisson Rates - Crime Data	59
C.1	Simulated Data A1 - MCMC Output Statistics	91
C.2	Simulated Data A2 - MCMC Output Statistics	92

C.3	Simulated Data A3 - MCMC Output Statistics	93
C.4	Simulated Data A4 - MCMC Output Statistics	94
C.5	Simulated Data B - MCMC Output Statistics	95
C.6	Simulated Data C - MCMC Output Statistics	96
C.7	Simulated Data D - MCMC Output Statistics	97
C.8	Simulated Data DC1 - MCMC Output Statistics	98
C.9	Simulated Data DC2 - MCMC Output Statistics	99

Chapter 1

Introduction

1.1 Background

As a consequence of the technological advancement of the last decades, an enormous amount of data, with different structures and characteristics, became available. This new vast amount of information foster the need for new modeling techniques to be developed, stimulating the construction of new tools for the proper analysis and treatment of this new type of information. Several sources of counting data are part of this new volume of information, e.g., the number of crimes in a particular region of a city and how its interpretation is carried out can have a severe influence in the manner that social policies are implemented. This is just an example that justifies why one can see the interest of the academy in modeling these counting type processes. [Cox *et al.* \[1981\]](#) have classified the approaches used for modeling these types of datasets in parameter-driven and observation-driven. The first concept considers the Poisson intensity evolving independent of the past data realizations, while the second one, assumes a dependence between the current Poisson intensity and the previous data realization. This work deals with the observation-driven case. Several authors have been working to propose an adequate framework to analyze this type of datasets. Regarding the analysis of univariate integer process and its properties, [Ferland *et al.* \[2006\]](#), [Fokianos *et al.* \[2009\]](#), [Franke \[2010\]](#), [Konstantinos and Tjøstheim \[2011\]](#), [Neumann \[2011\]](#), [Doukhan *et al.* \[2012\]](#), [Fokianos and Fried \[2012\]](#), [Davis and Liu \[2012\]](#), [Tjøstheim \[2012\]](#), [Blasques *et al.* \[2014\]](#), [Wang *et al.* \[2014\]](#), [Ahmad and Francq \[2016\]](#), and [Park and Jeong \[2017\]](#) have contributed analyzing and establishing several properties required for adequate inference when dealing with this type of models. [Ferland *et al.* \[2006\]](#) introduced the univariate INGARCH process as an analogous of the autoregressive conditional heteroskedastic GARCH(p,q) model, gave the necessary conditions for its existence, prove that the integer GARCH process is strictly stationary, present some convergence results and showed the second moment stationarity through a sequence of approximation processes. [Fokianos *et al.* \[2009\]](#) worked over the univariate Poisson observation driven autoregression posing some dynamics to the conditional mean in two ways: linearly and nonlinearly. The authors showed geometric ergodicity for both the observed values and the Poisson intensities using a perturbed version of the reference model, establishing its geometrical convergence and, by consequence, the consistency and asymptotic normality of the maximum likelihood estimates for the reference model. Likelihood inference is described in detail for the linear case. Additionally, this work is the first one in the literature to introduce nonlinear modeling to the evolution of rates. [Franke \[2010\]](#) showed the presence of weak dependence for a general class of

INGARCH models, under the hypothesis that the conditional mean follows a contraction with its constants summing up to a constant less than one. Doukhan and Louhichi [1999] introduced this concept. Konstantinos and Tjøstheim [2011], based on the generalized linear model theory, used a canonical link to introduce a counting process with loglinear autoregressive rates. This approach makes possible to deal with large positive and negative autocorrelation coefficients. The authors use the same technique as in Fokianos *et al.* [2009] to demonstrate that a perturbed model is geometrically ergodic and that it approaches a non-perturbed one under certain conditions. In addition, the properties of the maximum likelihood estimator were established. Neumann [2011] showed that the pair Poisson intensity and its respective data realization has a unique stationary distribution under a contraction condition. Additionally, the author establishes absolute regularity for the counting process under the used set of conditions. Doukhan *et al.* [2012] showed finiteness and stationarity for a general class of Poisson autoregressions based on the weak dependency property. The only exception presented was in the case of loglinear models, that do not present the contraction property. Differently from the perturbation technique used in Fokianos *et al.* [2009], where an approximation lemma is necessary to guarantee closeness between a perturbed and a non-perturbed model, is straight forward to show ergodicity through the weak dependency property. Besides, this framework permits to cover a large class of models. At Fokianos and Fried [2012], the authors analyze interventions for the loglinear Poisson autoregression. They propose test procedures to detect this type of outliers, implement joint maximum likelihood estimation of model parameters and outliers sizes, and derives a methodology that corrects the data when these types of deviances are detected. Davis and Liu [2012] applied the iterated random functions theory over a process that realizes from an integer univariate exponential family random variable. The authors showed that this type of process has its conditional mean process as a geometric moment contracting Markov Chain and that the observation process is absolutely regular with its coefficients presenting a geometrical decay under certain conditions. Additionally, the authors established an asymptotic theory for the maximum likelihood parameter estimates given the proposed set of assumptions. Tjøstheim [2012] did a review of the latest developments and techniques of count series autoregressions. The author addresses both the perturbation and weak dependence framework for convergence analysis and likelihood inference for estimation purposes. As types of autoregressive structures being treated are covered, the linear model, the nonlinear linking one, and the loglinear case. Blasques *et al.* [2014] considered a class of generalized autoregressive score(GAS) processes, which includes the INGARCH as a special case, and characterized the parameter's space region that guarantees the stationarity and ergodicity for the proposed process. Wang *et al.* [2014] introduced and defined the properties for self-excited threshold univariate Poisson autoregression model. Given an specific intensity level, this model has its rates behavior depending on if this threshold is crossed, resulting in a two regimes model. The stability of the self-excited rate process and maximum likelihood estimation were discussed in detail. Christou and Fokianos [2014] proposed an autoregression for negative binomial processes by suitably parameterizing its mean. In addition, probabilistic properties and the quasi-maximum likelihood inference were established for the proposed model. In Ahmad and Francq [2016], the asymptotic theory of the Poisson Quasi Maximum likelihood estimator is developed for a broad class of counting process. Park and Jeong [2017] introduced a Poisson autoregressive varying coefficient model where autocorrelation is treated through a latent process. In addition, the regression coefficients vary non-parametrically as a function of the underlying variables, through a linear

combination of a spline basis functions. The authors proposed efficient semi-parametric Bayesian inference using a collapsed Gibbs sampler for convergence improvement.

To extend the univariate case to a bivariate Poisson autoregression, first is necessary to establish the properties for a Bivariate Poisson model and later to incorporate a dynamical behavior to its intensities. Campbell [1934], Teicher [1954], and Holgate [1964] can be cited as first references for the multivariate Poisson model. Campbell [1934] derived the bivariate distribution using a generating function, Teicher [1954] extended the model to a p -dimensional one, and Holgate [1964] discuss estimation of bivariate case. The work of these authors covers only positive correlated Poisson random variables, which present a limitation when dealing with pairs that can compete, like in the case of biological populations. Lakshminarayana *et al.* [1999] proposed a formulation that can deal with a pair of negatively correlated Poisson random variables. The idea is to obtain a distribution combining two independent Poisson marginals along with a suitable function that contains a constant dependency parameter $\tilde{\delta}$. Under some restrictions over $\tilde{\delta}$ space, this combination can result in a valid probability function. Another possible approach to model this joint dependence is through the use of Copulas. Some authors propose this last approach, as is going to be seen in instants.

Regarding the dynamical behavior of bivariate Poisson intensities, Heinen and Rengifo [2007], Liu [2012], Andreassen *et al.* [2013], and Cui and Zhu [2017] have made their contribution to the field. Heinen and Rengifo [2007] applied the copula technique to propose a multivariate double Poisson autoregressive model. The authors selected a double Poisson model for the construction of each univariate autoregression, which regards the capacity of treating both under and over-dispersed data. The connection between these models was implemented through a multivariate normal copula. For robustness, the model was estimated using a t-copula, yielding to a similar result. Liu [2012] proposed an extension of the univariate autoregression to a bivariate one using a sum of Poisson random variables(trivariate reduction). Also, the necessary conditions for the intensities to have a unique stationary distribution are presented, and maximum likelihood estimation is discussed. A drawback of this technique is associated with the incapability of modeling negative cross-correlated data. Using the same formulation as in Liu [2012], Lee and Tjøstheim [2016] studied the problem of parameter change in the Bivariate Poisson INGARCH(1,1) framework. The authors verified that the maximum likelihood estimators of the model's parameters are asymptotically normal and based on this were proposed the CMLE- and residual-based CUMSUM tests for change point detection. Andreassen *et al.* [2013] discussed the modeling of bivariate time series of counts dividing the discussion into two parts. First, using the trivariate reduction technique as applied in Liu [2012], the authors showed that the stability results found before for bivariate Poisson autoregression could be generalized, substituting the Poisson distribution by a general one from the exponential family. Second, as in Heinen and Rengifo [2007], the copula framework was proposed to be an alternative for the modeling of negative cross-correlation between two Poisson autoregression. Also, were presented that stability results apply to the last framework. Cui and Zhu [2017] used the results from Lakshminarayana *et al.* [1999] to propose a process that extends Liu [2012], using a common joint random variable capable of modeling negative cross-correlated data. Also, the authors established the maximum likelihood procedure for this problem, but constraints in the space of δ are not correctly specified, resulting in an erroneous conclusion.

1.2 Contributions and Presentation of the Dissertation

This dissertation contributes to the field of counting processes/time series modeling. As main goals tackled can be pointed: a better comprehension and a precise delimitation of the constant linkage parameter δ_t proposed in [Lakshminarayana *et al.* \[1999\]](#), responsible for tying two independent Poisson random variables into a bivariate one. From adding a time index at δ_t , comes the second contribution of this work, where some dynamics to this parameter are tried to be implemented. [Cui and Zhu \[2017\]](#) propose the bivariate INGARCH with δ_t being a constant. Here, δ_t can have a constant component δ_t , one that possibly depends on its previous value, mapped through the parameter α , and other that transmits the influence of the previous data realization, mapped by the component β . This attempt try to give more flexibility to capture some dynamics in correlation. As the third contribution is presented the bayesian assessment to the estimation of this model. An exhaustive study was implemented using this framework, and its performance were compared with the results obtained in [Cui and Zhu \[2017\]](#).

The work is organized as follows. In chapter 2, we introduce the bivariate Poisson model proposed by [Lakshminarayana *et al.* \[1999\]](#) and better discuss what conditions the proposed linking parameter must respect for this model be a probability distribution. In chapter 3, the time dependence on Poisson rates is introduced as in [Cui and Zhu \[2017\]](#), but now possibly having a dynamic structure on the linking parameter. Additionally, the likelihood function was rebuilt in such a way to incorporate the restrictions imposed to the model being treated. The prior selection is made at this chapter too. At last, the chapter 4 is divided into two parts: at first, several simulation studies are performed for analysis of the model being proposed and for a better comprehension of the estimation method; Then, three applications of the model were executed: the first one, try to model the number of trades for the Euro future contract and British Pound future contract, both traded at CME - this is an example of dataset containing positive cross-correlation; a second one, that models some artificial generated data from a univariate INGARCH(1,1); and finally, the series of robbery and mischiefs from the census tract 101 of Pittsburgh were modeled, trying to give an example of a dataset containing negative cross-correlation.

Chapter 2

The Bivariate Poisson distribution

2.1 Introduction

As presented in [Holgate \[1964\]](#), one method to generate a bivariate correlated Poisson random variable is the *Trivariate Reduction*. The method consists in defining a random vector $Y = [Y_1, Y_2]$, where $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$ are the combination of the Poisson random variables, $X_1 \sim \text{Poisson}(\lambda_1 - \phi)$, $X_2 \sim \text{Poisson}(\lambda_2 - \phi)$, and $X_3 \sim \text{Poisson}(\phi)$, $\phi \in [0, \min(\lambda_1, \lambda_2)]$. This results in the following joint probability mass function

$$\begin{aligned} \mathbb{P}(Y = y) &= \mathbb{P}(Y_1 = y_1, X_3 = y_2 - X_2) \\ &= \exp(-(\lambda_1 + \lambda_2 - \phi)) \frac{(\lambda_1 - \phi)^{y_1} (\lambda_2 - \phi)^{y_2}}{y_1! y_2!} \\ &\times \sum_{k=0}^{\min(y_1, y_2)} \binom{y_1}{k} \binom{y_2}{k} k! \left(\frac{\phi}{(\lambda_1 - \phi)(\lambda_2 - \phi)} \right)^k \end{aligned} \quad (2.1)$$

Each marginal in eq.(2.1) $Y_i \sim \text{Poisson}(\lambda_i)$, $i \in \{1, 2\}$, and the covariance of the random vector Y is equal to ϕ . Since ϕ is a Poisson intensity and is limited below by zero, this incurs in the incapability of treating negative cross-correlations. There are several real counting problems where this limitation would compromise an adequate modeling, for example, biological growth models where generally variables have negative cross-correlation, or even the number of crimes in a city, where the decrease of one type of infraction, due to some public policy, can make another type of violation increase. This makes the multivariate counting problem modeling, with the flexibility to treat negative cross-correlation, very desirable. [Lakshminarayana et al. \[1999\]](#) show a method that can treat this characteristic. The idea is to combine two well known probability mass functions $\mathbb{P}(Y_i = y_i)$, $i \in \{1, 2\}$, through a function containing a linking dependency parameter $\tilde{\delta}$, which under certain constraints, results into a joint probability mass function

$$\mathbb{P}(Y_1 = y_1, Y_2 = y_2) = \mathbb{P}_1(Y_1 = y_1) \mathbb{P}_2(Y_2 = y_2) \left[1 + \tilde{\delta} \left(g_1(y_1) g_2(y_2) - E[g_1(y_1)] E[g_2(y_2)] \right) \right] \quad (2.2)$$

This proposition was first done at Pandit *et al.* [1988]. Additionally, $g_1(y_1)$ and $g_2(y_2)$ are defined over $\mathbb{R}^2$ and must be bounded functions. At Lakshminarayana *et al.* [1999], $Y_i \sim \text{Poisson}(\lambda_i)$, $i = 1, 2$, and g_i functions are a combination between a polynomial of order 0, $P_0(y_i)$, and a dump function, $D(y_i)$

$$g_i(y_i) = P_0(y_i) D(y_i) - E[P_0(Y_i) D(Y_i)] = 1 \cdot e^{-y_i} - E[e^{-Y_i}] = e^{-y_i} - e^{-c\lambda_i} \quad (2.3)$$

resulting into the joint bivariate Poisson model

$$P(Y_1 = y_1, Y_2 = y_2) = \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp\{-(\lambda_1 + \lambda_2)\} \left[1 + \tilde{\delta} \left(e^{-y_1} - e^{-c\lambda_1} \right) \left(e^{-y_2} - e^{-c\lambda_2} \right) \right] \quad (2.4)$$

Where

$$\lambda = [\lambda_1, \lambda_2]', \lambda_i \in \mathbb{R}^+, i \in \{1, 2\}$$

$$Y = [Y_1, Y_2]', Y_i \in \mathbb{Z}^+, i \in \{1, 2\}$$

and

$$c \stackrel{\text{def}}{=} 1 - e^{-1} \quad (2.5)$$

for a proper choice of the parameter $\tilde{\delta}$.

The session 2.2 discusses under what conditions this proper choice can be made. Assuming that a valid $\tilde{\delta}$ is picked and we have a valid joint probability mass function, statistics like the covariance matrix are possible of being calculated. For this purpose, at appendix C.1, we calculate the characteristic function for the bivariate Poisson model

$$\begin{aligned} \phi(t)_Y = \mathbb{E}[e^{it'Y}] &= e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} \\ &+ \tilde{\delta} \left(e^{\lambda_1(e^{it_1}-1-c)} - e^{\lambda_1(e^{it_1}-1)} \right) \left(e^{\lambda_2(e^{it_2}-1-c)} - e^{\lambda_2(e^{it_2}-1)} \right) \end{aligned} \quad (C.1)$$

where differentiating (C.1) and evaluating at $t = [0, 0]'$

$$\left. \frac{d^n}{dt^n} \phi(t)_Y \right|_{t=0} = i^n \mathbb{E}[Y^n] \quad (C.2)$$

permit to calculate the moments and the cross moments of the joint probability function (2.4). Using this result, the covariance matrix for this model is going to be

$$\text{Var}[Y] = \begin{bmatrix} \lambda_1 & \tilde{\delta} c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1 + \lambda_2)} \\ \tilde{\delta} c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1 + \lambda_2)} & \lambda_2 \end{bmatrix} \quad (2.6)$$

and, by consequence, the correlation coefficient for the bivariate Poisson model should be

$$\rho_{Y_1, Y_2} = \tilde{\delta} c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1 + \lambda_2)} \quad (\text{C.13})$$

2.2 The Delta Domain

For (2.4) be considered a probability measure, the choice of $\tilde{\delta}$ must be such that

$$1 + \tilde{\delta} \left(e^{-y_1} - e^{-c\lambda_1} \right) \left(e^{-y_2} - e^{-c\lambda_2} \right) \geq 0, \quad \forall (y_1, y_2) \in \Omega \quad (2.7)$$

as pointed in [Lakshminarayana et al. \[1999\]](#). If is chosen a $\tilde{\delta}$ that does not preserve the inequality (2.7), negative outputs from the equation (2.4) are possible of being calculated, infringing the first Kolmogorov axiom, not resulting in a probability measure. Taking this into consideration, it is possible to define the limits of the region that, for any chosen $\tilde{\delta}$, the inequality above is preserved, making the probability definition proposed by [Lakshminarayana et al. \[1999\]](#) valid. The authors argue that the boundary of $\tilde{\delta}$ must lie within the less restrictive region

$$\|\tilde{\delta}\| \leq \frac{1}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})} \quad (2.8)$$

which is not true. For example, consider the case where $\lambda_1 = 0.5$ and $\lambda_2 = 0.5$. This results that $\|\tilde{\delta}\| \leq 13.61793$. Let's select $\tilde{\delta} = -8$, which is inside the region of validity proposed by the authors. The probability calculated with (2.4), for the point $(Y_1, Y_2) = (2, 2)$, is -0.01045956, invalidating the condition (2.8) proposed by them. [Cui and Zhu \[2017\]](#) also use the boundary (2.8), making some of the obtained results for the proposed Bivariate INGARCH model not so accurate. This example is in them work, third example of table 1. See the figure 2.1 to visualize the misspecified probability function that had just been exemplified.

Figure 2.1: *Cui and Zhu [2017] third example of table 1 - Misspecified $\tilde{\delta}$ boundary resulting in not adequate $\tilde{\delta}$ selections*

Let's define the correct region of validity for $\tilde{\delta}$. Since the inequality (2.7) is extremely nonlinear, it is useful to divide the sample space Ω into smaller regions, transforming the target problem into ones that are easier to treat. Consider that $i, j \in \{1, 2\}$ and note that

$$e^{-y_i} - e^{-c\lambda_i} \geq 0 \iff y_i \leq c\lambda_i, \forall i, j \quad (2.9)$$

The case for positive and negative values of $\tilde{\delta}$ are going to be analyzed separately. For each one, let's divide the sample space into three regions: i) $y_i \leq c\lambda_i$ and $y_j \leq c\lambda_j$; ii) $y_i \leq c\lambda_i$ and $y_j > c\lambda_j$, and iii) $y_i > c\lambda_i$ and $y_j > c\lambda_j$, $i \neq j$. Then, for each case, it is possible to determine the conditions that $\tilde{\delta}$ must satisfy for (2.7) be respected on each specific region. Moreover, the set of all possible values that $\tilde{\delta}$ can assume, Δ_c , is a subset of the real line constrained by all the following restrictions.

- Case 1: $\tilde{\delta} > 0$, $y_i \leq c\lambda_i$ and $y_j \leq c\lambda_j$

$$1 + \underbrace{\tilde{\delta} \left(e^{-y_i} - e^{-c\lambda_i} \right) \left(e^{-y_j} - e^{-c\lambda_j} \right)}_{\geq 0 \forall \lambda_i, \lambda_j, y_i, y_j} \geq 0$$

In this situation any value of $\tilde{\delta}$ satisfy (2.7).

- Case 2: $\tilde{\delta} > 0$, $y_i \leq c\lambda_i$ and $y_j > c\lambda_j$

$$1 + \underbrace{\tilde{\delta} (e^{-y_i} - e^{-c\lambda_i})}_{\geq 0 \forall \lambda_i, y_i} \underbrace{(e^{-y_j} - e^{-c\lambda_j})}_{< 0 \forall \lambda_j, y_j} \geq 0$$

For this partition, $(e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j})$ is negative. Thus, the boundary of $\tilde{\delta}$ for this specific condition must be chosen such that, for the most negative value $(e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j})$, the inequality (2.7) remains valid.

$$\begin{aligned} 1 + \tilde{\delta} (e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j}) &\geq 1 + \tilde{\delta} (1 - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j}) \\ &\stackrel{y_j \rightarrow \infty}{\geq} 1 - \tilde{\delta} (1 - e^{-c\lambda_i}) e^{-c\lambda_j} \\ &\geq 0 \end{aligned}$$

This results that

$$\tilde{\delta} \leq \frac{1}{(1 - e^{-c\lambda_i}) e^{-c\lambda_j}}, \quad (2.10)$$

- Case 3: $\tilde{\delta} > 0$, $y_i > c\lambda_i$ and $y_j > c\lambda_j$

$$1 + \underbrace{\tilde{\delta} (e^{-y_i} - e^{-c\lambda_i})}_{0 < \forall \lambda_i, y_i} \underbrace{(e^{-y_j} - e^{-c\lambda_j})}_{< 0 \forall \lambda_j, y_j} \geq 0$$

As in case number 1, any positive selected $\tilde{\delta}$ in this region keeps the inequality (2.7) valid.

- Case 4: $\tilde{\delta} < 0$, $y_i \leq c\lambda_i$ and $y_j \leq c\lambda_j$. Since $(e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j})$ is positive and $\tilde{\delta}$ is negative, its selection must satisfy a condition analogous to the second one, where the multiplication of these two terms must be restricted to the region where its minimum value does not disrespected (2.7)

$$1 + \underbrace{\tilde{\delta}}_{< 0} \underbrace{(e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j})}_{\geq 0 \forall \lambda_i, y_i, \lambda_j, y_j} \geq 1 + \tilde{\delta} (1 - e^{-c\lambda_i}) (1 - e^{-c\lambda_j}) \geq 0$$

Therefore

$$\tilde{\delta} \geq \frac{-1}{(1 - e^{-c\lambda_1}) (1 - e^{-c\lambda_2})} \quad (2.11)$$

- Case 5: $\tilde{\delta} < 0$, $y_i \leq c\lambda_i$ and $y_j > c\lambda_j$. Here, both $(e^{-y_i} - e^{-c\lambda_i})$ and $(e^{-y_j} - e^{-c\lambda_j})$ are negative, resulting in a positive outcome for this multiplication. Consequently, any negative selected $\tilde{\delta}$ satisfies this case.
- Case 6: $\tilde{\delta} < 0$, $y_i > c\lambda_i$ and $y_j > c\lambda_j$. Since the terms $(e^{-y_i} - e^{-c\lambda_i})$ and $(e^{-y_j} - e^{-c\lambda_j})$ are negative, their multiplication results into a positive value. Hence, the selection of $\tilde{\delta}$ must be bounded by the condition where the most negative $\tilde{\delta} (e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j})$ does not disrespect (2.7)

$$1 + \tilde{\delta} (e^{-y_i} - e^{-c\lambda_i}) (e^{-y_j} - e^{-c\lambda_j}) \stackrel{y_i \rightarrow \infty}{y_j \rightarrow \infty} \geq 1 + \tilde{\delta} e^{-c(\lambda_i + \lambda_j)} \geq 0$$

Implying that

$$1 + \tilde{\delta} e^{-c(\lambda_i + \lambda_j)} \geq 0 \iff \tilde{\delta} \geq \frac{-1}{e^{-c(\lambda_i + \lambda_j)}} \quad (2.12)$$

Hence, $\forall y_1, y_2 \in \Omega$, a valid selection of $\tilde{\delta}$ must satisfy simultaneously the conditions (2.10), (2.11) and (2.12), resulting into

$$\Delta_c \stackrel{\text{def}}{=} \left\{ \tilde{\delta} \in \Delta : \max \left\{ \frac{-1}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})}, \frac{-1}{e^{-c(\lambda_1 + \lambda_2)}} \right\} \leq \tilde{\delta} \leq \min \left\{ \frac{1}{(1 - e^{-c\lambda_1})e^{-c\lambda_2}}, \frac{1}{(1 - e^{-c\lambda_2})e^{-c\lambda_1}} \right\} \right\} \quad (2.13)$$

the set of all values of $\tilde{\delta}$ that makes (2.4) a valid probability function for a given λ_1 and λ_2 .

For a better visualization between the difference of Δ_c proposed by Lakshminarayana *et al.* [1999] and this work, see figures 2.2 and 2.3. Conform Lakshminarayana *et al.* [1999] formulation, the Δ_c limits tend to extreme large values as λ_1 and λ_2 approximate to the origin and tend to be bounded between -1 and +1 as $\lambda_1, \lambda_2 \rightarrow \infty$. Additionally, the limiting surfaces are symmetric. Differently, the volume proposed by this work is not symmetric, it is more constrained as λ_1 and λ_2 tends to zero and, for $\lambda_1, \lambda_2 \rightarrow \infty$, Δ_c is limited below by -1 and tend to infinity for positive values of $\tilde{\delta}$.

Figure 2.2: Δ_c proposed by *Lakshminarayana et al. [1999]*

Figure 2.3: Δ_c proposed by this work

The difference between these regions is notable and has significant consequences for the correlation coefficient modeling. Consider the example that we have analyzed at the beginning of this

section, where $\lambda_1 = 0.5$ and $\lambda_2 = 0.5$. Even if it was a valid model, the correlation modeling does not connect with the proper sense of correlation due to the misspecification of Δ_c . By the definition proposed by [Lakshminarayana et al. \[1999\]](#), the absolute value of $\tilde{\delta}$ must be smaller than 13.61793, resulting by (C.13), that the model would be capable of incorrectly modeling a correlation coefficient between the interval $[-1.44, 1.44]$. [Cui and Zhu \[2017\]](#) propose to solve this problem imposing a hard limit to the range where the correlation coefficient would lie inside

$$\|\rho_{Y_1, Y_2}\| \leq \min \left\{ \frac{c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1 + \lambda_2)}}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})}, 1 \right\}$$

Differently, this job proposes that this hard limitation is not necessary since it is a consequence of a wrong definition of Δ_c . From (2.13), a valid $\tilde{\delta}$ for this example would lie inside the asymmetric interval $[-1.88, 5.06]$, and its associated range of possible correlation measures spread from -0.19 until 0.53.

Given the proposed region (2.13), it is possible to show that the correlation coefficient is naturally limited between the interval $[-1, 1]$. Moreover, the maximum possible correlation achieved is equal to the constant c in the positive side and $-c$ in the negative one. As just used in the definition of the region Δ_c , let $i \neq j$ and $i, j \in \{1, 2\}$.

For better comprehension, let's analyze first the upper limit for positive values of correlation and later the case for negative values. Let $\tilde{\delta} > 0$ and $\tilde{\delta} \in \Delta_c$. This results from (C.13)

$$\rho_{Y_i, Y_j} = \tilde{\delta} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \leq \min \left\{ \frac{1}{(1 - e^{-c\lambda_1}) e^{-c\lambda_2}}, \frac{1}{(1 - e^{-c\lambda_2}) e^{-c\lambda_1}} \right\} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \quad (2.14)$$

When $\lambda_i = \lambda_j$ and since λ_i is always greater than zero, using the Bernoulli inequality (2.14) becomes

$$\begin{aligned} \rho_{Y_i, Y_j} &\leq \frac{1}{(e^{c\lambda_1} - 1)} c^2 \lambda_i = \frac{c}{(e^{c\lambda_1} - 1)} [c\lambda_i \pm 1] \leq \frac{c}{(e^{c\lambda_1} - 1)} \left[\left(1 + \frac{c\lambda_i}{n}\right)^n - 1 \right], \forall n \in \mathbf{N} \\ &\xrightarrow{n \rightarrow \infty} \frac{c}{(e^{c\lambda_1} - 1)} = c \end{aligned} \quad (2.15)$$

For $\lambda_i > \lambda_j$, is demonstrated at appendix (B.0.1) that

$$\lambda_i > \lambda_j \iff \frac{1}{(1 - e^{-c\lambda_i}) e^{-c\lambda_j}} < \frac{1}{(1 - e^{-c\lambda_j}) e^{-c\lambda_i}} \quad (B.1)$$

By consequence, from (2.14) and using the Bernoulli inequality as in (2.15)

$$\begin{aligned} \rho_{Y_i, Y_j} &= \tilde{\delta} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \leq \frac{1}{(1 - e^{-c\lambda_i}) e^{-c\lambda_j}} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \\ &= \frac{c^2 \sqrt{\lambda_i \lambda_j}}{e^{c\lambda_i} - 1} < \frac{c}{e^{c\lambda_i} - 1} c\lambda_i < c \end{aligned}$$

Then, for positive values of $\tilde{\delta}$, the correlation coefficient is bounded by c , having the possibility of achieving equality only when $\lambda_i = \lambda_j$.

Now assume that $\tilde{\delta} < 0$. This results that the correlation coefficient is bounded inferiorly by

$$\rho_{Y_i, Y_j} = \tilde{\delta} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \geq \max \left\{ \frac{-1}{(1 - e^{-c\lambda_i})(1 - e^{-c\lambda_j})}, \frac{-1}{e^{-c(\lambda_i + \lambda_j)}} \right\} c^2 \sqrt{\lambda_i \lambda_j} e^{-c(\lambda_i + \lambda_j)} \quad (2.16)$$

Define $\tau \stackrel{\text{def}}{=} 1 - e^{-c\lambda_i} - e^{-c\lambda_j}$. If we observe that

$$\frac{-1}{(1 - e^{-c\lambda_i})(1 - e^{-c\lambda_j})} = \frac{-1}{\underbrace{1 - e^{-c\lambda_i} - e^{-c\lambda_j}}_{\tau} + e^{-c(\lambda_i + \lambda_j)}}$$

depending on the signal of τ , the nonlinearity in (2.16) is going to be activated differently. For positive values $\tau > 0$, as demonstrated in appendix (B.0.2), the inequality (2.16) becomes

$$\rho_{Y_i, Y_j} \geq \frac{-c}{\sqrt{(1 - e^{-c\lambda_i})(1 - e^{-c\lambda_j})}} \quad (B.2)$$

For $\lambda_i = \lambda_j$, only values of λ_i that respect

$$\lambda_i > c^{-1} \log 2 \quad (2.17)$$

are possible of being selected. Selecting $\lambda_i = c^{-1} \log 2 + \epsilon, \epsilon > 0$, (B.2) becomes

$$\rho_{Y_i, Y_j} \geq \frac{-c}{\sqrt{(1 - e^{\log 2 + \epsilon})(1 - e^{\log 2 + \epsilon})}}$$

and letting $\epsilon \rightarrow 0$, this results into

$$\rho_{Y_i, Y_j} \geq -c \quad (2.18)$$

At appendix (B.0.3) is proved that, for any pair of Poisson rates selected in the region defined by $\tau > 0$, the minimum correlation coefficient possible of being achieved is limited from below by $-c$ exclusive.

Analyzing the case when $\tau \leq 0$, the inequality (2.16), becomes

$$\rho_{Y_i, Y_j} \geq -c^2 \sqrt{\lambda_i \lambda_j} \quad (2.19)$$

When $\lambda_i = \lambda_j$, the selection of λ_i must be restricted to the following condition

$$1 - 2e^{-c\lambda_i} \leq 0 \iff \lambda_i \leq c^{-1} \log 2 \quad (2.20)$$

Implying that under this specific situation the correlation coefficient is limited by

$$\rho_{Y_i, Y_j} \geq -c^2 \sqrt{\lambda_i^2} = -c \log 2 \quad (2.21)$$

Moreover, at appendix (B.0.4), is showed that the correlation coefficient is going to be inferiorly bounded by $-c \log 2$ for any chosen λ_i and λ_j in the region defined by $\tau \leq 0$.

At figures 2.4 and 2.5 are plotted the possibility of correlation modeling given the Δ_c : for Cui and Zhu [2017], the volume where ρ_{Y_1, Y_2} would be valid is symmetric and is forced to be limited between -1 and 1 near $\lambda_1 = 0$ and $\lambda_2 = 0$. As these rates increase, the correlation between Y_1 and Y_2 tends to zero. Differently, the maximum positive correlation proposed by this job is of the order of +0.6 and only is possible for very small values of λ_1 and λ_2 . The minimum negative correlation has a magnitude of -0.4 and only can be achieved for Poisson rates of the order of 1. Additionally, even if $\tilde{\delta}$ have very large boundaries as the Poisson rates become very large, the correlation coefficient is limited and tends to zero. This nonlinearity and incapability of modeling a higher correlation coefficient for Poisson rates greater in magnitude is a serious issue and will be left for future work.

Figure 2.4: *Planes that limit where the correlation coefficient would lie between in accordance with Cui and Zhu [2017]*

Figure 2.5: *Planes that limit where the correlation coefficient would lie between in accordance with this work*

At the figures 2.6 were plotted the probability distributions (2.4) of the other first two examples found at table 1 in Cui and Zhu [2017]. Since these examples respect the region of validity of $\tilde{\delta}$ defined at this work by (2.13), (2.4) is a probability measure. The first example has $\Delta_c = [-2.97445, 2.622264]$ and the second one has its $\tilde{\delta}$ limits between -1.228726 and 6.956573.

Figure 2.6: Cui and Zhu [2017] - Table 1 - Probability density plots of the first and the second example

Note that even if the model is flexible sufficient to treat the positive and negative correlation cases, their magnitude is very constrained. For the first example, the minimum correlation coefficient

is equal to -0.2121, while its maximum is 0.2121. The second one, since the Poisson rates are higher and by consequence more restrictive, the minimum correlation coefficient is -0.0098, and the maximum is 0.0098.

2.3 Sampling From the Bivariate Poisson Distribution

To draw a random vector $Y = [Y_1, Y_2]$, $Y_i, i, j \in \{1, 2\}$ from the Bivariate Poisson model (2.4), this work applies the inverse transformation method over the conditional probability function (2.24). Since each marginal follows a Poisson distribution with intensity λ_i , the idea is first to draw a univariate data point from a Poisson random variable Y_i . Given this realization, together with a draw from uniform random variable U , the mapping to the target random variable can be implemented through the inverse of the conditional cumulative probability distribution function. As in Robert and Casella [2004], for a nondecreasing function F on $\mathfrak{R}$, the generalized inverse of F , F^- , is the function defined by

$$F^-(u) = \inf\{y : F(y) \geq u\} \quad (2.22)$$

Defining the random variable U following a uniform variable $\mathcal{U}(0, 1)$, then can be shown that the distribution of $Y = F^-(U)$ is going to be F . In our case, the distribution function F is the conditional cumulative distribution function F_{y_i}

$$F_{y_i}(Y_j = y_j) \stackrel{\text{def}}{=} F(Y_j = y_j | Y_i = y_i) = \sum_{l=0}^{y_j} \mathbb{P}(Y_j = l | Y_i = y_i) \quad (2.23)$$

where the conditional probability mass function for $Y_j | Y_i = y_i$ can be calculated by

$$\mathbb{P}(Y_j = y_j | Y_i = y_i) = \frac{\mathbb{P}(Y_i = y_i, Y_j = y_j)}{\mathbb{P}(Y_i = y_i)} = \frac{\lambda_j^{y_j}}{y_j!} e^{\lambda_j} \left[1 + \tilde{\delta} \left(e^{-y_i} - e^{-c\lambda_i} \right) \left(e^{-y_j} - e^{-c\lambda_j} \right) \right] \quad (2.24)$$

where $y_j \in \mathbb{Z}^+$.

The algorithm that implements the idea above is described by

Algorithm 1

- 1: **procedure** INVERSE TRANSFORMATION SAMPLING
  - 2:   Define the values of the parameters λ_i , λ_j and $\tilde{\delta}$
  - 3:   Draw an observation Y_i , which is distributed by a Poisson distribution with intensity λ_i
  - 4:   Sample an observation u from a uniform random variable $U \sim \mathcal{U}(0, 1)$
  - 5:   **if** $u < F_{y_i}(Y_j = 0)$ **then**
  - 6:   let $y_j = 0$
  - 7:   **else**
  - 8:   Do until the inequality below were not satisfied:
  - 9: $k = k + 1$
  - 10:   **if** $F_{y_i}(Y_j = k - 1) \leq u < F_{y_i}(Y_j = k)$ **then**
  - 11:   let $y_j = k$
-

Note that at steps 5 and 10 of the proposed algorithm, since it is difficult to invert the cumulative conditional probability distribution (2.24), this inversion is done iteratively given that this function is non-decreasing.

