AN ASSESSMENT OF COPULA-BASED REGRESSION MODELS FOR BIVARIATE COUNT DATA

By

Hilda Inés Calderón Cartagena

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

MATHEMATICAL SCIENCES (STATISTICS)

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS

2018

Approved by:

Dámaris Santana Morant, Ph.D Member, Graduate Committee	Date
Raúl E. Macchiavelli, Ph.D Member, Graduate Committee	Date
Pedro A. Torres Saavedra, Ph.D President, Graduate Committee	Date
Mercedes S. Ferrer Alameda, M.E. Representative of Graduate Studies	Date
Víctor Ocasio González, Ph.D Chairperson of the Department	Date

Abstract of Dissertation Presented to the Graduate School
of the University of Puerto Rico in Partial Fulfillment of the
Requirements for the Degree of Master of Science

AN ASSESSMENT OF COPULA-BASED REGRESSION MODELS FOR BIVARIATE COUNT DATA

By

Hilda Inés Calderón Cartagena

2018

Chair: Pedro A. Torres Saavedra
Major Department: Mathematical Sciences

It is known that analyzing correlated bivariate count data as independent in a regression context can lead to inefficient coefficients estimates. However, the number of parametric bivariate distributions that can be found in the literature to model bivariate counts are limited and not flexible enough to account for general correlation structures and different marginal distributions. Copula-based regression models provide a more flexible way of generating joint distributions for bivariate data by admitting different marginal distributions and various dependence structures. The purpose of this work was to evaluate the performance of copula-based regression models for bivariate counts under different scenarios, and to apply this approach to bivariate crash data in Puerto Rico highways. Scenarios with low, medium and high degrees of dependence were considered, as well as different sample sizes. In particular, the application of copulas when one of the marginal means was small was examined. Overall, if appropriate copulas are fitted, copula-based regression models provide more efficient estimators for the regression parameters when compared to modeling the counts independently, even when the data exhibits a degree

of association as low as a Kendall's $\tau = 0.3$, though we recommend a sample size of $N = 300$ or higher to assure an unbiased estimation of the copula parameter. The gain in efficiency increases with the degree of association. Also, traditional penalized likelihood-based criteria, such as AIC and BIC, seem to have a fairly good performance in selecting the best model among a set of candidate copula models. As a last note, interpretation of the copula parameter about the dependence structure is possible but should be made carefully since the range of its transformation to a dependence measure is narrower than $[-1, 1]$.

Resumen de Disertación Presentado a Escuela Graduada
de la Universidad de Puerto Rico como requisito parcial de los
Requerimientos para el grado de Maestría en Ciencias

EVALUACIÓN DE MODELOS DE REGRESIÓN USANDO CÓPULAS PARA DATOS DE CONTEOS BIVARIADOS

Por

Hilda Inés Calderón Cartagena

2018

Consejero: Pedro A. Torres Saavedra
Departamento: Ciencias Matemáticas

Es conocido que analizar datos de conteo bivariados correlacionados de manera independiente en un problema de regresión puede llevar a estimaciones de los coeficientes ineficientes. Sin embargo, las distribuciones bivariadas parámetricas que aparecen en la literatura para modelar conteos correlacionados tienen limitaciones y no son lo suficientemente flexibles como para admitir estructuras de correlación generales y distribuciones marginales diferentes. Los modelos basados en cópulas proveen una forma más flexible de generar distribuciones conjuntas para datos bivariados al admitir distribuciones marginales diferentes y varias estructuras de dependencia. El propósito de este trabajo fue hacer una evaluación del desempeño de los modelos de regresión basados en cópulas para datos de conteos bivariados bajo diferentes escenarios, y aplicar este método a datos de conteos de accidentes fatales y no fatales en autopistas de Puerto Rico. Escenarios con un bajo, moderado y alto grado de dependencia fueron considerados, así como diferentes tamaños de muestra. En particular, se examinó la aplicación de cópulas cuando una las medias marginales es pequeña. En general, si se ajustan cópulas apropiadas, los modelos

de regresión basados en cópulas proveen estimadores más eficientes para los coeficientes en comparación a ajustar modelos independientes a cada conteo, aún cuando los datos exhiben bajos grados de dependencia, Overall, if appropriate copulas are fitted, copula-based regression models provide more efficient estimators for the regression parameters when compared to modeling the counts independently, even when the data exhibits a degree of association as low as a Kendall's $\tau = 0.3$, aunque recomendamos un tamaño de muestra de $N = 300$ o superior para asegurar una estimación insesgada del parámetro de cópula. La ganancia en eficiencia aumenta con el grado de correlación. Además, los criterios tradicionales basados en verosimilitud, como AIC y BIC, parecen tener un buen desempeño en seleccionar el mejor modelo entre un conjunto de modelos de cópulas. Cabe señalar, finalmente, que la interpretación del parámetro de cópula sobre la estructura de dependencia es posible pero debe hacerse considerando que el intervalo de su transformación a una medida de dependencia es más estrecho que $[-1, 1]$.

To Mom, my biggest inspiration. Where would I be had it not been for your love, strength, and example?

ACKNOWLEDGMENTS

To Dr. Pedro A. Torres Saavedra for his support and encouragement as a mentor.

To Dr. Dámaris Santana Morrant and Dr. Raúl E. Macchiavelli, my committee members, for their collaboration.

To Dr. Erika Jaramillo Giraldo and Dr. Didier Valdés, for providing the data for the application in this work.

To the Department of Mathematical Sciences for providing the support to pursue this degree.

To the Center of Technologies of Information (CTI) for giving access to their servers.

To my mom and sister, for pushing me and for their support through this process.

To my friends and family for being there through all those tough times.

TABLE OF CONTENTS

LIST OF TABLES

Table page

LIST OF FIGURES

Figure page

LIST OF ABBREVIATIONS

pdf	probability density function.
pmf	probability mass function.
cdf	cumulative distribution function.
MCSE	Monte Carlo Standard Error.
IFM	Inference Function of Margins.
MSE	Mean Squared Error.
SE	Standard Error.
GLM	Generalized linear model.
MCMC	Markov Chain Monte Carlo.
AIC	Akaike Information Criteria.
BIC	Bayesian Information Criteria.
SD	Standard deviation.

LIST OF SYMBOLS

N	Sample size.
τ	Kendall's τ.
θ	Copula parameter.
$F(\cdot)$	Cumulative distribution function.
$f(\cdot)$	Probability density (mass) distribution.
$C(\cdot)$	Copula representation of the joint distribution.
l	log-likelihood.
μ	Mean.
ϕ	Dispersion parameter.
R	Statistical Software R.
ϕ	Dispersion parameter in Negative Binomial distribution.
$\|\| \cdot \|\|$	Euclidean norm.
y_i	Response variable i.
$\boldsymbol{\beta}$	Vector of regression parameters or coefficients.
$H(\cdot)$	Joint distribution.
$h(\cdot)$	Joint density (mass) distribution.
Φ_2	Bivariate normal distribution.

CHAPTER 1
INTRODUCTION

It is known that analyzing correlated bivariate count data as independent in a regression context can lead to inefficient coefficients estimates. Nevertheless, the number of parametric bivariate distributions that can be found in the literature to model bivariate counts are limited and not flexible enough to account for general correlation structures and different marginal distributions [30]. Copula-based models provide a more flexible way of generating joint distributions for bivariate data by admitting different marginal distributions and various dependence structures. Moreover, using copulas to model multivariate data allows to model the marginals and the dependence structure separately.

Although there exists theoretical knowledge about copulas for count data, surprisingly, there is a lack of comprehensive simulation studies on this subject. Therefore, there is not a clear understanding of the advantages and disadvantages of modeling correlated bivariate counts using copula-based regression models. In particular, the performance of copula-based regression models is unknown when the marginal distributions exhibit different means. Likewise, the gain of modeling bivariate counts with the proposed method has not been assessed for different correlation strengths (low, moderate and high).

This research builds on theory and other bivariate distributions for discrete variables to provide researchers an overview of the benefits and limitations of using copula-based regression models for bivariate count data. This work has been motivated by a real dataset based on the counts of fatal and non-fatal crashes in highways of Puerto Rico where the marginal distributions have very different means. We will focus on scenarios that have not been studied before.

This research motivating example refers to the number of fatal and non-fatal vehicle crashes. The data is for 144 segments of highways in Puerto Rico for a period of six years $(2004-2009)$ collected by the Highways and Transportation Authority of Puerto Rico as described in [14]. Detailed crash count data is obtained mostly for safety issues in different severity levels, e.g., Fatal, Injury, Property damage. Even though crash data by severity are innately multivariate, they are often analyzed separately without taking into account the dependence and shared variables that exist between severities. To approach crash frequencies as independent can lead to less precise and perhaps biased estimates [2]. In this context, it is worthwhile to consider bivariate count data copula-based models to improve the modeling and analysis of vehicle crashes data since copulas are a useful tool for multivariate modeling when the correlation among variables is difficult to incorporate into a joint distribution.

The purpose of this study is to evaluate the performance of copula-based regression models for bivariate counts. The specific objectives of this work are:

i) Examine the statistical properties of coefficients and standard errors estimates of copula-based regression models for bivariate count data,

ii) Determine the consequences of selecting the wrong copula on the statistical properties of the estimators,

iii) Assess the performance of copula-based regression models for bivariate count data with very different means for marginal distributions,

iv) Analyze the effect of sample size and correlation strengths on the performance of copula-based regression models for bivariate count data,

v) Evaluate the performance of likelihood-based model selection criteria, such as AIC and BIC, to recognize the true copula from a set of candidate copulas,

vi) Apply copula-based regression models to bivariate crash data in Puerto Rico highways.

The remaining of this thesis consists of five more chapters. In Chapter 2, a theoretical background on modeling bivariate count data is developed. Standard bivariate count models are introduced. This is followed by an overview on copula-based models as the proposed approach to jointly model multivariate count data. In Chapter 3, the performance of copula-based regression models for correlated counts is evaluated by several simulation studies considering different scenarios. The aim is to examine the statistical properties of coefficients and standard errors of copula-based regression models for count data in different scenarios considering sample size and magnitude of association. This leads to a guide of the scenarios where copula-based models are suited for modeling bivariate count data. Scenarios where the marginal means are different are included driven by our motivating example. The power of likelihood-based model selection criteria in distinguishing the true copula from a set of candidate copulas is also assessed by a power study in this chapter. In Chapter 4, a discussion of the results obtained from the simulations studies in Chapter 3 is presented. In Chapter 5, the copula-based regression models approach is applied to the number of fatal and non-fatal crashes in highways of Puerto Rico between 2004 and 2009. The data exhibits a marginal Kendall's τ association of 0.345. In the final chapter, Chapter 6, a conclusion of this thesis is made along with a discussion of future research topics related to the work presented here.

CHAPTER 2
LITERATURE REVIEW

Bivariate count data are two non-negative random variables that are often observed in different fields, e.g., the number of fatal and non-fatal crashes in a road segment, the number of points scored by each team in a game, purchases of two kinds of products by a customer. Consequently, both variables share the unit characteristics and are often correlated. Analyzing such data as independent, when they should be modeled jointly, can lead to coefficients estimates and standard errors that do not meet desired statistical properties. Failing to account for the association in multivariate data may produce inaccurate estimates of the regression parameters and their standard errors, resulting in inconsistent, inefficient, less precise and perhaps biased estimates [2, 6, 8, 31].

Joint modeling of multivariate data can improve the statistical properties of estimation and the ability to make inferences about the dependence between the variables. For bivariate data, a joint model can also allow probability statements about the conditional distribution of y_1 given y_2 [6].

A number of parametric bivariate distributions can be found in the literature to model bivariate counts, such as bivariate Poisson. However, it is common to encounter overdispersion in count data, i.e. the variance is larger than the mean. Therefore, assuming a bivariate Poisson is not adequate. Then, distributions like bivariate Negative Binomial and bivariate Poisson Inverse Gaussian, or models such as, bivariate Hurdle or Zero-inflated, are considered instead. Yet these attempts suffer from shortcomings. On one hand, these models do not allow for negative correlation between the counts. In many applications, the assumption of positive dependence might not be plausible. A bivariate Poisson lognormal distribution [3], for example, can handle both overdispersion and a more general dependence structure, including negative correlation, but there are several disadvantages to this approach. First, the bivariate Poisson lognormal model only applies when both margins are assume to have the same distribution. Second, the numerical integration can be time consuming. This distribution is commonly used to model crash data by severities, by using Bayesian methods [2, 10, 23, 24, 31]. A semiparametric bivariate model was developed in [16, 17] known as bivariate Poisson Laguerre polynomial model. This distribution has limitations in the range of possible correlations for counts with small means. In summary, there is a lack of parametric bivariate distributions for correlated count data with a flexible dependence structure and unequal marginal distributions.

In the literature, applications of copula-based regression models for discrete variables have gained popularity. For example, McHale and Scarf used copulas to generate a bivariate count regression model for the discrete pair shots-for and shots-against on 1048 soccer matches [25]. The data revealed negative correlation with a Kendall's $\tau = -0.191$. The authors considered Archimedian copulas and found that a model with Frank copula and negative binomial marginals was effective in capturing the negative dependence in the data, and chose it over the independent model.

Nikoloulopoulos and Karlis recommended the use of the covariates in the marginals and the copula parameter when the aim is to understand the dependence structure [30]. Illustrating this for count data, the authors presented an application in marketing, where they analyzed the number of purchases for food and non-food data based on customers' characteristics. The data used consisted of 2580 customers, and overdispersion was observed. They chose a model with Normal copula and negative binomial marginals over a group of other copula models. However, they did not include the independent model in the paper, and failed to demonstrate the gain of accounting for dependence using copulas.

Another practical use for copula models is when studying the distribution of the difference between two correlated counts. Although never used before for this mean, Cameron *et al.* used copula models to analyze the difference between reported number of doctor visits and actual number of doctor visits of 502 patients in order to determine sources of misreporting [5]. They found the joint distribution for the counts using the copula approach, and then derived the distribution of the difference. As a result, a Frank copula model with negative binomial marginals provided the best fit for the difference among the other bivariate but more restrictive models.

In a key textbook about copulas [33] the authors present a Monte Carlo experiment for a bivariate Poisson example. They simulated count data using a 500 sample size. The mean values were 2.01 and 1.35, replacing the mean values of the covariates. Both marginals had similar means and somewhat small, meaning that the identifiability issue encountered in copulas for discrete data could be troublesome. For this particular scenario, they found that the copula parameter was misestimated by 1%. Being this the only scenario considered, they failed to illustrate the consequences of having very different marginal distribution means as well as to consider various sample sizes and correlation strengths.

Surprisingly, little attention has been devoted to examining the statistical properties of copula-based estimators for count data under different scenarios. In addition, the consequences of not accounting for dependence in count data is not assessed by simulation studies. Another subject missing in the literature is the study of how likelihood-based model selection methods, such as AIC or BIC, work for copula-based regression models in the discrete case, in particular when weak dependence between two count variables is observed.

A variety of distribution models in the literature deal with bivariate distributions for counts. The most common model for correlated count data are the bivariate Poisson distribution and its generalizations. A less common approach is to derive joint distributions using copulas. In this chapter, an introduction to the most commonly used bivariate distributions for count data and copula-based models also for count data is presented. In particular, the focus will be on contrasting the properties of both approaches. The chapter begins with a brief discussion of Generalized Linear Models (GLM) in the univariate case. Section 2.2 focuses in GLMs for count data. A description of bivariate count data models and their properties is presented in Section 2.3. Sections 2.4 and 2.5 finalize this chapter with a discussion of bivariate copula-based models and their practical utility for count data.

2.1 Generalized Linear Models

Generalized linear models (GLMs) extend general linear models to admit response variables that have error distributions other than a Normal distribution. In these models the dependent variable y_i follows an exponential family distribution with mean μ_i. GLMs allow to model the mean μ_i as a function of $\boldsymbol{x}_i^t\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is a p-dimensional vector of unknown parameters and $\boldsymbol{x}_i$ is a p-dimensional row vector of covariates.

The response random variable y_i follows an exponential family distribution if its probability density function (pdf) or mass function (pmf) has the form

$$f(y_i;\theta_i,\phi) = \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{s(\phi)} + c(y_i,\phi)\right\}, \tag{2.1}$$

where, for $i = 1,\dots,n, \theta_i$ is known as the canonical parameter and ϕ as the dispersion parameter, and the functions $b(\cdot)$, $s(\cdot)$ and $c(\cdot)$ are known. There are some one-parameter distributions where ϕ is known, such as the Poisson or binomial ($\phi = 1$).

There are three components to any GLM:

1. Random component - the dependent (response) variable Y_i, for $i = 1, \ldots, n$, and its probability distribution which belong to the exponential family.
2. Systematic component - states the linear combination of the vector of covariates $\boldsymbol{x}_i$, creating the linear predictor $\eta_i = \boldsymbol{x}_i^t\boldsymbol{\beta}$.
3. Link function g - states the link between the random and the systematic component, i.e., establishes how the mean μ_i of the dependent variable is related to the linear predictor through $\eta_i = g(\mu_i) = \boldsymbol{x}_i^t\boldsymbol{\beta}$. The function g is monotonic and differentiable. Examples of link functions include the identity, log, reciprocal, logit and probit. When the canonical parameter $\theta = g(\mu_i)$, then $g(\mu_i)$ is the canonical link.

Estimation in GLM

Let $Y = (y_1, \ldots, y_n)^t$ be a vector of n independent random response variables, then the log-likelihood is

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n} \log f(y_i; \theta_i, \phi) = \sum_{i=1}^{n} \frac{y_i\theta_i - b(\theta_i)}{s(\phi)} + \sum_{i=1}^{n} c(y_i, \phi). \tag{2.2}$$

When fitting GLMs, where $\eta_i = g(\mu_i) = \boldsymbol{x}_i^t\boldsymbol{\beta}$ is modeled, it is necessary to find an estimation for the vector $\boldsymbol{\beta}$ and ϕ. The maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ and $\hat{\phi}$ are found by solving the likelihood equations derived from 2.2. These equations are usually nonlinear in $\boldsymbol{\beta}$, therefore, iterative methods such as Newton-Raphson or Fisher Scoring are needed to find the ML estimator $\hat{\boldsymbol{\beta}}$. For more details refer to [1].