This procedure is necessary for the next section, and in chapter 4, where the analysis of simulated data is performed.

2.4 Examples

For completeness, let's analyze some examples of the Bivariate Poisson model (2.4). Consider three simulated data samples, having $n=100$ realizations each, generated with the following three setups:

M1 $\lambda_1 = 1, \lambda_2 = 0.8$ and $\tilde{\delta} = 3$

M2 $\lambda_1 = 0.5, \lambda_2 = 0.8$ and $\tilde{\delta} = -1$

M3 $\lambda_1 = 0.5, \lambda_2 = 0.8$ and $\tilde{\delta} = -2.27$

Theoretically, the three respective Δ_c limiting boundaries are $\Delta_{c,M1} = [-3.1199, 3.538978]$ and $\Delta_{c,M2} = \Delta_{c,M3} = [-2.2744, 3.4559]$, which maps to correlation coefficients delimited by $\rho_{M1} \in [-0.3573, 0.4039]$ and $\rho_{M2}, \rho_{M3} \in [-0.2527, 0.3839]$. For these three examples, the true correlation coefficient calculated (C.13) are 0.34, -0.11 and -0.25. Redefining the model (2.4) to incorporate the restrictions on $\tilde{\delta}$, the bivariate model can be rewritten as

$$P(Y_1 = y_1, Y_2 = y_2 | \lambda_1, \lambda_2, \tilde{\delta}) = \begin{cases} 0, & \text{if } \tilde{\delta} \notin \Delta_c \\ \frac{\lambda_1^{y_1} \lambda_2^{y_2}}{y_1! y_2!} \exp\{-(\lambda_1 + \lambda_2)\} \left[1 + \tilde{\delta} (e^{-y_1} - e^{-c\lambda_1}) (e^{-y_2} - e^{-c\lambda_2}) \right], & \text{otherwise} \end{cases} \quad (2.25)$$

Additionally let's add the index m to the notation of Y to represent a specific drawn, Y_m , from a sample of size n . From now on this session, Y is going to be a vector that represents a sequence of draws Y_m 's,

$Y = \{Y_m\}_{m=1}^n = \{(Y_{1,m}, Y_{2,m})\}_{m=1}^n$, where Y_m follows the probability model (2.25) just defined. By consequence, since each draw is conditionally independent on the set of parameters (2.4), the likelihood of this experiment can be written as

$$\begin{aligned} \mathbb{P}(Y = y | \lambda_1, \lambda_2, \tilde{\delta}) &= \prod_{m=1}^n \mathbb{P}(Y_m = y_m | \lambda_1, \lambda_2, \tilde{\delta}) \\ &= \prod_{m=1}^n \begin{cases} 0, & \text{if } \tilde{\delta} \notin \Delta_c \\ \frac{\lambda_1^{y_{1,m}} \lambda_2^{y_{2,m}}}{y_{1,m}! y_{2,m}!} \exp\{-(\lambda_1 + \lambda_2)\} \phi_m, & \text{otherwise} \end{cases} \end{aligned} \quad (2.26)$$

where

$$\phi_m \stackrel{\text{def}}{=} 1 + \tilde{\delta} (e^{-y_{1,m}} - e^{-c\lambda_1}) (e^{-y_{2,m}} - e^{-c\lambda_2}) \quad (2.27)$$

Also, let λ_1 , λ_2 , and $\tilde{\delta}$ follow

$$\begin{cases} \lambda_1 \sim Weibull(\lambda_{\lambda_1}, k_{\lambda_2}) \\ \lambda_2 \sim Weibull(\lambda_{\lambda_2}, k_{\lambda_2}) \\ \tilde{\delta} \sim Normal(\mu_{\tilde{\delta}}, \sigma_{\tilde{\delta}}^2) \end{cases}$$

a priori. Thus, it is possible to implement the Random Walk Metropolis-Hastings algorithm 2 to obtain estimates of the posterior distribution. An explanation with more details of this procedure is discussed depth in appendix A. The selection of these Weibulls for λ_1 and λ_2 takes into consideration the characteristics of their support and, since no knowledge is taken in advance, would be a good practice to consider a large uncertainty for them. As a consequence, was picked $k_{\lambda_i} = 1$ for the weibull's shape parameter, $i \in \{1, 2\}$, which gives flexibility to the posterior concentrates mass near zero if it is the case, and $\lambda_{\lambda_i} = 10000$, what distributes the mass of the prior into a large region, translating the large uncertainty given the absence of a specialist information. For $\tilde{\delta}$ was used a normal distribution with zero mean and a variance equal 10000, following the same concept. All convergence diagnosis and output analysis in this work were made using the package CODA in R, implemented by Plummer *et al.* [2006]. At table 2.1, is presented the descriptive analysis of the MCMC outputs and, at figures 2.7, 2.8, and 2.9 are represented the trace charts of the estimation process.

Table 2.1: *Bivariate Poisson Simulation - Rates Without Time Dependence - MCMC sampling size: 100000. Burn in: 100000. Data Points:100*

Model	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
M1	40.6	λ_1	1.00	1.04	0.04	0.10	0.32	1.04	0.85	0.97	1.04	1.11	1.25
		λ_2	0.80	0.77	-0.03	0.09	0.28	0.87	0.61	0.71	0.77	0.83	0.95
		$\tilde{\delta}$	3.00	2.38	-0.62	0.59	1.84	27.85	1.07	2.03	2.45	2.81	3.36
		ρ	0.34	0.27	-0.07	0.07	0.21	3.06	0.12	0.23	0.28	0.32	0.38
M2	39.9	λ_1	0.50	0.55	0.05	0.07	0.23	0.63	0.41	0.50	0.55	0.60	0.70
		λ_2	0.80	0.75	-0.05	0.09	0.27	0.83	0.59	0.69	0.75	0.81	0.93
		$\tilde{\delta}$	-1.00	-1.10	0.10	0.67	2.10	39.27	-2.17	-1.60	-1.16	-0.69	0.39
		ρ	-0.11	-0.12	0.01	0.07	0.24	4.33	-0.24	-0.18	-0.13	-0.08	0.04
M3	40.1	λ_1	0.50	0.55	0.05	0.07	0.23	0.70	0.42	0.50	0.55	0.60	0.71
		λ_2	0.80	0.76	-0.04	0.09	0.27	0.94	0.60	0.70	0.76	0.82	0.94
		$\tilde{\delta}$	-2.27	-1.71	0.56	0.46	1.44	20.02	-2.36	-2.05	-1.80	-1.45	-0.61
		ρ	-0.25	-0.19	0.06	0.05	0.16	2.27	-0.27	-0.23	-0.20	-0.16	-0.07

For this specific sample generated, it is possible to verify the absence of bias and the small standard deviation for the mean of λ_1 and λ_2 . This high capacity of learning is associated with the high connection that these parameters have with the data since they represent the mean of each random variable Y_1 and Y_2 , respectively. Additionally, the standard error adjusted for correlation, present in table 2.1 indicates good performance in terms of convergence for these variables during the simulation process. This convergence also can be observed at the plot of the autocorrelation function(ACF) at Figure 2.10 for the case M1. For the cases M2 and M3, the ACF plots are present at appendix E. In the case of $\tilde{\delta}$, is verifiable that its uncertainty is much higher for the three

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

cases, despite the models were able to capture some information. Also is perceivable the of lack of convergence for this variable. This is a consequence of the low elasticity between the parameter value and what actually is measured from the data, the correlation. Comparing M2 and M3, since they have the same values for λ_1 and λ_2 , a $\tilde{\delta}$ equals -1, maps into a correlation coefficient equal -0.11, while a $\tilde{\delta}$ equals -2.27 in M3, results theoretically into a correlation coefficient equal -0.25. This is such a small change in the correlation coefficient, for a large variation of $\tilde{\delta}$. We can observe this nonlinear behavior at figures 2.3 and 2.5. As a consequence, a bias appeared in this example for values that are near its boundary limits and due to its high uncertainty, as in M3. In this case, the posterior of $\tilde{\delta}$ concentrates mass near the true parameter, but also spreads its mass through the rest of the random variable's support.

Figure 2.7: Trace Charts - Rates Without Time Dependence - M1 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

Figure 2.8: Trace Charts - Rates Without Time Dependence - M2 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

Figure 2.9: Trace Charts - Rates Without Time Dependence - M3 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

Figure 2.10: Trace Charts - Rates Without Time Dependence - M1 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

In section 4.3.2, the application of Bivariate INGARCH(1,1) to a Pittsburgh’s criminality dataset is discussed in detail, having as objective compare the performance between this work’s proposed model and other authors’ propositions. Since the model with time dependence on Poisson rates is applied in the next chapters, let us estimate the model without considering time evolution for this dataset. The prior’s setup and likelihood function are the same as just used. Observing figure 2.11, is possible verify the mixing of the Markov chains and the model’s capacity of learning from the data. As in the previous examples, it is observable the issue lack of convergence for the parameter $\tilde{\delta}$, present in table 2.2. The standard error of the Monte Carlo procedure adjusted for autocorrelation is much higher for $\tilde{\delta}$ when compared to its naive calculation, indicating lack of convergence. This result also can be verified at Figure 2.12.

Figure 2.11: Trace Charts - Rates Without Time Dependence - Pittsburgh Crime Data - Data Points: 120. Burn In: 10000. Sample Size: 100000.

The parameters λ_1 and λ_2 concentrates its means around 2 and 6 respectively, in accordance with the average of each variable realizations presented in 4.8. Another observed characteristic associated with these Poisson rates is their high dispersion, possibly linked to a dynamical behavior through the time. Regarding $\tilde{\delta}$, its distribution mass concentrates at the lower boundary of its support, suggesting this parameter has a negative value, but this model is incapable of adequate capture this characteristic of the data. For average Poisson rates of 2 and 6, their associated Δ_c is represented by the interval $[-1.4257, 3.6220]$, and the respective region where the correlation can be modeled is $[-0.012, 0.03]$. This suggests that the results presented by Cui and Zhu [2017], related to their model being able to treat the case of negative correlation, is not adequately presented. The high level of average Poisson rates is the cause, which makes this correlation modeling region narrow. At table 2.2 is presented the descriptive analysis of the MCMC procedure just discussed.

Table 2.2: *Bivariate Poisson Crime Data - Rates Without Time Dependence - MCMC sampling size: 100000. Burn in: 100000. Data Points:120*

Parameter	Mean	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
λ_1	2.09	0.13	0.42	1.60	1.84	2.01	2.09	2.18	2.36
λ_2	6.06	0.23	0.72	4.13	5.62	5.91	6.06	6.21	6.52
$\tilde{\delta}$	0.01	1.15	3.64	103.30	-1.36	-0.93	-0.28	0.71	2.78
ρ	0.00	0.01	0.03	0.85	-0.01	-0.01	0.00	0.01	0.02

Figure 2.12: *ACF Charts - Rates Without Time Dependence - Pittsburgh Crime Data - Data Points: 120. Burn In: 100000. Sample Size: 100000.*

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Chapter 3

A Bivariate Poisson INGARCH model

3.1 The Model

Consider the formulation of the INGARCH(1,1) based on the bivariate Poisson (2.4) proposed in Cui and Zhu [2017]. As in the reference work, denote $Y_t = [Y_{1,t}, Y_{2,t}]$, $Y_{i,t} \in \mathbb{Z}^+$, $i \in \{1, 2\}$, the bivariate observation at time t , $t \geq 1$, and $\lambda_t = [\lambda_{1,t}, \lambda_{2,t}]'$, $\lambda_{t,i} \in \mathbb{R}^{+*}$, $i \in \{1, 2\}$, the Poisson rates of each time series under consideration. The proposed model is defined as follows

$$\begin{cases} Y_t | \mathcal{F}_{t-1} \sim BP(\lambda_t, \tilde{\delta}) \\ \lambda_t = \omega + A\lambda_{t-1} + BY_{t-1} \end{cases} \quad (3.1)$$

where $\mathcal{F}_t = \sigma\{\lambda_0, Y_0, \dots, Y_t\}$, ω is a bivariate vector with entries greater than zero, A and B are two matrices with nonnegative entries a_{ij} and b_{ij} respectively. The vector ω represents the constant contribution in the Poisson rate, A is responsible for the effect of the previous rate λ_{t-1} into the current one, and B maps the feedback effect of the previous data realization into the present rate. The model (3.1) can be extended to incorporate dynamics into the parameter $\tilde{\delta}$, which would give flexibility for modeling the case in which the correlation parameter has some dynamics embedded. Let's suppose that $\tilde{\delta}$ in model (3.1) gives space for a parameter that evolves through the time, δ_t . This results in the dynamic bivariate INGARCH(1,1) model, which is defined as

$$\begin{cases} Y_t | \mathcal{F}_{t-1} \sim BP(\lambda_t, \delta_t) \\ \lambda_t = \omega_t + A\lambda_{t-1} + BY_{t-1} \\ \delta_t = (1 - \alpha - \beta)\tilde{\delta} + \alpha\delta_{t-1} + \beta u'_{t-1}u_{t-1} \end{cases} \quad (3.2)$$

given $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$, and

$$u_t \stackrel{\text{def}}{=} \begin{bmatrix} \frac{1}{\lambda_{1,t}^{1/2}} & 0 \\ 0 & \frac{1}{\lambda_{2,t}^{1/2}} \end{bmatrix} \begin{bmatrix} y_{1,t} - \lambda_{1,t} \\ y_{2,t} - \lambda_{2,t} \end{bmatrix} = \begin{bmatrix} \frac{y_{1,t} - \lambda_{1,t}}{\lambda_{1,t}^{1/2}} \\ \frac{y_{2,t} - \lambda_{2,t}}{\lambda_{2,t}^{1/2}} \end{bmatrix} \quad (3.3)$$

Analogous the autoregression for Poisson rates, δ_t is proposed to be a composition of three components: one that maps a constant contribution, $(1 - \alpha - \beta)\tilde{\delta}$, other that translates the effect of the previous δ_t into the current one, $\alpha\delta_{t-1}$, and a third one that tries to capture the impact that the last data point can have over δ_t , mapped by $\beta u'_{t-1}u_{t-1}$.

If α and β are equal zero, the model (3.1) is recovered, where the dependence parameter is the constant $\tilde{\delta}$. This does not imply that the correlation coefficient is constant under this situation since (C.13) is a function of $\lambda_{1,t}$ and $\lambda_{2,t}$, and they have a dynamic behavior over time. Actually it oscillates, having its magnitude depending on the values $\lambda_{1,t}$ and $\lambda_{2,t}$. For example, see Figure 3.1 related to the example D studied in next chapter. In this case, the elements of the matrices A and B are nonnegative greater than zero, with δ_t being equal $\tilde{\delta}$. For 10000 draws of the parameters during the MCMC sampling were plotted, for each time instant t , the box plot for $\lambda_{1,t}$, $\lambda_{2,t}$, and ρ_t . It is possible to verify the median of the correlation box plots having a dynamic behavior, a consequence of the Poisson rates evolving.

Figure 3.1: *Model D - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$*

Another verified characteristic is the dispersion of the correlation coefficient varying with the time. This occurs because the region of validity of δ_t , Δ_c , is changing as $\lambda_{1,t}$ and $\lambda_{2,t}$ are evolving. For higher levels of $\lambda_{1,t}$ and $\lambda_{2,t}$, smaller are the possible values that δ_t can assume and, by consequence, smaller is the region that a valid correlation can lie inside. On the opposite side, for smaller values of Poisson rates, larger is this region of validity of the parameter δ_t . At Figure 3.2 is possible to verify this behavior. For every sample of parameters generated during the MCMC simulation, is plotted the differences between the modulus of the boundary closest to δ_t and the modulus of δ_t , for every time instant t . For example, if δ_t is negative, the point plotted is going to be the difference between the absolute value of its inferior boundary to its current value. Now, if δ_t is positive, the point plotted is going to be the difference between its upper limit to its current value. The idea in this plotting is to observe that, for every t during this evolution, the current δ_t does not infringe the limits imposed by its inferior and superior boundaries. Since these limits are not symmetrical, as

observed in Figure 2.5, and we have both positive and negative values during the sampling process, then is possible to observe two different behaviors in the figure: one associated with the positive values of δ_t and other to the negative ones. Additionally, since now for each t we have a different Δ_c , consequence of the evolution of $\lambda_{1,t}$ and $\lambda_{2,t}$, this boundary will gain the index t , starting to be notated as $\Delta_{c,t}$.

Figure 3.2: *Model D - 10000 Simulated Scenarios.* $|\delta_{boundary,t}| - |\delta_t|, \forall t \in \{1, \dots, 200\}$

Figure 3.3: *Model DC1 - 100000 Simulated Scenarios.* $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$

The just presented example can be contrasted against one where δ_t is not constant. At this

other situation, the combination between the parameters α , β , and $\tilde{\delta}$ are responsible to configure a dynamic behavior to δ_t . As an example, see Figure 3.3, where the box plots of $\lambda_{1,t}$, $\lambda_{2,t}$, ρ_t , and δ_t are plotted for the simulated case DC1, that is going to be viewed with more detail in the next chapter. In this case, a dynamic behavior is present on the evolution of ρ_t , consequence of the dynamic structure imposed to δ_t , and also by the evolution in Poisson rates. Another characteristic perceivable is that, since the magnitude of $\lambda_{1,t}$ and $\lambda_{2,t}$ assume smaller values, the region where δ_t evolves are broader, implicating in a more flexible space for modeling the correlation coefficient.

As in the previous given example where δ_t were constant, at Figure 3.4 were plotted the differences between the modulus of the boundary closest to δ_t , and the modulus of δ_t , for every time instant t , for this other example. It is observable that the upper and inferior boundaries for δ_t are respected in this other example.

Figure 3.4: *Model DC1 - 10000 Simulated Scenarios.* $|\delta_{boundary,t}| - |\delta_t|, \forall t \in \{1, \dots, 200\}$

3.2 Estimation

Since (3.1) is a particular case of (3.2), let's write the problem in terms of the more general one. Consider $Y = \{Y_t\}_{t=1}^n = \{(Y_{1,t}, Y_{2,t})\}_{t=1}^n$, a sequence of Poisson bivariate random variables, $\lambda = \{\lambda_t\}_{t=1}^n = \{(\lambda_{1,t}, \lambda_{2,t})\}_{t=1}^n$, them associated sequence of intensities, and $\delta = \{\delta_t\}_{t=1}^n$ the vector containing the temporal evolution of the dependency parameter. Additionally, let's define the parameter vector $\theta \stackrel{\text{def}}{=} [\theta'_1, \theta'_2, \psi]'$, where $\theta_i = [\omega_i, a_{ij}, b_{ij}, \lambda_{0,i}]$, $i, j \in \{1, 2\}$ and $\psi = [\tilde{\delta}, \alpha, \beta, \delta_0]'$. The joint probability mass function for a valid sequence of bivariate Poisson draws Y can be calculated by

$$\mathbb{P}(Y = y, \Lambda = \lambda, \Delta = \delta | \theta, y_0) = \prod_{t=1}^n \left[\mathbb{P}(Y_t = y_t | \lambda_t, \delta_t) \mathbb{P}(\Delta_t = \delta_t | \psi, \delta_{t-1}, y_{t-1}) \mathbb{P}(\Lambda_t = \lambda_t | \theta, \lambda_{t-1}, y_{t-1}) \right] \quad (3.4)$$

Since given y_0 and θ , λ_t and δ_t propagate deterministically in this formulation.

$$\mathbb{P}(\Delta_t = \delta_t | \psi, \delta_{t-1}, y_{t-1}) = 1 \quad (3.5)$$

and

$$\mathbb{P}(\Lambda_t = \lambda_t | \theta, \lambda_{t-1}, y_{t-1}) = 1 \quad (3.6)$$

By consequence

$$\begin{aligned} \mathbb{P}(Y = y, \Lambda = \lambda, \Delta = \delta | \theta, y_0) &= \prod_{t=1}^n \frac{\lambda_{1,t}^{y_{1,t}} \lambda_{2,t}^{y_{2,t}}}{y_{1,t}! y_{2,t}!} \exp\{-(\lambda_{1,t} + \lambda_{2,t})\} \\ &\quad \times \left[1 + \delta_t \left(e^{-y_{1,t}} - e^{-c\lambda_{1,t}} \right) \left(e^{-y_{2,t}} - e^{-c\lambda_{2,t}} \right) \right] \\ &= \prod_{t=1}^n \frac{\lambda_{1,t}^{y_{1,t}} \lambda_{2,t}^{y_{2,t}}}{y_{1,t}! y_{2,t}!} \exp\{-(\lambda_{1,t} + \lambda_{2,t})\} \phi_t \end{aligned} \quad (3.7)$$

where

$$\phi_t \stackrel{\text{def}}{=} 1 + \delta_t \left(e^{-y_{1,t}} - e^{-c\lambda_{1,t}} \right) \left(e^{-y_{2,t}} - e^{-c\lambda_{2,t}} \right) \quad (3.8)$$

recovering the likelihood function as in [Cui and Zhu \[2017\]](#). We say valid since, as discussed at section 2.2, for each time instant t , δ_t must lie inside the interval $\Delta_{c,t}$ defined by $\lambda_{t,1}$ and $\lambda_{t,2}$. If $\exists t \in \{1, \dots, n\} : \delta_t \notin \Delta_{c,t}$, the probability calculated using equation (2.4) will lose its meaning for that specific instant t , implicating in the lost of sense of the joint probability function (3.7). To workaround this situation, we can define an auxiliary variable which indicates if δ_t is a valid one for each t

$$\mathcal{I}_{\delta_t} = \begin{cases} 1, & \text{if } \delta_t \in \Delta_{c,t} \\ 0, & \text{otherwise} \end{cases} \quad (3.9)$$

resulting in the random vector $\mathcal{I}_{\delta} = [\mathcal{I}_{\delta_1}, \dots, \mathcal{I}_{\delta_n}]'$, which specifies the validity of the sequence of bivariate Poisson draws Y . Then, the joint probability function (3.11) can be considered incomplete and be rewritten as

$$\mathbb{P}(Y = y, \Lambda = \lambda, \Delta = \delta, | \theta, y_0) \rightarrow \mathbb{P}(Y = y, \Lambda = \lambda, \Delta = \delta, \mathcal{I}_{\delta} = i_{\delta} | \theta, y_0)$$

defining a valid specification for the likelihood function that takes in consideration the evolution of the parameter δ_t . As a consequence, if $\exists t \in \{1, \dots, n\} : \delta_t \notin \Delta_{c,t}$, this is going to make the joint likelihood function have its value equals to zero, adequately treating the problem of the validity of δ_t for the entire sample.

The unconditional probability of the random variable $\mathcal{I}_{\delta_t}$ is difficult to calculate. However, its conditional probability given λ_t and δ_t is easily obtained by

$$\mathbb{P}(I_{\delta_t} = i_{\delta_t} | \delta_t, \lambda_t) = \begin{cases} 1, & \text{if } \delta_t \in \Delta_{c,t} \\ 0, & \text{otherwise} \end{cases} \quad (3.10)$$

Thus, the likelihood function can be properly written as

$$\begin{aligned} \mathbb{P}(Y = y, \Lambda = \lambda, \Delta = \delta, \mathcal{I}_\delta = i_\delta | \theta, y_0) &= \prod_{t=1}^n \left[\mathbb{P}(Y_t = y_t | \lambda_t, \delta_t) \mathbb{P}(\mathcal{I}_{\delta_t} = i_{\delta_t} | \lambda_t, \delta_t) \right. \\ &\quad \left. \times \mathbb{P}(\Delta_t = \delta_t | \psi, \delta_{t-1}, y_{t-1}) \mathbb{P}(\Lambda_t = \lambda_t | \theta, \lambda_{t-1}, y_{t-1}) \right] \\ &= \begin{cases} 0, & \text{if } \exists t \in \{1, \dots, n\} : \delta_t \notin \Delta_{c,t} \\ \prod_{t=1}^n \frac{\lambda_{1,t}^{y_{1,t}} \lambda_{2,t}^{y_{2,t}}}{y_{1,t}! y_{2,t}!} \exp\{-(\lambda_{1,t} + \lambda_{2,t})\} \phi_t, & \text{otherwise} \end{cases} \end{aligned} \quad (3.11)$$

Similar to the adjustment implemented to the likelihood function, an augmentation to the prior's parameter space is necessary to incorporate an indicator variable, as defined in (3.10), since $\lambda_{1,0}$, $\lambda_{2,0}$ and δ_0 were treated as parameters. Let π be the prior probability density of the parameter vector θ . Then, it can be rewritten as

$$\pi(\Theta = \theta) \rightarrow \pi(\Theta = \theta, I_{\delta_0} = i_{\delta_0}) \quad (3.12)$$

For simplicity, let us shorten the notation as

$$\pi(\Theta = \theta, I_{\delta_0} = i_{\delta_0}) = \pi(\theta, i_{\delta_0})$$

Via conditioning, (3.12) can be written as

$$\pi(\theta, i_{\delta_0}) = \pi(i_{\delta_0} | \delta_0, \lambda_0) \pi(\delta_0 | \lambda_0) \pi(\theta_1, \theta_2, \psi \setminus \delta_0) \quad (3.13)$$

For simplicity, this work fixed $\lambda_{1,0}$ and $\lambda_{2,0}$ equal to 0.1, values that are inside the support of each one of these variables. Since this model structure was proved to converge, as found in Liu [2012], the initial choice of these parameter values is not that important since it is expected the convergence of the Poisson rates after some realizations. Also, δ_0 was fixed for a value that lies inside its support definition. For δ_t exhibiting some dynamics, δ_0 was fixed equals to 0, since it is expected its convergence after some realizations. When δ_t were constant, δ_0 was also set equal to 0, since we are interested only in the behavior of $\tilde{\delta}$ in these cases. A natural choice, given the modeling proposed, would treat these parameters as random variables with significant variances. However, since the number of parameters being modeled is not that small, and we are not interested in the behavior of the model at t_0 , this author preferred to have them fixed.

As a consequence, $\pi(\delta_0 | \lambda_0) = 1$, and $\pi(i_{\delta_0} | \delta_0, \lambda_0) = 1$, resulting in

$$\pi(\theta, i_{\delta_0}) = \pi(\theta_1, \theta_2, \psi \setminus \delta_0) \quad (3.14)$$

By Bayes theorem, the posterior for (3.1) can be calculated by

$$\begin{aligned} \pi(\theta|y, \lambda, \delta, i_{\delta}, y_0) &\propto \mathbb{P}(y, \lambda, \delta, i_{\delta}|\theta, y_0) \pi(\theta, i_{\delta_0}) \\ &= \begin{cases} 0, & \text{if } \exists t \in \{1, \dots, n\} : \delta_t \notin \Delta_{c,t} \\ \left[\prod_{t=1}^n \frac{\lambda_{1,t}^{y_{1,t}} \lambda_{2,t}^{y_{2,t}}}{y_{1,t}! y_{2,t}!} \exp\{-(\lambda_{1,t} + \lambda_{2,t})\} \phi_t \right] \times \pi(\theta_1, \theta_2, \psi \setminus \delta_0), & \text{otherwise} \end{cases} \end{aligned} \quad (3.15)$$

3.2.1 Prior

Since the restrictions to assure the validity of (2.4) are treated by each $\mathcal{I}_{\delta_t}$ in the likelihood function, the prior elicitation only needs to concern with the natural constraints imposed by the model (3.2). By construction, α and β are dependent on each other. This is a consequence of their support definition, since $\alpha + \beta < 1$, $\alpha \geq 0$, and $\beta \geq 0$, implying that the effect of one over the other must be taken in account. All the other parameters do not require to be treated as dependent, since there are no hard ties into them support, like in the case of α and β . As a consequence, (3.14) become

$$\pi(\theta) \stackrel{ind}{=} \left[\prod_{i=1}^2 \pi(\omega_i) \right] \left[\prod_{j=1}^2 \prod_{i=1}^2 \pi(a_{i,j}) \right] \left[\prod_{j=1}^2 \prod_{i=1}^2 \pi(b_{i,j}) \right] \pi(\tilde{\delta}) \pi(\alpha, \beta) \quad (3.16)$$

Taking the support of each parameter in account, this work uses the following specification for each density in the prior (3.16)

$$\begin{cases} \omega_i \sim Weibull(\lambda_{\omega_i}, k_{\omega_i}) \\ a_{ij} \sim Weibull(\lambda_{a_{ij}}, k_{a_{ij}}) \\ b_{ij} \sim Weibull(\lambda_{b_{ij}}, k_{b_{ij}}) \\ \tilde{\delta} \sim Normal(\mu_{\tilde{\delta}}, \sigma_{\tilde{\delta}}^2) \\ (\alpha, \beta, 1 - \alpha - \beta) \sim Dirichlet(\nu_{\alpha}, \nu_{\beta}, \nu_{1-\alpha-\beta}) \end{cases} \quad (3.17)$$

where that $i \in \{1, 2\}$ and $j \in \{1, 2\}$. It is important to mention that, if any of the parameters above is fixed, its marginal probability distribution is going collapse to that fixed point, reducing the dimensionality of (3.16).

Chapter 4

Results

This chapter objectives to compute the performance of the Bayesian estimation of the parameters of the Bivariate INGARCH(1,1) proposed in chapter 3. First, a simulation study having as reference the work of Cui and Zhu [2017], was performed. For the simulated data cases implemented by the reference author, the same was executed in this work. Instead of using the MLE framework, the model parameters have their posterior distribution determined by the Random Walk Metropolis-Hastings Algorithm. The properly comparisons to the reference work also were executed. Additionally, simulated data were generated for the model's case where δ_t has some dynamic structure embedded. Second, a real data example modeling positive cross-correlation was treated. The number of trades of Euro Futures and British Pound Futures traded at CME was collected at Reuters Datascope and used as dataset in this section. The justification of not using the same as in Cui and Zhu [2017] is related to the incapability of this author in finding the same dataset at the NYSE website. They have available several data samples, but there is no Trades and Quotes(TAQ) database publicly regarding the same selected time frame. Third, two cases modeling negative cross-correlation were implemented as in Cui and Zhu [2017]. Like the first one, an artificial data example using an INGARCH(1,1) was generated and modeled. Secondly, the same crime dataset from Pittsburgh was obtained and treated using the Bayesian framework. Properly comparisons have been made. At Cui and Zhu [2017], for each application, the author have compared the Bivariate INGARCH(1,1) using the probability function (2.1), first proposed by Liu [2012], with the one implemented at the job. Both these estimations were realized by Cui and Zhu [2017]. For the comparison of applications with the ones in reference work, the reference results were copied for comparison purposes.

In addition, for each study analyzed in this chapter, were taken into consideration a prior in the form of (3.17), having its parameters incorporating a large uncertainty as done in section 2.4. For the Weibull part of (3.17), it was considered a shape parameter k equal 1 and a scale equal 10000. The parameter $\tilde{\delta}$ was considered to follow a normal distribution with zero mean, and variance equals 10000. The parameters $\nu_{1-\alpha-\beta}$, ν_α , and ν_β of the Dirichlet distribution associated with α , β and $1 - \alpha - \beta$ were taken equal 1, making the probability mass be uniformly distributed over the simplex where it was defined. Additionally, for the applications, only the model without a dynamic structure imposed to δ_t were implemented, since will be observed at the simulated data section that this imposition possibly generates a problem of confusion in the model.

4.1 Simulated Data Analysis

This first study implemented has as objective evaluate the performance of Bayesian estimation of Bivariate INGARCH(1,1) model. This assessment was divided into two parts: First, the evaluation of posteriors convergence was executed. For the same parameters configurations proposed in Cui and Zhu [2017], the analysis in terms of data points number, and in terms of the number of samples needed to be burned for convergence were implemented. Secondly, a replication study of several simulated datasets for each given parameter configuration was treated, objecting to analyze the performance of estimates.

Consider the configurations of vector $\theta = [\theta_1, \theta_2, \psi]$, where $\theta_i = [\omega_i, a_{ii}, a_{ij}, b_{ii}, b_{ij}, \lambda_{0,i}]'$, $i \neq j$, and $\psi = [\tilde{\delta}, \alpha, \beta, \delta_0]'$ displayed in table 4.1

Table 4.1: *Simulation - Parameters Configuration*

	Parameter	A1	A2	A3	A4	B	C	D	DC1	DC2
θ_1	ω_1	1.0	0.3	0.5	0.5	0.3	0.5	0.5	0.3	1.0
	a_{11}	0.4	0.2	0.1	0.6	0.2	0.2	0.3	0.3	0.1
	a_{12}						0.3	0.2		0.1
	b_{11}	0.2	0.5	0.8	0.1	0.5	0.2	0.1	0.2	0.1
	b_{12}					0.1		0.2		0.1
	λ_{01}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
θ_2	ω_2	0.5	0.5	0.5	0.3	0.5	0.3	0.3	0.2	0.1
	a_{22}	0.3	0.4	0.2	0.8	0.4	0.2	0.3	0.2	0.3
	a_{21}						0.4	0.1		0.5
	b_{22}	0.4	0.3	0.7	0.1	0.2	0.3	0.2	0.2	0.2
	b_{21}					0.3		0.3		0.1
	λ_{02}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
δ	$\tilde{\delta}$	0.5	0.7	-1	-0.5	0.7	-0.5	0.4	-1.0	1.0
	α								0.2	0.1
	β								0.2	0.4
	δ_0								-1.0	1.0

For each set proposed, a vector of 500 data points were generated. Then, estimation of model (3.2) were executed considering the first 50, 100, 200 and 500 data points. Additionally, the cases with a burn-in of 15000 and 500000 were taken into consideration for convergence assessment. Tables 4.2 and 4.3 present the situation where 500000 samples were burned. Other tables related to this part of the study can be found in appendix C.7.1.

At table 4.2 is illustrated the case where the matrix A and B are diagonal, which implies the absence of crossed effects from the previous Poisson rates and data realization respectively. It is possible to verify that the posteriors distributions present reasonable estimates, with a mean close to the actual value used to generate the datasets. As already observed in the case without time evolution, the standard deviation for the parameter $\tilde{\delta}$, are higher compared to uncertainty present in the posterior of other parameters, translating the lack of a direct connection between this parameter and the data, and also the complex nonlinear behavior of the geometry of space where δ_t is defined.

Another behavior that can be observed is present in the tables contained at C.7.1. A burn-in of 15000 was sufficient, in most cases, to the model being capable of learning the information contained in the data. Also, the sample size of 500 is too large and only helps to decrease the uncertainty of the posterior mean with marginal decreasing enhancements. This can be observed in the decreasing values in the standard deviation column. Another characteristic to be mentioned is that samples of size equal 50 were already capable of learning about the real parameters. About the chain's convergence, it is possible to observe the difficulty by the model in all cases to converge due to the high level of autocorrelation present in all autocorrelation plots, also seen in the standard errors adjusted by this effect.