2.2 Generalized Linear Models for Count Data

Most count models are based on the Poisson and negative binomial probability distributions. Both distributions belong to the exponential family, Poisson being a one-parameter distribution and negative binomial a two-parameter distribution. These distributions are introduced next.

Poisson Distribution

A random variable Y_i follows a Poisson distribution with parameter λ_i if its pmf can be expressed as

$$f(y_i;\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \quad y_i = 0, 1, \ldots. \tag{2.3}$$

The most frequently used link function g for the Poisson model is the log function. That is,

$$\log(\lambda_i) = \eta_i = \boldsymbol{x}_i^t\boldsymbol{\beta} \tag{2.4}$$

The Poisson distribution has the unique feature that the mean and variance are the same, i.e.,

$$E(Y_i) = Var(Y_i) = \lambda_i. \tag{2.5}$$

The variance is rarely equal to the mean when modeling real count data. When the variability of the data is greater than its mean this phenomenon is known as overdispersion. Overdispersion may occur due to different reasons. The problem with overdisperion is that it may cause incorrect estimates of the standard errors of coefficients. To account for overdispersion, an extension of the Poisson distribution is introduced next. For more details on what causes Poisson overdispersion and ways to handle it refer to [18].

Negative Binomial Distribution

The negative binomial distribution accounts for Poisson overdispersion by adding an extra parameter, the dispersion parameter ϕ. Y_i follows a negative binomial distribution with parameters μ_i and ϕ if its pmf can be expressed as

$$f(y_i;\mu_i,\phi) = \frac{\Gamma(y_i+\phi)}{\Gamma(\phi)\Gamma(y_i+1)}\left(\frac{\phi}{\mu_i+\phi}\right)^{\phi}\left(1-\frac{\phi}{\mu_i+\phi}\right)^{y_i}, \quad y_i = 0, 1, \ldots. \tag{2.6}$$

For this distribution

$$E(Y_i) = \mu_i, \quad Var(Y_i) = \mu_i + \frac{\mu_i^2}{\phi}.$$

For ϕ positive, $Var(Y_i) > E(Y_i)$. As for the Poisson model, the log link is the most frequently used in negative binomial regression models.

2.3 Bivariate Count Data Models

The first bivariate count model to arise in the literature is the bivariate Poisson model [22]. This distribution considers random variables z_{i1}, z_{i2}, z_3 to be Poisson and independently distributed with parameters $\lambda_{i1}, \lambda_{i2}, \lambda_3 > 0$, for $i = 1, 2, \ldots, n$, respectively. The resulting random variables $y_{i1} = z_{i1} + z_3$ and $y_{i2} = z_{i2} + z_3$ are then jointly distributed as bivariate Poisson. Other methods to obtain the bivariate Poisson distribution are described in [22]. The marginal distributions are Poisson, restricting the mean and variance to be the same for each response variable. The main drawback of this model is that it does not allow for overdispersion and can only describe non-negative correlation.

The bivariate negative binomial model is a generalization of the bivariate Poisson model, allowing for overdispersion using an additional parameter v_i called the unobserved heterogeneity component. As described in [8], the count variables y_{i1} and y_{i2} are conditionally independently distributed as

$$y_{ij}|\boldsymbol{x}_{ij}, v_i \sim Poisson(\lambda_{ij} v_i) \tag{2.7}$$

where

$$\lambda_{ij} = \exp(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}_j), \tag{2.8}$$

v_i is gamma distributed with mean $= 1$ and variance $= \frac{1}{a}$, and $\boldsymbol{x}_{ij}$ is the vector of covariates, for $j = 1, 2$ and $i = 1, \ldots, n$. The marginal distributions are negative binomial and v_i is restricted to be the same for both counts. Although this model does account for overdispersion, it is still limited in admitting different dependence structures because it only accommodates positive correlation and does not include independence [22].

Other distributions have been developed by compounding the bivariate Poisson distribution to obtain the desired flexibility to model correlation between counts, such as the Generalized Poisson and Poisson Inverse Gaussian. These distributions still do not allow for negative correlation [8].

The bivariate Poisson lognormal distribution is another example of a compound bivariate Poisson distribution [3], where the count variables y_{i1} and y_{i2} are distributed as

$$y_{ij}|\lambda_{ij}, v_{ij} \sim Poisson(\lambda_{ij} v_{ij}),$$
$$\lambda_{ij} = \exp(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}_j),$$

with

$$v_{ij} = \exp(b_{ij})$$

and

$$\boldsymbol{b}_i \sim N_j(0, \Sigma),$$

where Σ is the covariance matrix for $j = 1, 2$ and $i = 1, \ldots, n$. To account for correlation among the y_{ij}'s, the vector $\boldsymbol{v}_i = (v_{i1}, v_{i2})^t$ is distributed as bivariate lognormal with mean $\mu = 0$. As shown in [7], this model has an unrestricted correlation structure allowing for positive or negative correlation and overdispersion. However, this model still has some shortcomings. For instance, the joint probability density function of the bivariate Poisson lognormal model has no closed-form, therefore, numerical integration methods are required to obtain parameters estimation, such as the MCMC estimation described in [7]. Furthermore, the marginal distributions involve a complicated infinite series. With computational advances, this is probably the most used parametric multivariate distribution for counts that can be found in the literature nowadays using a Bayesian approach [2, 10, 23, 24, 31].

A semiparametric bivariate model based on an approximation by Laguerre polynomial expansion of the distribution of $\boldsymbol{v}_i$ in the bivariate Poisson lognormal was developed in [16, 17]. Unlike the Poisson lognormal bivariate model, this model has a closed form that allows to estimate the parameters by maximum likelihood. It is also suitable to model overdispersion and has an unrestricted correlation structure allowing for independence and positive or negative correlations. Nevertheless, this semiparametric bivariate model for count data has limitations for real applications due to the narrow range of possible correlations for counts with small means [17]. For more details on this and other bivariate Poisson mixed distributions and their properties, the reader can be refered to [8, 21, 22].

Our motivating example on fatal and non-fatal crashes introduced in Chapter 1 is an uncommon application of bivariate count data in that one of the response variables has a small mean and the other one has a large mean, and both counts exhibit overdispersion. In an attempt on obtaining more accurate estimates by accounting for correlations and overdispersion in crash counts by severity or type of collision, the Poisson lognormal distrubution was recently introduced [2, 10, 23, 24, 31]. Then again, a flexible bivariate model that admits overdispersion, an unrestricted correlation structure and different marginal distributions with very different means is desired. While the aforementioned models are an intent to attend these phenomena, none of them succeed in dealing with all of them simultaneously. Moreover, the interpretation of the parameters in these models are specific to each subject. With this in mind, we came across bivariate copula models for count data discussed in the next section.

2.4 Copula-based Regression Models

A copula is a multivariate probability distribution for which the marginal distributions follow a Uniform distribution on (0,1). In probability and statistics, copulas are useful because they offer a straightforward and flexible way of generating joint distributions for multivariate data by admitting different marginal distributions and various dependence structures between two or more random variables. Furthermore, sometimes the interest lies in understanding the dependence structure, and copula-based models allow to model the copula parameter using covariates as described in [30]. Additionally, copula-based models make data generation feasible for simulation purposes and provide marginal interpretation of the parameters. Copula theory is based on the Sklar's theorem [32] that describes functions that join two or more probability functions to form a multivariate pdf (see Theorem 1).

Theorem 1. *If H is a d-variate distribution with jth univariate marginal distributions F_j given a vector of covariates $\boldsymbol{x}_j$, then there is a function $C : [0,1]^d \to [0,1]$ so that*

$$H(\boldsymbol{y}; \boldsymbol{x}) = C(F_1(y_1; \boldsymbol{x}_1), \ldots, F_d(y_d; \boldsymbol{x}_d); \theta), \quad \boldsymbol{y} \in \mathbb{R}^d, \tag{2.9}$$

where θ is the copula parameter which measures the association between the marginals.

If F is continuous, then the copula C is unique, and is given by

$$C(\boldsymbol{u}) = H(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \quad \boldsymbol{u} \in [0,1]^d, \tag{2.10}$$

where $F_1^{-1}, \ldots, F_d^{-1}$ are the inverse quantile functions of the univariate marginals.

If F is discrete, then the copula C is unique only in the set $Range(F_1) \times \cdots \times Range(F_d)$.

A formal definition of copula is given in [20] as follows:

Definition 2.4.1. *A copula is a multivariate distribution with marginal distributions Uniform(0,1). That is, if C is a copula then it is the multivariate distribution* *of a vector of two or more dependent variables that follow a marginal Uniform distribution on (0,1). A d-dimensional copula has the following properties:*

(a) $C(u_1, \ldots, u_{i-1}, 0, u_{u+1}, \ldots, u_d) = 0$

(b) $C(1, \ldots, 1, u, 1, \ldots, 1) = u$

(c) C is d-increasing

From this definition one can see that the univariate approach assuming independence is the particular case of a copula representation where $C(\boldsymbol{u}) = u_1 \cdot u_2 \cdot \cdots \cdot u_d$. Copula families are given as cumulative density distributions. Then, for $H(\boldsymbol{y}; \boldsymbol{x}) = C(F_1(y_1; \boldsymbol{x}_1), \ldots, F_d(y_d; \boldsymbol{x}_d); \theta)$ with continuous univariate marginals $F_1, \ldots, F_d$ and densities $f_1, \ldots, f_d$, the density function is given by

$$h(\boldsymbol{y}; \boldsymbol{x}) = c(F_1(y_1; \boldsymbol{x}_1), \ldots, F_d(y_d; \boldsymbol{x}_d); \theta) \times \prod_{j=1}^{d} f_j(y_j; \boldsymbol{x}_j), \quad \boldsymbol{y} \in \mathbb{R}^d, \tag{2.11}$$

where $\boldsymbol{x}_j$ is the vector of covariates, $c(\cdot)$ is the density of the copula obtained by taking the mixed partial derivative of d-order, and θ is the copula parameter.

In the discrete case, probability distributions are not strictly increasing, therefore rectangular probabilities need to be obtained. For example, for the bivariate case of counts the pmf is given by

$$\begin{aligned} h(y_1, y_2; \boldsymbol{x}) &= P(Y_1 = y_1, Y_2 = y_2; \boldsymbol{x}) \\ &= P(y_1 - 1 < Y_1 \leq y_1, y_2 - 1 < Y_2 \leq y_2; \boldsymbol{x}) \\ &= H(y_1, y_2; \boldsymbol{x}) - H(y_1 - 1, y_2; \boldsymbol{x}) - H(y_1, y_2 - 1; \boldsymbol{x}) + H(y_1 - 1, y_2 - 1; \boldsymbol{x}) \\ &= C(F_1(y_1; \boldsymbol{x}_1), F_2(y_2; \boldsymbol{x}_2); \theta) - C(F_1(y_1 - 1; \boldsymbol{x}_1), F_2(y_2; \boldsymbol{x}_2); \theta) \\ &\quad - C(F_1(y_1; \boldsymbol{x}_1), F_2(y_2 - 1; \boldsymbol{x}_2); \theta) + C(F_1(y_1 - 1; \boldsymbol{x}_1), F_2(y_2 - 1; \boldsymbol{x}_2); \theta) \end{aligned} \tag{2.12}$$

A multivariate probability distribution can be represented as the composition of a copula and the univariate marginals of the multivariate distribution as shown in Eq. 2.11. In light of this, another useful feature of copulas is that the modeling of the univariate marginal distributions can be separated from the dependency structure, which is then modeled by the copula. Since often there is more information about the marginals rather than the joint distribution, this is probably the most interesting feature of copula models. To illustrate this, consider the Clayton copula, which will be discussed later, and marginal distributions F_1 and F_2. The Clayton copula has the form $(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$. Taking mixed partial derivatives it follows that the joint density of the random variables Y_1 and Y_2 is given by

$$\begin{aligned} h(y_1, y_2; \boldsymbol{x}) &= \frac{\partial^2}{\partial F_2 \partial F_1}(F_1(y_1)^{-\theta} + F_2(y_2)^{-\theta} - 1)^{-\frac{1}{\theta}} \\ &= (1+\theta)(F_1(y_1)^{-\theta} + F_2(y_2)^{-\theta} - 1)^{-\frac{1}{\theta}-2} \cdot \frac{dF_1(y_1)}{dy_1} \cdot \frac{dF_2(y_2)}{dy_2} \\ &= (1+\theta)(F_1(y_1)^{-\theta} + F_2(y_2)^{-\theta} - 1)^{-\frac{1}{\theta}-2} \cdot f_1(y_1) \cdot f_2(y_2) \end{aligned}$$

There are many different copula families and they vary in properties such as symmetry, tail dependence, and dependence structure. For example, some admit both positive and negative dependence, others allow stronger association on one of the tails (i.e. tail dependence). Examples of scatterplots and contour plots for some copula families are shown in Figures 2–1 and 2–2, respectively. These examples demonstrate the flexibility of a copula in that it can handle different dependence structures. For instance, both the Normal and Frank copula families have symmetrical dependence structures, but the Normal copula allows for stronger dependence in the tails, while Clayton and Galambos are copula families with asymmetrical dependence structures. The focus of this research is on these specific copulas. These copulas can be broken into three types: Archimedean copulas, extreme-value copulas, and elliptical copulas which will be discussed in detail next.

2.4.2 Parametric copula families

Many parametric copula families have been described in the literature [20, 27, 33]. Each complies with different desired properties when developing multivariate

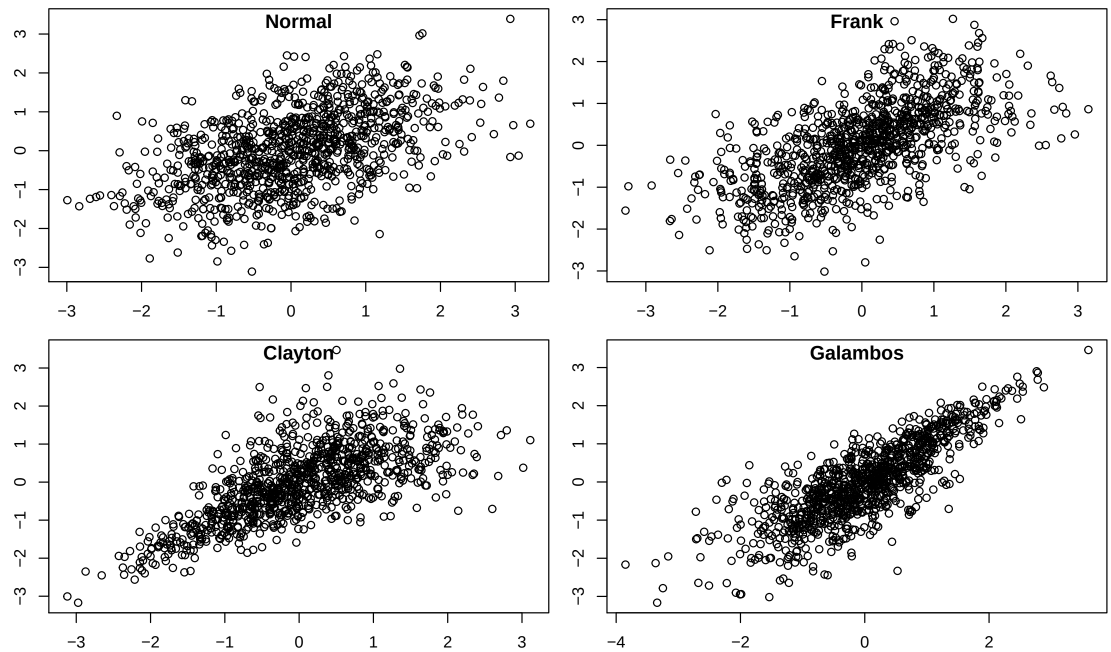

Figure 2–1: Scatterplots of simulated data for some copulas combined with N(0, 1) marginal distributions; Kendall's $\tau = 0.5$.

distributions through copulas. One desirable property is a wide range, admitting both positive and negative dependence, as well as tail dependence (i.e., measure of the degree of dependence on one or both of the extremes of the joint distribution). Another wanted property is a distribution function that makes likelihood estimation computationally feasible. In Table 2–1, the four different one-parameter copula families that will be used in this research are listed: Frank, Clayton, Normal and Galambos. In Figures 2–1 and 2–2, these copulas are presented graphically. This selection was based on their frequent application to model bivariate count data and for the different shapes and tail dependence they exhibit. In this section these parametric copulas and their properties are introduced. For more details on these and other copula families, we refer the reader to [20, 27, 29].

Normal Copulas

The Normal copula family belongs to the class of elliptical copulas and is commonly used for count data [5, 30]. It inherits the dependence structure of the multivariate Normal distribution, admitting both negative and positive dependence. This

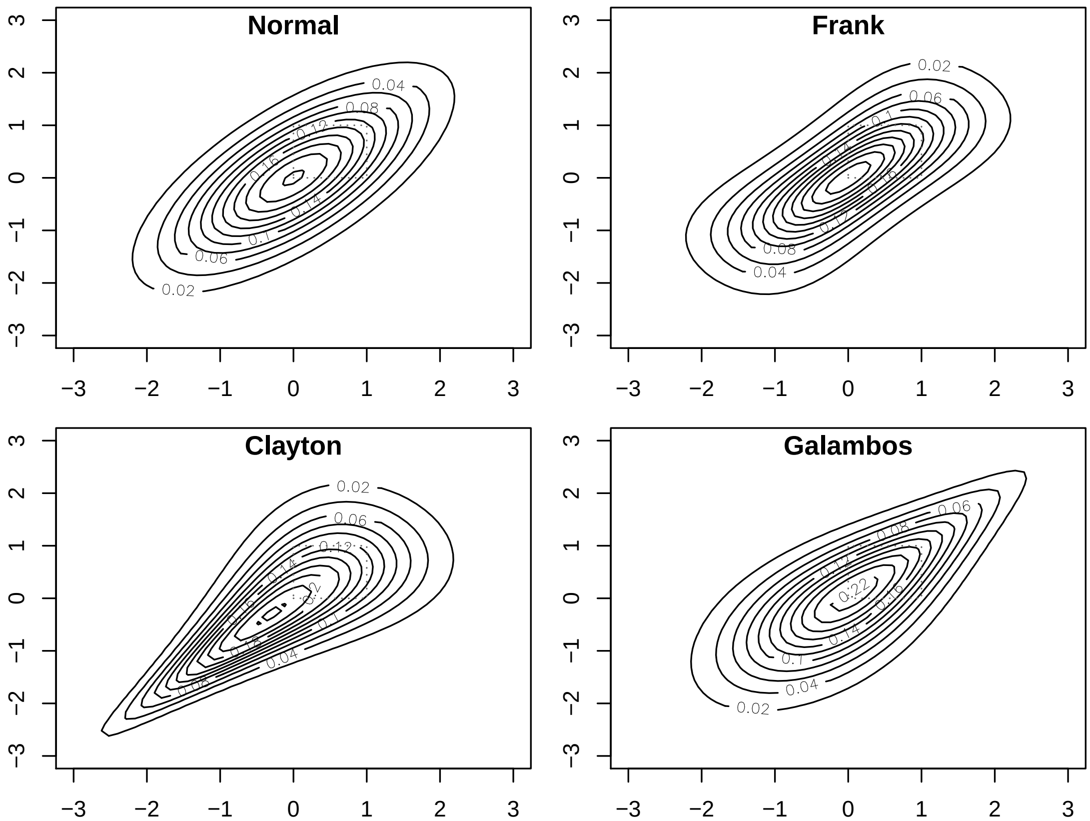

Figure 2–2: Contour plots of the density for some copulas combined with N(0, 1) marginal distributions; Kendall's $\tau = 0.5$.

copula family is probably the most computationally demanding because it lacks a closed form and rectangular probabilities need to be computed for maximum likelihood estimation. Nevertheless, the computation of these probabilities are already implemented in R in the package `mvtnorm` [13]. This copula family exhibits reflection symmetry.

Frank Copulas

The Frank copula family is part of the Archimedean class of copulas and has a closed form cumulative distribution function (cdf). It is the only Archimedean copula with reflection symmetry. It is frequently used in modeling bivariate data [5, 25, 30] because it allows flexible dependence, admitting both positive and negative dependence and independence. Though, for three or more dependent variables, the negative dependence is limited [20].

Table 2–1: Parametric copula families, their copula parameter range and the copula parameter link function $s(\theta)$. Note: Φ_2 is the bivariate standard Normal cdf with correlation θ; Φ^{-1} is the quantile function of $N(0,1)$.

Copula family	$C(u,v;\theta)$	$\theta \in$	$s(\theta)$
Clayton	$(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$	$(0,\infty)$	$\log\theta$
Frank	$-\frac{1}{\theta}\log\left[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right]$	$(-\infty,\infty)$	θ
Galambos	$uve^{[(1-u)^{-\theta}+(1-v)^{-\theta}]^{-\frac{1}{\theta}}}$	$[0,\infty)$	$\log\theta$
Normal	$\Phi_2(\Phi^{-1}(u),\Phi^{-1}(v))$	$(-1,1)$	$\log\frac{1+\theta}{1-\theta}$

Clayton Copulas

The Clayton copula family is also part of the Archimedean class of copulas. This copula family admits only positive dependence and has a closed form cdf making maximum likelihood estimation rather easy. It is specially useful when bivariate data shows lower tail dependence [5, 25, 30].

Galambos Copulas

The Galambos copula family belongs to the extreme-value class of copulas. It is the lower extreme value limit of an Archimedean copula. Extreme-value copulas are useful when modeling the dependence structure between rare events [15]. Another advantage is that they are not symmetric. The amount of dependence for the Galambos copula family ranges from independence to complete dependence and it shows upper tail dependence. Despite its properties, the Galambos copula family is the least used to model bivariate count data from our list [30].

2.4.3 Inference in copula-based regression models

Given $y_{1j},\ldots,y_{nj}$ independent and covariates $\boldsymbol{x}_{1j},\ldots,\boldsymbol{x}_{nj}$ for $j=1,2$, the log-likelihood for bivariate copula-based models is defined as

$$l=\sum_{i=1}^{n}\log h(y_{i1},y_{i2};\boldsymbol{x}_{i1},\boldsymbol{x}_{i2}), \tag{2.13}$$

where h is the pdf defined in Eq. 2.11 for the continuous case or the pmf defined in Eq. 2.12 for the bivariate discrete case.

After choosing appropriate univariate distributions, and a selection of copulas that could be a good fit for their properties, the optimization of the log-likelihood in Eq. 2.13 can be made using general optimization methods, such as quasi-Newton or Nelder-Mead. Therefore, maximum likelihood estimation can be used to obtain estimates of the parameters for the univariate marginal distributions and the copula parameter. Model comparisons between the independent, and the copula-based models can be made with the log likelihood ratio statistic. To choose the best copula model likelihood based criteria such as AIC and BIC can be used, which becomes a comparison of the log-likelihood values across different models if all of the competing models have the same covariates and a single copula parameter. In this respect, a study of how this likelihood-based model selection methods behave for copula-based regression models in the discrete case is not found in the literature. In addition, prior research [11] states that the conditions under which the estimation of the copula parameter can be found are not established yet for the discrete case. This research attempts to contribute on these two unattended aspects by a simulation study.

From Eq. 2.11, the estimation of the copula parameters is said to be separated from the estimation of univariate marginals parameters [33]. Thus, a two steps method for likelihood estimation as described in [20] can be used. The steps are as follows:

1. Find appropriate models for the marginal distributions and get estimates of the parameters by maximum likelihood.
2. Fix the parameters of the marginals and maximize Eq. 2.13 to estimate the copula parameters.

This approach is known as Inference Function of Margins (IFM) or two-stage estimation. When the model is correctly specified, this approach produces consistent parameter estimates with the traditional maximum likelihood, although less efficiently [33].

2.4.4 Dependence measured by Kendall's τ

Copula parameters for distinct families have different range making them incomparable. Therefore, it is convenient to convert the copula parameters to a concordance measure such as Kendall's τ.

For continuous data, Kendall's τ is defined as the difference between the probability of concordance and discordance, and takes the values in $[-1, 1]$. However, for the discrete case the probability of ties is positive and τ does not reach the values ± 1. Various generalization of the Kendall's τ for discrete data have been studied and derived [4, 9, 26, 28]. For bivariate count data the population version of Kendall's τ is given in [30] as

$$\begin{aligned}\tau(Y_1, Y_2) = &\sum_{y_1=0}^{\infty}\sum_{y_2=0}^{\infty} h(y_1, y_2; \boldsymbol{x}_1, \boldsymbol{x}_2)\{4C(F_1(y_1-1; \boldsymbol{x}_1), F_2(y_2-1; \boldsymbol{x}_2); \theta) \\ &- h(y_1, y_2; \boldsymbol{x}_1, \boldsymbol{x}_2)\} + \sum_{y_1=0}^{\infty} f_1^2(y_1; \boldsymbol{x}_1) + \sum_{y_2=0}^{\infty} f_2^2(y_2; \boldsymbol{x}_2) - 1,\end{aligned} \tag{2.14}$$

where Y_1, Y_2 are discrete random variables with joint pmf h as given in Eq. 2.12; F_1, F_2 are the marginal cdfs; f_1, f_2 are the pmfs; and C is the copula with copula parameter θ.

It is clear from Eq. 2.14 that the marginals have an effect on the Kendall's τ for the discrete case. To illustrate this effect, in Figure 2–3 Kendall's τ values as defined in Eq. 2.14 are plotted using Poisson marginal distributions with the same parameter λ, where λ is the mean and variance of a Poisson distribution, and varying the copula parameter θ for the four copula families listed in Table 2–1. It can be seen that the range of Kendall's τ is affected by the values of λ. Indeed, the range is narrower for small values of λ and almost reach ± 1 as λ goes to infinity. In order to better understand the behavior of Kendall's τ for discrete data and given our motivational example, in Figure 2–4 Kendall's τ values are also plotted but using Poisson marginal distributions with different λ values. The surface plots suggest that if at least one of the λ values is small then the Kendall's τ range is narrower than ± 1. Note that when $\lambda_1 = \lambda_2$ the values of Kendall's τ achieve its highest value for high values of θ according to the range of each copula family. In this context, it is necessary to account for true range when evaluating the dependence strength between the dependent variables using any concordance measure such as Kendall's τ.

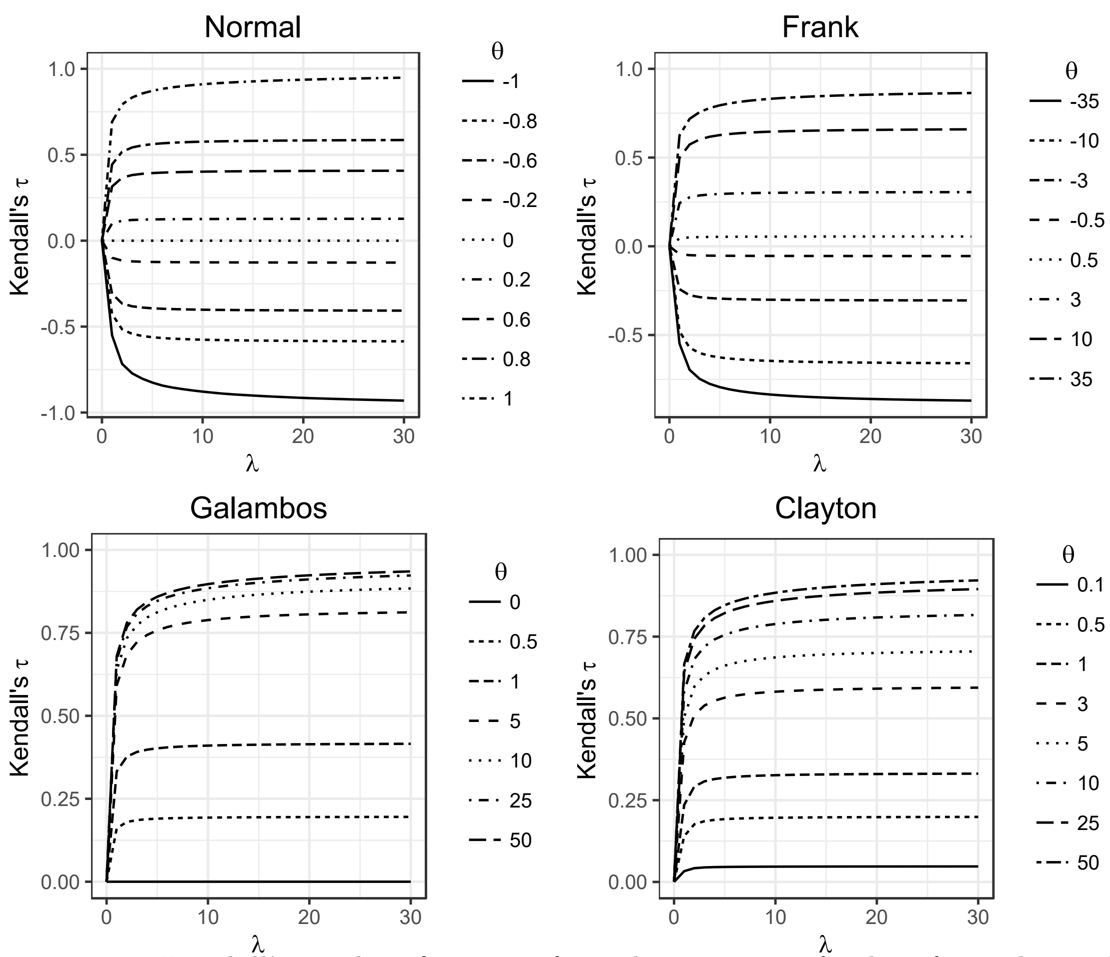

Figure 2–3: Kendall's τ values for a set of copula parameter θ values for each copula using Poisson marginal distributions with the same parameter $\lambda \in [1, 30]$.

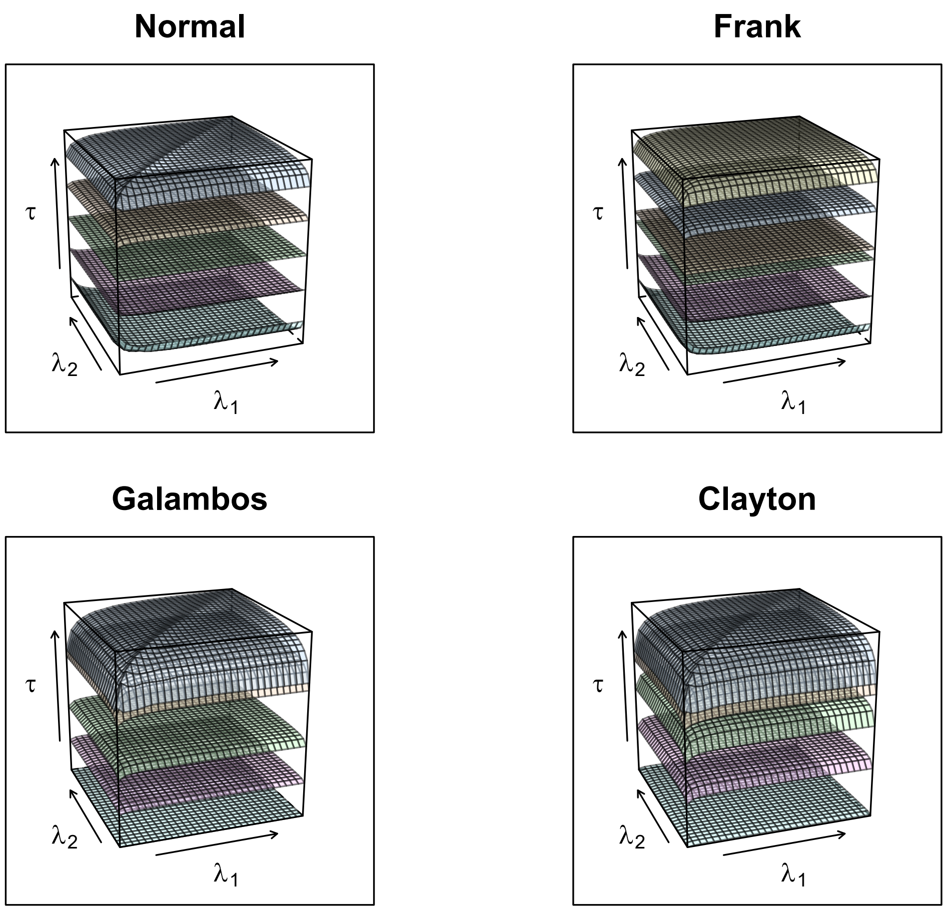

Figure 2–4: Kendall's τ values for a set of copula parameter θ values for each copula using Poisson marginal distributions with different parameter $\lambda_i \in [1, 30], i = 1, 2$.

2.5 Copula-based Models for Count Data

Modeling count data through copulas is still legitimate despite the issues mentioned in this chapter, such as the lack of uniqueness for C in Theorem 1 for the discrete case and the dependence of concordance measures like Kendall's τ on the marginal distributions. A deeper understanding of this approach's advantages and limitations are brought together in [11]. Briefly, some of the main conclusions follow.

First, the identifiability issue of Eq. 2.9 in the discrete case could be more disadvantageous when at least one of the variables has a small mean. This is because a small mean implies a smaller range of values for the response variable. Nevertheless, this should not be a problem if the range for the bivariate copula distribution is restricted to the $Range(Y_1) \times Range(Y_2)$. Secondly, as mentioned in Section 2.4.4, the range of dependence is now a function of the copula parameter and the marginal distributions, and does not necessarily attain the bounds ± 1. Therefore, the true range of any concordance measure needs to be considered when analyzing the degree of dependence between the variables. Also, in the continuous case, two random variables are independent if and only if $C(u,v) = uv$, where $u, v \in (0,1)$. But, although $C(u,v) = uv$ implies independence in the discrete case, the converse is not necessarily true. Additionally, the authors establish that the conditions in which the parameter θ is estimable by maximum likelihood estimation in the discrete case are unknown at the time. In this respect, we aim to contribute by doing a simulation study for count data for different scenarios.

As stated earlier, copula modeling for the discrete case is still valid since the most important dependence properties of the copula are passed to the multivariate distribution with discrete marginals, and the copula parameter θ can still be used to describe the dependence structure when converted to some concordance measure such as Kendall's τ if its range is accounted for. Moreover, not only this approach is less computationally demanding and more flexible than other bivariate count models found in the literature by admitting different marginals and a variety of dependence structures, but also copula-based models prove to be useful in special cases. For instance, there is often more information on the marginals rather than the bivariate distribution. Some examples that illustrate the use of copula models for multivariate count data can be found in the literature, such as [5, 25, 30]. These are related to marketing, econometrics, and sports statistics, respectively. The application presented in this research will add to these by providing an application under a scenario where one of the marginal distributions has a small mean and the other one has a large mean.

2.6 Regression to the copula parameter

The authors in [30] proposed introducing a regression part for the copula parameter θ. This is useful when the interest lies in studying the dependence structure and how any covariate affects the association between the counts.

Let $\boldsymbol{b}$ be the parameter vector for the copula parameter θ. Then, using the correct function $s(\cdot)$, the regression part for θ is given by $s(\theta_i) = \boldsymbol{b}^t \boldsymbol{x}_i$, for $i = 1, \ldots, n$.

CHAPTER 3
METHODOLOGY

Our motivating example, as described in Chapter 1, is related to the number of fatal and non-fatal vehicle crashes in different highway segments. Both counts share the characteristics of the segment, and therefore some association among the number of fatal and non-fatal crashes is expected. Since modeling bivariate count data independently can lead to coefficient's estimates and standard errors that do not meet desired statistical properties, such as unbiasedness, consistency and efficiency, we propose modeling the dependence among two counts using copula-based regression models. Moreover, modeling of bivariate count data allows the analyst to quantify and explain the dependence between counts. However, caution should be exercised given the limitations mentioned in Section 2.5. Interest in this approach arises from its flexibility and ease of implementation. For example, this approach admits different marginal distributions and the variety of dependence structures it offers is quite large by admitting both negative and positive correlations, and tail dependence. Moreover, there is often more information about the marginal distributions of two correlated variables than about their joint distribution. With the copula approach the joint distribution is derived given the marginals. Additionally, the coefficient's estimates are marginally interpretable which is more straightforward than other existing multivariate distributions.