A case that claims attention is A3, where the posterior average mean of $\tilde{\delta}$ has no connection with the actual value used for generating this respective dataset. This bias can have its cause illustrated at figure 4.1. During the evolution of the Poisson rates, exist some instants where $\lambda_{1,t}$ and $\lambda_{2,t}$ reach approximately 20 and 10 respectively. This leads to a $\Delta_{c,t} = [-1.001, 556.244]$, imposing a very constrained inferior barrier to the parameter δ_t . Besides, these higher Poisson rates make the correlation coefficient at these moments to collapse near zero. The associated interval that ρ_t can

Table 4.2: *Simulated Data - MCMC Output Statistics - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500*

Model	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
A1	39.4	ω_1	1.00	1.10	0.10	0.26	0.80	16.26	0.67	0.92	1.08	1.25	1.70
		a_{11}	0.40	0.37	-0.03	0.11	0.33	6.67	0.14	0.30	0.37	0.44	0.56
		b_{11}	0.20	0.22	0.02	0.04	0.12	2.04	0.15	0.19	0.22	0.25	0.30
		ω_2	0.50	0.56	0.06	0.10	0.30	6.37	0.38	0.49	0.56	0.62	0.76
		a_{22}	0.30	0.20	-0.10	0.08	0.23	3.97	0.05	0.15	0.20	0.25	0.35
		b_{22}	0.40	0.43	0.03	0.04	0.14	1.20	0.34	0.40	0.43	0.46	0.52
		$\tilde{\delta}$	0.50	0.61	0.11	0.51	1.61	28.44	-0.46	0.28	0.63	0.97	1.53
A2	38.2	ω_1	0.30	0.34	0.04	0.06	0.17	1.72	0.24	0.30	0.34	0.38	0.46
		a_{11}	0.20	0.18	-0.02	0.06	0.19	2.21	0.06	0.14	0.18	0.22	0.30
		b_{11}	0.50	0.54	0.04	0.05	0.15	1.74	0.45	0.51	0.54	0.57	0.63
		ω_2	0.50	0.56	0.06	0.12	0.38	4.75	0.35	0.47	0.55	0.63	0.82
		a_{22}	0.40	0.31	-0.09	0.09	0.29	4.04	0.12	0.25	0.31	0.37	0.48
		b_{22}	0.30	0.32	0.02	0.05	0.14	1.50	0.24	0.30	0.32	0.35	0.42
		$\tilde{\delta}$	0.70	0.44	-0.26	0.38	1.20	12.66	-0.33	0.18	0.44	0.70	1.17
A3	41.8	ω_1	0.50	0.49	-0.01	0.07	0.23	2.42	0.35	0.44	0.49	0.54	0.64
		a_{11}	0.10	0.08	-0.02	0.04	0.13	1.98	0.01	0.05	0.07	0.11	0.16
		b_{11}	0.80	0.82	0.02	0.04	0.12	1.72	0.74	0.80	0.82	0.85	0.90
		ω_2	0.50	0.50	0.00	0.09	0.30	4.06	0.32	0.43	0.49	0.56	0.70
		a_{22}	0.20	0.20	0.00	0.05	0.14	2.55	0.11	0.17	0.20	0.23	0.28
		b_{22}	0.70	0.72	0.02	0.04	0.13	2.10	0.64	0.69	0.72	0.75	0.80
		$\tilde{\delta}$	-1.00	-0.04	0.96	0.61	1.92	22.40	-0.95	-0.54	-0.11	0.42	1.20
A4	36.8	ω_1	0.50	0.71	0.21	0.23	0.74	16.05	0.33	0.54	0.69	0.85	1.25
		a_{11}	0.60	0.47	-0.13	0.14	0.44	9.18	0.17	0.38	0.48	0.57	0.71
		b_{11}	0.10	0.12	0.02	0.04	0.12	1.28	0.05	0.10	0.12	0.15	0.20
		ω_2	0.30	0.46	0.16	0.19	0.58	22.11	0.17	0.31	0.44	0.59	0.85
		a_{22}	0.80	0.68	-0.12	0.10	0.31	12.36	0.48	0.61	0.69	0.75	0.85
		b_{22}	0.10	0.15	0.05	0.04	0.13	1.88	0.07	0.12	0.15	0.18	0.23
		$\tilde{\delta}$	-0.50	-0.48	0.02	0.43	1.37	11.87	-1.25	-0.80	-0.51	-0.20	0.44

lie inside given these intensities and Δ_t is very constrained, having its limits delimited by -0.00000003289300724 and 0.00001826361822548.

Next are presented the results of table 4.3, where the matrices A and/or B are non-diagonal.

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

The B case has the matrix A diagonal and B non-diagonal. This implies that the current Poisson rates, λ_t , depends from its previous value, λ_{t-1} , without crossed effects, and from the previous data realization having $Y_{i,t-1}$ affecting $\lambda_{j,t}$, $i \in \{1, 2\}, i \neq j$; C has the matrix A non-diagonal and B diagonal, following the same idea; and D have crossed effect from the previous Poisson rates and from the most recent realized data, having the matrices A and B both non-diagonal. The same behavior present in the previous analysis can be checked in this case, having estimates that indicates the capability of leaning of the model, with a large uncertainty present on the parameter's posterior mean $\tilde{\delta}$. The lack convergence is also observed in this cases. Also, some bias is present at the constant components ω for the cases C and D.

Figure 4.1: *Model C - 100000 Simulated Scenarios. $\lambda_{1,t}, \lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ box-plots $\forall t \in \{1, \dots, 200\}$*

Additionally, for all cases, it is observed that the parameters from matrix B, associated with previous data contribution, have much less uncertainty if compared to the parameters that map the contribution of previous non-observed variable λ_{t-1} , represented by matrix A.

The next assessment is to the cases DC1 and DC2, where the parameter δ_t is not considered constant, having some structure governing its dynamics through the time. The components associated with the matrices A and B presented a similar behavior as in the previously discussed cases. A problem appears when considering the variables associated with the dynamic variable δ_t . The higher dispersion present in $\tilde{\delta}$ in all examples discussed until now is associated with the weak transmission from the information contained on data to this parameter, and not by model misspecification. By consequence, even if the data generator process has a structure embedded at the dynamics of δ_t , it would be difficult to estimate the posterior distribution of its components in such cases. This can be observed at 4.4. The variable δ_{level} represents the constant contribution $(1 - \alpha - \beta) \tilde{\delta}$ through the time. Its posterior mean, for both cases, were similar to the values used during the data generation,

but the means of the distribution of the parameters which composes δ_{level} , are quite different from the ones proposed. The posterior means of α and β differs quite a lot from the ones proposed, indicating some confusion by the model being estimated given the higher variability of the posterior. Intending to quantify if this could be a convergence problem, another analysis was implemented to DC1 and DC2, comparing their parameters posterior estimates, previous realized with a burn-in of 500000 and a sampling size of 100000, to a new procedure where was burned 5000000 samples and used to compute the posterior estimates 5000000 samples.

Table 4.3: *Simulated Data - MCMC Output Statistics (Cont.) - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500*

Model	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
B	39.7	ω_1	0.30	0.38	0.08	0.08	0.26	6.00	0.22	0.32	0.37	0.43	0.55
		a_{11}	0.20	0.17	-0.03	0.06	0.20	6.46	0.05	0.13	0.17	0.22	0.30
		b_{11}	0.50	0.54	0.04	0.04	0.14	3.48	0.47	0.51	0.54	0.57	0.64
		b_{12}	0.10	0.09	-0.01	0.03	0.09	1.36	0.03	0.07	0.09	0.11	0.15
		ω_2	0.50	0.40	-0.10	0.09	0.29	5.17	0.24	0.33	0.39	0.46	0.60
		a_{22}	0.40	0.42	0.02	0.06	0.19	4.82	0.30	0.38	0.42	0.46	0.54
		b_{22}	0.20	0.26	0.06	0.04	0.13	2.44	0.18	0.23	0.26	0.29	0.34
		b_{21}	0.30	0.23	-0.07	0.03	0.10	1.30	0.16	0.20	0.22	0.25	0.29
		$\tilde{\delta}$	0.70	0.71	0.01	0.46	1.46	46.30	-0.19	0.38	0.71	1.01	1.63
C	39.6	ω_1	0.50	0.78	0.28	0.21	0.68	10.98	0.40	0.63	0.77	0.93	1.22
		a_{11}	0.20	0.24	0.04	0.12	0.37	3.64	0.03	0.15	0.23	0.32	0.47
		a_{12}	0.30	0.17	-0.13	0.09	0.28	4.31	0.02	0.10	0.16	0.22	0.36
		b_{11}	0.20	0.20	0.00	0.04	0.13	1.34	0.13	0.17	0.20	0.23	0.28
		ω_2	0.30	0.32	0.02	0.20	0.62	5.74	0.03	0.17	0.30	0.45	0.76
		a_{22}	0.20	0.18	-0.02	0.10	0.31	4.26	0.02	0.11	0.17	0.25	0.38
		a_{21}	0.40	0.31	-0.09	0.11	0.35	4.41	0.08	0.23	0.31	0.38	0.52
		b_{22}	0.30	0.34	0.04	0.04	0.13	1.22	0.26	0.31	0.34	0.37	0.42
		$\tilde{\delta}$	-0.50	-0.64	-0.14	0.35	1.10	10.36	-1.25	-0.90	-0.67	-0.41	0.08
D	41.2	ω_1	0.50	0.81	0.31	0.19	0.59	13.56	0.45	0.68	0.80	0.93	1.20
		a_{11}	0.30	0.14	-0.16	0.11	0.34	9.39	0.00	0.05	0.11	0.20	0.39
		a_{12}	0.20	0.25	0.05	0.13	0.41	14.73	0.02	0.16	0.26	0.34	0.51
		b_{11}	0.10	0.11	0.01	0.04	0.13	1.91	0.03	0.08	0.11	0.14	0.19
		b_{12}	0.20	0.23	0.03	0.05	0.15	2.33	0.14	0.20	0.23	0.26	0.32
		ω_2	0.30	0.28	-0.02	0.15	0.49	8.40	0.02	0.17	0.27	0.38	0.61
		a_{22}	0.30	0.24	-0.06	0.13	0.40	12.75	0.02	0.14	0.24	0.33	0.48
		a_{21}	0.10	0.20	0.10	0.13	0.41	11.41	0.01	0.09	0.18	0.29	0.47
		b_{22}	0.20	0.22	0.02	0.04	0.13	1.73	0.14	0.19	0.22	0.25	0.30
		b_{21}	0.30	0.22	-0.08	0.04	0.13	2.23	0.14	0.19	0.22	0.25	0.30
		$\tilde{\delta}$	0.40	0.93	0.53	0.59	1.87	41.27	-0.18	0.50	0.93	1.34	2.10

The results of this analysis can be verified at tables 4.5 and 4.6. The posterior means of parameters and its respective standard deviations are quite similar when comparing the set burn-in 500000 / Sample 100000 to the new one burn-in 5000000 / Sample 5000000, for both DC1 and DC2. This regards that is not an issue of the estimation procedure and also indicates that all information that was contained in the data was already captured by the study (burn-in 500000 / Sample 100000). In the DC2 case, the average of $\tilde{\delta}$ was 0 for the (burn-in 5000000 / Sample 5000000) study, but this is a consequence of the left tail that appeared during the sampling scheme. If the median for this parameter were observed, it is possible to check that it is of the same order as in the first study (~ 1). At figures 4.2 and 4.4 are presented the trace charts for the parameters of DC2 case. In the first figure are drawn the traces for burn-in 500000 / Sample 100000 study and, at the second

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

one, are represented the ones for burn-in 5000000 / Sample 5000000.

Table 4.4: *Simulated Data - MCMC Output Statistics (Cont.) - MCMC sampling size: 100000. Burn in: 500000. Data Points: 500*

Model	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
DC1	39.8	ω_1	0.30	0.31	0.01	0.12	0.37	17.02	0.09	0.22	0.32	0.40	0.51
		a_{11}	0.30	0.35	0.05	0.20	0.64	31.10	0.03	0.19	0.33	0.50	0.75
		b_{11}	0.20	0.16	-0.04	0.04	0.14	2.75	0.07	0.12	0.15	0.18	0.25
		ω_2	0.20	0.15	-0.05	0.04	0.12	1.06	0.09	0.13	0.15	0.18	0.24
		a_{22}	0.20	0.28	0.08	0.11	0.36	3.66	0.05	0.19	0.27	0.35	0.50
		b_{22}	0.20	0.30	0.10	0.05	0.17	3.14	0.20	0.27	0.30	0.34	0.42
		$\tilde{\delta}$	-1.00	-1.11	-0.11	0.69	2.19	54.80	-2.85	-1.47	-1.05	-0.65	0.05
		α	0.20	0.38	0.18	0.24	0.77	9.97	0.02	0.18	0.35	0.55	0.89
		β	0.20	0.09	-0.11	0.06	0.20	2.44	0.00	0.04	0.09	0.14	0.23
		δ_{level}	-0.60	-0.51	-0.09	0.33	1.03	14.03	-1.19	-0.74	-0.49	-0.26	0.02
DC2	40.9	ω_1	1.00	1.01	0.01	0.13	0.40	14.57	0.75	0.92	1.01	1.10	1.24
		a_{11}	0.10	0.07	-0.03	0.05	0.16	5.12	0.00	0.03	0.06	0.11	0.18
		a_{12}	0.10	0.04	-0.06	0.03	0.09	2.05	0.00	0.01	0.03	0.05	0.12
		b_{11}	0.10	0.08	-0.02	0.04	0.12	1.31	0.01	0.05	0.08	0.11	0.16
		b_{12}	0.10	0.18	0.08	0.04	0.11	1.69	0.11	0.15	0.18	0.20	0.24
		ω_2	0.10	0.15	0.05	0.10	0.33	11.02	0.01	0.07	0.14	0.22	0.38
		a_{22}	0.30	0.45	0.15	0.08	0.25	10.24	0.28	0.40	0.46	0.51	0.59
		a_{21}	0.50	0.23	-0.27	0.06	0.20	11.07	0.12	0.19	0.23	0.28	0.36
		b_{22}	0.20	0.23	0.03	0.04	0.12	2.22	0.16	0.21	0.23	0.26	0.31
		b_{21}	0.10	0.10	0.00	0.04	0.12	2.72	0.02	0.07	0.10	0.12	0.17
		$\tilde{\delta}$	1.00	0.91	-0.09	0.78	2.45	78.60	-1.04	0.46	1.00	1.45	2.15
		α	0.10	0.35	0.25	0.22	0.69	7.19	0.02	0.17	0.33	0.50	0.81
		β	0.40	0.21	-0.19	0.12	0.38	4.18	0.01	0.11	0.21	0.30	0.44
		δ_{level}	0.50	0.46	-0.04	0.43	1.35	26.60	-0.21	0.13	0.39	0.72	1.42

The better resolution of the second figure is associated with the large sample used for computing these charts. Even if reducing the variance significantly for parameters a_{11} , a_{12} , and ω_2 , the format of the selected priors (Weibull with shape parameter equal 1) still appearing in the trace charts of the posteriors. This indicates these parameters difficulty of learning regarding the used dataset. Another feature possible to be observed is associated with the mixing difficulty of the parameter a_{21} . Comparing both figures 4.2 and 4.4, it is possible observe this parameter alternating between 0 and 0.5, its real value. This behavior is another indication of the confusion problem that arose when the flexibility of δ_t was allowed. Additionally, the posterior for parameters α and β seems to have incorporated almost none information from the data, given the lack of mass concentration at any point of their support and also the similarity of their posterior to them prior proposed. At last, the observed difference in the posterior means of $\tilde{\delta}$ computed at study burn-in 500000 / Sample 100000 and burn-in 5000000 / Sample 5000000 have its cause visible in the trace chart of this variable presented in figure 4.4: there is an instant that the Metropolis-Hastings procedure has its samples being accepted for values of $\tilde{\delta}$ until near -10, which have biased the posterior mean in the second case. Also the difficulty of convergence is observed for both cases in Also in Figure E.11 and Figure 4.5.

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Table 4.5: *Simulated Data - MCMC Output Statistics - DC1 Large BurnIn - Large Sample - Data Points: 500*

Burn In	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p97.5
500000	39.9	ω_1	0.30	0.31	0.01	0.12	0.37	17.02	0.09	0.22	0.32	0.51
		a_{11}	0.30	0.35	0.05	0.20	0.64	31.10	0.03	0.19	0.33	0.75
		b_{11}	0.20	0.16	-0.04	0.04	0.14	2.75	0.07	0.12	0.15	0.25
		ω_2	0.20	0.15	-0.05	0.04	0.12	1.06	0.09	0.13	0.15	0.24
		a_{22}	0.20	0.28	0.08	0.11	0.36	3.66	0.05	0.19	0.27	0.50
		b_{22}	0.20	0.30	0.10	0.05	0.17	3.14	0.20	0.27	0.30	0.42
		$\bar{\delta}$	-1.00	-1.11	-0.11	0.69	2.19	54.79	-2.85	-1.47	-1.05	0.05
		α	0.20	0.38	0.18	0.24	0.77	9.98	0.02	0.18	0.35	0.89
		β	0.20	0.09	-0.11	0.06	0.20	2.43	0.00	0.04	0.09	0.23
		δ_{level}	-0.60	-0.51	-0.09	0.33	1.03	14.03	-1.19	-0.74	-0.49	0.02
5000000	39.8	ω_1	0.30	0.32	0.02	0.12	0.05	2.65	0.11	0.23	0.32	0.54
		a_{11}	0.30	0.34	0.04	0.19	0.09	4.37	0.02	0.18	0.33	0.71
		b_{11}	0.20	0.16	-0.04	0.05	0.02	0.43	0.07	0.12	0.16	0.25
		ω_2	0.20	0.16	-0.04	0.04	0.02	0.18	0.09	0.13	0.15	0.24
		a_{22}	0.20	0.27	0.07	0.12	0.05	0.57	0.05	0.19	0.27	0.50
		b_{22}	0.20	0.30	0.10	0.05	0.02	0.46	0.20	0.26	0.30	0.42
		$\bar{\delta}$	-1.00	-1.11	-0.11	0.73	0.33	9.35	-2.87	-1.47	-1.02	0.03
		α	0.20	0.38	0.18	0.24	0.11	1.47	0.02	0.18	0.35	0.89
		β	0.20	0.10	-0.10	0.06	0.03	0.36	0.00	0.04	0.09	0.23
		δ_{level}	-0.60	-0.51	-0.09	0.32	0.14	1.96	-1.17	-0.73	-0.48	0.01

Table 4.6: *Simulated Data - MCMC Output Statistics - DC2 Large BurnIn - Large Sample - Data Points: 500*

Burn In	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p97.5
500000	40.9	ω_1	1.00	1.01	0.01	0.13	0.41	14.57	0.75	0.92	1.01	1.24
		a_{11}	0.10	0.07	-0.03	0.05	0.16	5.12	0.00	0.03	0.06	0.18
		a_{12}	0.10	0.04	-0.06	0.03	0.09	2.05	0.00	0.01	0.03	0.12
		b_{11}	0.10	0.08	-0.02	0.04	0.12	1.31	0.01	0.05	0.08	0.16
		b_{12}	0.10	0.18	0.08	0.04	0.11	1.69	0.11	0.15	0.18	0.24
		ω_2	0.10	0.15	0.05	0.10	0.33	11.02	0.01	0.07	0.14	0.38
		a_{22}	0.30	0.45	0.15	0.08	0.25	10.24	0.28	0.40	0.46	0.59
		a_{21}	0.50	0.23	-0.27	0.06	0.20	11.07	0.12	0.19	0.23	0.36
		b_{22}	0.20	0.23	0.03	0.04	0.12	2.22	0.16	0.21	0.23	0.31
		b_{21}	0.10	0.10	0.00	0.04	0.12	2.72	0.02	0.07	0.10	0.17
		$\bar{\delta}$	1.00	0.91	-0.09	0.78	2.45	78.60	-1.04	0.46	1.00	2.15
		α	0.10	0.35	0.25	0.22	0.69	7.19	0.02	0.17	0.33	0.81
		β	0.40	0.21	-0.19	0.12	0.38	4.18	0.01	0.11	0.21	0.44
		δ_{level}	0.50	0.46	-0.04	0.43	1.35	26.60	-0.21	0.13	0.39	1.42
5000000	37.3	ω_1	1.00	0.99	-0.01	0.15	0.06	3.15	0.68	0.90	1.00	1.26
		a_{11}	0.10	0.09	-0.01	0.07	0.03	1.83	0.00	0.03	0.07	0.27
		a_{12}	0.10	0.04	-0.06	0.04	0.02	0.56	0.00	0.01	0.03	0.14
		b_{11}	0.10	0.08	-0.02	0.04	0.02	0.24	0.01	0.05	0.07	0.16
		b_{12}	0.10	0.17	0.07	0.04	0.02	0.32	0.10	0.14	0.17	0.24
		ω_2	0.10	0.16	0.06	0.12	0.05	2.76	0.01	0.07	0.14	0.44
		a_{22}	0.30	0.46	0.16	0.14	0.06	5.31	0.15	0.38	0.48	0.68
		a_{21}	0.50	0.25	-0.25	0.17	0.08	13.87	0.01	0.11	0.22	0.65
		b_{22}	0.20	0.22	0.02	0.05	0.02	0.54	0.14	0.19	0.22	0.31
		b_{21}	0.10	0.09	-0.01	0.04	0.02	0.51	0.01	0.05	0.08	0.17
		$\bar{\delta}$	1.00	0.04	-0.06	2.59	1.16	143.40	-8.47	-0.05	0.83	2.21
		α	0.10	0.40	0.30	0.23	0.10	1.49	0.02	0.21	0.39	0.87
		β	0.40	0.22	-0.18	0.12	0.05	0.74	0.01	0.12	0.22	0.46
		δ_{level}	0.50	0.34	-0.16	0.47	0.21	5.49	-0.43	-0.01	0.28	1.39

¹Posterior Standard Deviation²Naive Standard Error $\times 10^{-3}$ ³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Figure 4.2: Trace Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Figure 4.3: ACF Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Figure 4.4: Trace Charts - DC2 - Data Points: 500. Burn In: 5000000. Sample Size: 5000000.

Figure 4.5: ACF Charts - DC2 - Data Points: 500. Burn In: 5000000. Sample Size: 5000000.

Replication study

A replication study using simulated data were implemented as in [Cui and Zhu \[2017\]](#). For each proposed case in 4.1, were generated $m = 200$ datasets containing 500 data realizations each. Two posterior computations were realized for each dataset generated: one containing the first $n = 200$ data points and other with $n = 500$. For evaluating the performance of the estimation procedure, the Mean Absolute Deviation Error(MADE) and the Mean Squared Error(MSE) were computed by

$$MADE = \frac{1}{m} \sum_{j=1}^m |\hat{v}_j - v^0| \quad (4.1)$$

$$MSE = \frac{1}{m} \sum_{j=1}^m (\hat{v}_j - v^0)^2 \quad (4.2)$$

where $\hat{v}_j$ is defined as the posterior mean of v parameter distribution at the j -th replication, and v^0 is its real value, the one used for data generation.

At tables 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12 are found summarized the average of m replication posterior means and its difference to the true value used for data generation, the standard deviation of m replication posterior means, the MADE and MSE as just defined, and the SMSE, the square root of the MSE. To work with the SMSE would be better than use the MSE since it is on the same scale as the data. In general, the statistics for most of the parameters associated with the contribution of previous data realizations presented a reasonable result, having no bias, lower standard deviation, and small MADE/MSE. As the sample size increased from $n = 200$ to $n = 500$, also were verified an enhancement in the performance of these statistics. Contrasting this results, the parameters that link the current Poisson rates to its previous realization, in general exhibit some bias, tending the parameters ω_1 and ω_2 to incorporate some of the contributions that would be expected be contained in parameters a_{11} , a_{12} , a_{21} and a_{22} . The SD, MADE, MSE, and SMSE for these parameters also present reasonable values, much of them smaller if compared to the ones described in [Cui and Zhu \[2017\]](#).

Returning to the case A3, an interesting comparison would be made analyzing the statistics of the parameter $\tilde{\delta}$. Even if the mean being unbiased in [Cui and Zhu \[2017\]](#), its standard deviation (0.9408) is more significant compared to the one determined in this work. This uncertainty can be associated with the high level of Poisson rates generated by this set of parameters, making the link between $\tilde{\delta}$ and the data no that strong to capture the intended information properly.

Table 4.7: *Simulation results for the proposed models A1, A2, A3 and A4. 200 Replications. MCMC sampling size: 100000. BurnIn: 500000.*

Model	N		ω_1	a_{11}	b_{11}	ω_2	a_{22}	b_{22}	$\bar{\delta}$
A1	200	Actual	1.0000	0.4000	0.2000	0.5000	0.3000	0.4000	0.5000
		Mean ⁴	1.2634	0.3001	0.1981	0.6042	0.245	0.3961	0.4433
		Difference	0.2634	-0.0999	-0.0019	0.1042	-0.055	-0.0039	-0.0567
		SD ⁵	0.2745	0.1039	0.0678	0.1354	0.0925	0.0751	0.5368
		MADE	0.3224	0.1242	0.0559	0.136	0.0901	0.0595	0.4267
		MSE	0.1443	0.0207	0.0046	0.0291	0.0115	0.0056	0.29
		SMSE ⁶	0.3799	0.1439	0.0678	0.1706	0.1072	0.0748	0.5385
	500	Actual	1.0000	0.4000	0.2000	0.5000	0.3000	0.4000	0.5000
		Mean	1.1818	0.3239	0.2055	0.5517	0.2633	0.4081	0.4837
		Difference	0.1818	-0.0761	0.0055	0.0517	-0.0367	0.0081	-0.0163
		SD	0.2416	0.1001	0.0444	0.0997	0.0761	0.0433	0.4449
		MADE	0.251	0.1032	0.0358	0.0916	0.0672	0.0354	0.3692
		MSE	0.0912	0.0158	0.002	0.0126	0.0071	0.0019	0.1973
		SMSE	0.302	0.1257	0.0447	0.1122	0.0843	0.0436	0.4442
A2	200	Actual	0.3000	0.2000	0.5000	0.5000	0.4000	0.3000	0.7000
		Mean	0.3342	0.1877	0.4803	0.6618	0.3069	0.3005	0.5211
		Difference	0.0342	-0.0123	-0.0197	0.1618	-0.0931	0.0005	-0.1789
		SD	0.074	0.0808	0.0852	0.1675	0.1055	0.0683	0.4421
		MADE	0.0639	0.0656	0.0714	0.1926	0.1201	0.0538	0.3661
		MSE	0.0066	0.0067	0.0076	0.0541	0.0197	0.0046	0.2265
		SMSE	0.0812	0.0819	0.0872	0.2326	0.1404	0.0678	0.4759
	500	Actual	0.3000	0.2000	0.5000	0.5000	0.4000	0.3000	0.7000
		Mean	0.3194	0.1824	0.5017	0.5807	0.3448	0.3093	0.6339
		Difference	0.0194	-0.0176	0.0017	0.0807	-0.0552	0.0093	-0.0661
		SD	0.052	0.0674	0.051	0.1265	0.0906	0.0411	0.3154
		MADE	0.0441	0.0582	0.0414	0.1188	0.084	0.0333	0.258
		MSE	0.0031	0.0048	0.0026	0.0224	0.0112	0.0018	0.1033
		SMSE	0.0557	0.0693	0.051	0.1497	0.1058	0.0424	0.3214
A3	200	Actual	0.5000	0.1000	0.8000	0.5000	0.2000	0.7000	-1.0000
		Mean	0.566	0.1024	0.7761	0.6104	0.1694	0.7029	-0.0168
		Difference	0.066	0.0024	-0.0239	0.1104	-0.0306	0.0029	0.9832
		SD	0.1433	0.043	0.0524	0.1557	0.0653	0.0628	0.2687
		MADE	0.1133	0.0345	0.0447	0.1482	0.0611	0.0514	0.9832
		MSE	0.0248	0.0018	0.0033	0.0363	0.0052	0.0039	1.0385
		SMSE	0.1575	0.0424	0.0574	0.1905	0.0721	0.0624	1.0191
	500	Actual	0.5000	0.1000	0.8000	0.5000	0.2000	0.7000	-1.0000
		Mean	0.5251	0.0964	0.7958	0.5643	0.1783	0.7077	-0.2312
		Difference	0.0251	-0.0036	-0.0042	0.0643	-0.0217	0.0077	0.7688
		SD	0.0871	0.0389	0.0372	0.1058	0.0476	0.0414	0.2694
		MADE	0.0717	0.0314	0.0304	0.097	0.0427	0.0338	0.7688
		MSE	0.0082	0.0015	0.0014	0.0153	0.0027	0.0018	0.6633
		SMSE	0.0906	0.0387	0.0374	0.1237	0.052	0.0424	0.8144
A4	200	Actual	0.5000	0.6000	0.1000	0.3000	0.8000	0.1000	-0.5000
		Mean	0.862	0.3652	0.1195	0.6462	0.6502	0.1277	-0.1879
		Difference	0.362	-0.2348	0.0195	0.3462	-0.1498	0.0277	0.3121
		SD	0.2158	0.1224	0.0494	0.3654	0.1492	0.0552	0.5403
		MADE	0.3753	0.2395	0.0407	0.3566	0.1572	0.046	0.4918
		MSE	0.1774	0.07	0.0028	0.2527	0.0446	0.0038	0.3879
		SMSE	0.4212	0.2646	0.0529	0.5027	0.2112	0.0616	0.6228
	500	Actual	0.5000	0.6000	0.1000	0.3000	0.8000	0.1000	-0.5000
		Mean	0.8103	0.4031	0.1129	0.5197	0.7075	0.1184	-0.3509
		Difference	0.3103	-0.1969	0.0129	0.2197	-0.0925	0.0184	0.1491
		SD	0.208	0.1233	0.0391	0.2763	0.1062	0.0317	0.4777
		MADE	0.3214	0.2026	0.0327	0.2353	0.1001	0.0281	0.3828
		MSE	0.1393	0.0539	0.0017	0.1242	0.0198	0.0013	0.2493
		SMSE	0.3732	0.2322	0.0412	0.3524	0.1407	0.0361	0.4993

Table 4.8: *Simulation results for the proposed model B. 200 Replications. BurnIn: 500000.*

Model	N		ω_1	a_{11}	b_{11}	b_{12}	ω_2	a_{22}	b_{22}	b_{21}	$\tilde{\delta}$
B	200	Actual	0.3000	0.2000	0.5000	0.1000	0.5000	0.4000	0.2000	0.3000	0.7000
		Mean	0.3647	0.1758	0.484	0.1055	0.6512	0.3213	0.2069	0.3224	0.6118
		Difference	0.0647	-0.0242	-0.016	0.0055	0.1512	-0.0787	0.0069	0.0224	-0.0882
		SD	0.1192	0.0715	0.0744	0.0411	0.1884	0.104	0.0682	0.0699	0.5427
		MADE	0.1076	0.0622	0.0591	0.0328	0.1937	0.1092	0.0539	0.0587	0.4515
		MSE	0.0183	0.0057	0.0058	0.0017	0.0582	0.017	0.0047	0.0054	0.3009
		SMSE	0.1353	0.0755	0.0762	0.0412	0.2412	0.1304	0.0686	0.0735	0.5485
	500	Actual	0.3000	0.2000	0.5000	0.1000	0.5000	0.4000	0.2000	0.3000	0.7000
		Mean	0.3394	0.1747	0.5027	0.1038	0.5681	0.3574	0.2111	0.3115	0.6595
		Difference	0.0394	-0.0253	0.0027	0.0038	0.0681	-0.0426	0.0111	0.0115	-0.0405
		SD	0.0826	0.0606	0.0475	0.0336	0.1206	0.0788	0.0429	0.048	0.4336
		MADE	0.0716	0.0539	0.0385	0.0254	0.1057	0.0723	0.0351	0.0392	0.3434
		MSE	0.0083	0.0043	0.0023	0.0011	0.0191	0.008	0.002	0.0024	0.1887
		SMSE	0.0911	0.0656	0.048	0.0332	0.1382	0.0894	0.0447	0.049	0.4344

Table 4.9: *Simulation results for the proposed model C. 200 Replications. BurnIn: 500000.*

Model	N		ω_1	a_{11}	a_{12}	b_{11}	ω_2	a_{22}	a_{21}	b_{22}	$\tilde{\delta}$
C	200	Actual	0.5000	0.2000	0.3000	0.2000	0.3000	0.2000	0.4000	0.3000	-0.5000
		Mean	0.5366	0.2121	0.2892	0.1856	0.493	0.1969	0.3115	0.2914	-0.2807
		Difference	0.0366	0.0121	-0.0108	-0.0144	0.193	-0.0031	-0.0885	-0.0086	0.2193
		SD	0.1864	0.0844	0.0953	0.0624	0.1808	0.0762	0.0987	0.0695	0.4528
		MADE	0.1545	0.0647	0.0787	0.0522	0.2076	0.0611	0.111	0.0561	0.3994
		MSE	0.0359	0.0072	0.0092	0.0041	0.0697	0.0058	0.0175	0.0049	0.2521
		SMSE	0.1895	0.0849	0.0959	0.064	0.264	0.0762	0.1323	0.07	0.5021
	500	Actual	0.5000	0.2000	0.3000	0.2000	0.3000	0.2000	0.4000	0.3000	-0.5000
		Mean	0.5158	0.204	0.2932	0.1987	0.4234	0.1959	0.335	0.3027	-0.4166
		Difference	0.0158	0.004	-0.0068	-0.0013	0.1234	-0.0041	-0.065	0.0027	0.0834
		SD	0.164	0.0832	0.0867	0.0415	0.1554	0.0874	0.1014	0.0428	0.3723
		MADE	0.1322	0.0666	0.0721	0.0324	0.1499	0.0704	0.0997	0.0344	0.2994
		MSE	0.027	0.0069	0.0075	0.0017	0.0393	0.0076	0.0145	0.0018	0.1448
		SMSE	0.1643	0.0831	0.0866	0.0412	0.1982	0.0872	0.1204	0.0424	0.3805

Table 4.10: *Simulation results for the proposed model D. 200 Replications. BurnIn: 500000.*

Model	N		ω_1	a_{11}	a_{12}	b_{11}	b_{12}	ω_2	a_{22}	a_{21}	b_{22}	b_{21}	$\tilde{\delta}$
D	200	Actual	0.5000	0.3000	0.2000	0.1000	0.2000	0.3000	0.3000	0.1000	0.2000	0.3000	0.4000
		Mean	0.5774	0.2144	0.2522	0.1116	0.1941	0.3923	0.1844	0.2029	0.1947	0.2862	0.5334
		Difference	0.0774	-0.0856	0.0522	0.0116	-0.0059	0.0923	-0.1156	0.1029	-0.0053	-0.0138	0.1334
		SD	0.2039	0.0591	0.0853	0.0501	0.0594	0.1645	0.0607	0.072	0.069	0.0659	0.6226
		MADE	0.1632	0.0923	0.0768	0.0387	0.0467	0.1405	0.1197	0.1034	0.0561	0.0546	0.5279
		MSE	0.0473	0.0108	0.01	0.0026	0.0035	0.0354	0.017	0.0157	0.0048	0.0045	0.4035
		SMSE	0.2175	0.1039	0.1	0.051	0.0592	0.1881	0.1304	0.1253	0.0693	0.0671	0.6352
	500	Actual	0.5000	0.3000	0.2000	0.1000	0.2000	0.3000	0.3000	0.1000	0.2000	0.3000	0.4000
		Mean	0.5596	0.2192	0.2554	0.1068	0.1973	0.3255	0.1957	0.1959	0.2059	0.295	0.4283
		Difference	0.0596	-0.0808	0.0554	0.0068	-0.0027	0.0255	-0.1043	0.0959	0.0059	-0.005	0.0283
		SD	0.1528	0.0702	0.0824	0.0391	0.046	0.1239	0.0601	0.0671	0.0417	0.0429	0.4991
		MADE	0.1302	0.0944	0.0779	0.0319	0.0364	0.0981	0.1092	0.0967	0.0336	0.0346	0.4018
		MSE	0.0268	0.0114	0.0098	0.0016	0.0021	0.0159	0.0145	0.0137	0.0018	0.0019	0.2486
		SMSE	0.1637	0.1068	0.099	0.04	0.0458	0.1261	0.1204	0.117	0.0424	0.0436	0.4986

Related to the cases where a dynamic structure were imposed to δ_t , at tables 4.11 and 4.12 is

⁴Mean: Average of the Replications' Posterior Mean

⁵SD: Standard Deviation of the Replications' Posterior Mean

⁶SMSE: Square Root of MSE

verifiable the difficulty presented by the posterior of the parameters α , β and $\tilde{\delta}$ to learn from data. The conjunction of these three, δ_{level} , even with some bias, had incorporated some information but was difficult to the model adequately specify from which parameter these contributions was coming. As visible in figure 4.4, the uncertainty associated with α and β are very high, and this result was confirmed again in this replication study.

Table 4.11: *Simulation results for the proposed model DC1. 200 Replications. BurnIn: 500000.*

Model	N		ω_1	a_{11}	b_{11}	ω_2	a_{22}	b_{22}	$\tilde{\delta}$	α	β	δ_{level}
DC1	200	Actual	0.3000	0.3000	0.2000	0.2000	0.2000	0.2000	-1.000	0.2000	0.2000	-0.6000
		Mean	0.3272	0.2649	0.1824	0.1934	0.2389	0.1948	-1.7547	0.471	0.1209	-0.3636
		Difference	0.0272	-0.0351	-0.0176	-0.0066	0.0389	-0.0052	-0.7547	0.271	-0.0791	0.2364
		SD	0.0716	0.1039	0.0655	0.0375	0.0756	0.0701	2.3516	0.101	0.0381	0.2519
		MADE	0.0609	0.0897	0.0561	0.03	0.0647	0.0574	1.1688	0.271	0.0816	0.2668
		MSE	0.0058	0.012	0.0046	0.0014	0.0072	0.0049	6.0721	0.0836	0.0077	0.119
		SMSE	0.0762	0.1095	0.0678	0.0374	0.0849	0.07	2.4642	0.2891	0.0877	0.345
	500	Actual	0.3000	0.3000	0.2000	0.2000	0.2000	0.2000	-1.000	0.2000	0.2000	-0.6000
		Mean	0.3157	0.2725	0.1885	0.1932	0.2079	0.2002	-1.6274	0.4607	0.1113	-0.3815
		Difference	0.0157	-0.0275	-0.0115	-0.0068	0.0079	0.0002	-0.6274	0.2607	-0.0887	0.2185
		SD	0.0649	0.1145	0.0451	0.0274	0.0755	0.0522	2.0002	0.1161	0.0296	0.2226
		MADE	0.0547	0.0991	0.0381	0.0221	0.0629	0.0411	0.963	0.2607	0.0887	0.246
		MSE	0.0044	0.0138	0.0022	0.0008	0.0057	0.0027	4.3745	0.0814	0.0087	0.097
		SMSE	0.0663	0.1175	0.0469	0.0283	0.0755	0.052	2.0915	0.2853	0.0933	0.3114

Table 4.12: *Simulation results for the proposed model DC2. 200 Replications. BurnIn: 500000.*

Model	N		ω_1	a_{11}	a_{12}	b_{11}	b_{12}	ω_2	a_{22}	a_{21}	b_{22}	b_{21}	$\tilde{\delta}$	α	β	δ_{level}
DC2	200	Actual	1	0.1	0.1	0.1	0.1	0.1	0.3	0.5	0.2	0.1	1	0.1	0.4	0.5
		Mean	0.7655	0.1619	0.1468	0.1118	0.1048	0.4193	0.2401	0.3833	0.1925	0.1235	-0.9024	0.4574	0.2236	0.1516
		Difference	-0.2345	0.0619	0.0468	0.0118	0.0048	0.3193	-0.0599	-0.1167	-0.0075	0.0235	-1.9024	0.3574	-0.1764	-0.3484
		SD	0.1799	0.0569	0.0703	0.0452	0.0403	0.2207	0.1021	0.196	0.0655	0.0574	4.6046	0.083	0.0698	0.3935
		MADE	0.2443	0.065	0.0548	0.0376	0.0328	0.3193	0.1005	0.1967	0.0532	0.0467	2.4607	0.3574	0.178	0.436
		MSE	0.0872	0.0071	0.0071	0.0022	0.0016	0.1504	0.014	0.0518	0.0043	0.0038	24.7155	0.1346	0.036	0.2755
		SMSE	0.2953	0.0843	0.0843	0.0469	0.04	0.3878	0.1183	0.2276	0.0656	0.0616	4.9715	0.3669	0.1897	0.5249
	500	Actual	1	0.1	0.1	0.1	0.1	0.1	0.3	0.5	0.2	0.1	1	0.1	0.4	0.5
		Mean	0.8402	0.1597	0.1161	0.1018	0.1062	0.3626	0.2696	0.3734	0.206	0.11	-0.3407	0.4542	0.1996	0.2961
		Difference	-0.1598	0.0597	0.0161	0.0018	0.0062	0.2626	-0.0304	-0.1266	0.006	0.01	-1.3407	0.3542	-0.2004	-0.2039
		SD	0.1637	0.068	0.0593	0.0358	0.0324	0.1882	0.119	0.1963	0.0469	0.042	3.1873	0.1054	0.0566	0.3404
		MADE	0.1848	0.0662	0.0427	0.0286	0.0257	0.2628	0.1041	0.1985	0.036	0.0338	1.7124	0.3542	0.2004	0.3107
		MSE	0.0522	0.0082	0.0038	0.0013	0.0011	0.1042	0.015	0.0544	0.0022	0.0019	11.9059	0.1365	0.0434	0.1569
		SMSE	0.2285	0.0906	0.0616	0.0361	0.0332	0.3228	0.1225	0.2332	0.0469	0.0436	3.4505	0.3695	0.2083	0.3961

4.2 Real Data Application - Positive Cross-Correlation

4.2.1 Number of Trades Data

Since the proposed model is not capable of modeling cross-correlation for high Poisson rates, this example was selected considering the necessary characteristics for the model have a good performance when learning from the data. The number of trades for the Euro future contract and for the British Pound future expiring in March 2019 was selected. This information was collected from 08:45:00 UTC until 09:15:00 UTC, at February 8th 2019, and is composed by bars of 1-second interval. A larger period, like 5 minutes, for example, was not chosen since the average number of trades per each bar would be too elevated, incurring in the model's incapability to treat this information. This dataset was collected from the Datascope Reuters Platform. At table 4.13 are

Table 4.13: *Number Trades data - Descriptive Analysis*

		N	Mean	SD	Median	Trimmed [*]	MAD [†]	Min	Max	Range	Skew	Kurtosis	SE [‡]
TradesEC1	1	361	4.3	4.563	3	3.446	3.0	0	26	26	1.8	3.783	0.2
TradesBP1	2	361	1.3	1.933	1	0.869	1.5	0	14	14	2.4	8.263	0.1

* Trimmed, trimmed sample mean with trim defaulting to 0.1

[†] MAD, mean absolute deviation

[‡] SE, standard error of the mean

Figure 4.6: *CME Futures Number of Trades Data - Descriptive Analysis - EC1, Euro Future Contract; BP1, British Pound Future Contract*

presented the descriptive analysis of this dataset. The number of trades variance for the Euro is almost five times its average, while the British Pound's is about three times its mean, indicating a case of overdispersion. In Figure 4.6 is plotted the data series, the ACF for each one, the cross-correlation function, the histograms for the data being analyzed, a historical dispersion chart for both series (historical variance / historical mean, window = 21 data points), and a moving historical correlation chart (window = 21 data points). From these charts can be observed the spikes present in the dataset, making the variance of each univariate time series not being coincident with its mean. Additionally, this dynamic behavior in data points was not capable of being captured by the ACF, indicating that it is not a linear dependence.

The trace plots resulting from the Metropolis-Hastings procedure are represented by Figure D.13 for the diagonal case, and Figure D.14 for the non-diagonal one, both present at appendix D. The parameters a_{22} in both situations, a_{21} , a_{12} , and b_{22} have their prior variance reduced significantly, even presenting some difficulties in learning since posteriors' format are quite similar the one proposed as priors. Additionally, at 4.7 are represented the Poisson rates' evolution and their respective

Figure 4.7: Model Trade - 100000 Simulated Scenarios. $\lambda_{1,t}$, $\lambda_{2,t}$ and $\rho_{Y_{1,t}, Y_{2,t}}$ histograms $\forall t \in \{1, \dots, 361\}$

correlation for the first 100000 posterior draws. This plot is of the model where both matrices A and B are diagonal. Even being the level of Poisson rates quite higher, which imposes a difficulty to the true level of correlation be captured by the model (sample correlation ~ 0.28), it is possible to verify the positive correlation being treated at some point at figure 4.7. In terms of convergence, it is possible to observe in the autocorrelation plots in appendix E, and also in table 4.14 the

difficulty of the model to converge. This can be observed in the standard errors accounting for the autocorrelation, and in the charts, where indicates the great dependence between the samples being drawn. Only in the cases of parameters b_{22} and b_{21} is observable a satisfactory result in terms of convergence.

Table 4.14: *Estimates - Number of Trades Data*

Parameters	Diagonal [†]			Non-Diagonal [‡]		
	Mean ¹	SE ²	TSSE ³	Mean	SE	TSSE
ω_1	2.7010	0.7900	55.4200	2.7185	0.5800	48.4300
ω_2	1.1820	0.2000	6.0600	0.8950	0.2500	17.2800
a_{11}	0.3041	0.1700	11.5900	0.1550	0.1300	8.5800
a_{12}				0.3750	0.1500	24.9600
a_{21}				0.0329	0.0400	1.7900
a_{22}	0.1209	0.1300	8.2400	0.1101	0.1100	14.9900
b_{11}	0.0713	0.0300	0.4200	0.0423	0.0300	0.9200
b_{12}				0.1800	0.0900	2.1100
b_{21}				0.0397	0.0200	0.1400
b_{22}	0.0341	0.0300	0.1700	0.0239	0.0300	0.1200
$\tilde{\delta}$	1.6814	0.4300	5.2100	1.5428	0.4300	12.5400
ϕ						
Log-Lik	-1902.812			-1894.384		
AIC	3850.415			3843.027		
BIC	3877.637			3885.805		

[†] BurnIn: 500000. Samples: 500000

[‡] BurnIn: 500000. Samples: 500000

Also in the referred table are compared the posterior means of the diagonal implementation against the non-diagonal one. The common estimates are quite similar, having the non-diagonal version transferring part of the λ_t dependency to the crossed contributions($a_{i,j}, b_{ij}, i \neq j$). Also, by AIC and the log-likelihood, it seems that a fitting improvement occurred with the increase in the number of parameters.

4.3 Real Data Application - Negative cross-correlation

4.3.1 Artificial Data

As in Cui and Zhu [2017], two artificial data examples generated using a univariate INGARCH(1,1) were implemented. The idea of this study is the generation of a dataset that exhibits negative cross-correlation, without setting parameters as in the simulation study. Since we do not have the same seed for the random number generation, the same data was not possible to be replicated. The intention here is to show that the proposed estimation procedure is capable of capture the same parameter features, as in work executed by the reference author.

¹Mean: Posterior Mean

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Consider the univariate Poisson INGARCH(1,1), i.e

$$Y_t | \mathcal{F}_{t-1} \sim \mathbb{P}(\lambda_t), \quad \lambda_t = \omega + a\lambda_{t-1} + bY_{t-1} \quad (4.3)$$

Two parameters sets for the model 4.3 where considered to generate both artificial data examples, E1 and E2. Let $(\omega, a, b)' = (0.50, 0.05, 0.70)$ and $(\omega, a, b)' = (1.00, 0.35, 0.45)$ the set of parameters which generate these examples respectively. At each case, Cui and Zhu [2017] sampled 1000 realizations from (4.3) and built a bivariate integer process setting the first 500 data points as the realizations of variable $Y_{1,t}$ and the other remaining drawn data as $Y_{2,t}$. The descriptive statistics of each dataset generated in this work are expressed at tables 4.15 and 4.16. In both cases, the presence of over dispersion is verified, since the averages for both cases are inferior to their correspondent variances. For E1 and E2, their variance vector are $(4.12, 2.88)'$ and $(9.28, 6.72)'$ respectively. In addition, for each case, the sample cross correlation is -0.027 and 0.100.

Table 4.15: *Artificial Data E1 - Descriptive Analysis*

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
$Y_{1,t}$	1	500	2.0	2.030	1	1.648	1.5	0	11	11	1.4	1.915	0.1
$Y_{2,t}$	2	500	1.9	1.698	2	1.685	1.5	0	9	9	1.1	1.534	0.1

* Trimmed, trimmed sample mean with trim defaulting to 0.1

† MAD, mean absolute deviation

‡ SE, standard error of the mean

Table 4.16: *Artificial Data E2 - Descriptive Analysis*

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
$Y_{1,t}$	1	500	5.0	3.047	5	4.785	3	0	19	19	1.0	1.582	0.1
$Y_{2,t}$	2	500	4.9	2.593	5	4.793	3	0	15	15	0.5	0.241	0.1

* Trimmed, trimmed sample mean with trim defaulting to 0.1

† MAD, mean absolute deviation

‡ SE, standard error of the mean

The trace plots for the case E1 are present in appendix D, Figure D.9 for the diagonal case, and Figure D.10 for the non-diagonal one. For E2, the correspondent traces can be found at Figure D.11 and Figure D.12, for the diagonal and non-diagonal cases respectively. Their autocorrelation plot are present in appendix E. Except for a_{11} in the configuration E1, and $\tilde{\delta}$ in E2, all parameters seem had learned from the data. In the first case, a_{11} reduced drastically its prior variance, but was not able to specify where on its support it should concentrate mass. On second the one, $\tilde{\delta}$ spread the mass of its prior density inside the region of validity of δ_t , but the model was not capable of leaning about this parameter. Also, as in the previous studies, the convergence of the Metropolis-Hastings method presented some issues in both these examples.

At table ?? is compared the results presented at Cui and Zhu [2017] with the ones obtained in this work for the case E1. Even being obtained from different datasets, the results are quite similar. The main difference between Cui and Zhu [2017]'s proposition and this work, regards the parameter

$\tilde{\delta}$, in both diagonal and non-diagonal cases. The reference authors have negative estimates of the order -0.7 in both circumstances, while this work has approximately -0.1 in the two situations. Additionally, Cui and Zhu [2017] argue that their model has an improvement compared to the Liu's implementation. However, observing the log-likelihood, AIC, and BIC is possible to say that this improvement is almost insignificant. At table ?? are compared the results for the configuration E2. Except the parameter $\tilde{\delta}$, it is possible to say that all other parameters are of the same order. Also, as appeared in the simulation cases recently presented, the level parameters ω_1 and ω_2 tended to have a greater contribution in the Poisson rates λ_t by the Bayesian framework, while the formulation presented by Cui and Zhu [2017] have a higher magnitude being transferred to the current Poisson rates from its previous realization. A strange occurrence is associated with the parameter $\tilde{\delta}$. For Cui and Zhu [2017], if we consider the sample average of the dataset used for estimation (approx. 5 and 4.9) and calculating the limits for $\tilde{\delta}$ using 2.8 proposed by the author, the region of validity of $\tilde{\delta}$ should be limited inferiorly by -1.09, contradicting the estimate obtained by the author. We will see this misdoing occurring again during the crime data estimation.

Table 4.17: Estimates for Liu's, Cui's and proposed model - Artificial data E1

Parameters	Diagonal		Proposed [†]			Non-Diagonal		Proposed [‡]		
	Liu's	Cui's	Mean ¹	SE ²	TSSE ³	Liu's	Cui's	Mean	SE	TSSE
ω_1	0.5588	0.5663	0.5442	0.0900	1.3300	0.5531	0.5606	0.4001	0.1100	1.8500
ω_2	0.5122	0.5167	0.5882	0.1200	2.6800	0.5069	0.5117	0.5171	0.1300	2.0900
a_{11}	0.0001	0.0001	0.0329	0.0400	0.8100	0.0001	0.0001	0.0307	0.0400	0.6600
a_{12}						0.0001	0.0001	0.0537	0.0500	0.9800
a_{21}						0.0001	0.0001	0.0208	0.0200	0.3100
a_{22}	0.0724	0.0724	0.0870	0.0700	3.1200	0.0726	0.0725	0.0836	0.0700	1.6500
b_{11}	0.7265	0.7241	0.6956	0.0500	0.8300	0.7270	0.7246	0.6849	0.0500	0.8100
b_{12}						0.0001	0.0001	0.0377	0.0300	0.4600
b_{21}						0.0001	0.0001	0.0196	0.0200	0.1900
b_{22}	0.7210	0.7190	0.6040	0.0600	0.9400	0.7214	0.7193	0.6040	0.0600	1.1500
$\tilde{\delta}$		-0.7310	-0.1011	0.5800	23.4200		-0.7285	-0.1499	0.6100	26.6000
ϕ	0.0001					0.0001				
Log-Lik	-1597.856	-1596.881	-1567.158			-1598.018	-1597.037	-1566.284		
AIC	3209.711	3207.762	3178.056			3218.036	3216.073	3189.401		
BIC	3239.115	3237.165	3207.558			3264.152	3262.189	3235.762		

[†] BurnIn: 500000. Samples: 500000

[‡] BurnIn: 500000. Samples: 500000

¹Mean: Posterior Mean

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Table 4.18: *Estimates for Liu's, Cui's and proposed model - Artificial data E2*

Parameters	Diagonal		Proposed [†]			Non-Diagonal		Proposed [‡]		
	Liu's	Cui's	Mean ¹	SE ²	TSSE ³	Liu's	Cui's	Mean	SE	TSSE
ω_1	1.0059	1.0304	1.0891	0.2800	7.0100	0.9664	1.0204	0.7583	0.2900	9.6900
ω_2	1.0204	1.0225	1.2936	0.3400	10.5700	1.0064	1.0119	1.2064	0.3600	14.8500
a_{11}	0.3912	0.3918	0.3041	0.0900	2.9900	0.3918	0.3918	0.2796	0.0900	4.5300
a_{12}						0.0001	0.0001	0.0487	0.0500	1.1300
a_{21}						0.0001	0.0001	0.0195	0.0200	0.4600
a_{22}	0.3737	0.3720	0.3159	0.0900	3.2000	0.3747	0.3721	0.2856	0.1000	4.2600
b_{11}	0.4180	0.4147	0.4812	0.0600	1.4700	0.4180	0.4147	0.4797	0.0600	1.6000
b_{12}						0.0001	0.0001	0.0467	0.0400	0.6500
b_{21}						0.0001	0.0001	0.0187	0.0200	0.2300
b_{22}	0.4261	0.4276	0.4210	0.0600	1.1400	0.4258	0.4275	0.4314	0.0500	1.5100
$\tilde{\delta}$		-11.7795	1.0511	1.3200	34.2100		-11.7205	0.9829	1.2600	47.3300
ϕ	0.0001					0.0001				
Log-Lik	-2188.743	-2185.975	-2207.447			-2188.768	-2186.043	-2207.119		
AIC	4391.487	4385.95	4461.365			4399.537	4394.086	4477.700		
BIC	4420.89	4415.354	4490.867			4445.653	4440.201	4524.061		

[†] BurnIn: 500000. Samples: 500000[‡] BurnIn: 500000. Samples: 500000

4.3.2 Crime Data

As in Cui and Zhu [2017], the dataset of Robbery(ROB) and Criminal mischief(CMIS), present in census tract 101 of Pittsburgh for 1990 through 1999, was analyzed in this work. This information can be found in <http://www.forecastingprinciples.com/index.php/crimedata>. The time frame used for analysis comprehends 120 data points, as used by the reference authors. At table 4.19 are presented the descriptive statistics from this dataset. The empirical mean of ROB and CMIS series are 2 and 6 respectively. Their associated variances are 2.53 and 14.4, indicating the presence of overdispersion. This suggests that some dynamic might be embedded in the data generation process, possibly the bivariate INGARCH being a good choice for modeling this dataset. Also, the sample correlation between these time series, for the entire period, is equal -0.131, suggesting that a model which treats negative correlation might have a better fitting. Considering what we have already seen in section 2.4, the problem in using the bivariate Poisson proposed by Cui and Zhu [2017] in this dataset is associated with the level of these Poisson rates (2,6). The associated Δ_c is restricted to $[-1.42, 3.62]$, having an strict correlation range $[-0.0125, 0.0319]$ where the model can adjust the data's information inside. At Figure 4.8 is plotted the data being modeled with some additional descriptive information, like in previous examples.

The traces of estimation procedure can be visualized at figures 4.9, for the diagonal case, and 4.11 for the non-diagonal one. Focusing on the diagonal case, it is possible to verify the model's capacity of learning, even that some parameters, like a_{11} , present a high level of uncertainty associated with their posterior distribution. Another characteristic observed is the mass concentration of the posterior density of $\tilde{\delta}$ near its inferior boundary, which seems that the model is not flexible sufficient to accommodate the data. Another characteristic already verified in other estimations is the smaller uncertainty associated with the parameters that map the effect from previous data into the Poisson

¹Mean: Posterior Mean²Naive Standard Error $\times 10^{-3}$ ³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

rate λ_t . Now looking into the non-diagonal one, is verifiable some parameters difficulty of mixing, like a_{21} and ω_2 . Additionally, others presented a reduction in its prior variance, but was not capable of learning, as verified in b_{12} , b_{21} , and a_{12} . Another behavior presented is associated with a_{11} . In the diagonal case, this parameter seemed to have incorporated some information from the data even with a higher variance but, with the addition of flexibility to the model, this information were captured by other parameters. The behavior of $\tilde{\delta}$ was similar as in the diagonal case. About the convergence of the chains, in both the autocorrelation plots 4.10 and 4.12 is possible to observe the difficulty of the chains in converging.

At table 4.20, are compared the estimates at Cui and Zhu [2017] with the ones presented in this work. A burn-in size of 500000, together with a sampling containing 500000 data points, were used for computing the posterior distribution in both cases. As already observed, in general, the parameters associated with the constant contribution to the Poisson rates tended to have a higher influence in λ_t when compared to the MLE estimates presented by the reference author. Also, the parameters associated with previous data contribution tended to be quite similar in both estimation procedures. Also, the behavior presented by the crossed-effect parameters a_{12} , a_{21} , b_{12} , and b_{21} exhibited the same behavior for both methodologies, having its values tending to zero. Analyzing the measures of fitting,

Table 4.19: *Crime data - Descriptive Analysis*

		N	Mean	SD	Median	Trimmed*	MAD [†]	Min	Max	Range	Skew	Kurtosis	SE [‡]
CMIS	1	120	6.0	3.799	5	5.708	3.0	0	18	18	0.8	0.235	0.3
ROB	2	120	2.1	1.591	2	1.927	1.5	0	8	8	1.0	0.793	0.1

* Trimmed, trimmed sample mean with trim defaulting to 0.1

[†] MAD, mean absolute deviation

[‡] SE, standard error of the mean

Figure 4.8: *Pittsburgh Crime Data - Descriptive Analysis - ROB, Robbery; CMIS Criminal mischief*

Table 4.20: *Estimates for Liu's, Cui's and proposed model - Crime Data*

Parameters	Diagonal		Proposed [†]			Non-Diagonal		Proposed [†]		
	Liu's	Cui's	Mean ¹	SE ²	TSSE ³	Liu's	Cui's	Mean	SE	TSSE
ω_1	0.4561	0.4930	1.1676	0.0006	0.0095	0.4443	0.4820	1.1132	0.0005	0.0195
ω_2	3.5322	3.5274	3.0658	0.0009	0.0159	3.5324	3.5313	2.7802	0.0007	0.0499
a_{11}	0.7051	0.6859	0.3635	0.0003	0.0041	0.7044	0.6848	0.2304	0.0002	0.0057
a_{12}						0.0001	0.0001	0.0495	0.0001	0.0014
a_{21}						0.0001	0.0001	0.1626	0.0001	0.0122
a_{22}	0.0929	0.0923	0.1789	0.0001	0.0021	0.0920	0.0909	0.1231	0.0001	0.0041
b_{11}	0.0804	0.0741	0.0839	0.0001	0.0010	0.0811	0.0747	0.0741	0.0001	0.0009
b_{12}						0.0001	0.0001	0.0118	0.0000	0.0001
b_{21}						0.0001	0.0001	0.1121	0.0001	0.0057
b_{22}	0.3288	0.3299	0.3194	0.0001	0.0014	0.3290	0.3300	0.3303	0.0001	0.0038
δ		-6.4350	-0.0975	0.0013	0.0120		-6.4363	-0.0527	0.0013	0.0261
ϕ	0.0001					0.0001				
Log-Lik	-537.391	-535.9	-541.977			-537.448	-535.95	-543.657		
AIC	1088.782	1085.801	1125.886			1096.897	1093.899	1138.649		
BIC	1107.874	1104.892	1145.399			1126.502	1123.504	1169.312		

[†] BurnIn: 500000. Samples: 500000[‡] BurnIn: 500000. Samples: 500000 **Figure 4.9:** *Trace Charts - Crime Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.*¹Mean: Posterior Mean²Naive Standard Error $\times 10^{-3}$ ³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Figure 4.10: ACF Charts - Crime Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.

Figure 4.11: Trace Charts - Crime Non-Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.

Figure 4.12: ACF Charts - Crime Non-Diagonal - Data Points: 120. Burn In: 500000. Sample Size: 500000.

the Bayesian procedure seemed to have a poor fit when compared to the MLE estimates. This behavior was observed both in the diagonal and in the non-diagonal case. As already mentioned, $\hat{\delta}$ should have a struggling constraint given the level of rates presented in both these crime time series, what makes the estimate presented by Cui and Zhu [2017] of the order -6.4 strange in both situations, even being Δ_c the less restrictive region, as in the reference work. Given the estimates

presented by Cui and Zhu [2017], a chart containing the evolution of $\lambda_{1,t}$, $\lambda_{2,t}$, the boundaries for δ_t used by Cui and Zhu [2017](black lines), and the ones proposed in this work(red lines) can be observed at Figure 4.13. As argued before, the estimate of $\tilde{\delta}$ (green line) presented by the authors has infringed even the limits proposed by them. That is why its construction presented a better adjustment to the data, but their estimates are not valid, since under this configuration, (2.4) is not a probability measure.

Figure 4.13: Cui and Zhu [2017] Pittsburgh Poisson Intensities given estimates. Diagonal case. $\lambda_{1,t}$ and $\lambda_{2,t}$ stands for the Robbery(ROB) and Criminal mischiefs(CMIS) respectively

4.3.3 Forecasting

An out-of-sample forecasting exercise were performed for comparing the performance of estimates exhibited in the reference work with the ones presented in this dissertation. For simplicity, were employed the conditional expectations to compute the predictions of one-step-ahead Poisson rates,

$$\hat{\lambda}_{t+1} = \mathbb{E} [\lambda_{t+1} | \mathcal{F}_t] = \omega + A\lambda_t + BY_t = \sum_{k=0}^{t-1} (A^k \omega + A^k BY_{t-k}) \quad (4.4)$$

To replicate the same exercise implemented by the reference authors, the first 108 data points of the crime dataset were used for estimating the parameters in both diagonal and non-diagonal cases. With the remaining data, a one-step-ahead forecast procedure was implemented using the conditional expectation formula 4.4 just introduced. For the realized forecasts, the MADE and the MSE were computed, having as v^0 in (4.1) and (4.2) the data point realization which the prediction was realized. At table 4.21 are compared the results obtained in this work with the ones achieved in Cui and Zhu [2017].

Table 4.21: *MADE and MSE of Predicted Poisson Rates - Crime Data*

	Diagonal						Non-Diagonal					
	ROB			CMIS			ROB			CMIS		
	Liu's	Cui's	Proposed [†]	Liu's	Cui's	Proposed [†]	Liu's	Cui's	Proposed [‡]	Liu's	Cui's	Proposed [‡]
MADE	1.348	1.333	1.366	2.144	2.139	2.041	1.416	1.386	1.345	1.884	1.87	2.048
MSE	3.06	3.009	3.035	7.295	7.273	7.079	3.051	2.983	2.93	6.468	6.409	6.939

[†] BurnIn: 500000. Samples: 500000

[‡] BurnIn: 500000. Samples: 500000

Since the implementation of bivariate Poisson permitting negative correlation in Cui and Zhu [2017] is not valid, only the comparison to Liu's implementation (permitting only positive correlation) was made. For the diagonal case, where the posteriors have seemed bo better learn from the data during the estimation procedure, the Bayesian method seemed to beat the MLE formulation. For the cases where the forecasting error was not smaller, this error was of the same order. For the non-diagonal case, the message was mixed: the bayesian model performs a better prediction for the ROB time series, but is worse for the CMIS data.

Chapter 5

Conclusion

This work focused on the analysis and in the Bayesian estimation of the bivariate INGARCH(1,1), proposed in the work of Cui and Zhu [2017]. First, we started analyzing the bivariate Poisson formulation proposed by Lakshminarayana *et al.* [1999], the one that Cui and Zhu [2017] used to formulate their version of the bivariate INGARCH. In section 2.2, we studied in depth the conditions necessary for (2.4) being considered a probability measure. The restriction (2.8) to the parameter $\tilde{\delta}$ that Cui and Zhu [2017] used was to be shown incorrect, since this does not respect the inequality (2.7) for all possible elements of the sample space of the bivariate pair (Y_1, Y_2) , given the triple $(\lambda_1, \lambda_2, \tilde{\delta})$. To infringe the inequality (2.7), results in points of negative probability, as shown in figure 2.1. Another consequence of this misspecification is associated with the correlation coefficient. Cui and Zhu [2017] needed to impose a hard limit to the parameter since, in their formulation, correlations greater in modulus than one are possible. In the same section, a proposition of what would be the right region where the parameter $\tilde{\delta}$ would lie inside, Δ_c , was constructed. Differently from the incorrect region 2.8, which is symmetric, the correct support of the linkage parameter (2.13) is not, and have different behavior when compared to the one developed by Cui and Zhu [2017]: is limited at the origin and, as the Poisson rates become large, the positive limit of Δ_c increases. A benefit from rightly specifying the correct region Δ_c is that no correlation coefficient constraints are necessary to be imposed, as presented by the reference authors. A problem verified with the formulation of the bivariate Poisson proposed by Lakshminarayana *et al.* [1999] is that this model works well for only lower Poisson rates, of the order 1. At Figure 2.5 is shown that, for a pair of lambdas equal (1.2, 0.97) approximately, the minimum correlation this model is capable of capturing is of the order -0.43, which is very restrictive. Additionally, even $\tilde{\delta}$ can grow indefinitely for positive values as λ_1 and λ_2 increase in magnitude, the formulation of correlation formula (C.13) makes this coefficient tends to zero, being a problem when applying the model to real data. A possible solution can be associated in proposing another functional form to the polynomials in equation (2.3). The combination of an order zero polynomial with the dump function e^{-x} makes this region Δ_c to be depended on the levels of λ_1 and λ_2 , creating a nonlinear mapping between $\tilde{\delta}$ and the correlation coefficient. Regarding the bivariate INGARCH(1,1) proposed by 3.2, its analysis was realized in depth, objecting fully understand its behavior during the estimation. Since a dynamic structure was imposed, some adjustments were implemented to the likelihood function to adequate the model to its constraints. In specific, the likelihood augmentation was performed to incorporate the indicator variable $\mathcal{I}_t$, which indicates if δ_t is in a valid region for that specific time instant. It is difficult to calculate the unconditional probability of this variable but, conditioning to the current

δ_t and λ_t , its probability collapses to a binary and deterministic variable. Additionally, the form in which the likelihood function (3.11) was written, permits the extension of the model to incorporate stochasticity to δ_t , since is explicit its dependence on $\mathbb{P}(\Delta_t = \delta_t | \psi, \delta_{t-1}, y_{t-1})$. When, given the parameter vector θ and the data, δ_t is fully determined for each time instant, this represents the deterministic case, and this probability collapses to one. Several studies were executed, mainly based on replicating the ones implemented by Cui and Zhu [2017]. First, a study using simulated data were implemented, targeting comprehend the performance and convergence behavior of the proposed estimation method. This first part was subdivided into two analyses: the first one tried to access the behavior of the posterior convergence varying the burn-in, the sample size, and the number of data points used during the process. The second part intended to assess the performance of the estimators. Some conclusions were obtained from this study. First, samples of size greater than 50 are larger enough and almost not bring new information to the model. Additionally, as the sample size increased, a slightly marginal decrease occurred in the posterior standard deviation. For the simulated cases, where the data generation process is stable, a burn-in of 15000 and a sample size of 100000 showed to be sufficient to compute the posterior distribution of the parameters. When evaluating performance, the parameters that map the contribution of previous data realizations tended to learn with much less uncertainty when compared to the other parameters. Also, the constant vector ω tended to incorporate part of the information that elements of matrix A should have captured, implicating in some situations the presence of bias. When a dynamic structure was imposed to δ_t , a problem of confusion seemed to have appeared, since the model exhibited a difficulty to define from which parameter the information contained in the data should be associated. This behavior can be linked to the weak transmission that the data have to the parameter δ_t . This behavior was observed by the large uncertainty present on its posterior distribution, in almost all posterior estimations. To finalize, some applications were executed to real data. This demonstrates that the Bayesian framework is adequately capable of capture information contained in several types of datasets. A necessary improvement would be the proposition of another format for the bivariate probability function function (2.4), one that adequately treats the correlation coefficient and does not subjugate this statistic to the level of the Poisson rates. Also, the implementation of an estimation procedure that monitor the issue of convergence would result in better estimates.

Appendix A

Markov Chain Monte Carlo

At this appendix are covered the concepts that guarantee the conditions required for Markov Chain sampling. The topic is examined for the discrete case, which is sufficient to give the intuition for the Metropolis-Hastings Algorithm. A discussion in depth can be found at [Gamerman and Lopes \[2006\]](#), [Robert and Casella \[2004\]](#) and [Hoel *et al.* \[1972\]](#).

The big idea behind the Markov Chain Monte Carlo method(MCMC) is to obtain a sample from a target distribution π using a Markov Chain, $X_t, t \geq 0$, which has stationary distribution π . Regarding the work of a Bayesian statistician, π is going to be the distribution of the parameters used for Bayesian inference. For a Markov Chain to have this limiting characteristic, some properties must be satisfied to guarantee its convergence. Let us start by defining the Markovian Property, what is a stationary distribution, and what are these necessary characteristics for the MCMC method.

A.0.1 Markov Chains

Markovian Property

Definition A.0.1 *Consider that we are in an environment with finite or with a countably infinite number of states, X_t . The set of all possible values that X_t can assume is called State Space, S . If the probability of transitions from any state, $X_t \in S$, to another one, $X_{t+1} \in S$, present the following property*

$$\mathbb{P}(Y_t = y_t | Y_{t-1} = y_{t-1}, \dots, Y_1 = y_1) = \mathbb{P}(Y_t = y_t | Y_{t-1} = y_{t-1}) \quad (\text{A.1})$$

is said that this system presents the Markovian property and by consequence this is going to be a Markov Chain. For simplicity of notation, the probability of transition from the state y_{t-1} to the state y_t from now on will be noted as $\mathbb{P}(Y_t = y_t | Y_{t-1} = y_{t-1}) = Q(y_{t-1}, y_t)$.

The intuition behind (A.1) relates to the fact that the former state of the chain X_{t-1} only can affect the current one X_t , not the future. This presents a very interesting simplification for modeling complex systems since the diffusion probability depends only locally on its most recent realization. Additionally, if the transition probability (A.1) is independent of its index t , it is said that this Markov chain is homogeneous, meaning that the transitions do not depend on the time instant, only from which state the chain is in, and where the process is moving to.

Stationarity

Definition A.0.2 Let $X_t, t \geq 0$ be a Markov Chain defined over the state space S with probability of transition defined as in (A.1). Additionally, consider the function $\pi(x)$ defined over S such that

$$\sum_{x \in S} \pi(x) = 1$$

If π presents the following property

$$\sum_{x \in S} \pi(x) Q(x, y) = \pi(y), \quad y \in S$$

then π is called stationary distribution.

Stationarity is related to the invariance of the probability measure at different time instants. In simple words, the unconditional probability distribution at time $t > 0$ is the same as in time $t' \neq t$ for a chain having a stationary distribution. For one that does not present this property, the unconditional probability distribution vary on each transition, being a problem thinking in terms of building a sampling scheme by means of a Markov chain. For systems that are known to have an invariant measure π , the question that appears is if exists and under which conditions the chain's evolution is going to converge to the invariant measure

$$\lim_{t \rightarrow \infty} Q^t(x, y) = \lim_{t \rightarrow \infty} \sum_{x' \in S} Q^{t-1}(x, x') Q(x', y) = \pi(y), \quad y \in S \quad (\text{A.2})$$

If the convergence presented in (A.2) occurs, this means that, independent of the initial state the chain is situated, is possible to compute the invariant measure at state y , making this chain a potent tool for sampling purposes. The theorem 7/chapter 2 in Hoel *et al.* [1972] guarantee under what conditions this happens.

Theorem A.0.1 (Hoel *et al.* [1972] Chapter 2/Theorem 7) Let $X_t, t > 0$ be an irreducible positive recurrent Markov Chain having stationary distribution π . If the chain is aperiodic

$$\lim_{t \rightarrow \infty} Q^t(x, y) = \pi(y), \quad x, y \in S$$

If the chain is periodic with a period d , then for each pair x, y of states in S there is an integer $r, 0 \leq r < d$, such that $Q^t(x, y) = 0$ unless $t = nd + r$ for some nonnegative integer n and

$$\lim_{t \rightarrow \infty} Q^{td+r}(x, y) = \pi(y), \quad x, y \in S$$

The second part of the theorem has the convergence only under certain specific conditions, different from the aperiodic case where the convergence always occurs. Thus, the consequent evolutions

of a chain that is irreducible, positive recurrent, aperiodic, and having a stationary distribution π have its realizations being distributed unconditionally by the invariant measure π , being a powerful tool for simulation purposes. Over chains with these characteristics that the Metropolis-Hastings algorithm is built. On the next subsections, the missing concepts for the full comprehension of this theorem are going to be announced.

Recurrence

The hitting time T_A of a subset $A \subset S$ is defined as the minimum time required for the state X_t be an element of A .

$$T_A = \min(t > 0 : X_t \in A)$$

If $\nexists t > 0$ such that $X_t \in A$, then $T_A = \infty$, meaning the chain X_n will never hit any state of the subset A . This concept permits quantify the chances of transitions taking in consideration the time evolution of the chain. At a finite time instant $t' > t$, the probability of a Markovian Process move from $X_n = x$ to $X_{t'} = y$ can be noted as

$$\rho_{xy} = \mathbb{P}_x(T_y < \infty) = \mathbb{P}(T_y < \infty | X_t = x)$$

As a consequence, a chain departing from a state x has its probability of returning in finite time to the same state denoted as ρ_{xx} . If this return measure is less than one, this means that it exists a chance that the chain will never come back to its departure state. Formally, the definition of recurrence can be stated by

Definition A.0.3 *Let X_t be a Markov Chain defined over a state space S , $\forall t \geq 0$. Consider that at the time t , the chain is at the state $X_t = x$ and that we want to compute the probability of the chain return to its initial state $X_{t'} = x$, for some time $t' > t$.*

1. *if $\rho_{xx} = 1$, we say that x is a recurrent state. Moreover, if the expected return time to the state x is finite, x is going to said a positive recurrent state. If not, x is going to be called a null recurrent state.*
2. *else $\rho_{xx} < 1$, we say that x is a transient state*

Recurrence is important since it delivers the idea that a state can be drawn with some frequency, irrespective of whether the return time to the same state is high or low. The figure [A.1](#) represents a transient chain. For example, consider a process having as its initial state A . The probability ρ_{AA} equals 0, indicates that this chain, after having done its first transition, will never return back to A . For sampling purposes, to have a chain with transient states is a problem, since this can lead the process to get stuck at some subset of its state space. This makes the full coverage of the sampled variable support being compromised. Additionally, a chain with null recurrent states has the average return time to these states not finite, incurring into non-stationarity of the limiting distribution of the chain. Hence, positive recurrence is a necessary property for all states present in a Markov Chain used as a sampling scheme.