In this chapter a description of the implementation of copula-based regression models for bivariate count data is presented in Section 3.1. In Section 3.2 the Mean Squared Error (MSE) and likelihood-based selection criteria are introduced as the tools used to evaluate the performance of the set of candidate models and to select the one with the best fit. A set of simulation studies is included in Section 3.3 to assess the statistical properties of the parameter estimates obtained from fitting copula models to account for dependence in comparison to modeling the counts independently. A discussion on the performance of likelihood-based criteria, such as AIC or BIC, in selecting the true copula from a set of candidate copulas is also covered in these studies.

3.1 Copula-based Regression Model for Bivariate Count Data

Consider two correlated non-negative random variables Y_1 and Y_2 with univariate cumulative distribution functions (cdf's), F_1 and F_2, respectively, and one covariate x. For $j = 1, 2$ and $i = 1, \dots, n$, let Y_{ij} be independently distributed as $F_j(\cdot)$. For instance, the cdf's can be chosen to be negative binomial $F_j(\mu_{ij}, \phi_j)$, where

$$\mu_{ij} = \exp(\boldsymbol{x}_i^t \boldsymbol{\beta}_j),$$

$\boldsymbol{x}_i = (1, x)^t$, $\boldsymbol{\beta}_j = (a_j, b_j)^t$ and ϕ_j is the dispersion parameter (see Section 2.2).

To obtain their joint distribution the cdf's are linked by a copula. Provided that both cdf's are a good fit, the joint distribution given by the copula representation is

$$C(F_1(y_1; x), F_2(y_2; x); \theta), \tag{3.1}$$

where θ is the copula parameter and x is a covariate assumed, but not restricted, to be the same for both counts. Then, bivariate probabilities of the form $P(Y_1 = y_1, Y_2 = y_2)$ can be obtained from the pmf defined in Eq. 2.12.

As described in Section 2.4, the application of copulas to generate the joint distribution for bivariate count data requires that the marginal distributions are selected and estimated first. For this step several parametric models are considered and evaluated using goodness of fit tests. In Section 2.2 we described two of the most used distributions to model univariate count data: Poisson and negative binomial. The second step is to specify the copula. A set of appropriate copulas is selected for their dependence structure given prior information about the data. That is to say that previous exploration of the data is needed in order to select a good set of candidate copulas in this step. The copula families considered in this research are often used in the literature and exhibit different dependence structures and properties, namely, no tail dependence, lower and upper tail dependence, and negative dependence. These copulas include the Clayton, Frank, Galambos and Normal copula which were described briefly in Section 2.4.2. Then, the joint distribution parameters are estimated via full maximum likelihood or the IFM approach (see Section 2.4.3), and the copula with the best fit is chosen using likelihood-based criteria, such as AIC or BIC.

The log-likelihood to be maximized to obtain the parameter estimates via full maximum likelihood for a copula-based regression model for count data with one covariate assuming negative binomial marginals is given as

$$
\begin{aligned}
l &= \sum_{i=1}^{n} \log h(y_{i1}, y_{i2}; x_i, \boldsymbol{\beta}_1, \phi_1, \boldsymbol{\beta}_2, \phi_2) \\
&= \sum_{i=1}^{n} \log \Big[C(F_1(y_{i1}; x_i, \boldsymbol{\beta}_1, \phi_1), F_2(y_{i2}; x_i, \boldsymbol{\beta}_2, \phi_2); \theta) \\
&\quad - C(F_1(y_{i1} - 1; x_i, \boldsymbol{\beta}_1, \phi_1), F_2(y_{i2}; x_i, \boldsymbol{\beta}_2, \phi_2); \theta) \\
&\quad - C(F_1(y_{i1}; x_i, \boldsymbol{\beta}_1, \phi_1), F_2(y_{i2} - 1; x_i, \boldsymbol{\beta}_2, \phi_2); \theta) \\
&\quad + C(F_1(y_{i1} - 1; x_i, \boldsymbol{\beta}_1, \phi_1), F_2(y_{i2} - 1; x_i, \boldsymbol{\beta}_2, \phi_2); \theta) \Big],
\end{aligned}
\tag{3.2}
$$

where θ is the copula parameter. The implementation of the copula models included in this research was done in the statistical software R. There are various packages in R dedicated to copulas. For example, the package copula [19] includes methods for density, distribution, random generation and fitting of common copula families, but the available functions for fitting copulas do not allow to include covariates and were not useful in our example. However, the estimation method for the copula families included in this work was relatively easy to implement (See Appendix C). The log-likelihood in Eq. 3.2 for each bivariate copula model for count data with negative binomial marginals was programmed in R and maximized using the optim function. In this function the default method is an implementation of Nelder and Mead. Although this method is relatively slow, it works better than the quasi-Newton method "BFGS" in our scenarios due to its robustness. To fit the negative binomial independently, the glm.nb function from the MASS package was used. This function is a modification of the glm function as it includes the estimation of the dispersion parameter ϕ. Also, the pmvnorm function from the mvtnorm package was used in the implementation of the Normal copula.

To analyze the performance of copula-based models approach in different scenarios, this research includes simulation studies which will be discussed later. In particular, the aim is to provide a guide of when copulas could be considered to account for correlation between two counts.

3.2 Performance Measures and Selection Criteria

After estimating the parameters in a copula-based regression model with several copula families, one needs to select the copula model with the best performance. For the analyses here, the model that provides the best fit to the data among the set of candidate models was chosen based on likelihood-based criteria, which measure the loss of information of the fitted model. The Mean Squared Error (MSE) was also calculated to characterize the performance of the estimator of the vector of mean and dispersion parameters for the marginal distributions, separately.

Let $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2)^t$ be the estimator of the vector of mean parameters for the marginals F_1 and F_2, where $\hat{\boldsymbol{\beta}}_j = (\hat{a}_j, \hat{b}_j)^t$, $j = 1, 2$. Also, let $\hat{\boldsymbol{\phi}} = (\hat{\phi}_1, \hat{\phi}_2)^t$ be the estimator of the vector of dispersion parameters. Then, the MSE of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\phi}}$ can be defined as

$$MSE(\hat{\boldsymbol{\beta}}) = E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^2 = trace(\boldsymbol{Var(\hat{\beta})}) + ||\boldsymbol{Bias(\hat{\beta})}||^2, \tag{3.3}$$

and

$$MSE(\hat{\boldsymbol{\phi}}) = E(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})^2 = trace(\boldsymbol{Var(\hat{\phi})}) + ||\boldsymbol{Bias(\hat{\phi})}||^2 \tag{3.4}$$

where

$$trace(\boldsymbol{Var(\hat{\beta})}) = \sum_{j=1}^{2}(Var(\hat{a}_j) + Var(\hat{b}_j)),$$

$$trace(\boldsymbol{Var(\hat{\phi})}) = \sum_{j=1}^{2} Var(\hat{\phi}_j),$$

$$\boldsymbol{Bias(\hat{\beta})} = E(\hat{\boldsymbol{\beta}}) - \boldsymbol{\beta},$$

$$\boldsymbol{Bias(\hat{\phi})} = E(\hat{\boldsymbol{\phi}}) - \boldsymbol{\phi},$$

and $||\cdot||$ is the Euclidean norm. The MSE describes two statistical properties of any estimator: variability and bias. As a consequence, getting a small value of MSE sometimes involves a trade-off between variance and bias. Both quantities should be small for the estimator to have a good performance. When comparing two unbiased (i.e., Bias = 0) estimators, the MSE is equal to their variance and the most efficient estimator is the one with the smallest variance. Therefore, the MSE can be used as a criterion to select the model that provides the best estimator based on the compliance of desirable statistical properties.

On the other hand, penalized likelihood-base selection criteria, such as AIC and BIC, can be used to choose the best copula model from a set of candidate models. The AIC and BIC for a model is defined as $-2l + kp$, where l is the log-likelihood, p is the number of parameters in the model, and k is 2 for AIC and $\log(n)$ for BIC.

The model with the smallest AIC/BIC value gives a better fit. Since the number of parameters is the same across the different copula models in our implementations, the log-likelihood can be used as the model selection criterion, where the preferred model is the one with the highest log-likelihood value. For comparisons between the independent model and copula models, a penalized likelihood-base criterion should be used. An assessment of the ability of these criteria to select the true copula is also included in the simulation studies.

3.3 Simulation Studies

Although copula-based regression models have gained popularity in the discrete case, there is little knowledge about their performance and suitability in different scenarios. In order to better understand the statistical properties of coefficients estimates when jointly modeling bivariate count data using copulas, a set of simulation studies considering different sample sizes, degrees of dependence and dependence structures where conducted. Driven by our motivating example on fatal and non-fatal vehicle crashes, the aim of these simulation studies is to assess the performance of copula-based models for overdispersed bivariate count data and to establish the scenarios where their application is more advantageous compared to the alternative of independent models. Moreover, as shown in Figures 2–3 and 2–4, the dependence structure between two counts is determined by the copula parameter and the marginal means. With this in mind, scenarios where both counts had different marginal means were explored. Also, the performance of likelihood-based criteria, such as AIC and BIC, in selecting the true model among a set of candidate models was examined. In this section, a description of the simulation studies is provided.

3.3.1 STUDY 1 Copula-based regression models for overdispersed bivariate count data with the same marginal means

In the present simulation study, we address the statistical properties of coefficients estimates of copula-based regression models for overdispersed bivariate count data when both marginal means (20.08 when calculated at the mean value of the covariate x) and dispersion parameters ($\phi = 2$) are the same. Bivariate count data were generated from a bivariate copula distribution as defined in Eq. 3.1 for each of the four copula models in Table 2–1 assuming that the marginals, F_1 and F_2, are negative binomial. Five models were fitted: the four copula models in Table 2–1 and the independent model. Three degrees of association between the response variables were considered with values of Kendall's τ: 0.3, 0.5 and 0.7, to reflect low, medium and high association strengths. Three sample sizes N: 100, 300 and 500, were also included. A scenario where the response variables were generated independently to examine the consequences of modeling counts using copula-based models when little or no association is observed is also taken into consideration. For the simulation studies herein, the covariate x is distributed as Normal ($\mu = 10, \sigma = 0.5$). Each of these Monte Carlo experiments involved 1,000 replications. The simulation scheme used was as follows:

1. Select the sample size N, degree of association τ and true copula.
2. Set the parameters for both marginal models: $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = (a, b) = (-2, 0.5)$ and $\phi_1 = \phi_2 = 2$.
3. Generate the data.
   (a) Simulate N covariates x from a Normal ($\mu = 10, \sigma = 0.5$) distribution.
   (b) Obtain N means for each count: $\mu_1 = \mu_2 = \exp(a + bx)$.
   (c) Generate N vectors (y_1, y_2) from the true copula model. Different models were considered:
      i. Copula model with negative binomial marginals with mean μ and dispersion parameter ϕ, and corresponding copula parameter θ that meets an association of τ as defined in Eq. 2.14.
      ii. Independent negative binomials assuming no association between both counts, i.e., $\tau = 0$.

4. Fit negative binomial regression models to both counts independently.
5. Fit the four copula models in Table 2–1 by maximizing the log-likelihood in Eq. 3.2 using the estimates obtained in step 4 as the initial values.
6. Repeat steps 1-5 1,000 times for each copula: Clayton, Frank, Galambos, Normal; N: 100, 300, 500; and τ: 0.3, 0.5, 0.7, for a total of 36 scenarios. Do the same for the independent model, for 3 additional scenarios.

3.3.2 STUDY 2 Copula-based regression models for bivariate count data with different marginal means

In this study, the statistical properties of coefficients estimates of copula-based regression models for overdispersed bivariate count data considering different marginal means (2.72 and 20.08 when calculated at the mean value of the covariate x) and dispersion parameters ($\phi_1 = 5, \phi_2 = 2$) are assessed. These values were selected to represent our motivating example. Bivariate count data was generated from the four copula models in Table 2–1 assuming negative binomial marginals. That is, the realizations of the response variables (y_1, y_2) were taken from a bivariate copula distribution as defined in Eq. 3.1. Five models were fitted: the four copula models in Table 2–1 and the independent model. Three degrees of association between the response variables were considered with values of Kendall's τ: 0.3, 0.5 and 0.7, to reflect low, medium and high associations. Three sample sizes N: 100, 300 and 500, were also included. We also included a scenario where the response variables were generated independently to examine how the copula-based models deal with little or no association. For the simulation studies herein, the covariate x is distributed as Normal ($\mu = 10$, $\sigma = 0.5$). Each of these Monte Carlo experiments involved 1,000 replications. The simulation scheme used was as follows:

1. Select the sample size N, degree of association τ and true copula.
2. Set the parameters for both marginal models: $\boldsymbol{\beta}_1 = (a_1, b_1) = (-4, 0.5)$, $\phi_1 = 5$, $\boldsymbol{\beta}_2 = (a_2, b_2) = (-2, 0.5)$, and $\phi_2 = 2$.

3. Generate the data.
   (a) Simulate N covariates x from a Normal ($\mu = 10$, $\sigma = 0.5$) distribution.
   (b) Obtain N means for each count: $\mu_1 = \exp(a_1 + b_1 x)$ and $\mu_2 = \exp(a_2 + b_2 x)$.
   (c) Generate N vectors (y_1, y_2) from the true copula model. Different models were considered:
      i. Copula with negative binomial marginals, and corresponding copula parameter θ that meets an association of τ as defined in Eq. 2.14.
      ii. Independent negative binomials assuming no association between both counts, i.e., $\tau = 0$.
4. Fit negative binomial regression models to both counts separately.
5. Fit the four copula models in Table 2–1 by maximizing the log-likelihood in Eq. 3.2 using the estimates obtained in step 4 as the initial values.
6. Repeat steps 1-5 1,000 times for each copula: Clayton, Frank, Galambos, Normal; N: 100, 300, 500; and τ: 0.3, 0.5, 0.7, for a total of 36 scenarios. Do the same for the independent model, for 3 additional scenarios.

3.4 Simulation Studies Results

The results of Studies 1 and 2 are presented next. Boxplots of the bias and Mean Squared Error (MSE) for the mean and dispersion parameters estimators are presented to check for desirable properties of any estimator: unbiasedness, minimum variance and efficiency in the different scenarios based on 1,000 simulations. Tables of the average bias are included in Appendices A and B. Another objective of this work was to assess the ability of likelihood-based criteria in selecting the true copula model in different scenarios. With this in mind, the percentage of times each copula is selected for each scenario based on BIC is also reported below.

3.4.1 STUDY 1 Results

Marginal parameters results

When both marginal means (20.08 when calculated at the mean value of the covariate x) and dispersion parameters ($\phi = 2$) are the same, Figures 3–1 through 3–5 indicate that the estimators $\hat{\boldsymbol{\beta}}$ were unbiased for the Independent model and the true copula model in every scenario. In contrast, fitting the wrong copula gave biased estimates for the mean parameters in most scenarios. For instance, Figure 3–1 shows that, when Clayton was the true copula, Frank, Galambos and Normal copulas gave biased estimates. Recall that the Clayton copula is the only one with lower tail dependence in our set of copulas and, by previous examination of the data, the other copulas should not be considered, which supports the importance of this step when fitting copula-based models. For the scenarios with high degree of association between the response variables where Frank and Normal were the true copulas (Figure 3–2 and 3–4), these two gave unbiased estimates along with the Independent model for the three sample sizes included in this work (remember that these two copulas share similar dependence structure). When Galambos (copula with upper tail dependence) was the true copula, the estimation for the mean parameters from other copulas was unbiased except for scenarios with a high degree of dependence.

In addition, the results show that the main gain of taking into account the dependence structure for the mean parameters using copula-based models was in precision. In other words, variance of the mean parameters estimators was significantly smaller when accounting for the association between counts by fitting an adequate copula model even when small association is observed. For a fixed sample size, the gain in precision is greater for stronger associations. Moreover, as shown in Figure 3–5, the Independent model, and the Frank and Normal copulas gave unbiased estimates for the mean parameters, and similar precision when no dependence was observed between the response variables for all sample sizes. This suggests that fitting either of these copulas might be an option in case small association is present and more precision can be obtained without losing accuracy.

On the other hand, the results show that the dispersion parameter estimator had a right skewed distribution in every scenario and there was no evidence that there was a difference in bias when the true copula was fitted versus the Independent model. In fact, previous research exists on the estimation of the negative binomial dispersion parameter ϕ, and it is known that maximum likelihood estimates of ϕ can be biased upward. In this respect, several estimating methods have been proposed in the literature [34, 35]. There was no gain in precision for the dispersion parameters in any of the scenarios in this study.

Figures 3–6 through 3–10 show the distribution of the MSE of the vector of coefficients and dispersion parameters, $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\phi}}$, under the different scenarios. In general, the MSE for $\hat{\boldsymbol{\beta}}$ is smaller when the true copula models are fitted when compared to the Independent model and other copulas. As a result, the estimator of the parameters of the mean from the true copula was more efficient than the estimator from the Independent model. The gain in efficiency was greater as sample size decreased for a given degree of dependence, and as the degree of association became stronger for a fixed N.