Figure A.1: Example: Graph diagram of a transient chain.

Irreducibility

To might access any state, given an initial reference, is another feature that is important for a Markov based sampling scheme to have. This is related to the idea of irreducibility.

Definition A.0.4 A closed set $C \subset S$ is said to irreducible if

$$\rho_{xy} = 1, \forall x, y \in S$$

Moreover, If all states of the state space S are irreducible, then the Markov Chain is said to be irreducible.

The Figure A.2 represents an irreducible Markov Chain. With probability 1, is possible to move from any state to another one at a finite time.

Figure A.2: Example: Graph diagram of a irreducible chain.

Periodicity

Periodicity of a state regards the idea of how many transitions a Markov chain must have to be possible return to a specific state departing from the same one.

Definition A.0.5 Consider $X_t, t > 0$, a irreducible Markov Chain defined over S . Additionally, as in [Hoel et al. \[1972\]](#), the period d of any state $x \in S$, where $Q_t(x, x) > 0$, is given by

$$d = \text{g.c.d}\{n \geq 1 : Q^n(x, x) > 0\} \quad (\text{A.3})$$

Note that the period is common for every state $x \in S$ since X_t is an irreducible chain. Furthermore, if d is equal 1, X_n is said to be an aperiodic chain. Otherwise, the Markov Chain is said to be periodic with period d .

Analyzing example [A.2](#), it is possible to verify that the period of the state A is equal 2. Departing from A is possible to return to the same state after $\{2, 4, 6, \dots\}$ transitions, implying by definition [A.0.5](#) that the great common division for this set is 2. Now at the chain in [A.3](#), returns to the state A can happen after $\{1, 2, 3, \dots\}$ transitions, making this state aperiodic. Additionally, since the chain [A.3](#) is irreducible, every state of this chain is aperiodic, resulting in an aperiodic chain.

Figure A.3: Example: Graph diagram of an aperiodic chain.

Reversibility

A reversible Markov Chain, with invariant measure π , is such that

$$\pi(x) Q(x, y) = \pi(y) Q(y, x), \quad \forall x, y \in S \quad (\text{A.4})$$

The intuition behind this concept is related to the rates that a chain moves from the state x to y , and vice versa in the equilibrium. Actually, they are the same for a reversible Markov chain.

This concept is important since it assures positive recurrence to a Markov process defined over an irreducible state space, making the convergence observed by theorem (A.0.1) valid for a set up with these characteristics.

A.0.2 The Random Walk Metropolis-Hastings Algorithm

As well described in [Gamerman and Lopes \[2006\]](#), consider a state space S , where each state x have probability measure p_x . The idea proposed by [Metropolis *et al.* \[1953\]](#) has its essence in assembling a Markov chain with stationary distribution π , such that $\pi(x) = p_x, \forall x \in S$. For this construction, the authors proposed construct a Markovian process $X_t \in S$ defined using an irreducible transition symmetric matrix $Q(x, y) = Q(y, x), \forall x, y \in S$. In addition, the proposed value y for the next state X_{t+1} is accepted with probability $\min(1, p_y/p_x)$. Otherwise, the chain remains at the same state x . Under these conditions, the authors showed that, if the probabilities p_x satisfy the reversibility condition (A.4), π is a stationary distribution. Additionally, if Q is aperiodic, π is also the limiting distribution of the chain.

[Hastings \[1970\]](#) propose a generalization of the work of [Metropolis *et al.* \[1953\]](#) defining the acceptance probability in a form that, when combined with an arbitrary transition kernel $q(\cdot, \cdot)$, it defines a reversible chain. Let ϕ be a proposition for a new state and θ the current one of this chain. The acceptance probability proposed by [Hastings \[1970\]](#) is defined as

$$\alpha(\theta, \phi) = \min\left(1, \frac{\pi(\phi) q(\phi, \theta)}{\pi(\theta) q(\theta, \phi)}\right) \quad (\text{A.5})$$

The algorithm based on chains with acceptance probability defined by (A.5), together with a symmetric transition kernel, as discussed before, is called the Metropolis-Hastings algorithm.

Consider the Random Walk process

$$\theta^{(t)} = \theta^{(t-1)} + \epsilon_t \quad (\text{A.6})$$

where the disturbances ϵ_t have distribution f symmetric around zero. This process have a symmetric transition matrix and, by the announced above, its limiting distribution is the invariant measure π . As a consequence, we can set up a sampling scheme based on a random walk to obtain samples of the limiting distribution of the parameters that we are interested in perform inference. If we select a transition kernel $q(\theta, \phi) = f_{\epsilon_t}(\phi - \theta)$ for this random walk process, the random walk Metropolis-Hastings can be performed following the steps in algorithm 2.

Algorithm 2

-
- 1: **procedure** RANDOM WALK METROPOLIS-HASTINGS ALGORITHM
  - 2:   Set a initial state for the chain θ_0
  - 3: *Do until convergence:*
  - 4:   Draw ϕ from the proposal transition kernel $q(\theta^{(t-1)}, \cdot)$
  - 5:   Draw $U \sim \text{Unif}(0, 1)$ and evaluate the acceptance probability $\alpha(\theta^{(t-1)}, \phi)$ defined in (A.5):
  - 6:   **if** $\alpha(\theta^{(t-1)}, \phi) < u$ **then**
  - 7:   Accept the move of the chain: $\theta^{(t)} \leftarrow \phi$
  - 8:   **else**
  - 9:   Reject the move of the chain: $\theta^{(t)} \leftarrow \theta^{(t-1)}$
-

As proposal for the disturbances ϵ_t , this work uses a t distribution centered at zero, as first introduced by Geweke [1991]. Here, since it deals with the multivariate case, the kernel $t_\nu(\mu, \Sigma)$ proposed is

$$f_\nu(\epsilon_t | \mu, \Sigma) \Big|_{\mu=0} = \frac{\Gamma[(\nu + p)/2]}{\Gamma(\nu/2) \nu^{p/2} \pi^{p/2} |\Sigma|^{1/2}} \left[1 + \frac{1}{\nu} \epsilon_t' \Sigma^{-1} \epsilon_t \right] \quad (\text{A.7})$$

where ν is the number of degrees of freedom, Σ is a $p \times p$ positive definite shape matrix, and μ is a p -dimensional location parameter. The selection of Σ was made like in the method three proposed in Bennett *et al.* [1996], where the authors used for the multivariate normal case the covariance matrix being the inverse of the observed information at the current value $\theta^{(t-1)}$

$$\Sigma = \mathcal{I}(\theta^{(t-1)})^{-1} = - \left[\nabla \nabla' l(\theta | \mathbf{Y} = \mathbf{y}) \right]_{\theta=\theta^{(t-1)}}^{-1} \quad (\text{A.8})$$

$l(\theta | \mathbf{Y} = \mathbf{y})$ is the logarithm of the likelihood function (3.7). At Bennett *et al.* [1996], the Σ calculation was done at each accepted step. Here, Σ was selected only at the initial instant of the chain's evolution θ^0 and tuned by a constant c , as proposed by Tierney and Mira [1999]. The objective in doing so is associated with having a method for selection of the shape matrix Σ , with the necessity of choosing only the parameter c instead of p tuning parameters. If the initial state of the chain has much information being measured, the steps during the sampling process are going to be smaller, since the chain would be in a region where the parameters capture more information from the data. If the initial point of the chain is in a region distant from the more informative one, a larger covariance Σ is going to be selected, increasing the chances of full coverage of the sample space. Another treatment implemented during the selection of Σ , is related to its necessity of being positive definite. Since exists several parameters being estimated, they could possibly depend on each other. After inverting the information matrix, this could be possibly non-positive definite. To adjust this, Σ was decomposed as proposed in Rousseeuw and Molenberghs [1993], having its negative eigenvalues replaced by its absolute values before start the sampling process.

Appendix B

Some Results

B.0.1 Demonstration - Inequality (B.1)

Let $\lambda_i, \lambda_j \in \mathbb{R}^{*+}$, $i, j \in \{1, 2\}$, and $i \neq j$. For $\lambda_i > \lambda_j$

$$\begin{aligned}
 \lambda_i > \lambda_j &\iff -c\lambda_i < -c\lambda_j \\
 &\iff e^{-c\lambda_i} < e^{-c\lambda_j} \\
 &\iff 1 - e^{-c\lambda_i} > 1 - e^{-c\lambda_j} \\
 &\iff (1 - e^{-c\lambda_i}) e^{-c\lambda_j} > (1 - e^{-c\lambda_j}) e^{-c\lambda_j} > (1 - e^{-c\lambda_j}) e^{-c\lambda_i} \\
 &\iff \frac{1}{(1 - e^{-c\lambda_i}) e^{-c\lambda_j}} < \frac{1}{(1 - e^{-c\lambda_j}) e^{-c\lambda_i}}
 \end{aligned}$$

Then

$$\lambda_i > \lambda_j \iff \frac{1}{(1 - e^{-c\lambda_i}) e^{-c\lambda_j}} < \frac{1}{(1 - e^{-c\lambda_j}) e^{-c\lambda_i}} \quad (\text{B.1})$$

B.0.2 Demonstration - Correlation Boundary - $\tau > 0$ and $\tilde{\delta} < 0$

Consider the case for positive values of $\tau > 0$. The inequality (2.16) becomes

$$\begin{aligned}
 \rho_{Y_1, Y_2} = \tilde{\delta} c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1 + \lambda_2)} &\geq \frac{-c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1 + \lambda_2)}}{(1 - e^{-c\lambda_1}) (1 - e^{-c\lambda_2})} \\
 &= \frac{-c \sqrt{(c\lambda_1 \pm 1) (c\lambda_2 \pm 1)} e^{-c(\lambda_1 + \lambda_2)}}{(1 - e^{-c\lambda_1}) (1 - e^{-c\lambda_2})} \\
 &= \frac{-c \sqrt{((1 + c\lambda_1) - 1) ((1 + c\lambda_2) - 1)} e^{-c(\lambda_1 + \lambda_2)}}{(1 - e^{-c\lambda_1}) (1 - e^{-c\lambda_2})}
 \end{aligned}$$

Using the Bernoulli inequality

$$\rho_{Y_1, Y_2} \geq \frac{-c \sqrt{\left(\left(1 + c \frac{\lambda_1}{n}\right)^n - 1 \right) \left(\left(1 + c \frac{\lambda_2}{m}\right)^m - 1 \right) e^{-c(\lambda_1 + \lambda_2)}}{(1 - e^{-c\lambda_1})(1 - e^{-c\lambda_2})}, \forall n, m \in \mathbb{N}$$

$$\xrightarrow{n, m \rightarrow \infty} \frac{-c}{\sqrt{(1 - e^{c\lambda_1})(1 - e^{c\lambda_2})}}$$

Then

$$\rho_{Y_1, Y_2} \geq \frac{-c}{\sqrt{(1 - e^{c\lambda_1})(1 - e^{c\lambda_2})}}, \quad (\text{B.2})$$

B.0.3 Demonstration - Correlation Boundary - $\tau > 0$ and $\tilde{\delta} < 0$

Let $\lambda_i > \lambda_j$ constrained to $1 - e^{-c\lambda_i} - e^{-c\lambda_j} > 0$ associated with ρ_{Y_i, Y_j} . For $\lambda_i = \lambda_j$, we have already showed in (2.18) that $\rho_{Y_i, Y_j} \geq -c$. From the restriction $\tau > 0$, we can derive the following relationship

$$1 - e^{-c\lambda_i} - e^{-c\lambda_j} > 0 \iff \frac{e^{c\lambda_j}}{e^{c\lambda_j} - 1} < e^{c\lambda_i} \quad (\text{B.3})$$

From (B.3) and (B.2)

$$\rho_{Y_i, Y_j} \geq \frac{-c}{\sqrt{(1 - e^{c\lambda_i})(1 - e^{c\lambda_j})}} > \frac{-c}{\sqrt{\left(1 - \frac{e^{c\lambda_j}}{e^{c\lambda_j} - 1}\right)(1 - e^{c\lambda_j})}} = -c \quad (\text{B.4})$$

Then, for any pair of Poisson rates selected in the region defined by $\tau > 0$, the minimum correlation coefficient possible of being achieved is limited from below by -c exclusive.

B.0.4 Demonstration - Correlation Boundary - $\tau \leq 0$ and $\tilde{\delta} < 0$

Let $\lambda_i = \lambda_j$, subject to $1 - e^{-c\lambda_i} - e^{-c\lambda_j} \leq 0$, and $\lambda'_i > \lambda'_j$ constrained to $1 - e^{-c\lambda'_i} - e^{-c\lambda'_j} \leq 0$. Additionally, take $\tilde{\delta}$ and $\tilde{\delta}'$,

$$\tilde{\delta} = \frac{-1}{e^{-2c(\lambda_i)}}$$

$$\tilde{\delta}' = \frac{-1}{e^{-c(\lambda'_i + \lambda'_j)}}$$

associated respectively to the correlation coefficients ρ and ρ' . Assume that $\rho' < \rho$. By (2.19)

$$-c^2 \sqrt{\lambda'_i \lambda'_j} < -c \log 2 \iff \lambda'_i > \frac{c^{-2}}{\lambda'_j} \log^2 2$$

Resulting that

$$1 - e^{-c \left(\frac{c^{-2}}{\lambda'_j} \log^2 2 \right)} - e^{-c \lambda'_j} < 1 - e^{-c \lambda'_i} - e^{-c \lambda'_j} \leq 0 \quad (\text{B.5})$$

As presented in (2.21), ρ is bounded by $-c \log 2$. To prove that a correlation coefficient ρ' , associated with $\tilde{\delta}'$, λ'_i , and λ'_j s.t. $\tau' \leq 0$ is limited inferiorly by $-c \log 2$, by a valid selection of λ'_i and λ'_j , the inequality (B.5) must be contradicted. Let's select $\lambda'_i = c^{-1} \log 3$. This implies that λ'_j must be chosen such that

$$1 - e^{-\log 3} - e^{-c \lambda'_j} \leq 0 \iff \lambda'_j \leq c^{-1} \log \frac{3}{2}$$

Selecting $\lambda'_j = c^{-1} \log \frac{3}{2}$ and evaluating the condition (B.5)

$$1 - e^{-(\log \frac{3}{2})^{-1} \log^2 2} - e^{-\log \frac{3}{2}} = 0.0275697$$

results in a contradiction of (B.5). This implies that under constraint $\tau \leq 0$, the minimum possible correlation coefficient achieved, for any valid λ_i and λ_j , is going to be bounded by $-c \log 2 \approx -0.4381526$.

Appendix C

Moments Calculation

C.1 Characteristic Function

If $Y \sim BP(\lambda, \tilde{\delta})$ (2.4), its characteristic function is going to be

$$\begin{aligned}
 \phi(t)_Y &= \mathbb{E} \left[e^{it'Y} \right] = \mathbb{E} \left[\mathbb{E} \left[e^{i(t_1 Y_1 + t_2 Y_2)} \middle| Y_1 \right] \right] \\
 &= \mathbb{E} \left[\sum_{k=0}^{\infty} e^{i(t_1 Y_1 + t_2 k)} \frac{\lambda_2^k}{k!} e^{-\lambda_2} \left[1 + \tilde{\delta} \left(e^{-Y_1} - e^{-c\lambda_1} \right) \left(e^{-k} - e^{-c\lambda_2} \right) \right] \right] \\
 &= \mathbb{E} \left[e^{it_1 Y_1 - \lambda_2} \sum_{k=0}^{\infty} \frac{(\lambda_2 e^{it_2})^k}{k!} \right] \\
 &\quad + \mathbb{E} \left[\tilde{\delta} \left(e^{-Y_1} - e^{-c\lambda_1} \right) e^{it_1 Y_1} e^{-\lambda_2} \left(\sum_{k=0}^{\infty} \frac{(\lambda_2 e^{it_2-1})^k}{k!} - e^{-c\lambda_2} \sum_{k=0}^{\infty} \frac{(\lambda_2 e^{it_2})^k}{k!} \right) \right] \\
 &= \mathbb{E} \left[e^{it_1 Y_1} \right] e^{\lambda_2(e^{it_2}-1)} + \tilde{\delta} \left(\mathbb{E} \left[e^{Y_1(it_1-1)} \right] - e^{-c\lambda_1} \mathbb{E} \left[e^{it_1 Y_1} \right] \right) \left(e^{\lambda_2(e^{it_2-1}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \\
 &= e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} + \tilde{\delta} \left(e^{\lambda_1(e^{it_1-1}-1)} - e^{\lambda_1(e^{it_1}-(1+c))} \right) \left(e^{\lambda_2(e^{it_2-1}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right)
 \end{aligned} \tag{C.1}$$

□

Given C.1, we are able to calculate the Bivariate Poisson first and second moments by taking its derivatives

$$\left. \frac{d^n}{dt^n} \phi(t)_Y \right|_{t=0} = i^n \mathbb{E} [Y^n] \tag{C.2}$$

C.2 First Order Moment

Using the characteristic function C.1 and taking its first derivative

$$\left. \frac{d}{dt} \phi(t)_Y \right|_{t=0} = i\mathbb{E}[Y] \implies \mathbb{E}[Y] = \frac{1}{i} \left. \frac{d}{dt} \phi(t)_Y \right|_{t=0}$$

Taking the derivative in terms of t_1 :

$$\begin{aligned} \frac{d}{dt_1} \phi(t)_Y &= \frac{d}{dt_1} \left[e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} + \tilde{\delta} \left(e^{\lambda_1(e^{it_1-1}-1)} - e^{\lambda_1(e^{it_1}-(1+c))} \right) \left(e^{\lambda_2(e^{it_2-1}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \right] \\ &= i\lambda_1 e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{it_1} \\ &\quad + i\tilde{\delta}\lambda_1 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) \left(e^{\lambda_2(e^{it_2-1}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \end{aligned} \quad (\text{C.3})$$

Evaluating C.3 at $t = 0$

$$\begin{aligned} \left. \frac{d}{dt_1} \phi(t)_Y \right|_{t=0} &= i\lambda_1 + i\tilde{\delta}\lambda_1 \left(e^{-\lambda_1(1-e^{-1})-1} - e^{\lambda_1(1-(1+c))} \right) \left(e^{-\lambda_2(1-e^{-1})} - e^{\lambda_2(1-(1+c))} \right) \\ &= i\lambda_1 + i\tilde{\delta}\lambda_1 \left(e^{-\lambda_1 c-1} - e^{-\lambda_1 c} \right) \left(e^{-\lambda_2 c} - e^{-\lambda_2 c} \right) \\ &= i\lambda_1 \end{aligned}$$

Then the expected value for the component Y_1 is going to be

$$\mathbb{E}[Y_1] = \lambda_1 \quad (\text{C.4})$$

The analogous can be done in terms of the second component of the vector t to determine the first moment of the vector Y .

C.3 Second Order Moment

Taking the second order derivative of C.1, we are able to calculate the second order moment of the probability distribution 2.4 used in this work

$$\left. \frac{d^2}{dt^2} \phi(t)_Y \right|_{t=0} = -\mathbb{E}[YY'] \implies \mathbb{E}[YY'] = - \left. \frac{d^2}{dt^2} \phi(t)_Y \right|_{t=0}$$

Calculating the derivative of C.3 in terms of t_1

$$\begin{aligned}
\frac{d^2}{dt_1^2} \phi(t)_Y &= \frac{d}{dt_1} \left[i\lambda_1 e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{it_1} \right] \\
&+ \frac{d}{dt_1} \left[i\tilde{\delta}\lambda_1 \left(e^{\lambda_1(e^{it_1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) \left(e^{\lambda_2(e^{it_2}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \right] \\
&= i\lambda_1 \frac{d}{dt_1} \left[e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{it_1} \right] \\
&+ i\tilde{\delta}\lambda_1 \left(e^{\lambda_2(e^{it_2}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \frac{d}{dt_1} \left[e^{\lambda_1(e^{it_1}-1)} e^{it_1-1} \right] \\
&- i\tilde{\delta}\lambda_1 \left(e^{\lambda_2(e^{it_2}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \frac{d}{dt_1} \left[e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right] \\
&= -\lambda_1 e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{it_1} \left(e^{it_1} + 1 \right) \\
&- \lambda_1 \tilde{\delta} \left(e^{\lambda_2(e^{it_2}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) e^{\lambda_1(e^{it_1}-1+it_1-1)} \left(\lambda_1 e^{it_1-1} + 1 \right) \\
&+ 2\lambda_1 \tilde{\delta} \left(e^{\lambda_2(e^{it_2}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \left(e^{e^{2it_1}-(1+c)} \right)
\end{aligned} \tag{C.5}$$

Evaluating C.5 at $t = 0$

$$\begin{aligned}
\left. \frac{d^2}{dt_1^2} \phi(t)_Y \right|_{t=0} &= -2\lambda_1 - \lambda_1 \tilde{\delta} \left(e^{-c\lambda_2} - e^{-c\lambda_2} \right) e^{-c\lambda_1} \left(\lambda_1 e^{-1} + 1 \right) + 2\lambda_1 \tilde{\delta} \left(e^{-c\lambda_2} - e^{-c\lambda_2} \right) \left(e^{-c} \right) \\
&= -2\lambda_1
\end{aligned} \tag{C.6}$$

Hence, the second order moment for Y_1 is

$$\mathbb{E} \left[Y_1^2 \right] = 2\lambda_1 \tag{C.7}$$

And, by definition, we can calculate the variance of Y_1

$$\begin{aligned}
Var[Y_1] &= \mathbb{E} \left[Y_1^2 \right] - \mathbb{E} \left[Y_1 \right]^2 \\
&= 2\lambda_1 - \lambda_1 \\
&= \lambda_1
\end{aligned} \tag{C.8}$$

Taking the derivative of C.3 in terms of t_2

$$\begin{aligned}
\frac{\partial^2}{\partial t_2 \partial t_1} \phi(t)_Y &= \frac{\partial}{\partial t_2} \left[i\lambda_1 e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{it_1} \right] \\
&+ \frac{\partial}{\partial t_2} \left[i\tilde{\delta}\lambda_1 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) \left(e^{\lambda_2(e^{it_2-1}-1)} - e^{\lambda_2(e^{it_2}-(1+c))} \right) \right] \\
&= i\lambda_1 e^{\lambda_1(e^{it_1}-1)} e^{it_1} \frac{\partial}{\partial t_2} \left[e^{\lambda_2(e^{it_2}-1)} \right] \\
&+ i\tilde{\delta}\lambda_1 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) \frac{\partial}{\partial t_2} \left[e^{\lambda_2(e^{it_2-1}-1)} \right] \\
&- i\tilde{\delta}\lambda_1 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) \frac{\partial}{\partial t_2} \left[e^{\lambda_2(e^{it_2}-(1+c))} \right] \\
&= -\lambda_1 \lambda_2 e^{\lambda_1(e^{it_1}-1)} e^{\lambda_2(e^{it_2}-1)} e^{i(t_1+t_2)} \\
&- \tilde{\delta}\lambda_1 \lambda_2 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) e^{\lambda_2(e^{it_2-1}-1)} e^{it_2-1} \\
&+ \tilde{\delta}\lambda_1 \lambda_2 \left(e^{\lambda_1(e^{it_1-1}-1)} e^{it_1-1} - e^{\lambda_1(e^{it_1}-(1+c))} e^{it_1} \right) e^{\lambda_2(e^{it_2}-(1+c))} e^{it_2}
\end{aligned} \tag{C.9}$$

Evaluating C.9 at $t = 0$

$$\left. \frac{\partial^2}{\partial t_2 \partial t_1} \phi(t)_Y \right|_{t=0} = -\lambda_1 \lambda_2 + \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)-1} - \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} \tag{C.10}$$

Resulting in the first order cross-moment

$$\mathbb{E}[Y_1 Y_2] = \lambda_1 \lambda_2 - \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)-1} + \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} \tag{C.11}$$

By definition, the covariance between Y_1 and Y_2 is going to be

$$\begin{aligned}
Cov[Y_1, Y_2] &= \mathbb{E}[Y_1 Y_2] - \mathbb{E}[Y_1] \mathbb{E}[Y_2] \\
&= \lambda_1 \lambda_2 - \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)-1} + \tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} - \lambda_1 \lambda_2 \\
&= +\tilde{\delta} c \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)} (1 - e^{-1}) \\
&= +\tilde{\delta} c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)}
\end{aligned} \tag{C.12}$$

Then the correlation coefficient can be calculated by its definition

$$\rho_{Y_1, Y_2} = \frac{Cov[Y_1, Y_2]}{\sqrt{Var[Y_1]} \sqrt{Var[Y_2]}} = \frac{\tilde{\delta} c^2 \lambda_1 \lambda_2 e^{-c(\lambda_1+\lambda_2)}}{\sqrt{\lambda_1} \sqrt{\lambda_2}} = \tilde{\delta} c^2 \sqrt{\lambda_1 \lambda_2} e^{-c(\lambda_1+\lambda_2)} \tag{C.13}$$

C.4 Log-Likelihood Derivatives

Consider the Likelihood function (3.11). Then the Log-Likelihood function is

$$l(\theta|Y = Y) = \sum_{t=1}^n l_t(\theta|Y_t = Y_t) \quad (\text{C.14})$$

where

$$l_t(\theta|Y_t = Y_t) = y_{1,t} \log(\lambda_{1,t}) + y_{2,t} \log(\lambda_{2,t}) - \log(y_{1,t}! y_{2,t}!) - (\lambda_{1,t} + \lambda_{2,t}) + \log(\phi_t) \quad (\text{C.15})$$

C.5 First Order Derivatives

We have that $\theta = [\theta'_1, \theta'_2, \psi]'$, where $\theta_i = [\omega_i, a_{ii}, a_{ij}, b_{ii}, b_{ij}, \lambda_{0,i}]$, $i, j \in \{1, 2\}$, $i \neq j$ and $\psi = [\tilde{\delta}, \alpha, \beta, \delta_0]'$. Calculating the gradient of (C.15)

$$\begin{aligned} \frac{\partial l_t(\theta|Y_t = Y_t)}{\partial \theta_i} &= \left[\frac{y_{i,t}}{\lambda_{i,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_i} \\ &+ \left[\frac{y_{j,t}}{\lambda_{j,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} \right] \frac{\partial \lambda_{j,t}}{\partial \theta_i} \\ &+ \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial \psi_t}{\partial \theta_i}, \quad i \neq j \end{aligned} \quad (\text{C.16})$$

$$\frac{\partial l_t(\theta|Y_t = Y_t)}{\partial \psi} = \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial \psi_t}{\partial \psi}, \quad i \neq j \quad (\text{C.17})$$

where,

$$\begin{aligned} \frac{\partial \psi_t}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i} \left[(1 - \alpha - \beta) \tilde{\delta} + \alpha \psi_{t-1} + \beta u'_{t-1} u_{t-1} \right] \\ &= \alpha \frac{\partial \psi_{t-1}}{\partial \theta_i} + \beta \frac{\partial u'_{t-1} u_{t-1}}{\partial \theta_i} \\ &= \alpha \frac{\partial \psi_{t-1}}{\partial \theta_i} + \beta \frac{(\lambda_{i,t-1}^2 - y_{i,t-1}^2)}{\lambda_{i,t-1}^2} \frac{\partial \lambda_{i,t-1}}{\partial \theta_i} + \beta \frac{(\lambda_{j,t-1}^2 - y_{j,t-1}^2)}{\lambda_{j,t-1}^2} \frac{\partial \lambda_{j,t-1}}{\partial \theta_i} \end{aligned} \quad (\text{C.18})$$

- $\frac{\partial \psi_t}{\partial \psi}$ components:

$$\frac{\partial \psi_t}{\partial \tilde{\delta}} = (1 - \alpha - \beta) + \alpha \frac{\partial \psi_{t-1}}{\partial \tilde{\delta}} \quad (\text{C.19})$$

$$\frac{\partial \psi_t}{\partial \alpha} = -\tilde{\delta} + \psi_{t-1} + \alpha \frac{\partial \psi_{t-1}}{\partial \alpha} \quad (\text{C.20})$$

$$\frac{\partial \psi_t}{\partial \beta} = -\tilde{\delta} + \alpha \frac{\partial \psi_{t-1}}{\partial \beta} + u'_{t-1} u_{t-1} \quad (\text{C.21})$$

$$\frac{\partial \psi_t}{\partial \delta_0} = \alpha \frac{\partial \psi_{t-1}}{\partial \delta_0} \quad (\text{C.22})$$

- $\frac{\partial \lambda_{i,t}}{\partial \theta_i}$ and $\frac{\partial \lambda_{j,t}}{\partial \theta_i}$ components:

$$\frac{\partial \lambda_{i,t}}{\partial \omega_i} = 1 + a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial \omega_i} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial \omega_i} \quad (\text{C.23})$$

$$\frac{\partial \lambda_{j,t}}{\partial \omega_i} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial \omega_i} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial \omega_i} \quad (\text{C.24})$$

$$\frac{\partial \lambda_{i,t}}{\partial a_{ii}} = \lambda_{i,t-1} + a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial a_{ii}} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial a_{ii}} \quad (\text{C.25})$$

$$\frac{\partial \lambda_{j,t}}{\partial a_{ii}} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial a_{ii}} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial a_{ii}}$$

$$\frac{\partial \lambda_{i,t}}{\partial a_{ij}} = \lambda_{j,t-1} + a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial a_{ij}} \quad (\text{C.26})$$

$$\frac{\partial \lambda_{j,t}}{\partial a_{ij}} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial a_{ij}} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}}$$

$$\frac{\partial \lambda_{i,t}}{\partial b_{ii}} = y_{i,t-1} + a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial b_{ii}} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial b_{ii}} \quad (\text{C.27})$$

$$\frac{\partial \lambda_{j,t}}{\partial b_{ii}} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial b_{ii}} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial b_{ii}} \quad (\text{C.28})$$

$$\frac{\partial \lambda_{i,t}}{\partial b_{ij}} = y_{j,t-1} + a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial b_{ij}} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial b_{ij}} \quad (\text{C.29})$$

$$\frac{\partial \lambda_{j,t}}{\partial b_{ij}} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial b_{ij}} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial b_{ij}} \quad (\text{C.30})$$

$$\frac{\partial \lambda_{i,t}}{\partial \lambda_{0,i}} = a_{ii} \frac{\partial \lambda_{i,t-1}}{\partial \lambda_{0,i}} + a_{ij} \frac{\partial \lambda_{j,t-1}}{\partial \lambda_{0,i}} \quad (\text{C.31})$$

$$\frac{\partial \lambda_{j,t}}{\partial \lambda_{0,i}} = a_{jj} \frac{\partial \lambda_{j,t-1}}{\partial \lambda_{0,i}} + a_{ji} \frac{\partial \lambda_{i,t-1}}{\partial \lambda_{0,i}} \quad (\text{C.32})$$

In the case where ψ_t is a constant, α and β are equal zero in (C.18) and the crossed terms a_{ij} and b_{ij} are equal zero, we are able to recover the derivative calculated in [Cui and Zhu \[2017\]](#)

$$\frac{\partial l_t(\theta|Y_t = Y_t)}{\partial \theta_i} = \left[\frac{y_{i,t}}{\lambda_{i,t}} - 1 + \frac{c\tilde{\delta}e^{-c\lambda_{i,t}}(e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{1 + \tilde{\delta}(e^{-y_{1,t}} - e^{-c\lambda_{1,t}})(e^{-y_{2,t}} - e^{-c\lambda_{2,t}})} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_i}$$

C.6 Second Order Derivatives

Taking the derivative of (C.16)

$$\begin{aligned}
\frac{\partial^2 l_t(\theta|Y_t = Y_t)}{\partial \theta_i \partial \theta'_i} = & \left[\frac{y_{i,t}}{\lambda_{i,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial^2 \lambda_{i,t}}{\partial \theta_i \partial \theta'_i} \\
& - \left[\frac{y_{i,t}}{\lambda_{i,t}^2} + \frac{c^2 \psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} + \left(\frac{c\psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right)^2 \right] \frac{\partial \lambda_{i,t}}{\partial \theta_i} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\left(\frac{c}{\phi_t} \right)^2 \psi_t e^{-c(\lambda_{i,t} + \lambda_{j,t})} \right] \frac{\partial \lambda_{j,t}}{\partial \theta_i} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \theta_i} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\frac{y_{j,t}}{\lambda_{j,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} \right] \frac{\partial^2 \lambda_{j,t}}{\partial \theta_i \partial \theta'_i} \\
& - \left[\frac{y_{j,t}}{\lambda_{j,t}^2} + \frac{c^2 \psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} + \left(\frac{c\psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} \right)^2 \right] \frac{\partial \lambda_{j,t}}{\partial \theta_i} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\left(\frac{c}{\phi_t} \right)^2 \psi_t e^{-c(\lambda_{i,t} + \lambda_{j,t})} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_i} \frac{\partial \lambda_{j,t}}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \theta_i} \frac{\partial \lambda_{j,t}}{\partial \theta'_i} \\
& + \left[\frac{(e^{-y_{1,t}} - e^{-c\lambda_{1,t}})(e^{-y_{2,t}} - e^{-c\lambda_{2,t}})}{\phi_t} \right] \frac{\partial^2 \psi_t}{\partial \theta_i \partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t^2} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_i} \frac{\partial \psi_t}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t^2} \right] \frac{\partial \lambda_{j,t}}{\partial \theta_i} \frac{\partial \psi_t}{\partial \theta'_i} \\
& - \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}})(e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right]^2 \frac{\partial \psi_t}{\partial \theta_i} \frac{\partial \psi_t}{\partial \theta'_i}
\end{aligned} \tag{C.33}$$

$$\begin{aligned}
\frac{\partial^2 l_t(\theta|Y_t = Y_t)}{\partial \theta_j \partial \theta'_i} = & \left[\frac{y_{i,t}}{\lambda_{i,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial^2 \lambda_{i,t}}{\partial \theta_j \partial \theta'_i} \\
& - \left[\frac{y_{i,t}}{\lambda_{i,t}^2} + \frac{c^2 \psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} + \left(\frac{c\psi_t e^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right)^2 \right] \frac{\partial \lambda_{i,t}}{\partial \theta_j} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\left(\frac{c}{\phi_t} \right)^2 \psi_t e^{-c(\lambda_{i,t} + \lambda_{j,t})} \right] \frac{\partial \lambda_{j,t}}{\partial \theta_j} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \theta_j} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\frac{y_{j,t}}{\lambda_{j,t}} - 1 + \frac{c\psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} \right] \frac{\partial^2 \lambda_{j,t}}{\partial \theta_j \partial \theta'_i} \\
& - \left[\frac{y_{j,t}}{\lambda_{j,t}^2} + \frac{c^2 \psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} + \left(\frac{c\psi_t e^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t} \right)^2 \right] \frac{\partial \lambda_{j,t}}{\partial \theta_j} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
& + \left[\left(\frac{c}{\phi_t} \right)^2 \psi_t e^{-c(\lambda_{i,t} + \lambda_{j,t})} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_j} \frac{\partial \lambda_{j,t}}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \theta_j} \frac{\partial \lambda_{j,t}}{\partial \theta'_i} \\
& + \left[\frac{(e^{-y_{1,t}} - e^{-c\lambda_{1,t}}) (e^{-y_{2,t}} - e^{-c\lambda_{2,t}})}{\phi_t} \right] \frac{\partial^2 \psi_t}{\partial \theta_j \partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t^2} \right] \frac{\partial \lambda_{i,t}}{\partial \theta_j} \frac{\partial \psi_t}{\partial \theta'_i} \\
& + \left[\frac{ce^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t^2} \right] \frac{\partial \lambda_{j,t}}{\partial \theta_j} \frac{\partial \psi_t}{\partial \theta'_i} \\
& - \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right]^2 \frac{\partial \psi_t}{\partial \theta_j} \frac{\partial \psi_t}{\partial \theta'_i}
\end{aligned} \tag{C.34}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \psi \partial \theta'_i} l_t (\theta | Y_t = Y_t) &= \left[\frac{ce^{-c\lambda_{i,t}} (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \psi} \frac{\partial \lambda_{i,t}}{\partial \theta'_i} \\
&+ \left[\frac{ce^{-c\lambda_{j,t}} (e^{-y_{i,t}} - e^{-c\lambda_{i,t}})}{\phi_t^2} \right] \frac{\partial \psi_t}{\partial \psi} \frac{\partial \lambda_{j,t}}{\partial \theta'_i} \\
&+ \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial^2 \psi_t}{\partial \psi \partial \theta'_i} \\
&- \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right]^2 \frac{\partial \psi_t}{\partial \psi} \frac{\partial \psi_t}{\partial \theta'_i} \quad (C.35)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \psi \partial \psi'} l_t (\theta | Y_t = Y_t) &= \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right] \frac{\partial^2 \psi_t}{\partial \psi \partial \psi'} \\
&- \left[\frac{(e^{-y_{i,t}} - e^{-c\lambda_{i,t}}) (e^{-y_{j,t}} - e^{-c\lambda_{j,t}})}{\phi_t} \right]^2 \frac{\partial \psi_t}{\partial \psi} \frac{\partial \psi_t}{\partial \psi'} \quad (C.36)
\end{aligned}$$

where,

$$\begin{aligned}
\frac{\partial^2 \psi_t}{\partial \theta_i \partial \theta'_i} &= \alpha \frac{\partial \psi_{t-1}^2}{\partial \theta_i \partial \theta'_i} + \frac{2\beta y_{i,t-1}^2}{\lambda_{i,t-1}^3} \frac{\partial \lambda_{i,t-1}}{\partial \theta_i} \frac{\partial \lambda_{i,t-1}}{\partial \theta'_i} + \beta \left(\frac{\lambda_{i,t-1}^2 - y_{i,t-1}^2}{\lambda_{i,t-1}^2} \right) \frac{\partial^2 \lambda_{i,t-1}}{\partial \theta_i \partial \theta'_i} \\
&+ \frac{2\beta y_{j,t-1}^2}{\lambda_{j,t-1}^3} \frac{\partial \lambda_{j,t-1}}{\partial \theta_i} \frac{\partial \lambda_{j,t-1}}{\partial \theta'_i} + \beta \left(\frac{\lambda_{j,t-1}^2 - y_{j,t-1}^2}{\lambda_{j,t-1}^2} \right) \frac{\partial^2 \lambda_{j,t-1}}{\partial \theta_i \partial \theta'_i} \quad (C.37)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \psi_t}{\partial \theta_j \partial \theta'_i} &= \alpha \frac{\partial \psi_{t-1}^2}{\partial \theta_j \partial \theta'_i} + \frac{2\beta y_{i,t-1}^2}{\lambda_{i,t-1}^3} \frac{\partial \lambda_{i,t-1}}{\partial \theta_j} \frac{\partial \lambda_{i,t-1}}{\partial \theta'_i} + \beta \left(\frac{\lambda_{i,t-1}^2 - y_{i,t-1}^2}{\lambda_{i,t-1}^2} \right) \frac{\partial^2 \lambda_{i,t-1}}{\partial \theta_j \partial \theta'_i} \\
&+ \frac{2\beta y_{j,t-1}^2}{\lambda_{j,t-1}^3} \frac{\partial \lambda_{j,t-1}}{\partial \theta_j} \frac{\partial \lambda_{j,t-1}}{\partial \theta'_i} + \beta \left(\frac{\lambda_{j,t-1}^2 - y_{j,t-1}^2}{\lambda_{j,t-1}^2} \right) \frac{\partial^2 \lambda_{j,t-1}}{\partial \theta_j \partial \theta'_i} \quad (C.38)
\end{aligned}$$

- $\frac{\partial^2 \psi_t}{\partial \psi \partial \psi'}$ components:

$$\frac{\partial^2 \psi_t}{\partial \tilde{\delta}^2} = 0 \quad (C.39)$$

$$\frac{\partial^2 \psi_t}{\partial \alpha \partial \tilde{\delta}} = -1 + \frac{\partial \psi_{t-1}}{\partial \tilde{\delta}} + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \alpha \partial \tilde{\delta}} \quad (\text{C.40})$$

$$\frac{\partial^2 \psi_t}{\partial \beta \partial \tilde{\delta}} = -1 + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \beta \partial \tilde{\delta}} \quad (\text{C.41})$$

$$\frac{\partial^2 \psi_t}{\partial \delta_0 \partial \tilde{\delta}} = \alpha \frac{\partial^2 \psi_{t-1}}{\partial \delta_0 \partial \tilde{\delta}} = 0 \quad (\text{C.42})$$

$$\frac{\partial^2 \psi_t}{\partial \alpha^2} = 2 \frac{\partial \psi_{t-1}}{\partial \alpha} + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \alpha^2} \quad (\text{C.43})$$

$$\frac{\partial^2 \psi_t}{\partial \beta \partial \alpha} = \frac{\partial \psi_{t-1}}{\partial \beta} + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \beta \partial \alpha} \quad (\text{C.44})$$

$$\frac{\partial^2 \psi_t}{\partial \delta_0 \partial \alpha} = \frac{\partial \psi_{t-1}}{\partial \delta_0} + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \delta_0 \partial \alpha} \quad (\text{C.45})$$

$$\frac{\partial^2 \psi_t}{\partial \beta^2} = 0 \quad (\text{C.46})$$

$$\frac{\partial^2 \psi_t}{\partial \delta_0 \partial \beta} = \alpha \frac{\partial^2 \psi_{t-1}}{\partial \delta_0 \partial \beta} = 0 \quad (\text{C.47})$$

$$\frac{\partial^2 \psi_t}{\partial \delta_0^2} = \alpha \frac{\partial^2 \psi_{t-1}}{\partial \psi_0^2} = 0 \quad (\text{C.48})$$

- $\frac{\partial^2 \lambda_{i,t}}{\partial \theta_i \partial \theta'_i}$ and $\frac{\partial^2 \lambda_{j,t}}{\partial \theta_i \partial \theta'_i}$ components:

$$\frac{\partial^2 \lambda_{i,t}}{\partial \omega_i^2} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \omega_i^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \omega_i^2} = 0 \quad (\text{C.49})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \omega_i^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \omega_i^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \omega_i^2} = 0 \quad (\text{C.50})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ii} \partial \omega_i} = \frac{\partial \lambda_{i,t-1}}{\partial \omega_i} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ii} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ii} \partial \omega_i} \quad (\text{C.51})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ii} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ii} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ii} \partial \omega_i} \quad (\text{C.52})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ij} \partial \omega_i} = \frac{\partial \lambda_{j,t-1}}{\partial \omega_i} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij} \partial \omega_i} \quad (\text{C.53})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ij} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij} \partial \omega_i} \quad (\text{C.54})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ii} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial \omega_i} = 0 \quad (\text{C.55})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ii} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial \omega_i} = 0 \quad (\text{C.56})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ij} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial \omega_i} = 0 \quad (\text{C.57})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ij} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial \omega_i} = 0 \quad (\text{C.58})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial \omega_i} \quad (\text{C.59})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial \omega_i} \quad (\text{C.60})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ii}^2} = 2 \frac{\partial \lambda_{i,t-1}}{\partial a_{ii}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ii}^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ii}^2} \quad (\text{C.61})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ii}^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ii}^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ii}^2} \quad (\text{C.62})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ij} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}} + \frac{\partial \lambda_{j,t-1}}{\partial a_{ii}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij} \partial a_{ii}} \quad (\text{C.63})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ij} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij} \partial a_{ii}} \quad (\text{C.64})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ii} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial b_{ii}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial a_{ii}} \quad (\text{C.65})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ii} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial a_{ii}} \quad (\text{C.66})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ij} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial b_{ij}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial a_{ii}} \quad (\text{C.67})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ij} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial a_{ii}} \quad (\text{C.68})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial \lambda_{0,i}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial a_{ii}} \quad (\text{C.69})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial a_{ii}} \quad (\text{C.70})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ij}^2} = 2 \frac{\partial \lambda_{j,t-1}}{\partial a_{ij}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij}^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij}^2} \quad (\text{C.71})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ij}^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ij}^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ij}^2} \quad (\text{C.72})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ii} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial b_{ii}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial a_{ij}} \quad (\text{C.73})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ii} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii} \partial a_{ij}} \quad (\text{C.74})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ij} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial b_{ij}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial a_{ij}} \quad (\text{C.75})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ij} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial a_{ij}} \quad (\text{C.76})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial \lambda_{0,i}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial a_{ij}} \quad (\text{C.77})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial a_{ij}} \quad (\text{C.78})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ii}^2} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii}^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii}^2} = 0 \quad (\text{C.79})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ii}^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ii}^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ii}^2} = 0 \quad (\text{C.80})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ij} \partial b_{ii}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial b_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial b_{ii}} = 0 \quad (\text{C.81})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ij} \partial b_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij} \partial b_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij} \partial b_{ii}} = 0 \quad (\text{C.82})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i} \partial b_{ii}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial b_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial b_{ii}} = 0 \quad (\text{C.83})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i} \partial b_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial b_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial b_{ii}} = 0 \quad (\text{C.84})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ij}^2} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij}^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij}^2} = 0 \quad (\text{C.85})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ij}^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ij}^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ij}^2} = 0 \quad (\text{C.86})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i} \partial b_{ij}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial b_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial b_{ij}} = 0 \quad (\text{C.87})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i} \partial b_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i} \partial b_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i} \partial b_{ij}} = 0 \quad (\text{C.88})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,i}^2} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i}^2} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i}^2} = 0 \quad (\text{C.89})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,i}^2} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,i}^2} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,i}^2} = 0 \quad (\text{C.90})$$

• $\frac{\partial^2 \lambda_{i,t}}{\partial \theta_j \partial \theta'_i}$ and $\frac{\partial^2 \lambda_{j,t}}{\partial \theta_j \partial \theta'_i}$ components:

$$\frac{\partial^2 \lambda_{i,t}}{\partial \omega_j \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \omega_j \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \omega_j \partial \omega_i} = 0 \quad (\text{C.91})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \omega_j \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \omega_j \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \omega_j \partial \omega_i} = 0 \quad (\text{C.92})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{jj} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{jj} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{jj} \partial \omega_i} \quad (\text{C.93})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{jj} \partial \omega_i} = \frac{\partial \lambda_{j,t-1}}{\partial \omega_i} + a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{jj} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{jj} \partial \omega_i} \quad (\text{C.94})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ji} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial \omega_i} \quad (\text{C.95})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ji} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial \omega_i} + \frac{\partial \lambda_{i,t-1}}{\partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial \omega_i} \quad (\text{C.96})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{jj} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial \omega_i} = 0 \quad (\text{C.97})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{jj} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial \omega_i} = 0 \quad (\text{C.98})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ji} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial \omega_i} = 0 \quad (\text{C.99})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ji} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial \omega_i} = 0 \quad (\text{C.100})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial \omega_i} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial \omega_i} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial \omega_i} = 0 \quad (\text{C.101})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial \omega_i} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial \omega_i} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial \omega_i} = 0 \quad (\text{C.102})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{jj} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial a_{jj}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{jj} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{jj} \partial a_{ii}} \quad (\text{C.103})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{jj} \partial a_{ii}} = \frac{\partial \lambda_{j,t-1}}{\partial a_{jj}} + a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{jj} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{jj} \partial a_{ii}} \quad (\text{C.104})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ji} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial a_{ji}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial a_{ii}} \quad (\text{C.105})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ji} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial a_{ii}} + a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial a_{ii}} \quad (\text{C.106})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{jj} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial b_{jj}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial a_{ii}} \quad (\text{C.107})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{jj} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial a_{ii}} \quad (\text{C.108})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ji} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial b_{ji}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial a_{ii}} \quad (\text{C.109})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ji} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial a_{ii}} \quad (\text{C.110})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial a_{ii}} = \frac{\partial \lambda_{i,t-1}}{\partial \lambda_{0,j}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial a_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial a_{ii}} \quad (\text{C.111})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial a_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial a_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial a_{ii}} \quad (\text{C.112})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial a_{ji} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial a_{ji}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial a_{ij}} \quad (\text{C.113})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial a_{ji} \partial a_{ij}} = \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}} + a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial a_{ji} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial a_{ji} \partial a_{ij}} \quad (\text{C.114})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{jj} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial b_{jj}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial a_{ij}} \quad (\text{C.115})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{jj} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial a_{ij}} \quad (\text{C.116})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ji} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial b_{ji}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial a_{ij}} \quad (\text{C.117})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ji} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial a_{ij}} \quad (\text{C.118})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial a_{ij}} = \frac{\partial \lambda_{j,t-1}}{\partial \lambda_{0,j}} + a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial a_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial a_{ij}} \quad (\text{C.119})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial a_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial a_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial a_{ij}} \quad (\text{C.120})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{jj} \partial b_{ii}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial b_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial b_{ii}} = 0 \quad (\text{C.121})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{jj} \partial b_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{jj} \partial b_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{jj} \partial b_{ii}} = 0 \quad (\text{C.122})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ji} \partial b_{ii}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial b_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial b_{ii}} = 0 \quad (\text{C.123})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ji} \partial b_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial b_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial b_{ii}} = 0 \quad (\text{C.124})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial b_{ii}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial b_{ii}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial b_{ii}} = 0 \quad (\text{C.125})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial b_{ii}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial b_{ii}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial b_{ii}} = 0 \quad (\text{C.126})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial b_{ji} \partial b_{ij}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial b_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial b_{ij}} = 0 \quad (\text{C.127})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial b_{ji} \partial b_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial b_{ji} \partial b_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial b_{ji} \partial b_{ij}} = 0 \quad (\text{C.128})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial b_{ij}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial b_{ij}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial b_{ij}} = 0 \quad (\text{C.129})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial b_{ij}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial b_{ij}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial b_{ij}} = 0 \quad (\text{C.130})$$

$$\frac{\partial^2 \lambda_{i,t}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} = a_{ii} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} + a_{ij} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} = 0 \quad (\text{C.131})$$

$$\frac{\partial^2 \lambda_{j,t}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} = a_{jj} \frac{\partial^2 \lambda_{j,t-1}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} + a_{ji} \frac{\partial^2 \lambda_{i,t-1}}{\partial \lambda_{0,j} \partial \lambda_{0,i}} = 0 \quad (\text{C.132})$$

- $\frac{\partial^2 \psi_t}{\partial \psi \partial \theta'_i}$ components:

$$\frac{\partial^2 \psi_t}{\partial \tilde{\delta} \partial \theta_i} = \alpha \frac{\partial^2 \psi_{t-1}}{\partial \tilde{\delta} \partial \theta_i} = 0 \quad (\text{C.133})$$

$$\frac{\partial^2 \psi_t}{\partial \alpha \partial \theta_i} = \frac{\partial \psi_{t-1}}{\partial \theta_i} + \alpha \frac{\partial^2 \psi_{t-1}}{\partial \alpha \partial \theta_i} \quad (\text{C.134})$$

$$\frac{\partial^2 \psi_t}{\partial \beta \partial \theta_i} = \alpha \frac{\partial^2 \psi_{t-1}}{\partial \beta \partial \theta_i} + \frac{(\lambda_{i,t-1}^2 - y_{i,t-1}^2)}{\lambda_{i,t-1}^2} \frac{\partial \lambda_{i,t-1}}{\partial \theta_i} + \frac{(\lambda_{j,t-1}^2 - y_{j,t-1}^2)}{\lambda_{j,t-1}^2} \frac{\partial \lambda_{j,t-1}}{\partial \theta_i} \quad (\text{C.135})$$

C.7 Additional Tables

C.7.1 Simulated Data - Changing the number of data points/ Changing samples burned

Table C.1: *Simulated Data A1 - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
A1	15000	50	43.3	ω_1	1.00	1.14	0.14	0.42	1.34	21.23	0.46	0.82	1.09	1.41	2.07
				a_{11}	0.40	0.37	-0.03	0.17	0.54	10.59	0.04	0.24	0.37	0.49	0.70
				b_{11}	0.20	0.21	0.01	0.12	0.37	6.65	0.02	0.12	0.20	0.29	0.47
				ω_2	0.50	0.56	0.06	0.18	0.58	6.16	0.25	0.43	0.55	0.67	0.96
				a_{22}	0.30	0.15	-0.15	0.11	0.34	4.73	0.01	0.06	0.13	0.21	0.40
				b_{22}	0.40	0.47	0.07	0.12	0.39	4.51	0.24	0.39	0.46	0.55	0.73
				$\tilde{\delta}$	0.50	0.80	0.30	0.73	2.29	21.32	-0.85	0.34	0.95	1.35	1.88
		100	39.6	ω_1	1.00	1.36	0.36	0.39	1.25	19.41	0.65	1.08	1.35	1.63	2.17
				a_{11}	0.40	0.29	-0.11	0.16	0.50	5.96	0.02	0.17	0.28	0.40	0.61
				b_{11}	0.20	0.16	-0.04	0.08	0.26	2.88	0.02	0.10	0.16	0.22	0.33
				ω_2	0.50	0.60	0.10	0.15	0.46	7.06	0.32	0.49	0.59	0.69	0.89
				a_{22}	0.30	0.11	-0.19	0.09	0.27	2.64	0.00	0.05	0.10	0.17	0.32
				b_{22}	0.40	0.51	0.11	0.09	0.29	4.26	0.34	0.45	0.51	0.57	0.69
				$\tilde{\delta}$	0.50	0.53	0.03	0.75	2.37	22.22	-0.98	-0.01	0.59	1.13	1.75
		200	39.3	ω_1	1.00	1.34	0.34	0.34	1.08	22.29	0.70	1.10	1.33	1.58	2.01
				a_{11}	0.40	0.28	-0.12	0.14	0.45	8.37	0.03	0.17	0.28	0.38	0.58
				b_{11}	0.20	0.19	-0.01	0.06	0.20	2.62	0.06	0.14	0.18	0.23	0.31
				ω_2	0.50	0.58	0.08	0.11	0.34	2.57	0.38	0.51	0.58	0.66	0.80
				a_{22}	0.30	0.10	-0.20	0.07	0.22	2.36	0.00	0.04	0.09	0.14	0.26
				b_{22}	0.40	0.49	0.09	0.07	0.21	3.03	0.37	0.45	0.49	0.54	0.63
				$\tilde{\delta}$	0.50	0.32	-0.18	0.66	2.08	29.72	-0.93	-0.15	0.31	0.81	1.55
		500	50.3	ω_1	1.00	1.09	0.09	0.26	0.83	20.47	0.64	0.89	1.06	1.26	1.65
				a_{11}	0.40	0.37	-0.03	0.11	0.35	8.54	0.15	0.30	0.37	0.45	0.57
				b_{11}	0.20	0.22	0.02	0.05	0.14	3.04	0.14	0.19	0.22	0.25	0.31
				ω_2	0.50	0.57	0.07	0.11	0.33	8.75	0.38	0.49	0.56	0.64	0.79
				a_{22}	0.30	0.20	-0.10	0.08	0.25	4.73	0.05	0.14	0.20	0.25	0.35
				b_{22}	0.40	0.43	0.03	0.05	0.14	1.54	0.34	0.39	0.43	0.46	0.51
				$\tilde{\delta}$	0.50	0.60	0.10	0.48	1.52	29.27	-0.33	0.26	0.61	0.95	1.50
	500000	50	44.8	ω_1	1.00	1.18	0.18	0.45	1.43	25.54	0.47	0.84	1.12	1.45	2.19
				a_{11}	0.40	0.35	-0.05	0.17	0.53	10.01	0.03	0.22	0.35	0.47	0.65
				b_{11}	0.20	0.22	0.02	0.11	0.36	5.96	0.02	0.13	0.21	0.29	0.45
				ω_2	0.50	0.56	0.06	0.17	0.55	5.33	0.26	0.43	0.54	0.67	0.93
				a_{22}	0.30	0.14	-0.16	0.10	0.33	4.36	0.01	0.06	0.12	0.21	0.39
				b_{22}	0.40	0.48	0.08	0.13	0.40	4.86	0.24	0.39	0.47	0.56	0.74
				$\tilde{\delta}$	0.50	0.80	0.30	0.70	2.22	20.93	-0.81	0.35	0.92	1.34	1.84
		100	53.2	ω_1	1.00	1.31	0.31	0.39	1.22	20.22	0.62	1.03	1.28	1.56	2.10
				a_{11}	0.40	0.30	-0.10	0.16	0.50	6.30	0.03	0.18	0.30	0.42	0.62
				b_{11}	0.20	0.17	-0.03	0.08	0.26	3.43	0.02	0.11	0.16	0.22	0.35
				ω_2	0.50	0.57	0.07	0.14	0.45	7.09	0.32	0.47	0.56	0.66	0.87
				a_{22}	0.30	0.13	-0.17	0.09	0.29	3.31	0.01	0.06	0.11	0.19	0.34
				b_{22}	0.40	0.52	0.12	0.09	0.29	4.81	0.34	0.45	0.52	0.58	0.70
				$\tilde{\delta}$	0.50	0.53	0.03	0.73	2.32	24.70	-0.95	-0.01	0.58	1.11	1.73
		200	38.7	ω_1	1.00	1.37	0.37	0.35	1.10	23.16	0.73	1.11	1.35	1.61	2.07
				a_{11}	0.40	0.26	-0.14	0.14	0.44	8.15	0.02	0.16	0.26	0.36	0.55
				b_{11}	0.20	0.19	-0.01	0.07	0.21	3.06	0.07	0.14	0.19	0.23	0.32
				ω_2	0.50	0.58	0.08	0.11	0.34	2.53	0.38	0.51	0.58	0.65	0.80
				a_{22}	0.30	0.09	-0.21	0.07	0.22	2.43	0.00	0.04	0.08	0.14	0.26
				b_{22}	0.40	0.50	0.10	0.07	0.23	3.56	0.36	0.45	0.49	0.54	0.65
				$\tilde{\delta}$	0.50	0.33	-0.17	0.65	2.07	29.96	-0.93	-0.16	0.34	0.82	1.53
		500	39.4	ω_1	1.00	1.10	0.10	0.26	0.81	16.27	0.67	0.92	1.08	1.25	1.70
				a_{11}	0.40	0.37	-0.03	0.11	0.33	6.67	0.14	0.30	0.37	0.44	0.56
				b_{11}	0.20	0.22	0.02	0.04	0.13	2.05	0.15	0.19	0.22	0.25	0.30
				ω_2	0.50	0.56	0.06	0.10	0.30	6.37	0.38	0.49	0.56	0.62	0.76
				a_{22}	0.30	0.20	-0.10	0.08	0.24	3.97	0.05	0.15	0.20	0.25	0.35
				b_{22}	0.40	0.43	0.03	0.04	0.14	1.20	0.34	0.40	0.43	0.46	0.52
				$\tilde{\delta}$	0.50	0.61	0.11	0.51	1.61	28.44	-0.46	0.28	0.63	0.97	1.53

¹Posterior Standard Deviation

²Naive Standard Error $\times 10^{-3}$

³Time Series Standard Error $\times 10^{-3}$ - Adjusted for autocorrelation

Table C.2: *Simulated Data A2 - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
A2	15000	50	42.0	ω_1	0.30	0.41	0.11	0.16	0.52	5.53	0.14	0.29	0.39	0.51	0.78
				a_{11}	0.20	0.23	0.03	0.13	0.40	4.98	0.02	0.14	0.22	0.31	0.50
				b_{11}	0.50	0.48	-0.02	0.13	0.41	5.42	0.25	0.40	0.48	0.57	0.74
				ω_2	0.50	0.66	0.16	0.22	0.68	9.21	0.28	0.50	0.64	0.80	1.11
				a_{22}	0.40	0.19	-0.21	0.13	0.40	4.37	0.01	0.09	0.17	0.27	0.48
				b_{22}	0.30	0.37	0.07	0.12	0.38	4.24	0.14	0.28	0.36	0.44	0.62
				$\tilde{\delta}$	0.70	0.92	0.22	0.75	2.37	21.93	-0.78	0.44	1.06	1.50	2.01
		100	39.2	ω_1	0.30	0.39	0.09	0.12	0.39	6.08	0.18	0.30	0.37	0.46	0.66
				a_{11}	0.20	0.20	0.00	0.11	0.35	3.48	0.02	0.12	0.19	0.27	0.45
				b_{11}	0.50	0.45	-0.05	0.10	0.32	3.67	0.26	0.38	0.44	0.51	0.65
				ω_2	0.50	0.67	0.17	0.19	0.61	6.20	0.33	0.53	0.65	0.79	1.08
				a_{22}	0.40	0.19	-0.21	0.12	0.38	2.94	0.01	0.10	0.18	0.28	0.44
				b_{22}	0.30	0.40	0.10	0.10	0.30	3.04	0.22	0.34	0.40	0.47	0.59
				$\tilde{\delta}$	0.70	0.84	0.14	0.53	1.69	11.94	-0.41	0.52	0.92	1.26	1.61
		200	37.1	ω_1	0.30	0.38	0.08	0.09	0.28	2.71	0.22	0.32	0.37	0.44	0.57
				a_{11}	0.20	0.19	-0.01	0.09	0.28	3.20	0.03	0.13	0.19	0.25	0.37
				b_{11}	0.50	0.46	-0.04	0.07	0.21	2.32	0.34	0.42	0.46	0.51	0.60
				ω_2	0.50	0.64	0.14	0.14	0.44	4.65	0.39	0.54	0.64	0.74	0.92
				a_{22}	0.40	0.17	-0.23	0.10	0.31	2.99	0.01	0.10	0.16	0.24	0.38
				b_{22}	0.30	0.39	0.09	0.07	0.22	2.07	0.26	0.35	0.39	0.44	0.53
				$\tilde{\delta}$	0.70	0.45	-0.25	0.54	1.72	16.95	-0.67	0.07	0.47	0.85	1.42
		500	35.4	ω_1	0.30	0.35	0.05	0.05	0.17	1.61	0.24	0.31	0.34	0.38	0.46
				a_{11}	0.20	0.18	-0.02	0.06	0.19	1.99	0.06	0.14	0.18	0.22	0.30
				b_{11}	0.50	0.55	0.05	0.05	0.15	1.63	0.46	0.51	0.55	0.58	0.64
				ω_2	0.50	0.56	0.06	0.11	0.36	4.20	0.36	0.48	0.55	0.63	0.80
				a_{22}	0.40	0.31	-0.09	0.09	0.27	3.39	0.14	0.26	0.31	0.37	0.47
				b_{22}	0.30	0.32	0.02	0.04	0.14	1.25	0.24	0.29	0.32	0.35	0.41
				$\tilde{\delta}$	0.70	0.43	-0.27	0.39	1.24	12.54	-0.36	0.17	0.44	0.71	1.18
	500000	50	37.1	ω_1	0.30	0.39	0.09	0.16	0.50	5.28	0.14	0.27	0.37	0.48	0.76
				a_{11}	0.20	0.24	0.04	0.13	0.41	4.97	0.02	0.14	0.23	0.33	0.52
				b_{11}	0.50	0.48	-0.02	0.13	0.42	5.89	0.24	0.39	0.48	0.57	0.75
				ω_2	0.50	0.63	0.13	0.20	0.63	7.20	0.27	0.49	0.62	0.76	1.04
				a_{22}	0.40	0.20	-0.20	0.13	0.40	4.56	0.01	0.10	0.19	0.28	0.48
				b_{22}	0.30	0.37	0.07	0.12	0.37	3.86	0.15	0.29	0.37	0.45	0.61
				$\tilde{\delta}$	0.70	0.95	0.25	0.69	2.19	16.82	-0.64	0.51	1.06	1.49	1.98
		100	41.9	ω_1	0.30	0.37	0.07	0.11	0.35	4.93	0.17	0.29	0.36	0.44	0.61
				a_{11}	0.20	0.21	0.01	0.11	0.36	3.56	0.02	0.12	0.20	0.28	0.45
				b_{11}	0.50	0.46	-0.04	0.10	0.32	3.99	0.28	0.39	0.45	0.52	0.67
				ω_2	0.50	0.66	0.16	0.18	0.58	5.58	0.32	0.52	0.65	0.79	1.03
				a_{22}	0.40	0.20	-0.20	0.12	0.39	3.24	0.01	0.10	0.19	0.28	0.46
				b_{22}	0.30	0.40	0.10	0.09	0.29	2.85	0.22	0.34	0.40	0.46	0.58
				$\tilde{\delta}$	0.70	0.85	0.15	0.52	1.64	11.67	-0.34	0.53	0.93	1.25	1.59
		200	39.6	ω_1	0.30	0.38	0.08	0.09	0.29	3.35	0.21	0.31	0.37	0.44	0.57
				a_{11}	0.20	0.19	-0.01	0.09	0.29	3.62	0.03	0.13	0.19	0.25	0.39
				b_{11}	0.50	0.46	-0.04	0.07	0.22	2.45	0.33	0.41	0.46	0.51	0.60
				ω_2	0.50	0.64	0.14	0.14	0.44	4.78	0.38	0.54	0.63	0.73	0.93
				a_{22}	0.40	0.17	-0.23	0.10	0.33	3.66	0.01	0.09	0.16	0.24	0.39
				b_{22}	0.30	0.39	0.09	0.07	0.22	2.25	0.26	0.35	0.39	0.44	0.53
				$\tilde{\delta}$	0.70	0.44	-0.26	0.52	1.64	15.79	-0.57	0.08	0.45	0.81	1.41
		500	38.2	ω_1	0.30	0.34	0.04	0.06	0.18	1.72	0.24	0.30	0.34	0.38	0.46
				a_{11}	0.20	0.18	-0.02	0.06	0.19	2.21	0.06	0.14	0.18	0.22	0.30
				b_{11}	0.50	0.54	0.04	0.05	0.15	1.74	0.45	0.51	0.54	0.57	0.63
				ω_2	0.50	0.56	0.06	0.12	0.38	4.75	0.35	0.47	0.55	0.63	0.82
				a_{22}	0.40	0.31	-0.09	0.09	0.29	4.05	0.12	0.25	0.31	0.37	0.48
				b_{22}	0.30	0.32	0.02	0.05	0.14	1.51	0.24	0.30	0.32	0.35	0.42
				$\tilde{\delta}$	0.70	0.44	-0.26	0.38	1.21	12.66	-0.33	0.18	0.44	0.70	1.17

Table C.3: *Simulated Data A3 - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
A3	15000	50	39.0	ω_1	0.50	0.83	0.33	0.29	0.93	15.29	0.34	0.62	0.80	1.01	1.48
				a_{11}	0.10	0.19	0.09	0.10	0.33	4.48	0.02	0.11	0.18	0.26	0.41
				b_{11}	0.80	0.66	-0.14	0.11	0.34	3.42	0.45	0.59	0.67	0.74	0.87
				ω_2	0.50	0.68	0.18	0.22	0.70	9.03	0.29	0.52	0.66	0.82	1.16
				a_{22}	0.20	0.13	-0.07	0.11	0.34	3.53	0.00	0.04	0.10	0.18	0.40
				b_{22}	0.70	0.51	-0.19	0.14	0.43	3.90	0.25	0.42	0.51	0.60	0.78
				$\tilde{\delta}$	-1.00	0.40	1.40	0.79	2.50	14.15	-1.02	-0.24	0.45	1.06	1.71
		100	40.7	ω_1	0.50	0.70	0.20	0.17	0.55	7.55	0.38	0.58	0.70	0.82	1.06
				a_{11}	0.10	0.09	-0.01	0.07	0.22	3.05	0.00	0.04	0.08	0.13	0.26
				b_{11}	0.80	0.71	-0.09	0.08	0.25	2.59	0.55	0.66	0.72	0.77	0.87
				ω_2	0.50	0.60	0.10	0.18	0.56	7.60	0.29	0.48	0.59	0.71	0.97
				a_{22}	0.20	0.10	-0.10	0.08	0.24	2.57	0.00	0.04	0.08	0.15	0.29
				b_{22}	0.70	0.75	0.05	0.08	0.27	3.16	0.58	0.70	0.76	0.81	0.91
				$\tilde{\delta}$	-1.00	0.32	1.32	0.71	2.26	17.79	-0.94	-0.28	0.36	0.94	1.47
		200	44.6	ω_1	0.50	0.64	0.14	0.13	0.41	5.52	0.39	0.55	0.64	0.73	0.90
				a_{11}	0.10	0.09	-0.01	0.06	0.19	2.45	0.00	0.04	0.08	0.12	0.22
				b_{11}	0.80	0.78	-0.02	0.06	0.18	1.90	0.66	0.75	0.79	0.82	0.88
				ω_2	0.50	0.54	0.04	0.13	0.42	5.93	0.31	0.45	0.54	0.63	0.83
				a_{22}	0.20	0.17	-0.03	0.07	0.22	3.26	0.04	0.12	0.17	0.22	0.31
				b_{22}	0.70	0.69	-0.01	0.06	0.20	2.40	0.56	0.64	0.69	0.73	0.81
				$\tilde{\delta}$	-1.00	0.37	1.37	0.67	2.12	21.13	-0.89	-0.16	0.41	0.92	1.43
		500	39.8	ω_1	0.50	0.49	-0.01	0.08	0.24	2.67	0.35	0.44	0.49	0.54	0.64
				a_{11}	0.10	0.08	-0.02	0.04	0.13	2.04	0.01	0.04	0.07	0.10	0.17
				b_{11}	0.80	0.82	0.02	0.04	0.13	1.72	0.74	0.79	0.82	0.85	0.89
				ω_2	0.50	0.49	-0.01	0.09	0.30	4.28	0.32	0.43	0.49	0.56	0.69
				a_{22}	0.20	0.20	0.00	0.05	0.14	2.31	0.11	0.16	0.20	0.23	0.28
				b_{22}	0.70	0.72	0.02	0.04	0.13	2.02	0.64	0.69	0.72	0.75	0.80
				$\tilde{\delta}$	-1.00	-0.05	0.95	0.61	1.93	22.33	-0.96	-0.55	-0.12	0.41	1.20
	500000	50	41.8	ω_1	0.50	0.80	0.30	0.30	0.94	16.38	0.31	0.58	0.77	0.98	1.46
				a_{11}	0.10	0.19	0.09	0.11	0.33	4.85	0.01	0.10	0.18	0.26	0.41
				b_{11}	0.80	0.67	-0.13	0.11	0.35	3.79	0.45	0.59	0.67	0.75	0.89
				ω_2	0.50	0.65	0.15	0.21	0.66	8.30	0.28	0.50	0.63	0.78	1.11
				a_{22}	0.20	0.13	-0.07	0.11	0.35	4.04	0.00	0.05	0.10	0.19	0.41
				b_{22}	0.70	0.52	-0.18	0.14	0.44	4.31	0.26	0.43	0.52	0.61	0.80
				$\tilde{\delta}$	-1.00	0.39	1.39	0.78	2.46	14.28	-1.02	-0.24	0.46	1.05	1.64
		100	40.4	ω_1	0.50	0.69	0.19	0.18	0.56	7.18	0.37	0.56	0.68	0.80	1.06
				a_{11}	0.10	0.09	-0.01	0.06	0.20	2.40	0.00	0.04	0.07	0.12	0.23
				b_{11}	0.80	0.72	-0.08	0.08	0.24	2.39	0.57	0.67	0.72	0.77	0.87
				ω_2	0.50	0.58	0.08	0.17	0.53	6.27	0.27	0.46	0.57	0.68	0.93
				a_{22}	0.20	0.10	-0.10	0.08	0.26	2.88	0.00	0.04	0.09	0.15	0.30
				b_{22}	0.70	0.75	0.05	0.08	0.27	3.24	0.57	0.70	0.75	0.81	0.91
				$\tilde{\delta}$	-1.00	0.29	1.29	0.70	2.20	16.53	-0.94	-0.28	0.31	0.88	1.43
		200	41.3	ω_1	0.50	0.64	0.14	0.13	0.40	4.85	0.40	0.56	0.64	0.73	0.91
				a_{11}	0.10	0.09	-0.01	0.06	0.18	1.98	0.00	0.04	0.08	0.12	0.21
				b_{11}	0.80	0.79	-0.01	0.06	0.18	1.81	0.67	0.75	0.79	0.83	0.89
				ω_2	0.50	0.54	0.04	0.14	0.43	5.70	0.30	0.44	0.53	0.62	0.83
				a_{22}	0.20	0.18	-0.02	0.08	0.24	3.89	0.03	0.12	0.17	0.23	0.33
				b_{22}	0.70	0.68	-0.02	0.07	0.21	2.84	0.55	0.64	0.69	0.73	0.81
				$\tilde{\delta}$	-1.00	0.40	1.40	0.66	2.07	19.88	-0.87	-0.10	0.46	0.94	1.43
		500	41.8	ω_1	0.50	0.49	-0.01	0.07	0.24	2.43	0.35	0.44	0.49	0.54	0.64
				a_{11}	0.10	0.08	-0.02	0.04	0.13	1.98	0.01	0.05	0.07	0.11	0.16
				b_{11}	0.80	0.82	0.02	0.04	0.13	1.72	0.74	0.80	0.82	0.85	0.90
				ω_2	0.50	0.50	0.00	0.09	0.30	4.07	0.32	0.43	0.49	0.56	0.70
				a_{22}	0.20	0.20	0.00	0.05	0.15	2.55	0.11	0.17	0.20	0.23	0.28
				b_{22}	0.70	0.72	0.02	0.04	0.13	2.11	0.64	0.69	0.72	0.75	0.80
				$\tilde{\delta}$	-1.00	-0.04	0.96	0.61	1.92	22.39	-0.95	-0.54	-0.11	0.42	1.20