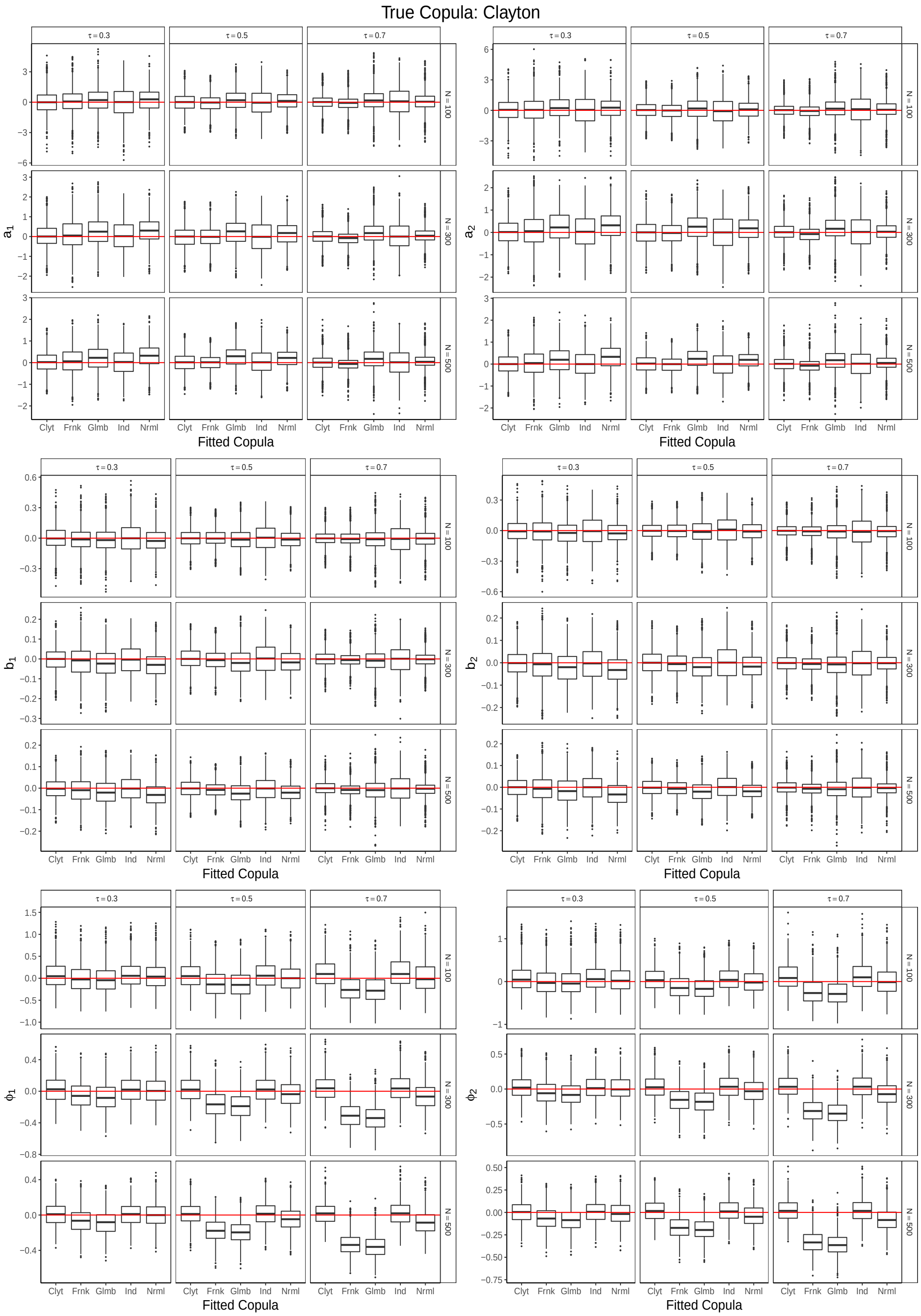

Figure 3–1: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with$\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

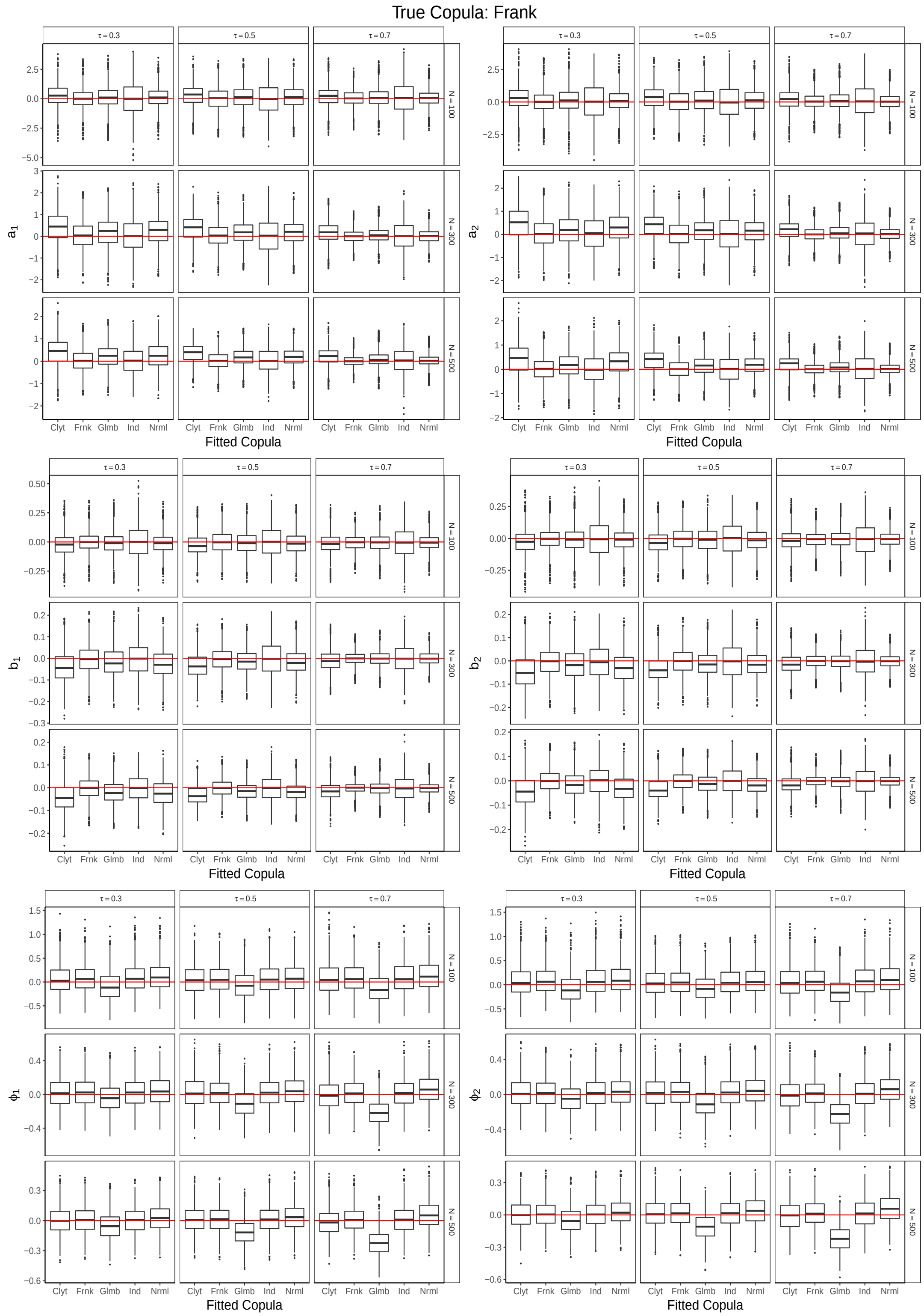

Figure 3–2: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with$\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

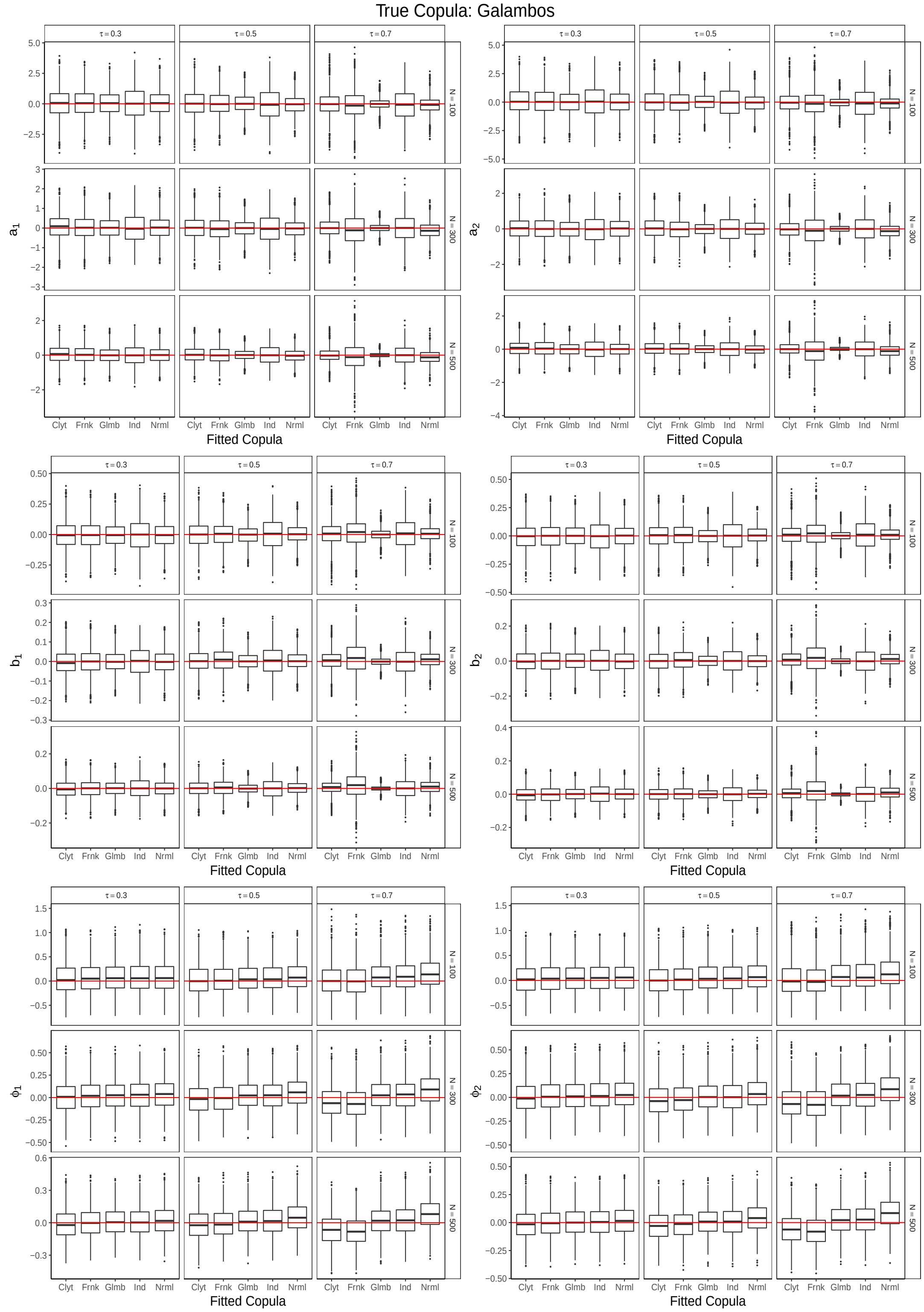

Figure 3–3: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with$\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

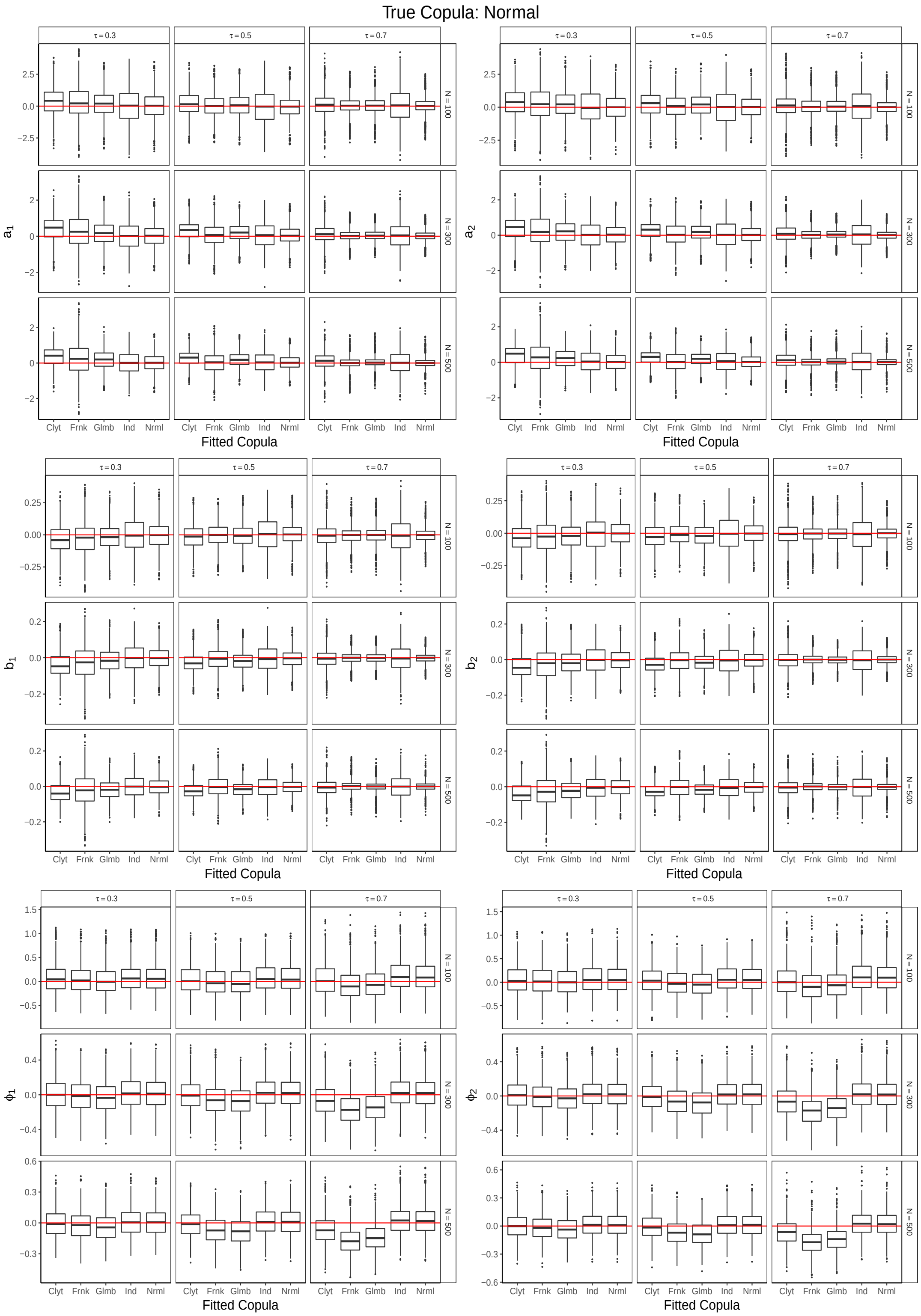

Figure 3–4: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with$\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

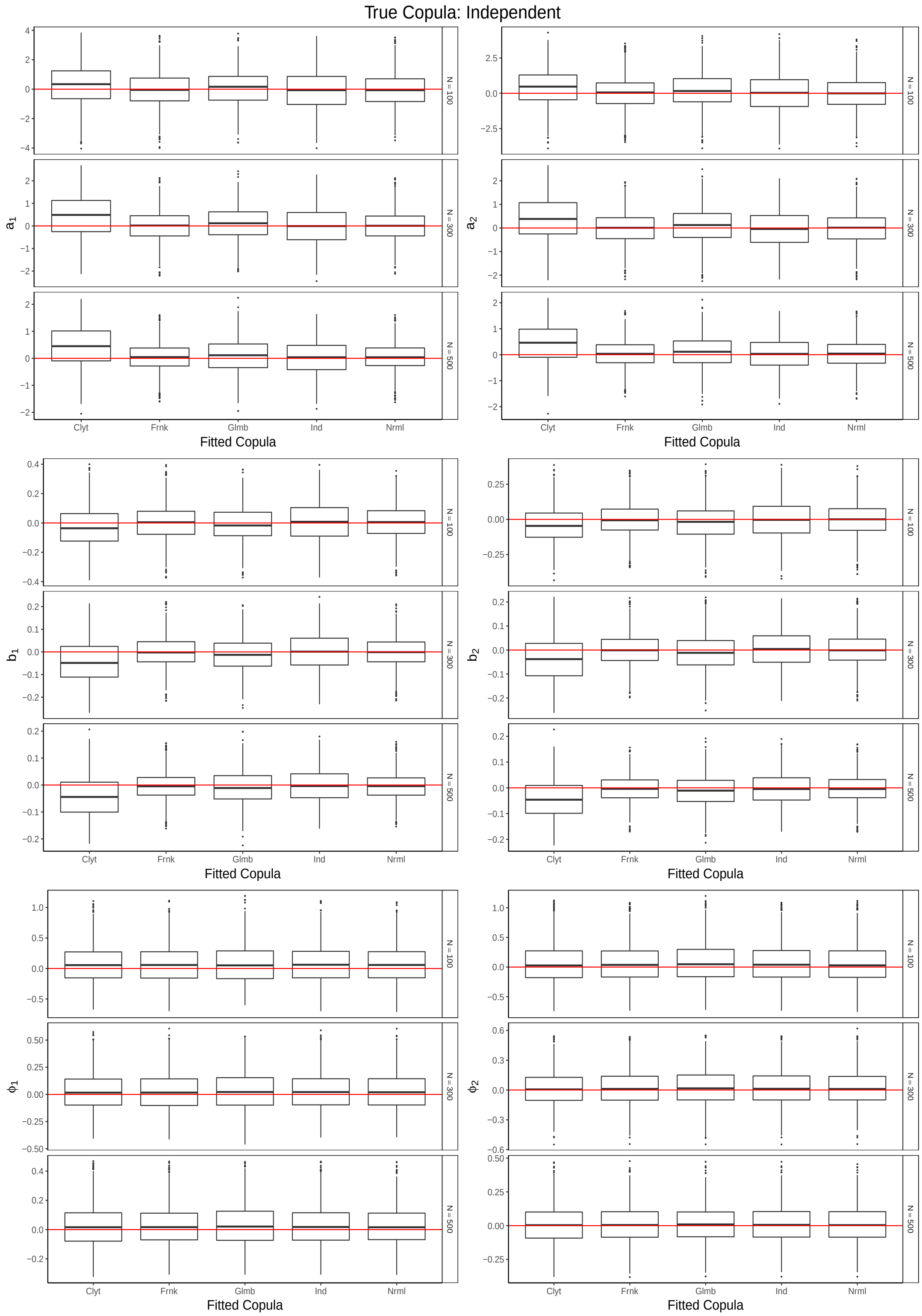

Figure 3–5: Boxplots of the bias of estimators based on 1,000 simulations for two independent counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with $\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t \boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

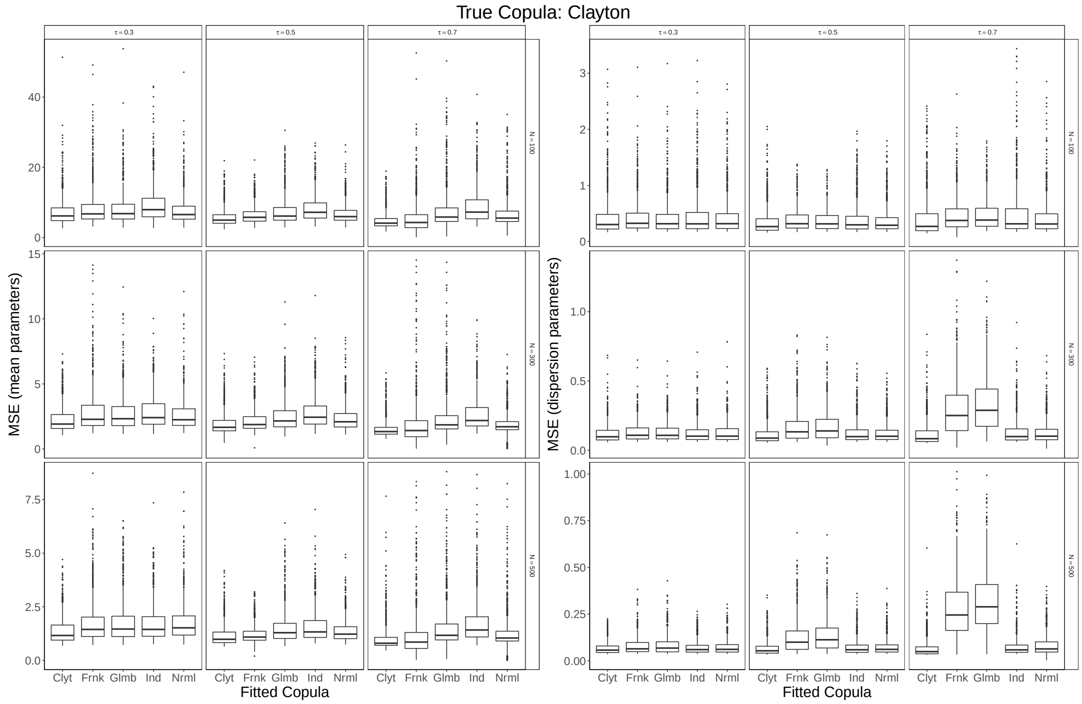

Figure 3–6: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

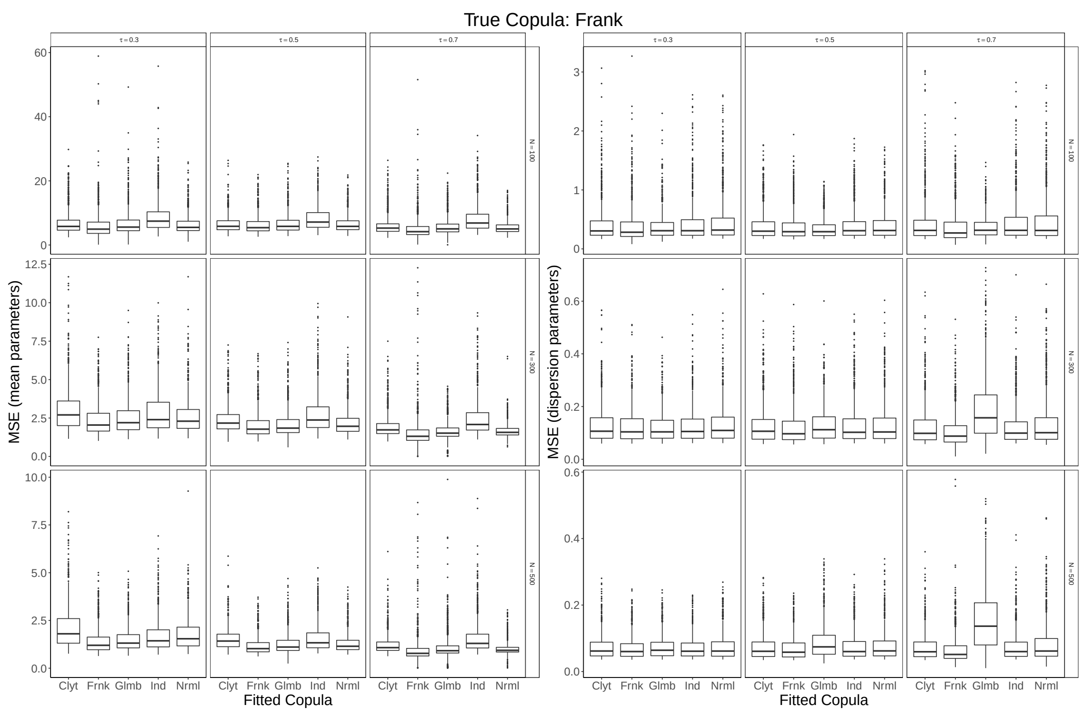

Figure 3–7: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

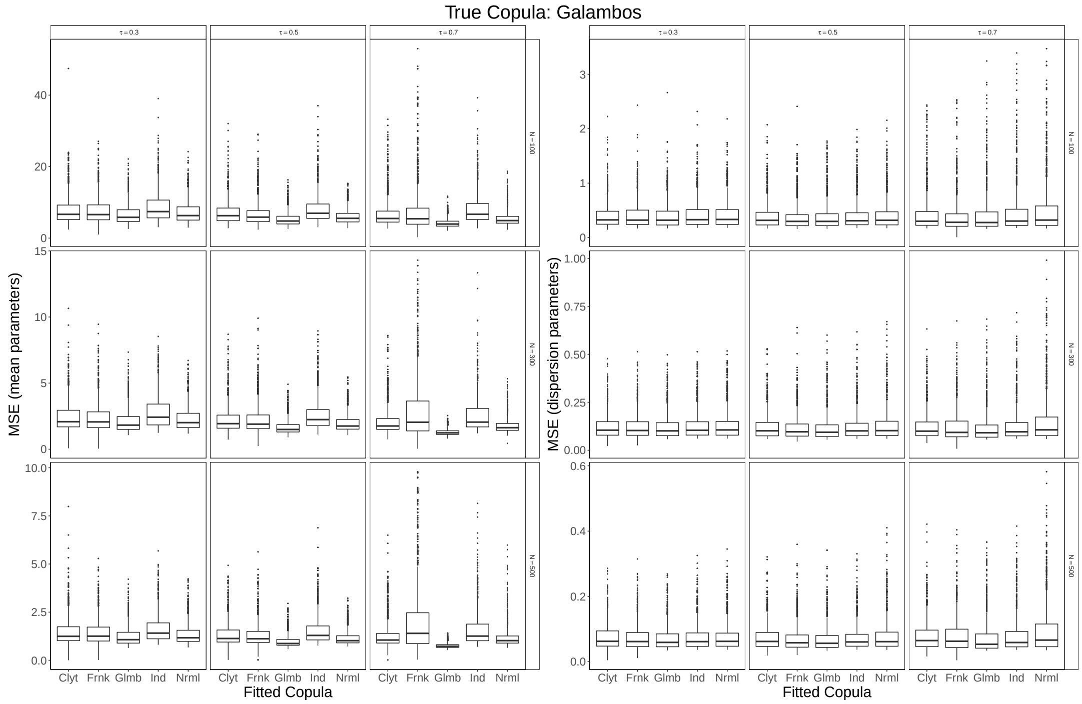

Figure 3–8: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

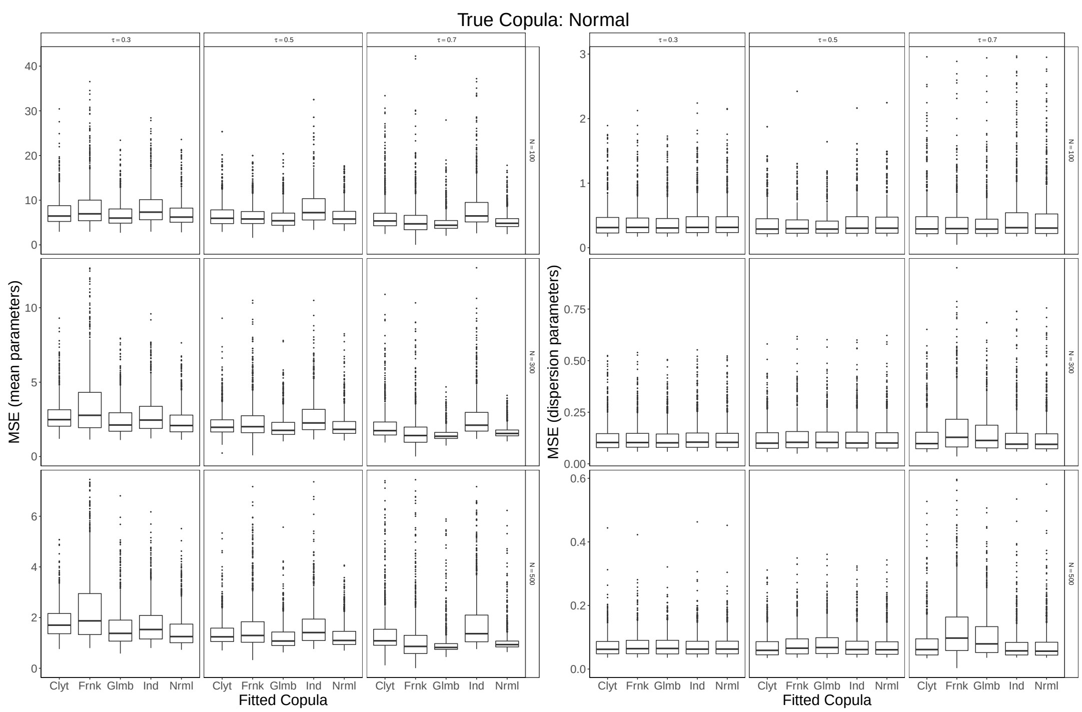

Figure 3–9: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

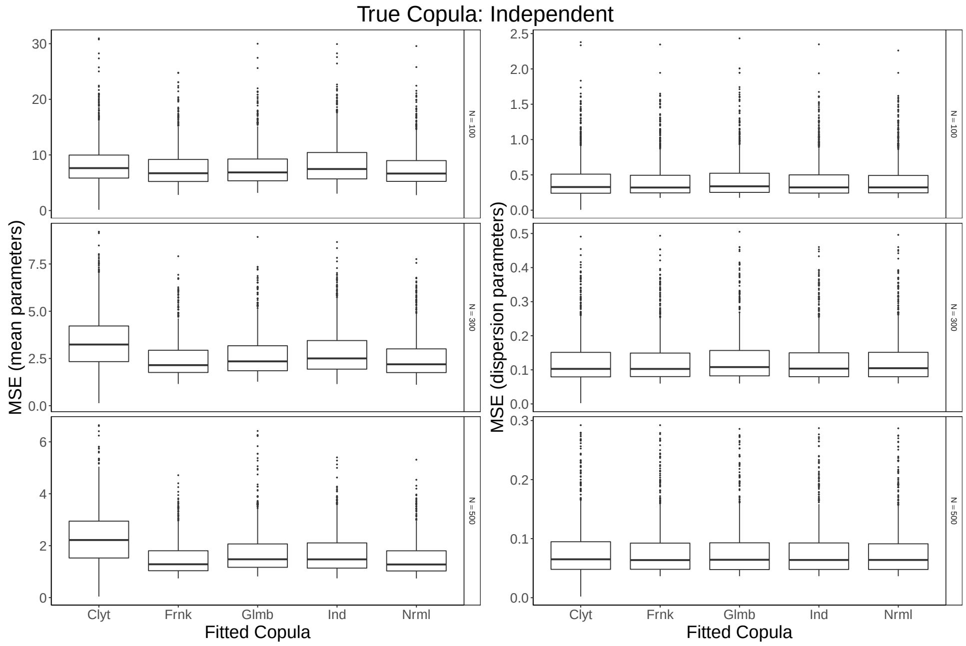

Figure 3–10: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two independent counts, Y_1 and Y_2.

Copula parameter results

Figure 3–11 shows the results of the copula parameter estimator for each copula model for different sample sizes and degrees of dependence, when both means and dispersion parameters are the same. In general, precision is gained with an increase in sample size for a fixed degree of association. However, the results suggest that the copula estimator might be biased when $N = 100$ for the Clayton, Frank and Galambos copulas, and for the Normal copula when the data exhibit a high degree of association.

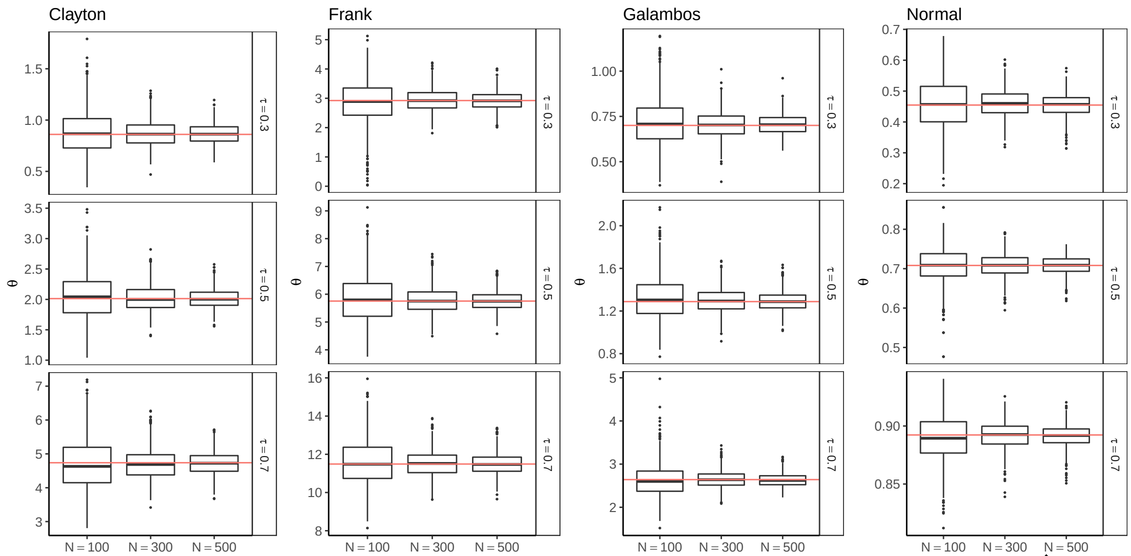

Figure 3–11: Boxplots of the estimates of the copula parameter estimator $\hat{\theta}$ when data comes from the copula in the title based on 1,000 simulations. Horizontal line: true copula parameter value.

Likelihood-based criteria performance results

Tables 3–1 and 3–2 and Figures 3–12 and 3–13summarize the percentage of times the true copula model was selected in each scenario. In general, selection of the true copula more than 99% of the times was achieved for high degrees of association ($\tau = 0.7$) and samples size $N \geq 300$ for all copula models. For sample size equal to 100 and small degree of association ($\tau = 0.3$), BIC failed to select the true copula between 14%-43% of the times depending on the true copula model. Interestingly, the Normal copula was commonly selected over the true copula model, as shown in Figure 3–12 under $\tau = 0.3$ or $N = 100$. As presented in Table 3–2 and Figure 3–13, BIC correctly selected the Independent model more than 95% of the times. The performance of likelihood-based criteria was better when the degree of association and sample size increased.

Table 3–1: Percentage of times the true copula is selected using BIC based on 1,000 simulations.

	$N = 100$			$N = 300$			$N = 500$		
True Copula	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$
Clayton	85.9	97.0	99.8	91.7	98.4	100.0	88.6	98.6	100.0
Frank	83.3	81.4	95.5	75.0	95.2	99.9	78.0	96.1	99.9
Galambos	78.8	90.1	94.9	93.6	98.4	99.6	97.6	99.9	99.9
Normal	56.6	75.7	91.7	85.3	97.2	99.8	92.7	99.4	100.0

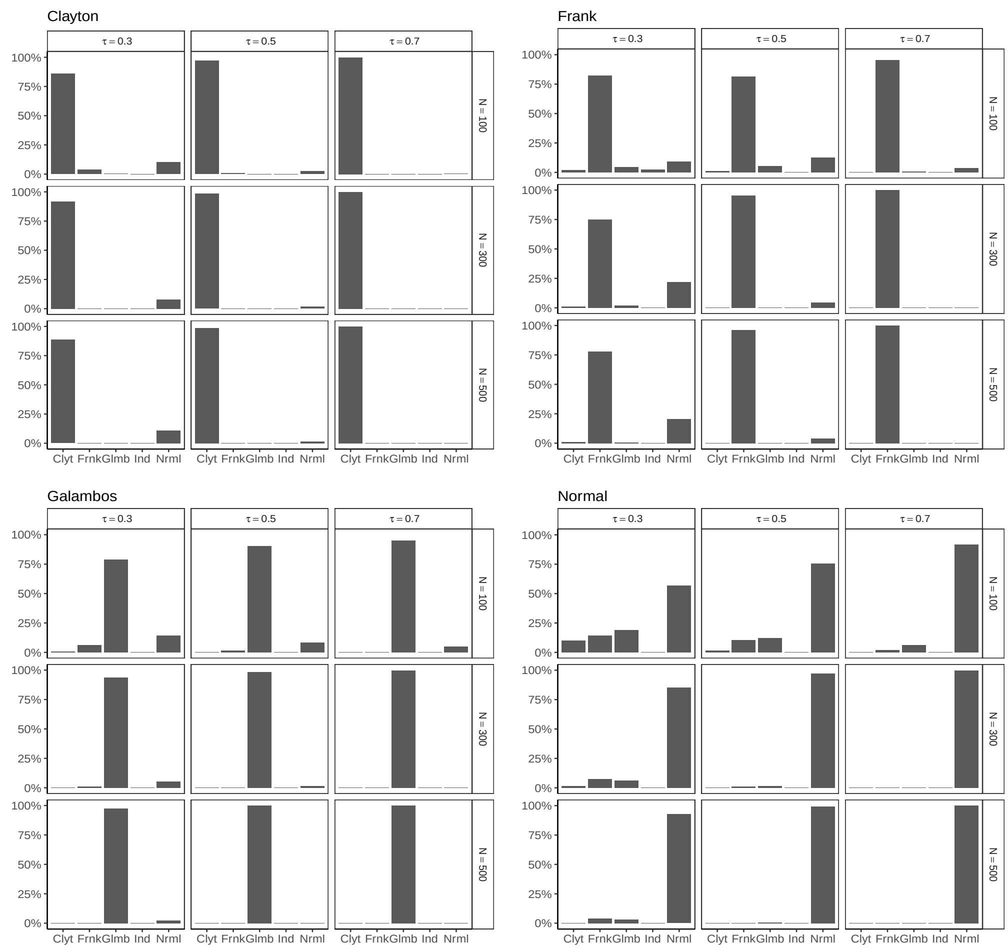

Figure 3–12: Percentage of times each copula is selected when data comes from the copula in the title using BIC based on 1,000 simulations.