Table C.4: *Simulated Data A4 - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
A4	15000	50	41.0	ω_1	0.50	0.95	0.45	0.38	1.19	16.77	0.32	0.65	0.93	1.23	1.69
				a_{11}	0.60	0.37	-0.23	0.20	0.64	8.75	0.03	0.21	0.37	0.53	0.75
				b_{11}	0.10	0.10	0.00	0.08	0.25	1.59	0.00	0.04	0.09	0.15	0.29
				ω_2	0.30	0.75	0.45	0.35	1.10	30.30	0.21	0.49	0.70	0.97	1.53
				a_{22}	0.80	0.38	-0.42	0.18	0.56	14.52	0.05	0.24	0.38	0.50	0.72
				b_{22}	0.10	0.29	0.19	0.11	0.34	3.54	0.09	0.21	0.29	0.37	0.51
				$\tilde{\delta}$	-0.50	0.09	0.59	0.99	3.15	23.30	-1.44	-0.71	-0.02	0.78	2.19
		100	40.8	ω_1	0.50	1.01	0.51	0.35	1.10	17.76	0.37	0.74	1.01	1.28	1.62
				a_{11}	0.60	0.33	-0.27	0.20	0.65	10.02	0.02	0.16	0.32	0.49	0.73
				b_{11}	0.10	0.06	-0.04	0.05	0.16	1.19	0.00	0.02	0.05	0.09	0.19
				ω_2	0.30	0.71	0.41	0.35	1.10	50.44	0.21	0.46	0.64	0.89	1.59
				a_{22}	0.80	0.45	-0.35	0.17	0.53	14.74	0.09	0.35	0.47	0.57	0.75
				b_{22}	0.10	0.28	0.18	0.10	0.30	3.77	0.10	0.21	0.28	0.35	0.47
				$\tilde{\delta}$	-0.50	-0.32	0.18	0.83	2.63	26.48	-1.53	-0.96	-0.44	0.20	1.57
		200	40.6	ω_1	0.50	1.00	0.50	0.33	1.03	18.03	0.39	0.76	1.02	1.26	1.57
				a_{11}	0.60	0.33	-0.27	0.20	0.63	9.40	0.02	0.17	0.31	0.48	0.72
				b_{11}	0.10	0.06	-0.04	0.04	0.14	1.09	0.00	0.03	0.05	0.09	0.17
				ω_2	0.30	1.06	0.76	0.32	1.02	69.68	0.57	0.82	1.00	1.26	1.77
				a_{22}	0.80	0.32	-0.48	0.12	0.39	28.07	0.07	0.23	0.33	0.40	0.54
				b_{22}	0.10	0.26	0.16	0.07	0.21	4.38	0.14	0.22	0.26	0.31	0.39
				$\tilde{\delta}$	-0.50	-0.49	0.01	0.54	1.72	24.45	-1.41	-0.89	-0.54	-0.13	0.67
		500	41.7	ω_1	0.50	0.73	0.23	0.26	0.81	20.64	0.31	0.53	0.70	0.88	1.31
				a_{11}	0.60	0.47	-0.13	0.15	0.49	13.16	0.13	0.37	0.48	0.58	0.73
				b_{11}	0.10	0.12	0.02	0.04	0.12	1.34	0.05	0.09	0.12	0.14	0.20
				ω_2	0.30	0.49	0.19	0.20	0.64	29.50	0.19	0.34	0.45	0.58	1.01
				a_{22}	0.80	0.67	-0.13	0.10	0.32	13.93	0.43	0.61	0.68	0.74	0.83
				b_{22}	0.10	0.15	0.05	0.04	0.12	1.64	0.08	0.12	0.15	0.18	0.23
				$\tilde{\delta}$	-0.50	-0.49	0.01	0.44	1.39	12.66	-1.25	-0.81	-0.51	-0.20	0.44
	500000	50	39.8	ω_1	0.50	0.91	0.41	0.37	1.18	16.15	0.31	0.62	0.88	1.18	1.67
				a_{11}	0.60	0.39	-0.21	0.20	0.63	8.57	0.03	0.24	0.40	0.55	0.76
				b_{11}	0.10	0.10	0.00	0.08	0.26	1.75	0.00	0.04	0.09	0.15	0.30
				ω_2	0.30	0.69	0.39	0.35	1.11	33.46	0.17	0.42	0.62	0.90	1.50
				a_{22}	0.80	0.41	-0.39	0.20	0.64	19.24	0.03	0.27	0.42	0.56	0.79
				b_{22}	0.10	0.28	0.18	0.12	0.38	4.63	0.07	0.20	0.28	0.36	0.54
				$\tilde{\delta}$	-0.50	0.07	0.57	0.95	3.02	21.10	-1.43	-0.68	-0.04	0.73	2.09
		100	40.1	ω_1	0.50	0.97	0.47	0.35	1.09	16.63	0.36	0.70	0.95	1.22	1.62
				a_{11}	0.60	0.36	-0.24	0.20	0.64	9.32	0.02	0.19	0.36	0.51	0.74
				b_{11}	0.10	0.06	-0.04	0.05	0.16	1.10	0.00	0.02	0.05	0.09	0.19
				ω_2	0.30	0.75	0.45	0.37	1.18	55.08	0.20	0.46	0.68	0.99	1.55
				a_{22}	0.80	0.43	-0.37	0.18	0.56	18.12	0.07	0.30	0.44	0.55	0.75
				b_{22}	0.10	0.29	0.19	0.10	0.30	3.83	0.11	0.22	0.28	0.35	0.49
				$\tilde{\delta}$	-0.50	-0.35	0.15	0.81	2.55	24.48	-1.53	-0.97	-0.47	0.14	1.49
		200	38.8	ω_1	0.50	0.98	0.48	0.33	1.04	18.39	0.38	0.72	0.98	1.24	1.57
				a_{11}	0.60	0.34	-0.26	0.20	0.64	9.83	0.02	0.18	0.33	0.50	0.72
				b_{11}	0.10	0.06	-0.04	0.04	0.14	1.09	0.00	0.03	0.05	0.09	0.16
				ω_2	0.30	0.94	0.64	0.36	1.15	97.02	0.24	0.65	1.01	1.21	1.54
				a_{22}	0.80	0.37	-0.43	0.19	0.61	76.71	0.10	0.24	0.32	0.53	0.79
				b_{22}	0.10	0.25	0.15	0.08	0.26	7.74	0.09	0.20	0.25	0.31	0.41
				$\tilde{\delta}$	-0.50	-0.49	0.01	0.56	1.76	26.69	-1.41	-0.91	-0.55	-0.14	0.74
		500	36.8	ω_1	0.50	0.71	0.21	0.23	0.74	16.05	0.33	0.54	0.69	0.85	1.25
				a_{11}	0.60	0.47	-0.13	0.14	0.44	9.18	0.17	0.38	0.48	0.57	0.71
				b_{11}	0.10	0.12	0.02	0.04	0.12	1.28	0.05	0.10	0.12	0.15	0.20
				ω_2	0.30	0.46	0.16	0.19	0.59	22.11	0.17	0.31	0.44	0.59	0.85
				a_{22}	0.80	0.68	-0.12	0.10	0.31	12.36	0.48	0.61	0.69	0.75	0.85
				b_{22}	0.10	0.15	0.05	0.04	0.13	1.89	0.07	0.12	0.15	0.18	0.23
				$\tilde{\delta}$	-0.50	-0.48	0.02	0.43	1.37	11.88	-1.25	-0.80	-0.51	-0.20	0.44

Table C.5: *Simulated Data B - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
B	15000	50	39.0	ω_1	0.30	0.48	0.18	0.23	0.74	11.96	0.11	0.31	0.45	0.61	1.04
				a_{11}	0.20	0.24	0.04	0.12	0.37	6.70	0.03	0.15	0.23	0.32	0.47
				b_{11}	0.50	0.45	-0.05	0.12	0.38	8.07	0.22	0.37	0.45	0.53	0.69
				b_{12}	0.10	0.10	0.00	0.07	0.22	2.10	0.00	0.04	0.09	0.14	0.26
				ω_2	0.50	0.71	0.21	0.32	1.01	15.88	0.21	0.47	0.66	0.91	1.43
				a_{22}	0.40	0.35	-0.05	0.16	0.50	7.98	0.05	0.24	0.35	0.46	0.65
				b_{22}	0.20	0.30	0.10	0.12	0.38	7.46	0.08	0.22	0.30	0.38	0.54
				b_{21}	0.30	0.11	-0.19	0.07	0.24	2.48	0.01	0.05	0.10	0.15	0.28
				$\hat{\delta}$	0.70	0.81	0.11	0.94	2.98	54.40	-0.98	0.06	0.91	1.57	2.35
		100	39.0	ω_1	0.30	0.56	0.26	0.18	0.57	12.67	0.26	0.43	0.54	0.67	0.98
				a_{11}	0.20	0.22	0.02	0.10	0.33	7.86	0.03	0.15	0.22	0.30	0.43
				b_{11}	0.50	0.42	-0.08	0.09	0.28	5.20	0.24	0.36	0.42	0.48	0.59
				b_{12}	0.10	0.03	-0.07	0.03	0.10	0.56	0.00	0.01	0.03	0.05	0.11
				ω_2	0.50	0.65	0.15	0.27	0.86	18.92	0.22	0.46	0.62	0.82	1.27
				a_{22}	0.40	0.32	-0.08	0.13	0.40	7.10	0.07	0.24	0.32	0.41	0.56
				b_{22}	0.20	0.30	0.10	0.09	0.28	8.74	0.14	0.24	0.29	0.36	0.49
				b_{21}	0.30	0.17	-0.13	0.08	0.24	3.44	0.03	0.12	0.17	0.22	0.33
				$\hat{\delta}$	0.70	0.56	-0.14	0.77	2.42	60.68	-0.89	-0.02	0.59	1.14	1.96
		200	40.2	ω_1	0.30	0.48	0.18	0.12	0.39	5.70	0.25	0.39	0.48	0.56	0.73
				a_{11}	0.20	0.18	-0.02	0.09	0.28	4.54	0.02	0.12	0.18	0.24	0.37
				b_{11}	0.50	0.46	-0.04	0.06	0.21	2.75	0.34	0.42	0.46	0.50	0.59
				b_{12}	0.10	0.07	-0.03	0.04	0.14	1.41	0.00	0.03	0.06	0.09	0.16
				ω_2	0.50	0.54	0.04	0.17	0.53	7.77	0.24	0.42	0.52	0.64	0.90
				a_{22}	0.40	0.28	-0.12	0.10	0.31	5.42	0.09	0.22	0.28	0.35	0.48
				b_{22}	0.20	0.35	0.15	0.07	0.22	3.58	0.22	0.30	0.35	0.40	0.49
				b_{21}	0.30	0.19	-0.11	0.06	0.19	2.38	0.07	0.15	0.19	0.23	0.31
				$\hat{\delta}$	0.70	0.29	-0.41	0.62	1.96	33.50	-0.86	-0.16	0.27	0.73	1.50
		500	39.9	ω_1	0.30	0.39	0.09	0.08	0.24	4.71	0.25	0.34	0.39	0.44	0.56
				a_{11}	0.20	0.18	-0.02	0.06	0.18	5.58	0.06	0.14	0.18	0.22	0.29
				b_{11}	0.50	0.54	0.04	0.04	0.13	2.98	0.46	0.52	0.54	0.57	0.63
				b_{12}	0.10	0.08	-0.02	0.03	0.10	1.57	0.02	0.06	0.08	0.11	0.15
				ω_2	0.50	0.42	-0.08	0.09	0.30	5.13	0.26	0.36	0.42	0.48	0.63
				a_{22}	0.40	0.41	0.01	0.06	0.19	4.58	0.29	0.37	0.41	0.45	0.53
				b_{22}	0.20	0.26	0.06	0.04	0.14	2.55	0.17	0.23	0.26	0.29	0.34
				b_{21}	0.30	0.23	-0.07	0.03	0.10	1.39	0.16	0.21	0.23	0.25	0.29
				$\hat{\delta}$	0.70	0.73	0.03	0.52	1.64	58.18	-0.30	0.38	0.74	1.10	1.67
B	500000	50	42.3	ω_1	0.30	0.44	0.14	0.20	0.63	8.05	0.11	0.29	0.42	0.56	0.88
				a_{11}	0.20	0.24	0.04	0.11	0.36	6.20	0.04	0.16	0.24	0.32	0.47
				b_{11}	0.50	0.45	-0.05	0.11	0.35	6.73	0.23	0.37	0.45	0.52	0.67
				b_{12}	0.10	0.11	0.01	0.08	0.25	3.02	0.01	0.05	0.10	0.16	0.30
				ω_2	0.50	0.67	0.17	0.28	0.90	12.01	0.21	0.46	0.63	0.84	1.32
				a_{22}	0.40	0.36	-0.04	0.15	0.48	8.09	0.07	0.25	0.36	0.46	0.66
				b_{22}	0.20	0.32	0.12	0.13	0.40	8.42	0.09	0.23	0.32	0.41	0.58
				b_{21}	0.30	0.10	-0.20	0.07	0.23	2.21	0.01	0.05	0.09	0.15	0.28
				$\hat{\delta}$	0.70	0.71	0.01	0.92	2.90	50.92	-0.95	0.00	0.76	1.46	2.30
		100	41.7	ω_1	0.30	0.55	0.25	0.16	0.50	10.12	0.27	0.44	0.54	0.65	0.89
				a_{11}	0.20	0.23	0.03	0.10	0.30	6.47	0.05	0.16	0.23	0.30	0.42
				b_{11}	0.50	0.40	-0.10	0.08	0.25	3.94	0.26	0.35	0.40	0.45	0.57
				b_{12}	0.10	0.04	-0.06	0.03	0.11	0.77	0.00	0.01	0.03	0.06	0.13
				ω_2	0.50	0.61	0.11	0.25	0.78	15.51	0.21	0.43	0.59	0.76	1.16
				a_{22}	0.40	0.33	-0.07	0.13	0.41	7.54	0.07	0.24	0.33	0.42	0.58
				b_{22}	0.20	0.32	0.12	0.10	0.32	10.27	0.14	0.25	0.32	0.39	0.52
				b_{21}	0.30	0.15	-0.15	0.07	0.23	3.01	0.02	0.10	0.15	0.20	0.30
				$\hat{\delta}$	0.70	0.58	-0.12	0.80	2.54	65.14	-0.93	-0.02	0.57	1.18	2.05
		200	40.3	ω_1	0.30	0.47	0.17	0.13	0.40	6.16	0.24	0.38	0.46	0.55	0.73
				a_{11}	0.20	0.19	-0.01	0.09	0.29	4.66	0.02	0.12	0.19	0.25	0.38
				b_{11}	0.50	0.45	-0.05	0.06	0.20	2.48	0.33	0.41	0.45	0.49	0.58
				b_{12}	0.10	0.07	-0.03	0.04	0.14	1.47	0.00	0.04	0.06	0.10	0.17
				ω_2	0.50	0.53	0.03	0.16	0.51	7.61	0.24	0.41	0.51	0.63	0.88
				a_{22}	0.40	0.28	-0.12	0.10	0.32	5.55	0.08	0.21	0.28	0.35	0.48
				b_{22}	0.20	0.36	0.16	0.07	0.23	4.08	0.22	0.31	0.36	0.41	0.50
				b_{21}	0.30	0.18	-0.12	0.06	0.19	2.14	0.07	0.14	0.18	0.22	0.30
				$\hat{\delta}$	0.70	0.25	-0.45	0.65	2.05	36.97	-0.91	-0.23	0.23	0.70	1.59
		500	39.6	ω_1	0.30	0.38	0.08	0.08	0.26	6.00	0.22	0.32	0.37	0.43	0.55
				a_{11}	0.20	0.17	-0.03	0.06	0.20	6.46	0.05	0.13	0.17	0.22	0.30
				b_{11}	0.50	0.54	0.04	0.04	0.14	3.48	0.47	0.51	0.54	0.57	0.64
				b_{12}	0.10	0.09	-0.01	0.03	0.10	1.36	0.03	0.07	0.09	0.11	0.15
				ω_2	0.50	0.40	-0.10	0.09	0.29	5.18	0.24	0.33	0.39	0.46	0.60
				a_{22}	0.40	0.42	0.02	0.06	0.19	4.82	0.30	0.38	0.42	0.46	0.54
				b_{22}	0.20	0.26	0.06	0.04	0.13	2.44	0.18	0.23	0.26	0.29	0.34
				b_{21}	0.30	0.23	-0.07	0.03	0.10	1.30	0.16	0.20	0.22	0.25	0.29
				$\hat{\delta}$	0.70	0.71	0.01	0.46	1.46	46.29	-0.19	0.38	0.71	1.01	1.63

Table C.6: Simulated Data C - MCMC Output Statistics

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
C	15000	50	40.7	ω_1	0.50	0.62	0.12	0.39	1.24	31.89	0.04	0.31	0.56	0.86	1.51
				a_{11}	0.20	0.19	-0.01	0.14	0.45	4.64	0.01	0.08	0.17	0.28	0.52
				a_{12}	0.30	0.43	0.13	0.27	0.84	21.64	0.03	0.23	0.40	0.59	1.06
				b_{11}	0.20	0.15	-0.05	0.10	0.30	3.95	0.01	0.07	0.14	0.21	0.36
				ω_2	0.30	0.63	0.33	0.26	0.82	11.84	0.14	0.44	0.62	0.80	1.15
				a_{22}	0.20	0.17	-0.03	0.12	0.38	4.66	0.01	0.07	0.15	0.24	0.44
				a_{21}	0.40	0.13	-0.27	0.11	0.34	3.45	0.00	0.04	0.10	0.18	0.40
				b_{22}	0.30	0.34	0.04	0.10	0.33	4.84	0.14	0.26	0.33	0.41	0.55
				δ	-0.50	-0.04	0.46	0.82	2.58	32.08	-1.26	-0.68	-0.16	0.48	1.77
		100	40.5	ω_1	0.50	0.75	0.25	0.34	1.07	17.09	0.11	0.51	0.75	0.99	1.42
				a_{11}	0.20	0.18	-0.02	0.14	0.44	4.69	0.01	0.07	0.15	0.26	0.51
				a_{12}	0.30	0.31	0.01	0.18	0.56	10.57	0.03	0.17	0.29	0.42	0.70
				b_{11}	0.20	0.11	-0.09	0.07	0.22	2.50	0.01	0.05	0.10	0.15	0.27
				ω_2	0.30	0.61	0.31	0.28	0.87	10.77	0.10	0.41	0.61	0.81	1.17
				a_{22}	0.20	0.15	-0.05	0.11	0.34	3.31	0.01	0.06	0.13	0.21	0.40
				a_{21}	0.40	0.18	-0.22	0.14	0.43	5.03	0.01	0.07	0.15	0.27	0.50
				b_{22}	0.30	0.36	0.06	0.09	0.28	3.79	0.20	0.30	0.36	0.42	0.54
				δ	-0.50	-0.48	0.02	0.61	1.94	20.90	-1.34	-0.96	-0.58	-0.10	0.94
		200	40.3	ω_1	0.50	0.86	0.36	0.27	0.87	13.29	0.34	0.66	0.86	1.05	1.39
				a_{11}	0.20	0.21	0.01	0.14	0.44	5.52	0.01	0.10	0.19	0.30	0.52
				a_{12}	0.30	0.18	-0.12	0.11	0.36	4.86	0.01	0.09	0.16	0.25	0.45
				b_{11}	0.20	0.15	-0.05	0.06	0.19	1.91	0.04	0.11	0.15	0.19	0.27
				ω_2	0.30	0.52	0.22	0.24	0.77	9.30	0.07	0.34	0.52	0.70	0.99
				a_{22}	0.20	0.12	-0.08	0.08	0.27	2.23	0.01	0.05	0.10	0.17	0.32
				a_{21}	0.40	0.19	-0.21	0.13	0.41	5.38	0.01	0.09	0.17	0.28	0.48
				b_{22}	0.30	0.41	0.11	0.07	0.21	2.21	0.28	0.36	0.41	0.45	0.54
				δ	-0.50	-0.64	-0.14	0.46	1.44	12.90	-1.32	-0.99	-0.70	-0.36	0.38
	500	39.5		ω_1	0.50	0.81	0.31	0.23	0.73	13.15	0.39	0.65	0.81	0.96	1.29
				a_{11}	0.20	0.24	0.04	0.12	0.39	4.03	0.02	0.14	0.23	0.32	0.48
				a_{12}	0.30	0.16	-0.14	0.08	0.26	3.66	0.02	0.09	0.15	0.21	0.33
				b_{11}	0.20	0.20	0.00	0.04	0.12	1.29	0.13	0.17	0.20	0.23	0.28
				ω_2	0.30	0.33	0.03	0.20	0.63	5.94	0.03	0.18	0.31	0.45	0.79
				a_{22}	0.20	0.18	-0.02	0.09	0.30	3.92	0.02	0.11	0.18	0.25	0.38
				a_{21}	0.40	0.29	-0.11	0.11	0.34	4.72	0.08	0.22	0.30	0.37	0.49
				b_{22}	0.30	0.34	0.04	0.04	0.14	1.36	0.26	0.31	0.34	0.37	0.43
				δ	-0.50	-0.64	-0.14	0.35	1.11	10.58	-1.25	-0.90	-0.65	-0.40	0.09

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
C	500000	50	40.9	ω_1	0.50	0.56	0.06	0.35	1.11	25.46	0.03	0.28	0.53	0.80	1.33
				a_{11}	0.20	0.19	-0.01	0.14	0.45	4.83	0.01	0.08	0.17	0.29	0.53
				a_{12}	0.30	0.44	0.14	0.25	0.79	18.58	0.05	0.25	0.41	0.60	0.98
				b_{11}	0.20	0.15	-0.05	0.10	0.32	4.37	0.01	0.07	0.14	0.22	0.38
				ω_2	0.30	0.62	0.32	0.25	0.80	12.08	0.15	0.45	0.62	0.80	1.13
				a_{22}	0.20	0.17	-0.03	0.12	0.39	5.12	0.01	0.07	0.15	0.25	0.46
				a_{21}	0.40	0.13	-0.27	0.11	0.34	3.33	0.00	0.04	0.10	0.18	0.40
				b_{22}	0.30	0.34	0.04	0.10	0.33	5.01	0.15	0.26	0.33	0.41	0.55
				δ	-0.50	0.00	0.50	0.83	2.62	33.46	-1.24	-0.65	-0.11	0.55	1.83
		100	40.8	ω_1	0.50	0.71	0.21	0.34	1.07	17.74	0.09	0.47	0.70	0.95	1.40
				a_{11}	0.20	0.18	-0.02	0.14	0.45	4.49	0.01	0.07	0.15	0.27	0.52
				a_{12}	0.30	0.32	0.02	0.19	0.60	11.67	0.03	0.17	0.29	0.44	0.74
				b_{11}	0.20	0.11	-0.09	0.07	0.22	2.37	0.01	0.06	0.10	0.15	0.26
				ω_2	0.30	0.61	0.31	0.27	0.86	10.38	0.11	0.41	0.60	0.80	1.14
				a_{22}	0.20	0.15	-0.05	0.11	0.34	3.33	0.01	0.06	0.13	0.21	0.40
				a_{21}	0.40	0.19	-0.21	0.14	0.43	5.25	0.01	0.07	0.16	0.27	0.50
				b_{22}	0.30	0.36	0.06	0.08	0.27	3.40	0.20	0.30	0.36	0.42	0.53
				δ	-0.50	-0.45	0.05	0.64	2.03	22.66	-1.35	-0.96	-0.56	-0.04	1.04
		200	40.6	ω_1	0.50	0.86	0.36	0.25	0.80	11.90	0.35	0.68	0.86	1.03	1.34
				a_{11}	0.20	0.21	0.01	0.13	0.42	4.63	0.01	0.10	0.19	0.29	0.49
				a_{12}	0.30	0.19	-0.11	0.11	0.36	4.69	0.02	0.10	0.17	0.26	0.45
				b_{11}	0.20	0.15	-0.05	0.06	0.20	2.08	0.04	0.11	0.15	0.19	0.28
				ω_2	0.30	0.54	0.24	0.24	0.77	9.64	0.09	0.35	0.54	0.71	1.00
				a_{22}	0.20	0.12	-0.08	0.09	0.28	2.43	0.01	0.05	0.10	0.17	0.33
				a_{21}	0.40	0.19	-0.21	0.12	0.38	5.12	0.01	0.09	0.17	0.27	0.45
				b_{22}	0.30	0.40	0.10	0.07	0.21	2.15	0.28	0.36	0.40	0.45	0.54
				δ	-0.50	-0.64	-0.14	0.45	1.43	12.41	-1.32	-0.99	-0.70	-0.36	0.37
	500	39.6		ω_1	0.50	0.78	0.28	0.21	0.68	10.98	0.40	0.63	0.77	0.93	1.22
				a_{11}	0.20	0.24	0.04	0.12	0.38	3.64	0.03	0.15	0.23	0.32	0.47
				a_{12}	0.30	0.17	-0.13	0.09	0.28	4.31	0.02	0.10	0.16	0.22	0.36
				b_{11}	0.20	0.20	0.00	0.04	0.13	1.34	0.13	0.17	0.20	0.23	0.28
				ω_2	0.30	0.32	0.02	0.20	0.62	5.75	0.03	0.17	0.30	0.45	0.76
				a_{22}	0.20	0.18	-0.02	0.10	0.31	4.26	0.02	0.11	0.17	0.25	0.38
				a_{21}	0.40	0.31	-0.09	0.11	0.35	4.41	0.08	0.23	0.31	0.38	0.52
				b_{22}	0.30	0.34	0.04	0.04	0.13	1.22	0.26	0.31	0.34	0.37	0.42
				δ	-0.50	-0.64	-0.14	0.35	1.10	10.36	-1.25	-0.90	-0.67	-0.41	0.08

Table C.7: *Simulated Data D - MCMC Output Statistics*

	Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
D	15000	50	40.4	ω_1	0.50	0.71	0.21	0.38	1.21	34.90	0.08	0.44	0.66	0.93	1.63	
				a_{11}	0.30	0.15	-0.15	0.13	0.42	8.87	0.00	0.05	0.12	0.22	0.49	
				a_{12}	0.20	0.37	0.17	0.18	0.58	25.31	0.04	0.22	0.37	0.51	0.69	
				b_{11}	0.10	0.09	-0.01	0.07	0.22	3.00	0.00	0.03	0.07	0.13	0.26	
				b_{12}	0.20	0.21	0.01	0.12	0.38	6.48	0.02	0.12	0.20	0.29	0.48	
				ω_2	0.30	0.59	0.29	0.34	1.08	29.17	0.04	0.33	0.56	0.82	1.32	
				a_{22}	0.30	0.16	-0.14	0.12	0.38	8.44	0.00	0.06	0.13	0.23	0.43	
				a_{21}	0.10	0.25	0.15	0.16	0.50	11.23	0.01	0.12	0.22	0.35	0.60	
				b_{22}	0.20	0.19	-0.01	0.11	0.35	5.23	0.01	0.10	0.18	0.26	0.44	
				b_{21}	0.30	0.14	-0.16	0.09	0.28	4.86	0.01	0.07	0.13	0.20	0.33	
				δ	0.40	1.56	1.16	0.98	3.11	75.47	-0.40	0.84	1.63	2.34	3.24	
		100	41.3	ω_1	0.50	0.99	0.49	0.41	1.30	90.33	0.22	0.66	1.01	1.27	1.76	
				a_{11}	0.30	0.22	-0.08	0.18	0.56	24.46	0.01	0.08	0.18	0.34	0.64	
				a_{12}	0.20	0.25	0.05	0.16	0.50	16.99	0.01	0.13	0.24	0.36	0.59	
				b_{11}	0.10	0.07	-0.03	0.05	0.16	2.47	0.00	0.03	0.06	0.10	0.19	
				b_{12}	0.20	0.07	-0.13	0.06	0.18	1.82	0.00	0.03	0.06	0.11	0.22	
				ω_2	0.30	0.54	0.24	0.30	0.93	44.96	0.04	0.31	0.52	0.75	1.15	
				a_{22}	0.30	0.15	-0.15	0.11	0.33	8.07	0.01	0.07	0.13	0.22	0.38	
				a_{21}	0.10	0.23	0.13	0.15	0.47	11.39	0.01	0.11	0.22	0.33	0.56	
				b_{22}	0.20	0.25	0.05	0.09	0.29	7.55	0.07	0.19	0.25	0.31	0.44	
				b_{21}	0.30	0.17	-0.13	0.07	0.22	4.45	0.05	0.13	0.17	0.22	0.32	
				δ	0.40	1.30	0.90	1.07	3.40	173.96	-0.80	0.51	1.47	2.08	3.08	
		200	41.5	ω_1	0.50	0.97	0.47	0.27	0.86	16.47	0.48	0.78	0.95	1.14	1.54	
				a_{11}	0.30	0.15	-0.15	0.11	0.36	6.52	0.00	0.05	0.12	0.21	0.41	
				a_{12}	0.20	0.20	0.00	0.13	0.40	9.40	0.01	0.10	0.18	0.29	0.47	
				b_{11}	0.10	0.06	-0.04	0.04	0.13	1.20	0.00	0.02	0.05	0.08	0.16	
				b_{12}	0.20	0.23	0.03	0.08	0.24	3.62	0.08	0.18	0.23	0.28	0.38	
				ω_2	0.30	0.41	0.11	0.23	0.74	11.08	0.03	0.22	0.39	0.57	0.90	
				a_{22}	0.30	0.13	-0.17	0.10	0.30	5.33	0.01	0.06	0.12	0.19	0.36	
				a_{21}	0.10	0.18	0.08	0.12	0.37	8.74	0.01	0.09	0.17	0.27	0.43	
				b_{22}	0.20	0.29	0.09	0.07	0.23	3.27	0.15	0.24	0.29	0.34	0.43	
				b_{21}	0.30	0.21	-0.09	0.06	0.19	3.30	0.09	0.17	0.21	0.25	0.33	
				δ	0.40	0.74	0.34	0.70	2.21	36.03	-0.54	0.24	0.71	1.23	2.14	
		500	41.9	ω_1	0.50	0.78	0.28	0.19	0.60	14.01	0.43	0.65	0.78	0.91	1.17	
				a_{11}	0.30	0.12	-0.18	0.10	0.32	7.94	0.00	0.04	0.09	0.18	0.38	
				a_{12}	0.20	0.29	0.09	0.13	0.40	14.31	0.03	0.20	0.30	0.38	0.52	
				b_{11}	0.10	0.10	0.00	0.04	0.13	1.91	0.02	0.07	0.10	0.13	0.18	
				b_{12}	0.20	0.23	0.03	0.04	0.14	1.97	0.14	0.20	0.23	0.26	0.31	
				ω_2	0.30	0.28	-0.02	0.16	0.52	9.19	0.02	0.15	0.26	0.38	0.64	
				a_{22}	0.30	0.19	-0.11	0.12	0.39	12.78	0.01	0.09	0.18	0.28	0.44	
				a_{21}	0.10	0.24	0.14	0.13	0.41	11.29	0.02	0.13	0.24	0.33	0.49	
				b_{22}	0.20	0.22	0.02	0.04	0.13	1.81	0.14	0.19	0.22	0.25	0.31	
				b_{21}	0.30	0.22	-0.08	0.04	0.13	2.20	0.14	0.19	0.22	0.25	0.30	
				δ	0.40	1.06	0.66	0.58	1.83	39.41	-0.04	0.64	1.06	1.47	2.17	
D	500000	50	41.2	ω_1	0.50	0.77	0.27	0.37	1.17	33.45	0.07	0.44	0.66	0.93	1.60	
				a_{11}	0.30	0.19	-0.11	0.14	0.43	9.05	0.01	0.08	0.16	0.28	0.51	
				a_{12}	0.20	0.28	0.08	0.16	0.51	18.34	0.02	0.16	0.26	0.36	0.67	
				b_{11}	0.10	0.10	0.00	0.08	0.25	3.78	0.00	0.04	0.08	0.14	0.29	
				b_{12}	0.20	0.22	0.02	0.12	0.39	6.53	0.02	0.13	0.21	0.31	0.48	
				ω_2	0.30	0.63	0.33	0.39	1.24	38.72	0.04	0.34	0.58	0.86	1.60	
				a_{22}	0.30	0.18	-0.12	0.14	0.44	10.57	0.01	0.07	0.15	0.27	0.50	
				a_{21}	0.10	0.21	0.11	0.13	0.41	6.76	0.01	0.10	0.20	0.30	0.48	
				b_{22}	0.20	0.19	-0.01	0.11	0.35	4.69	0.02	0.10	0.18	0.26	0.42	
				b_{21}	0.30	0.15	-0.15	0.10	0.30	5.47	0.01	0.07	0.13	0.21	0.37	
				δ	0.40	1.52	1.12	1.05	3.32	85.10	-0.62	0.77	1.59	2.32	3.30	
		100	41.8	ω_1	0.50	0.87	0.37	0.31	0.98	51.07	0.36	0.65	0.83	1.06	1.60	
				a_{11}	0.30	0.25	-0.05	0.17	0.54	21.68	0.01	0.10	0.22	0.38	0.60	
				a_{12}	0.20	0.27	0.07	0.17	0.54	18.63	0.01	0.13	0.25	0.38	0.63	
				b_{11}	0.10	0.07	-0.03	0.05	0.16	2.88	0.00	0.03	0.06	0.10	0.18	
				b_{12}	0.20	0.07	-0.13	0.06	0.18	1.69	0.00	0.03	0.06	0.10	0.21	
				ω_2	0.30	0.43	0.13	0.25	0.79	32.76	0.02	0.23	0.44	0.62	0.93	
				a_{22}	0.30	0.18	-0.12	0.12	0.38	10.12	0.01	0.09	0.17	0.27	0.43	
				a_{21}	0.10	0.22	0.12	0.15	0.47	10.81	0.01	0.10	0.20	0.33	0.54	
				b_{22}	0.20	0.26	0.06	0.09	0.29	7.00	0.08	0.20	0.25	0.31	0.44	
				b_{21}	0.30	0.17	-0.13	0.07	0.22	4.68	0.04	0.12	0.17	0.22	0.31	
				δ	0.40	0.64	0.24	0.95	3.00	140.00	-0.95	-0.10	0.61	1.32	2.54	
		200	40.9	ω_1	0.50	0.92	0.42	0.28	0.88	16.42	0.40	0.72	0.91	1.11	1.47	
				a_{11}	0.30	0.16	-0.14	0.12	0.39	7.23	0.01	0.06	0.13	0.24	0.46	
				a_{12}	0.20	0.22	0.02	0.13	0.42	9.60	0.01	0.11	0.20	0.31	0.50	
				b_{11}	0.10	0.06	-0.04	0.04	0.13	1.08	0.00	0.02	0.05	0.08	0.15	
				b_{12}	0.20	0.22	0.02	0.07	0.23	3.41	0.08	0.17	0.22	0.27	0.37	
				ω_2	0.30	0.39	0.09	0.23	0.73	10.06	0.03	0.21	0.36	0.54	0.88	
				a_{22}	0.30	0.15	-0.15	0.10	0.32	5.87	0.01	0.07	0.13	0.21	0.39	
				a_{21}	0.10	0.19	0.09	0.12	0.37	8.24	0.01	0.09	0.18	0.27	0.43	
				b_{22}	0.20	0.28	0.08	0.07	0.23	3.21	0.15	0.23	0.28	0.33	0.42	
				b_{21}	0.30	0.20	-0.10	0.06	0.19	3.18	0.09	0.16	0.21	0.25	0.32	
				δ	0.40	0.58	0.18	0.71	2.26	36.98	-0.68	0.05	0.54	1.07	2.02	
		500	41.2	ω_1	0.50	0.81	0.31	0.19	0.60	13.56	0.45	0.68	0.80	0.93	1.20	
				a_{11}	0.30	0.14	-0.16	0.11	0.34	9.40	0.00	0.05	0.11	0.20	0.39	
				a_{12}	0.20	0.25	0.05	0.13	0.41	14.73	0.02	0.16	0.26	0.34	0.51	