Table 3–2: Percentage of times the independent model is correctly selected using BIC based on 1,000 simulations.

True Model	$N = 100$	$N = 300$	$N = 500$
Independent	94.9	96.1	97.4

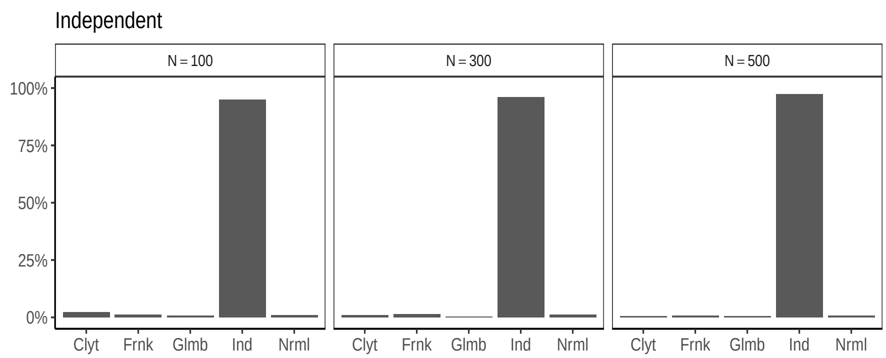

Figure 3–13: Percentage of times each copula model is selected when data comes from the independent model using BIC based on 1,000 simulations.

3.4.2 STUDY 2 Results

The results for Study 2, which considered different marginal means (2.72 and 20.08 when calculated at the mean value of the covariate x) and different dispersion parameters ($\phi_1 = 5$ and $\phi_2 = 2$, respectively), were fairly similar to those of Study 1. A brief discussion of the results is presented next.

Marginal parameters results

As in Study 1, the results show that estimators $\hat{\boldsymbol{\beta}}$ from the Independent model and the true copula model in every scenario produced unbiased estimates, and fitting the wrong copula gave biased estimates in some scenarios (Figures 3–14 through 3–18). Nevertheless, unlike in Study 1, when the association between the responses was low ($\tau = 0.3$), the estimates produced by every candidate copula were close to the true value. Therefore, the wrong copula models gave biased estimates mostly under medium ($\tau = 0.5$) and high ($\tau = 0.7$) correlation between counts. In contrast to Study 1, where the five fitted models gave unbiased estimates of the mean parameters when the data was generated from the Galambos copula with medium association between counts, biased estimates were given by the Clayton and Frank copulas (Figure 3–16). The results for this study show that, when one of the count variables has a small mean and the degree of association between the counts is medium or high, taking into account the dependence structure provides significantly more precision in the estimation of the regression coefficients. Once again, the gain in precision increases as the degree of dependence gets stronger.

The dispersion parameter estimator also had a right skewed distribution in every scenario in this study, being more so for the count with small mean and dispersion parameter $\phi = 5$, particularly when $N = 100$. Similarly, no gain in precision was observed and the wrong copula produced even more biased estimates.

Finally, as shown in Figures 3–19 through 3–23, MSE is smaller for the true copula model when compared to the Independent model. Moreover, Figure 3–23 suggests that the Frank copula gave more efficient estimators when no dependence is observed.

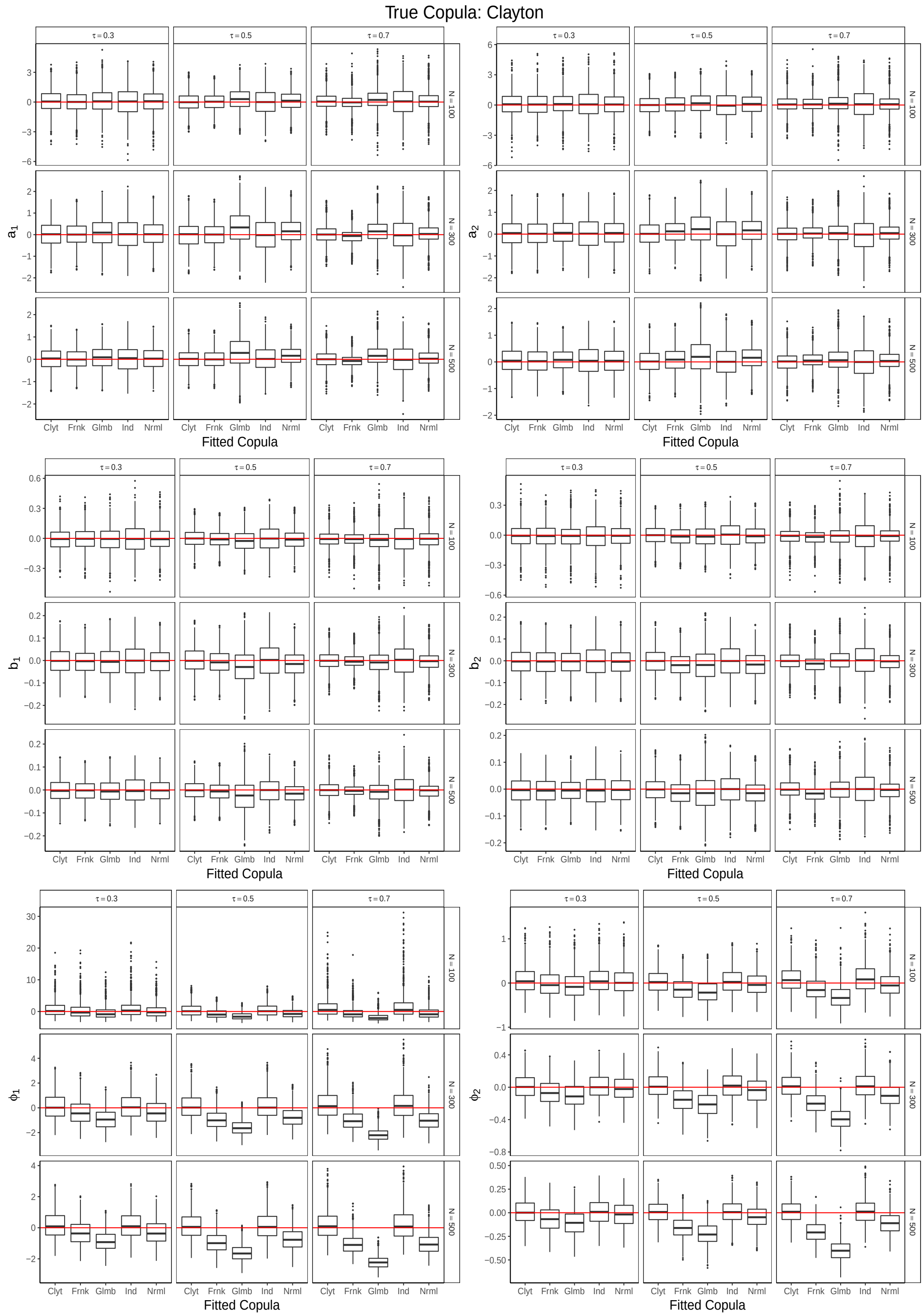

Figure 3–14: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with $\mu_j = \exp(\boldsymbol{x}^t\boldsymbol{\beta}_j)$ and dispersion parameter ϕ_j, for $j = 1, 2$.

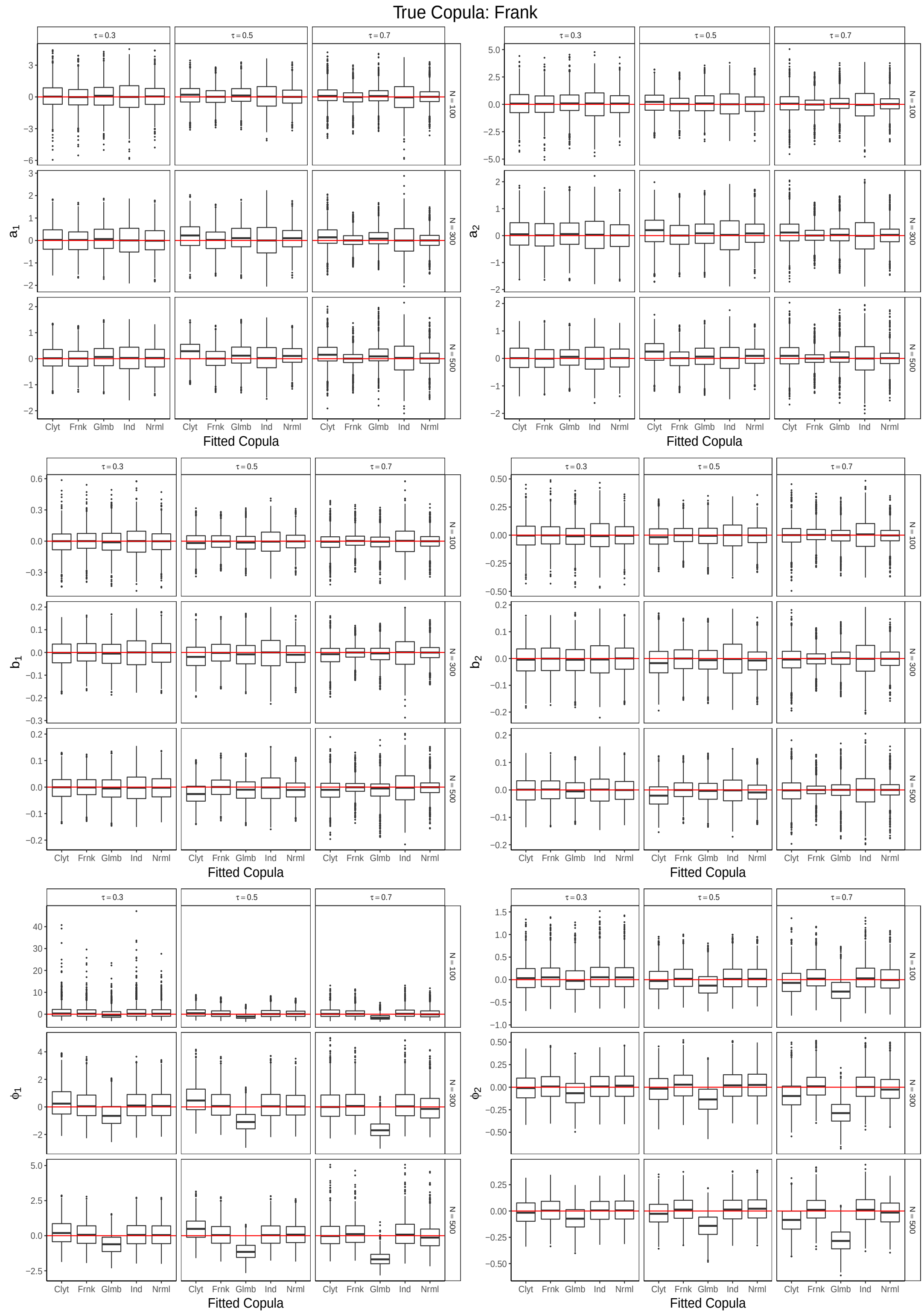

Figure 3–15: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with $\mu_j = \exp(\boldsymbol{x}^t \boldsymbol{\beta}_j)$ and dispersion parameter ϕ_j, for $j = 1, 2$.

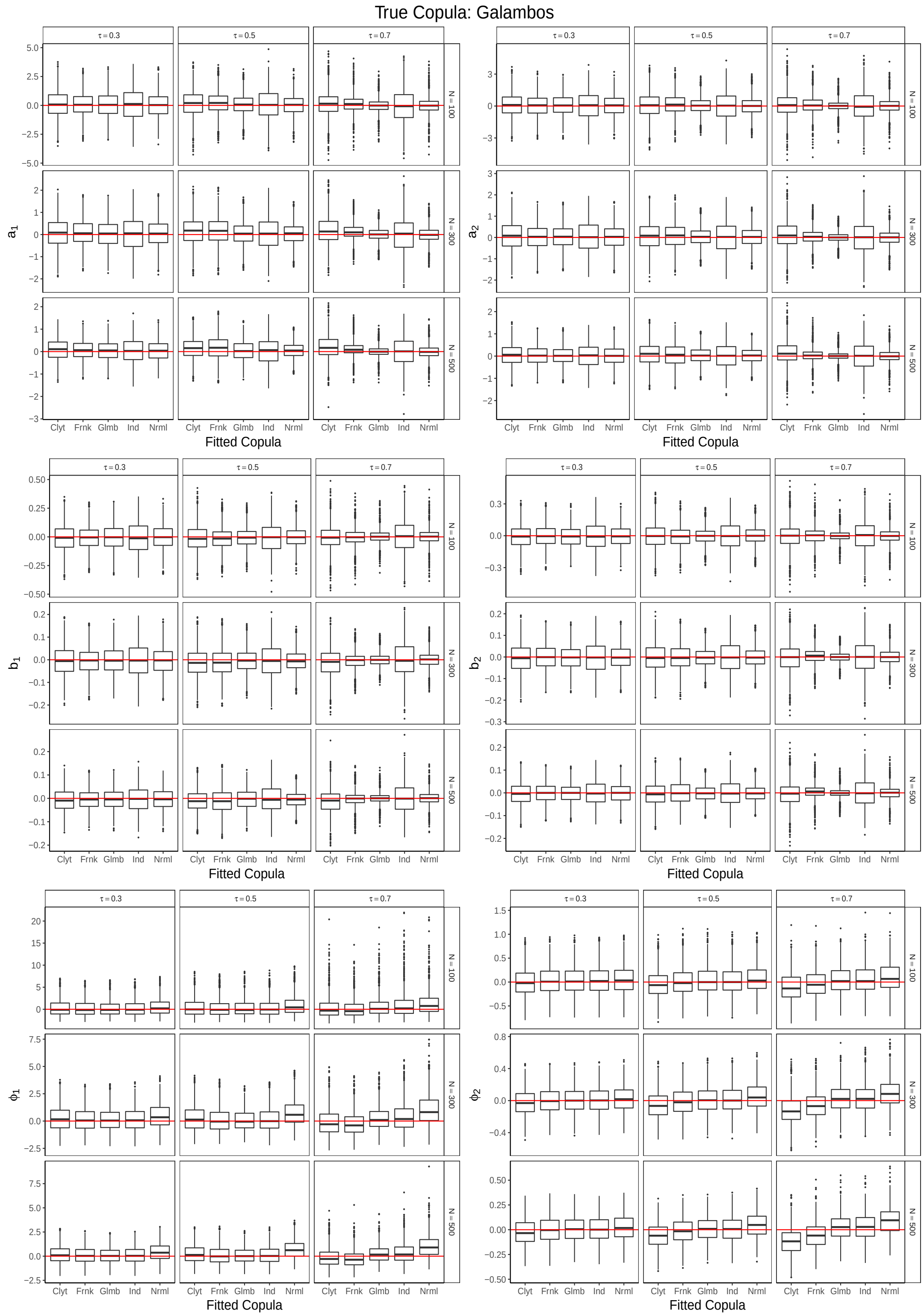

Figure 3–16: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with $\mu_j = \exp(\boldsymbol{x}^t \boldsymbol{\beta}_j)$ and dispersion parameter ϕ_j, for $j = 1, 2$.

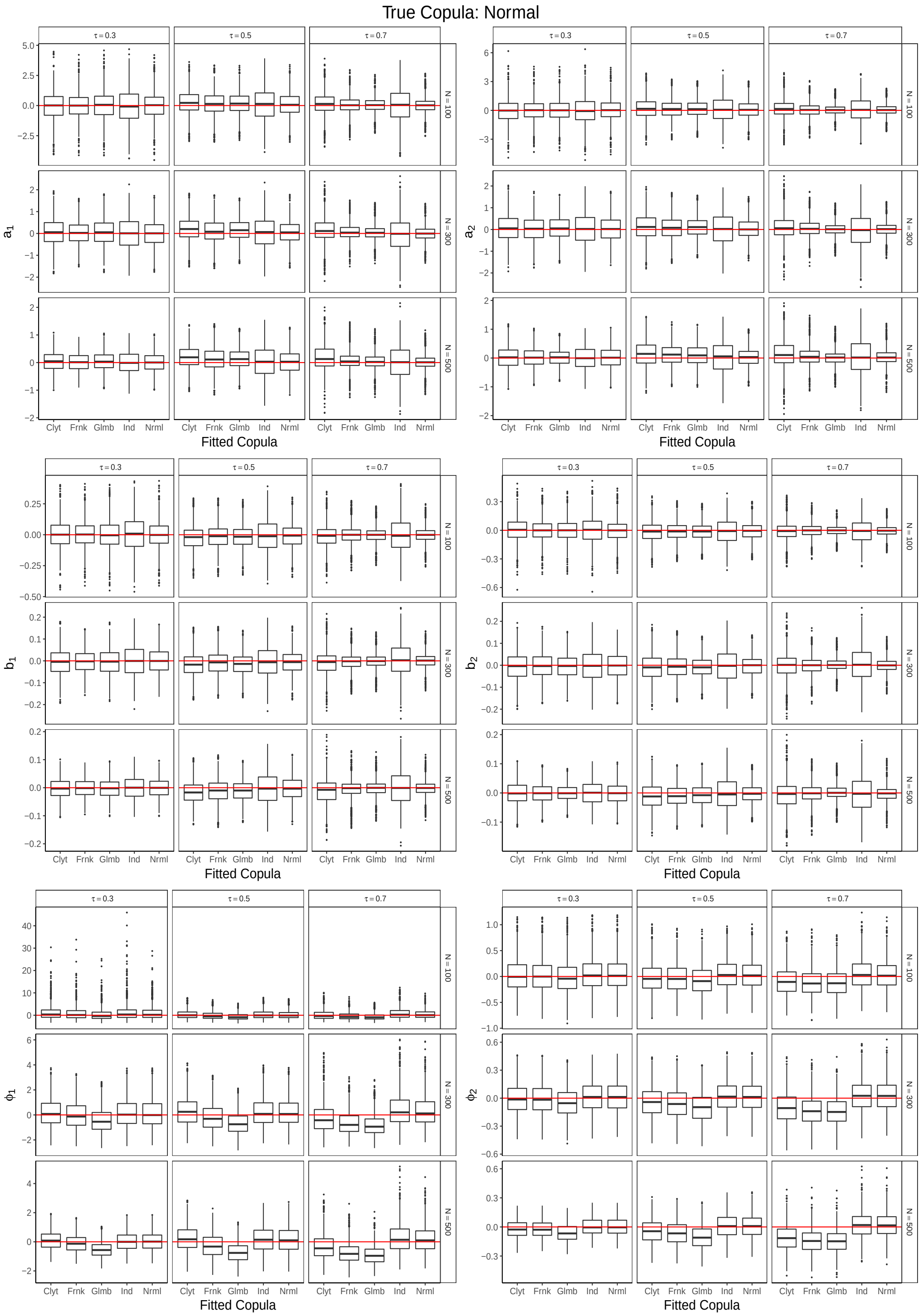

Figure 3–17: Boxplots of the bias of estimators based on 1,000 simulations for two jointly distributed counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with $\mu_j = \exp(\boldsymbol{x}^t\boldsymbol{\beta}_j)$ and dispersion parameter ϕ_j, for $j = 1, 2$.