Table C.8: *Simulated Data DC1 - MCMC Output Statistics*

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5
DC1	15000	50	41.5	ω_1	0.30	0.43	0.13	0.14	0.44	21.60	0.19	0.33	0.41	0.49	0.75
				a_{11}	0.30	0.16	-0.14	0.13	0.42	10.41	0.00	0.05	0.12	0.23	0.49
				b_{11}	0.20	0.21	0.01	0.15	0.47	25.86	0.01	0.10	0.19	0.29	0.56
				ω_2	0.20	0.14	-0.06	0.08	0.25	5.47	0.02	0.08	0.13	0.20	0.32
				a_{22}	0.20	0.38	0.18	0.23	0.73	18.09	0.02	0.19	0.36	0.55	0.82
				b_{22}	0.20	0.24	0.04	0.13	0.40	11.43	0.04	0.15	0.22	0.32	0.53
				$\bar{\delta}$	-1.00	-1.65	-0.65	1.72	5.42	329.16	-4.91	-3.01	-1.47	-0.36	1.40
				α	0.20	0.50	0.30	0.25	0.79	7.94	0.04	0.31	0.52	0.70	0.91
				β	0.20	0.15	-0.05	0.11	0.34	6.79	0.01	0.07	0.13	0.22	0.40
				δ_{level}	-0.60	-0.36	0.24	0.47	1.49	29.73	-1.23	-0.68	-0.34	-0.07	0.66
		100	41.2	ω_1	0.30	0.37	0.07	0.13	0.41	13.78	0.12	0.27	0.37	0.46	0.61
				a_{11}	0.30	0.27	-0.03	0.17	0.55	16.97	0.01	0.13	0.25	0.38	0.63
				b_{11}	0.20	0.28	0.08	0.11	0.34	11.49	0.06	0.20	0.27	0.35	0.49
				ω_2	0.20	0.13	-0.07	0.08	0.26	4.85	0.02	0.07	0.12	0.18	0.32
				a_{22}	0.20	0.43	0.23	0.22	0.70	14.71	0.03	0.26	0.44	0.61	0.82
				b_{22}	0.20	0.22	0.02	0.10	0.31	5.38	0.05	0.15	0.21	0.28	0.43
				$\bar{\delta}$	-1.00	-1.63	-0.63	1.26	4.00	135.39	-4.39	-2.35	-1.45	-0.77	0.56
				α	0.20	0.43	0.23	0.24	0.77	8.71	0.02	0.23	0.44	0.62	0.87
				β	0.20	0.14	-0.06	0.09	0.27	3.27	0.01	0.07	0.14	0.20	0.32
				δ_{level}	-0.60	-0.53	0.07	0.40	1.25	18.49	-1.27	-0.81	-0.52	-0.25	0.26
		200	40.0	ω_1	0.30	0.39	0.09	0.12	0.37	13.33	0.14	0.31	0.39	0.47	0.60
				a_{11}	0.30	0.26	-0.04	0.16	0.52	18.31	0.02	0.13	0.25	0.37	0.62
				b_{11}	0.20	0.18	-0.02	0.07	0.21	5.72	0.06	0.13	0.17	0.22	0.32
				ω_2	0.20	0.15	-0.05	0.06	0.17	1.89	0.06	0.11	0.15	0.19	0.27
				a_{22}	0.20	0.33	0.13	0.15	0.48	5.45	0.04	0.22	0.33	0.44	0.63
				b_{22}	0.20	0.25	0.05	0.08	0.25	5.16	0.11	0.20	0.25	0.30	0.41
				$\bar{\delta}$	-1.00	-1.51	-0.51	1.12	3.54	124.62	-4.22	-2.17	-1.34	-0.72	0.32
				α	0.20	0.47	0.27	0.26	0.83	10.66	0.03	0.24	0.48	0.68	0.93
				β	0.20	0.11	-0.09	0.07	0.23	2.56	0.00	0.05	0.10	0.16	0.26
				δ_{level}	-0.60	-0.48	0.12	0.36	1.13	14.56	-1.23	-0.71	-0.44	-0.22	0.16
		500	38.9	ω_1	0.30	0.34	0.04	0.12	0.38	17.66	0.11	0.25	0.35	0.43	0.54
				a_{11}	0.30	0.30	0.00	0.19	0.61	29.08	0.02	0.15	0.26	0.45	0.70
				b_{11}	0.20	0.16	-0.04	0.05	0.16	3.57	0.06	0.12	0.16	0.20	0.26
				ω_2	0.20	0.15	-0.05	0.04	0.12	1.14	0.08	0.13	0.15	0.18	0.24
				a_{22}	0.20	0.28	0.08	0.11	0.36	3.52	0.06	0.20	0.28	0.36	0.50
				b_{22}	0.20	0.30	0.10	0.05	0.16	2.89	0.20	0.26	0.30	0.33	0.41
				$\bar{\delta}$	-1.00	-1.25	-0.25	0.82	2.61	78.63	-3.32	-1.69	-1.12	-0.67	0.00
				α	0.20	0.41	0.21	0.25	0.78	10.20	0.02	0.20	0.39	0.60	0.88
				β	0.20	0.10	-0.10	0.06	0.20	2.45	0.00	0.04	0.09	0.14	0.23
				δ_{level}	-0.60	-0.51	0.09	0.32	1.00	13.01	-1.18	-0.73	-0.49	-0.27	0.00
DC1	500000	50	40.9	ω_1	0.30	0.31	0.01	0.12	0.39	14.67	0.08	0.22	0.31	0.40	0.56
				a_{11}	0.30	0.29	-0.01	0.21	0.68	22.37	0.01	0.11	0.24	0.42	0.76
				b_{11}	0.20	0.18	-0.02	0.10	0.32	10.94	0.02	0.10	0.17	0.25	0.39
				ω_2	0.20	0.14	-0.06	0.07	0.22	3.63	0.03	0.08	0.13	0.18	0.29
				a_{22}	0.20	0.33	0.13	0.20	0.64	12.34	0.02	0.16	0.31	0.47	0.76
				b_{22}	0.20	0.26	0.06	0.15	0.48	13.50	0.04	0.15	0.24	0.35	0.60
				$\bar{\delta}$	-1.00	-0.98	0.02	1.42	4.49	188.29	-3.79	-1.82	-1.01	-0.10	2.00
				α	0.20	0.49	0.29	0.27	0.84	7.98	0.03	0.27	0.50	0.72	0.94
				β	0.20	0.12	-0.08	0.09	0.28	4.26	0.01	0.05	0.10	0.16	0.34
				δ_{level}	-0.60	-0.24	0.36	0.51	1.61	32.91	-1.15	-0.55	-0.25	-0.02	0.98
		100	41.2	ω_1	0.30	0.32	0.02	0.14	0.44	15.06	0.09	0.21	0.30	0.41	0.62
				a_{11}	0.30	0.34	0.04	0.18	0.57	17.55	0.03	0.21	0.34	0.47	0.73
				b_{11}	0.20	0.24	0.04	0.10	0.32	9.37	0.04	0.17	0.23	0.30	0.46
				ω_2	0.20	0.13	-0.07	0.07	0.24	4.10	0.02	0.07	0.12	0.18	0.30
				a_{22}	0.20	0.42	0.22	0.20	0.64	12.28	0.04	0.27	0.43	0.57	0.80
				b_{22}	0.20	0.23	0.03	0.10	0.33	5.35	0.05	0.16	0.22	0.29	0.46
				$\bar{\delta}$	-1.00	-1.41	-0.41	1.18	3.74	112.15	-4.17	-2.03	-1.31	-0.66	0.75
				α	0.20	0.42	0.22	0.24	0.75	7.92	0.02	0.23	0.42	0.60	0.87
				β	0.20	0.13	-0.07	0.08	0.26	3.15	0.01	0.06	0.12	0.19	0.31
				δ_{level}	-0.60	-0.50	0.10	0.42	1.32	20.55	-1.23	-0.80	-0.51	-0.23	0.42
		200	41.4	ω_1	0.30	0.37	0.07	0.13	0.41	16.32	0.14	0.27	0.37	0.48	0.61
				a_{11}	0.30	0.29	-0.01	0.18	0.58	21.89	0.01	0.13	0.27	0.43	0.66
				b_{11}	0.20	0.17	-0.03	0.07	0.21	5.36	0.05	0.12	0.16	0.21	0.31
				ω_2	0.20	0.15	-0.05	0.05	0.17	1.80	0.06	0.11	0.15	0.19	0.27
				a_{22}	0.20	0.32	0.12	0.15	0.48	5.22	0.04	0.21	0.31	0.43	0.63
				b_{22}	0.20	0.26	0.06	0.09	0.27	6.01	0.12	0.20	0.26	0.31	0.46
				$\bar{\delta}$	-1.00	-1.34	-0.34	0.89	2.82	74.65	-3.31	-1.84	-1.28	-0.73	0.24
				α	0.20	0.44	0.24	0.24	0.77	7.75	0.03	0.23	0.43	0.63	0.89
				β	0.20	0.11	-0.09	0.07	0.24	2.52	0.00	0.04	0.09	0.16	0.27
				δ_{level}	-0.60	-0.51	0.09	0.36	1.14	14.02	-1.25	-0.76	-0.49	-0.25	0.12
		500	39.8	ω_1	0.30	0.31	0.01	0.12	0.37	17.03	0.09	0.22	0.32	0.40	0.51
				a_{11}	0.30	0.35	0.05	0.20	0.64	31.10	0.03	0.19	0.33	0.50	0.75
				b_{11}	0.20	0.16	-0.04	0.04	0.14	2.75	0.07	0.12	0.15	0.18	0.25
				ω_2	0.20	0.15	-0.05	0.04	0.12	1.06	0.09	0.13	0.15	0.18	0.24
				a_{22}	0.20	0.28	0.08	0.11	0.36	3.67	0.05	0.19	0.27	0.35	0.50
				b_{22}	0.20	0.30	0.10	0.05	0.17	3.14	0.20	0.27	0.30	0.34	0.42
				$\bar{\delta}$	-1.00	-1.11	-0.11	0.69	2.19	54.79	-2.85	-1.47	-1.05	-0.65	0.05
				α	0.20	0.38	0.18	0.24	0.77	9.98	0.02	0.18	0.35	0.55	0.89
				β	0.20	0.09	-0.11	0.06	0.20	2.44	0.00	0.04	0.09	0.14	0.23
				δ_{level}	-0.60	-0.51	0.09	0.33	1.03	14.03	-1.19	-0.74	-0.49	-0.26	0.05

Table C.9: Simulated Data DC2 - MCMC Output Statistics

Model	Burn In	N	Accepted(%)	Parameter	Actual	Mean	Difference	SD ¹	SE ²	TSSE ³	p2.5	p25	p50	p75	p97.5		
DC2	15000	50	39.6	ω_1	1.00	0.81	-0.19	0.20	0.63	21.32	0.40	0.67	0.82	0.95	1.18		
				a_{11}	0.10	0.13	0.03	0.10	0.30	10.88	0.00	0.05	0.11	0.19	0.35		
				a_{12}	0.10	0.10	0.00	0.10	0.31	7.96	0.00	0.03	0.07	0.14	0.38		
				b_{11}	0.10	0.14	0.04	0.11	0.33	5.43	0.01	0.05	0.11	0.20	0.39		
				b_{12}	0.10	0.12	0.02	0.08	0.26	4.02	0.01	0.06	0.11	0.17	0.31		
				ω_2	0.10	0.29	0.19	0.21	0.68	19.73	0.01	0.12	0.24	0.43	0.81		
				a_{22}	0.30	0.20	-0.10	0.17	0.52	17.37	0.01	0.07	0.16	0.29	0.62		
				a_{21}	0.50	0.25	-0.25	0.17	0.53	24.19	0.01	0.09	0.23	0.38	0.57		
				b_{22}	0.20	0.27	0.07	0.13	0.40	14.08	0.05	0.17	0.26	0.36	0.52		
				b_{21}	0.10	0.14	0.04	0.11	0.34	10.28	0.01	0.06	0.12	0.20	0.41		
				δ	1.00	1.31	0.31	2.64	8.35	735.56	-3.35	-0.74	1.37	2.82	6.09		
				α	0.10	0.58	0.48	0.26	0.82	13.66	0.05	0.38	0.63	0.79	0.95		
				β	0.40	0.15	-0.25	0.11	0.34	4.93	0.01	0.07	0.13	0.21	0.40		
				δ_{level}	0.50	0.26	-0.24	0.61	1.94	51.00	-0.84	-0.06	0.14	0.50	1.86		
		100	40.9	ω_1	1.00	0.82	-0.18	0.24	0.76	47.24	0.35	0.64	0.86	1.00	1.22		
				a_{11}	0.10	0.18	0.08	0.17	0.53	36.78	0.01	0.05	0.11	0.25	0.57		
				a_{12}	0.10	0.10	0.00	0.08	0.26	4.34	0.00	0.04	0.09	0.15	0.31		
				b_{11}	0.10	0.12	0.02	0.07	0.23	3.26	0.01	0.06	0.11	0.17	0.28		
				b_{12}	0.10	0.17	0.07	0.08	0.25	3.34	0.02	0.11	0.17	0.22	0.33		
				ω_2	0.10	0.30	0.20	0.18	0.58	23.56	0.02	0.15	0.28	0.41	0.71		
				a_{22}	0.30	0.26	-0.04	0.16	0.51	17.99	0.01	0.12	0.25	0.38	0.59		
				a_{21}	0.50	0.21	-0.29	0.10	0.32	10.30	0.02	0.13	0.21	0.28	0.40		
				b_{22}	0.20	0.26	0.06	0.09	0.28	8.03	0.10	0.20	0.26	0.33	0.43		
				b_{21}	0.10	0.15	0.05	0.08	0.25	6.54	0.01	0.08	0.14	0.21	0.31		
				δ	1.00	-0.18	-1.18	2.22	7.03	689.81	-4.66	-1.78	0.13	1.34	4.28		
				α	0.10	0.47	0.37	0.24	0.76	11.59	0.03	0.28	0.48	0.66	0.89		
				β	0.40	0.24	-0.16	0.14	0.45	8.81	0.02	0.12	0.22	0.34	0.53		
				δ_{level}	0.50	0.07	-0.43	0.53	1.67	46.73	-0.94	-0.20	0.01	0.31	1.35		
		200	40.5	ω_1	1.00	0.78	-0.22	0.14	0.44	10.57	0.51	0.69	0.78	0.87	1.07		
				a_{11}	0.10	0.11	0.01	0.09	0.29	11.36	0.00	0.04	0.09	0.18	0.33		
				a_{12}	0.10	0.07	-0.03	0.06	0.19	5.79	0.00	0.02	0.05	0.09	0.22		
				b_{11}	0.10	0.10	0.00	0.06	0.20	3.63	0.01	0.05	0.09	0.14	0.24		
				b_{12}	0.10	0.21	0.11	0.06	0.18	3.20	0.10	0.17	0.21	0.24	0.32		
				ω_2	0.10	0.21	0.11	0.16	0.49	17.75	0.01	0.09	0.18	0.30	0.62		
				a_{22}	0.30	0.35	0.05	0.14	0.43	18.88	0.06	0.27	0.36	0.45	0.60		
				a_{21}	0.50	0.32	-0.18	0.09	0.30	16.48	0.15	0.23	0.32	0.39	0.48		
				b_{22}	0.20	0.26	0.06	0.07	0.23	7.19	0.13	0.21	0.26	0.31	0.40		
				b_{21}	0.10	0.09	-0.01	0.06	0.18	4.06	0.00	0.04	0.08	0.12	0.21		
				δ	1.00	-0.57	-1.57	1.93	6.09	355.52	-5.52	-1.61	-0.27	0.85	2.26		
				α	0.10	0.43	0.33	0.23	0.72	8.11	0.03	0.25	0.43	0.60	0.86		
				β	0.40	0.27	-0.13	0.14	0.44	5.59	0.02	0.16	0.26	0.36	0.55		
				δ_{level}	0.50	0.02	-0.48	0.48	1.51	35.48	-0.82	-0.27	-0.04	0.25	1.16		
		500	40.7	ω_1	1.00	0.98	-0.02	0.16	0.50	22.13	0.58	0.89	0.99	1.08	1.24		
				a_{11}	0.10	0.07	-0.03	0.06	0.19	6.80	0.00	0.03	0.06	0.10	0.22		
				a_{12}	0.10	0.05	-0.05	0.05	0.15	5.16	0.00	0.02	0.04	0.08	0.18		
				b_{11}	0.10	0.08	-0.02	0.04	0.12	1.39	0.01	0.05	0.08	0.11	0.16		
				b_{12}	0.10	0.17	0.07	0.03	0.10	1.52	0.11	0.15	0.17	0.19	0.24		
				ω_2	0.10	0.13	0.03	0.10	0.31	10.65	0.00	0.05	0.11	0.19	0.37		
				a_{22}	0.30	0.33	0.03	0.08	0.27	11.58	0.13	0.27	0.34	0.39	0.45		
				a_{21}	0.50	0.44	-0.06	0.07	0.24	15.62	0.31	0.39	0.43	0.50	0.59		
				b_{22}	0.20	0.23	0.03	0.04	0.14	3.01	0.14	0.20	0.23	0.26	0.32		
				b_{21}	0.10	0.07	-0.03	0.04	0.13	2.80	0.01	0.04	0.07	0.10	0.16		
				δ	1.00	0.37	-0.63	1.20	3.80	191.63	-2.43	-0.18	0.67	1.20	2.07		
				α	0.10	0.38	0.28	0.22	0.71	7.77	0.02	0.20	0.36	0.53	0.84		
				β	0.40	0.24	-0.16	0.13	0.42	4.99	0.01	0.13	0.24	0.34	0.49		
				δ_{level}	0.50	0.28	-0.22	0.45	1.41	32.82	-0.44	-0.03	0.22	0.54	1.32		
		DC2	500000	50	41.9	ω_1	1.00	0.77	-0.23	0.23	0.73	23.98	0.33	0.62	0.77	0.92	1.25
						a_{11}	0.10	0.16	0.06	0.12	0.38	14.96	0.01	0.06	0.14	0.24	0.43
						a_{12}	0.10	0.12	0.02	0.11	0.36	9.66	0.00	0.03	0.08	0.17	0.43
						b_{11}	0.10	0.11	0.01	0.09	0.28	3.70	0.00	0.04	0.09	0.16	0.33
						b_{12}	0.10	0.13	0.03	0.09	0.28	4.26	0.01	0.06	0.11	0.18	0.33
						ω_2	0.10	0.27	0.17	0.22	0.70	20.34	0.01	0.10	0.21	0.38	0.84
						a_{22}	0.30	0.17	-0.13	0.14	0.43	10.89	0.00	0.06	0.13	0.24	0.50
						a_{21}	0.50	0.34	-0.16	0.19	0.59	27.92	0.04	0.19	0.33	0.46	0.74
						b_{22}	0.20	0.27	0.07	0.13	0.40	14.49	0.05	0.17	0.26	0.36	0.52
						b_{21}	0.10	0.12	0.02	0.09	0.30	8.07	0.00	0.05	0.10	0.17	0.36
						δ	1.00	1.10	0.10	1.72	5.45	320.50	-3.17	0.12	1.36	2.28	4.29
						α	0.10	0.50	0.40	0.26	0.81	12.16	0.03	0.29	0.52	0.71	0.92
						β	0.40	0.15	-0.25	0.11	0.35	4.66	0.01	0.06	0.12	0.21	0.41
						δ_{level}	0.50	0.41	-0.09	0.64	2.03	51.19	-0.74	0.01	0.26	0.77	1.90
				100	41.0	ω_1	1.00	0.72	-0.28	0.22	0.70	39.70	0.24	0.58	0.74	0.88	1.09
						a_{11}	0.10	0.16	0.06	0.11	0.36	16.09	0.01	0.07	0.14	0.23	0.42
						a_{12}	0.10	0.14	0.04	0.12	0.37	8.84	0.00	0.05	0.11	0.21	0.43
						b_{11}	0.10	0.12	0.02	0.08	0.24	4.38	0.01	0.06	0.11	0.17	0.28
						b_{12}	0.10	0.19	0.09	0.09	0.28	4.22	0.03	0.13	0.19	0.25	0.38
						ω_2	0.10	0.35	0.25	0.29	0.91	55.99	0.01	0.13	0.27	0.48	1.09
						a_{22}	0.30	0.23	-0.07	0.14	0.43	11.21	0.01	0.13	0.23	0.33	0.50
						a_{21}	0.50	0.20	-0.30	0.13	0.41	15.17	0.01	0.09	0.18	0.29	0.47
						b_{22}	0.20	0.27	0.07	0.09	0.29	9.65	0.11	0.21	0.27	0.33	0.47
						b_{21}	0.10	0.12	0.02	0.07	0.23	6.13	0.01	0.07	0.12	0.17	0.29
						δ	1.00	1.02	0.02	1.41	4.46	294.26	-1.41	0.04	0.89	1.99	3.91
						α	0.10	0.45	0.35	0.25	0.78	13.23	0.03	0.24	0.45	0.64	0.89
						β	0.40	0.20	-0.20	0.13	0.41	6.76	0.01	0.10	0.19	0.29	0.48
						δ_{level}	0.50	0.32	-0.18	0.50	1.59	37.89	-0.45	0.00	0.19	0.56	1.60
				200	41.4	ω_1	1.00	0.78	-0.22	0.17	0.53	14.30	0.49	0.65	0.77	0.89	1.14
						a_{11}	0.10	0.12	0.02	0.09	0.27	8.72	0.00	0.04	0.10	0.18	0.32
						a_{12}	0.10	0.06	-0.04	0.05	0.16	3.21	0.00	0.02	0.05	0.09	0.18
						b_{11}	0.10	0.10	0.00	0.06	0.18	3.18	0.01	0.05	0.10	0.14	0.22
						b_{12}	0.10	0.21	0.11	0.05	0.17	3.05	0.11	0.18	0.21	0.25	0.32
						ω_2	0.10	0.17	0.07	0.13	0.43	13.17	0.01	0.06	0.15	0.25	0.51
						a_{22}	0.30	0.28	-0.02	0.12	0.37	12.86	0.04	0.20	0.28	0.36	0.49
						a_{21}	0.50	0.40	-0.10	0.09	0.30	14.89	0.24	0.33	0.39	0.48	0.59
						b_{22}	0.20	0.27	0.07	0.06	0.20	5.48	0.16	0.23	0.27	0.31	0.42
						b_{21}	0.10	0.10	0.00	0.06	0.19	4.73	0.00	0.05	0.09	0.14	0.22
						δ	1.00	-0.61	-1.61	1.91	6.03	328.73	-5.28	-1.71	-0.28	0.81	2.23
						α	0.10	0.42	0.32	0.22	0.71	7.39	0.03	0.24	0.42	0.59	0.84
						β	0.40	0.28	-0.12	0.14	0.46	6.50	0.02	0.17	0.28	0.38	0.57
						δ_{level}	0.50	0.01	-0.49	0.48	1.53	34.23	-0.85	-0.28	-0.04	0.24	1.19
				500	40.9	ω_1	1.00	1.01	0.01	0.13	0.41	14.57	0.75	0.92	1.01	1.10	1.24
						a_{11}	0.10	0.07	-0.03	0.05	0.16	5.13	0.00	0.03	0.06	0.11	0.18
						a_{12}	0.10	0.04	-0.06	0.03	0.10	2.06	0.00	0.01	0.03	0.05	0.12
						b_{11}	0.10	0.08	-0.02	0.04	0.12	1.31	0.01	0.05	0.08	0.11	0.16
						b_{12}	0.10	0.18	0.08	0.04	0.11	1.69	0.11	0.15	0.18	0.20	0.24
						ω_2	0.10	0.15	0.05	0.10	0.33	11.03	0.01	0.07	0.14	0.22	0.38
						a_{22}	0.3										

Appendix D

Random Walk Metropolis Hasting - Trace Charts

Procedure A1

Figure D.1: Trace Charts - A1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A2

Figure D.2: Trace Charts - A2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A3

Figure D.3: Trace Charts - A3 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A4

Figure D.4: Trace Charts - A_4 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure B

Figure D.5: Trace Charts - B - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure C

Figure D.6: Trace Charts - C - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure D

Figure D.7: Trace Charts - D - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure DC1

Figure D.8: Trace Charts - DC1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure E1

Diagonal

Figure D.9: Trace Charts - E1 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Non-Diagonal

Figure D.10: Trace Charts - E1 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Procedure E2

Diagonal

Figure D.11: Trace Charts - E2 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Non-Diagonal

Figure D.12: Trace Charts - E2 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Trade Data

Diagonal

Figure D.13: Trace Charts - Trades Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.

Non-Diagonal

Figure D.14: Trace Charts - Trades Non-Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.

Appendix E

Random Walk Metropolis Hasting - ACF Plots

Procedure M2

Figure E.1: ACF Charts - Rates Without Time Dependence - M2 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

Procedure M3

Figure E.2: ACF Charts - Rates Without Time Dependence - M3 - Data Points: 100. Burn In: 100000. Sample Size: 100000.

Procedure A1

Figure E.3: ACF Charts - A1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A2

Figure E.4: ACF Charts - A2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A3

Figure E.5: ACF Charts - A3 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure A4

Figure E.6: ACF Charts - A4 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure B

Figure E.7: ACF Charts - B - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure C

Figure E.8: ACF Charts - C - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure D

Figure E.9: ACF Charts - D - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure DC1

Figure E.10: ACF Charts - DC1 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure DC2

Figure E.11: ACF Charts - DC2 - Data Points: 500. Burn In: 500000. Sample Size: 100000.

Procedure E1

Diagonal

Figure E.12: ACF Charts - E1 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Non-Diagonal

Figure E.13: ACF Charts - E1 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Procedure E2

Diagonal

Figure E.14: ACF Charts - E2 Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Non-Diagonal

Figure E.15: ACF Charts - E2 Non-Diagonal - Data Points: 500. Burn In: 500000. Sample Size: 500000.

Trade Data

Diagonal

Figure E.16: ACF Charts - Trades Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.

Non-Diagonal

Figure E.17: ACF Charts - Trades Non-Diagonal - Data Points: 361. Burn In: 500000. Sample Size: 100000.

E.1 Random Walk Metropolis Hasting - $\lambda_{1,t}$, $\lambda_{2,t}$, δ_t and ρ_t evolution

Procedure A1 - Sample size: 200

Procedure A2 - Sample size: 200

Procedure A3 - Sample size: 200 **Procedure A4 - Sample size: 200**

Procedure B - Sample size: 200 **Procedure C - Sample size: 200**

Procedure D - Sample size: 200

Procedure DC1 - Sample size: 200

Procedure DC2 - Sample size: 200

Procedure E1 - Sample size: 500**Diagonal** **Non-Diagonal**

Procedure E2 - Sample size: 500**Diagonal** **Non-Diagonal**

Bibliography

- Ahmad and Francq(2016)** Ali Ahmad and Christian Francq. Poisson qmle of count time series models. *Journal of Time Series Analysis*, 37:291–314. ISSN 0143-9782,1467-9892. doi: 10.1111/jtsa.12167. URL <http://doi.org/10.1111/jtsa.12167>. Citado na pág. 1, 2
- Andreassen et al.(2013)** C. M. Andreassen, J. L. Jensen and R. A. Davis. On bivariate time series of counts, 2013. Citado na pág. 3
- Bennett et al.(1996)** James E. Bennett, Amy Racine-Poon and Jon C. Wakefield. *MCMC for nonlinear hierarchical models*, chapter 19. Springer US. ISBN 978-0-412-05551-5,978-1-4899-4485-6. Citado na pág. iii, iv, 68
- Blasques et al.(2014)** F. Blasques, S. J. Koopman and A. Lucas. Stationarity and ergodicity of univariate generalized autoregressive score processes. *Electronic Journal of Statistics*, 8:1088–1112. ISSN 1935-7524. doi: 10.1214/14-EJS924. URL <http://doi.org/10.1214/14-EJS924>. Citado na pág. 1, 2
- Campbell(1934)** J. T. Campbell. The poisson correlation function. *Proceedings of the Edinburgh Mathematical Society*, 4:18. ISSN 0013-0915,1464-3839. doi: 10.1017/S0013091500024135. Citado na pág. 3
- Christou and Fokianos(2014)** Vasiliki Christou and Konstantinos Fokianos. Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis*, 35:55–78. ISSN 0143-9782,1467-9892. doi: 10.1111/jtsa.12050. URL <http://doi.org/10.1111/jtsa.12050>. Citado na pág. 2
- Cox et al.(1981)** D. R. Cox, Gudmundur Gudmundsson, Georg Lindgren, Lennart Bondesson, Erik Harsaae, Petter Laake, Katarina Juselius and Steffen L. Lauritzen. Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics*, 8. doi: 10.2307/4615819. Citado na pág. 1
- Cui and Zhu(2017)** Yan Cui and Fukang Zhu. A new bivariate integer-valued garch model allowing for negative cross-correlation. *TEST*. doi: 10.1007/s11749-017-0552-4. Citado na pág. iii, iv, viii, ix, 3, 4, 7, 8, 12, 14, 15, 16, 22, 24, 28, 31, 32, 42, 48, 49, 50, 51, 52, 57, 58, 59, 60, 61, 78
- Davis and Liu(2012)** R. Davis and H. Liu. Theory and inference for a class of nonlinear models with application to time series of counts. *arXiv*. Citado na pág. 1, 2
- Doukhan et al.(2012)** P. Doukhan, K. Fokianos and D. Tjøstheim. On weak dependence conditions for poisson autoregressions. *Statistics and Probability Letters*, 82:0–948. ISSN 0167-7152. doi: 10.1016/j.spl.2012.01.015. URL <http://doi.org/10.1016/j.spl.2012.01.015>. Citado na pág. 1, 2
- Doukhan and Louhichi(1999)** Paul Doukhan and Sana Louhichi. A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84:313–342. ISSN 0304-4149. doi: 10.1016/s0304-4149(99)00055-1. Citado na pág. 2
- Ferland et al.(2006)** René Ferland, Alain Latour and Driss Oraichi. Integer-valued garch process. *Journal of Time Series Analysis*, 27. doi: 10.1111/j.1467-9892.2006.00496.x. Citado na pág. 1

- Fokianos and Fried(2012)** K. Fokianos and R. Fried. Interventions in log-linear poisson autoregression. *Statistical Modelling*, 12:299–322. ISSN 1471-082X,1477-0342. doi: 10.1177/1471082x1201200401. URL <http://doi.org/10.1177/1471082x1201200401>. Citado na pág. 1, 2
- Fokianos et al.(2009)** Konstantinos Fokianos, Anders Rahbek and Dag Tjøstheim. Poisson autoregression. *Journal of the American Statistical Association*, 104. doi: 10.1198/jasa.2009.tm08270. Citado na pág. 1, 2
- Franke(2010)** J. Franke. Weak dependence of functional ingarch process. Citado na pág. 1
- Gamerman and Lopes(2006)** Dani Gamerman and Hedibert Freitas Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Texts in Statistical Science. Chapman and Hall/CRC, 2 edição. ISBN 1584885874, 9781584885870. Citado na pág. 62, 67
- Geweke(1991)** John F. Geweke. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Staff Report 148, Federal Reserve Bank of Minneapolis. URL <https://ideas.repec.org/p/fip/fedmsr/148.html>. Citado na pág. 68
- Hastings(1970)** W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109. ISSN 00063444. URL <http://www.jstor.org/stable/2334940>. Citado na pág. 67
- Heinen and Rengifo(2007)** A. Heinen and E. Rengifo. Multivariate autoregressive modeling of time series count data using copulas. *Journal of Empirical Finance*, 14:0–583. ISSN 0927-5398. doi: 10.1016/j.jempfin.2006.07.004. Citado na pág. 3
- Hoel et al.(1972)** Paul G. Hoel, Sidney C. Port and Charles J. Stone. *Introduction to stochastic processes*. The Houghton Mifflin series in statistics. Houghton Mifflin. ISBN 9780395120767,0395120764. Citado na pág. 62, 63, 66
- Holgate(1964)** P. Holgate. Estimation for the bivariate poisson distribution. *Biometrika*, 51(1/2): 241–245. ISSN 00063444. URL <http://www.jstor.org/stable/2334210>. Citado na pág. 3, 5
- Konstantinos and Tjøstheim(2011)** F. Konstantinos and D. Tjøstheim. Log-linear poisson autoregression. *Journal of Multivariate Analysis*, 102:563–578. ISSN 0047-259X,1095-7243. doi: 10.1016/j.jmva.2010.11.002. URL <http://doi.org/10.1016/j.jmva.2010.11.002>. Citado na pág. 1, 2
- Lakshminarayana et al.(1999)** J. Lakshminarayana, S.N.N. Pandit and K. Srinivasa Rao. On a bivariate poisson distribution. *Communication in Statistics- Theory and Methods*, 28. doi: 10.1080/03610929908832297. Citado na pág. iii, iv, viii, 3, 4, 5, 6, 7, 10, 11, 12, 60
- Lee and Tjøstheim(2016)** Lee S. Lee, Y. and D. Tjøstheim. Asymptotic normality and parameter change test for bivariate poisson ingarch models. *TEST*. ISSN 1133-0686,1863-8260. doi: 10.1007/s11749-016-0510-6. URL <http://doi.org/10.1007/s11749-016-0510-6>. Citado na pág. 3
- Liu(2012)** H. Liu. Some models for time series of counts. *Dissertation*. URL <https://academiccommons.columbia.edu/doi/10.7916/D8CN7B2P/download>. Citado na pág. iii, 3, 29, 31
- Metropolis et al.(1953)** Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092. doi: 10.1063/1.1699114. URL <https://doi.org/10.1063/1.1699114>. Citado na pág. 67
- Neumann(2011)** Michael H. Neumann. Absolute regularity and ergodicity of poisson count processes. *Bernoulli*, 17(4):1268–1284. doi: 10.3150/10-BEJ313. Citado na pág. 1, 2
- Pandit et al.(1988)** S.N.N. Pandit, K. Srinivasa Rao and Y.V.S.S. Sarma. On a bivariate binominal distribution. *Assam Statistical Review*, 2. Citado na pág. 6

- Park and Jeong(2017)** Taeyoung Park and Seonghyun Jeong. Analysis of poisson varying-coefficient models with autoregression. *Statistics*, páginas 1–16. ISSN 0233-1888,1029-4910. doi: 10.1080/02331888.2017.1353514. URL <http://doi.org/10.1080/02331888.2017.1353514>. Citado na pág. 1, 2
- Plummer et al.(2006)** Martyn Plummer, Nicky Best, Kate Cowles and Karen Vines. Coda: Convergence diagnosis and output analysis for mcmc. *R News*, 6(1):7–11. URL <https://journal.r-project.org/archive/>. Citado na pág. 19
- Robert and Casella(2004)** Christian Robert and George Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 2nd edição. ISBN 0387212396,9780387212395. Citado na pág. 17, 62
- Rousseeuw and Molenberghs(1993)** Peter J. Rousseeuw and Geert Molenberghs. Transformation of non positive semidefinite correlation matrices. *Communication in Statistics- Theory and Methods*, 22. doi: 10.1080/03610928308831068. Citado na pág. 68
- Teicher(1954)** Henry Teicher. On the multivariate poisson distribution. *Scandinavian Actuarial Journal*, 1954:1–9. ISSN 0346-1238. doi: 10.1080/03461238.1954.10414190. Citado na pág. 3
- Tierney and Mira(1999)** Luke Tierney and Antonietta Mira. Some adaptive monte carlo methods for bayesian inference. *Statistics in Medicine*, 18(17-18):2507–2515. doi: 10.1002/(SICI)1097-0258(19990915/30)18:17/18<2507::AID-SIM272>3.0.CO;2-J. Citado na pág. 68
- Tjøstheim(2012)** D. Tjøstheim. Some recent theory for autoregressive count time series. *TEST*, 21:413–438. ISSN 1133-0686,1863-8260. doi: 10.1007/s11749-012-0296-0. URL <http://doi.org/10.1007/s11749-012-0296-0>. Citado na pág. 1, 2
- Wang et al.(2014)** Chao Wang, Heng Liu, Jian-Feng Yao, Richard A. Davis and Wai Keung Li. Self-excited threshold poisson autoregression. *Journal of the American Statistical Association*, 109:777–787. ISSN 0162-1459,1537-274X. doi: 10.1080/01621459.2013.872994. URL <http://doi.org/10.1080/01621459.2013.872994>. Citado na pág. 1, 2