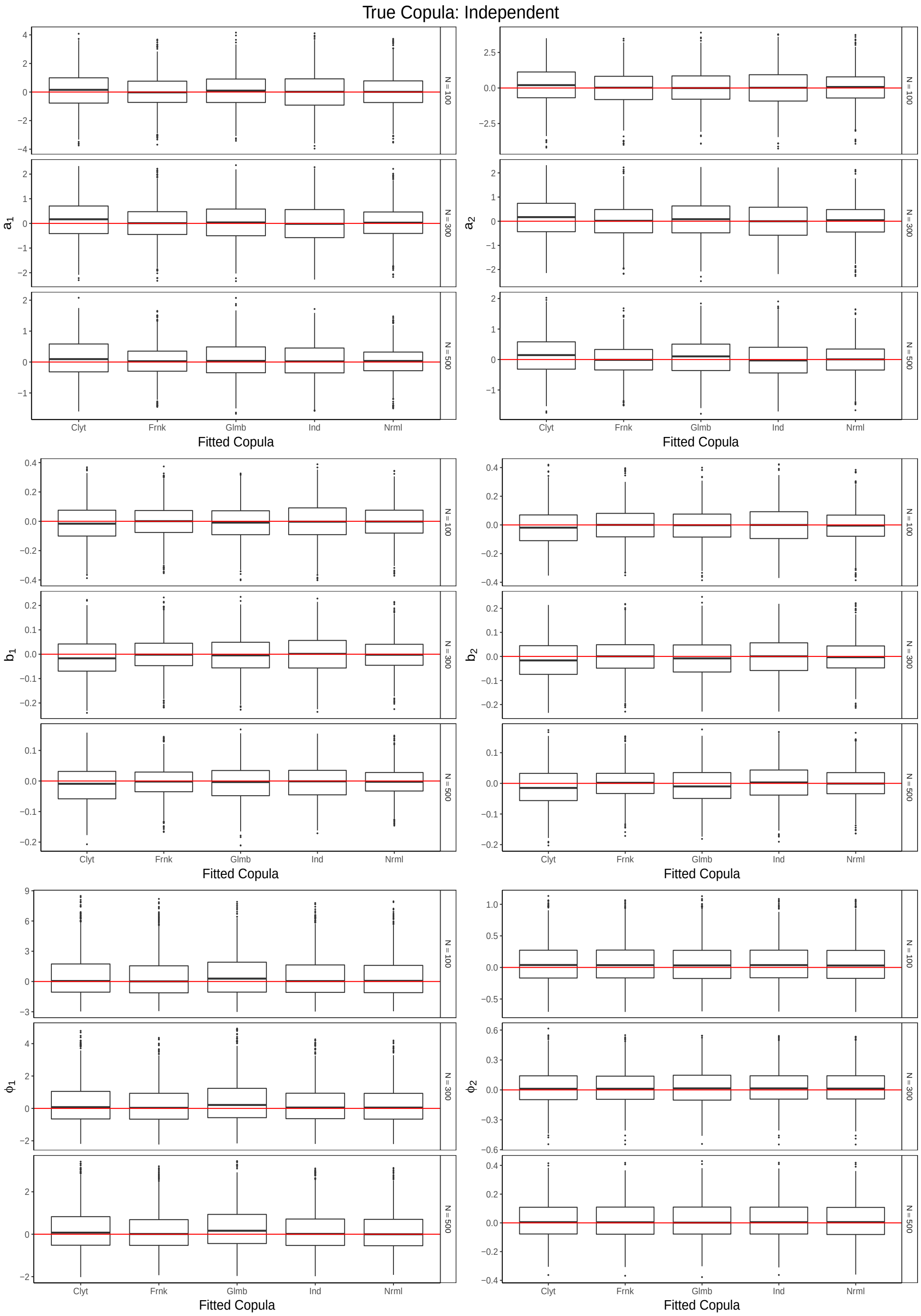

Figure 3–18: Boxplots of the bias of estimators based on 1,000 simulations for two independent counts, Y_1 (left) and Y_2 (right), where each follows a negative binomial distribution with$\mu_1 = \mu_2 = \exp(\boldsymbol{x}^t\boldsymbol{\beta})$ and dispersion parameter $\phi_1 = \phi_2$.

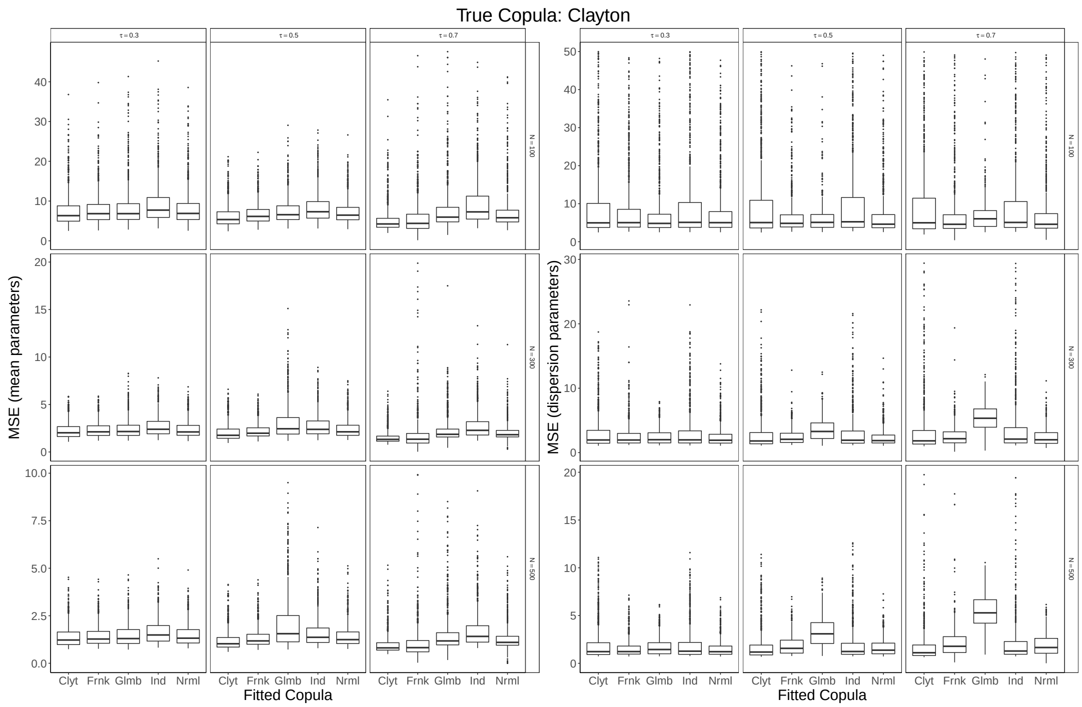

Figure 3–19: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

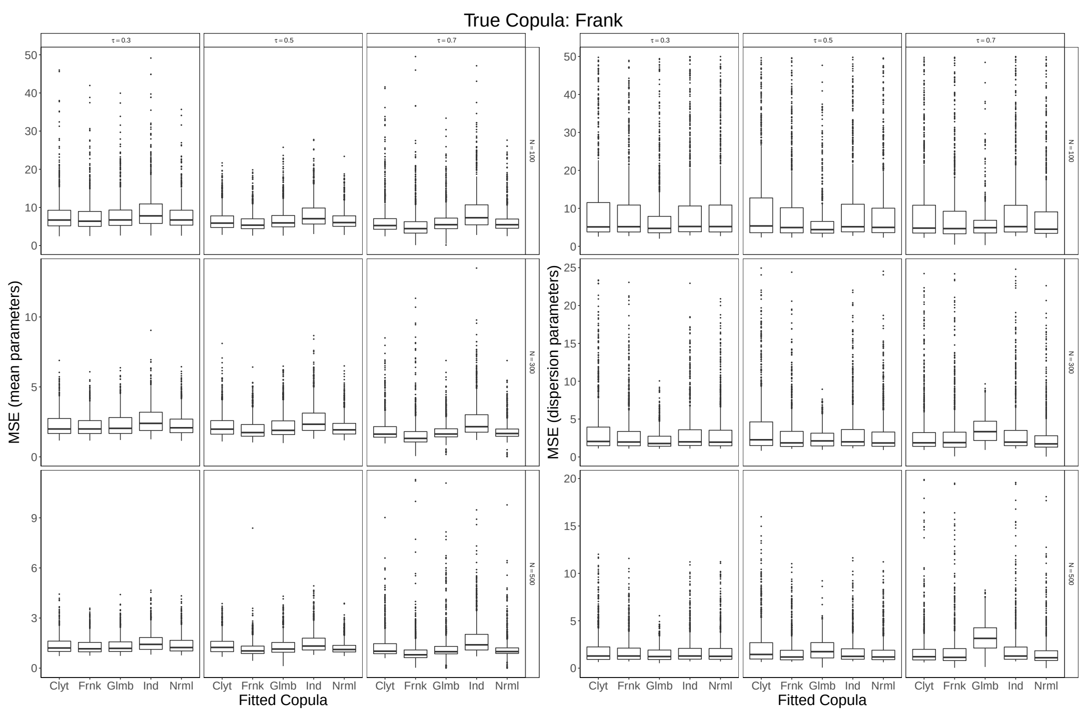

Figure 3–20: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

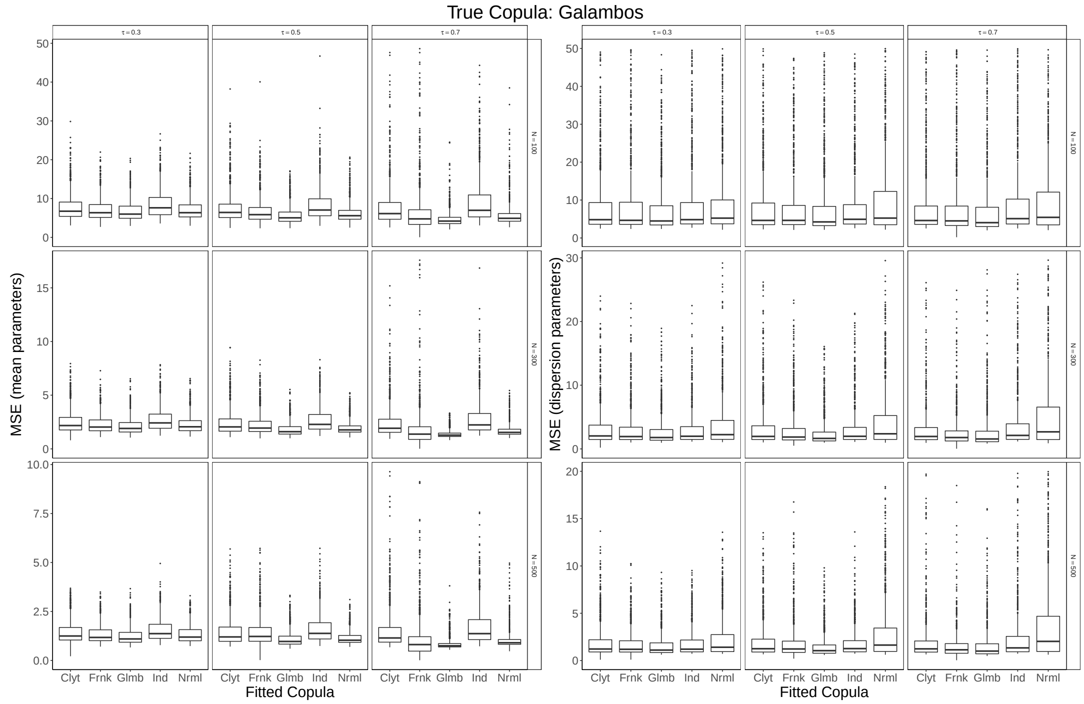

Figure 3–21: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

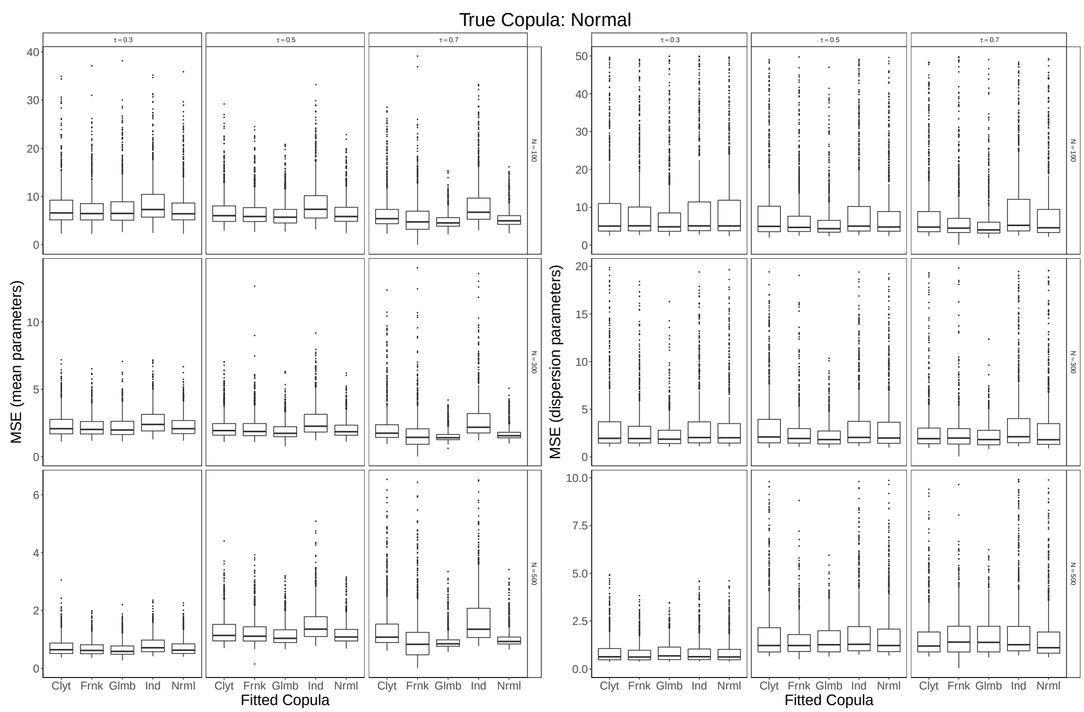

Figure 3–22: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

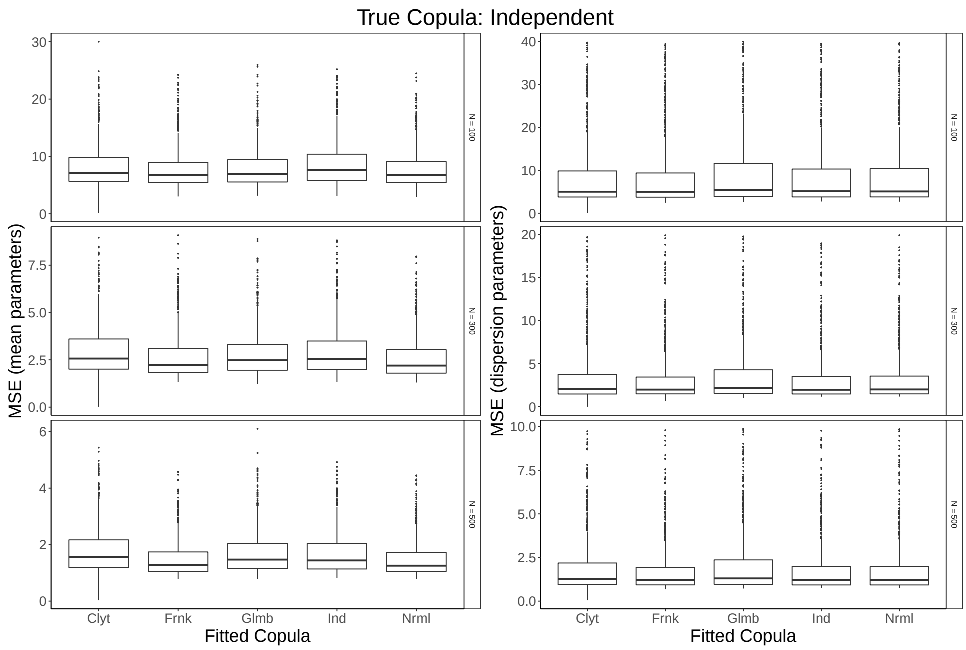

Figure 3–23: Boxplots of the MSE of $\hat{\boldsymbol{\beta}}$ (left) and $\hat{\boldsymbol{\phi}}$ (right) estimators based on 1,000 simulations for two jointly distributed counts, Y_1 and Y_2.

Copula parameter results

Figure 3–24 shows the results of the copula parameter estimator for each copula model for different sample sizes and degrees of dependence, when the means and dispersion parameters are very different. Overall, precision is gained with an increase in sample size for a fixed degree of association. As in Study 1, the results suggest that the copula estimator ($\hat{\theta}$) might be biased when $N = 100$ for the Clayton, Frank and Galambos copulas. For the Normal copula, $\hat{\theta}$ resulted to be biased when the data exhibited a medium and high degree of association and $N = 100$.

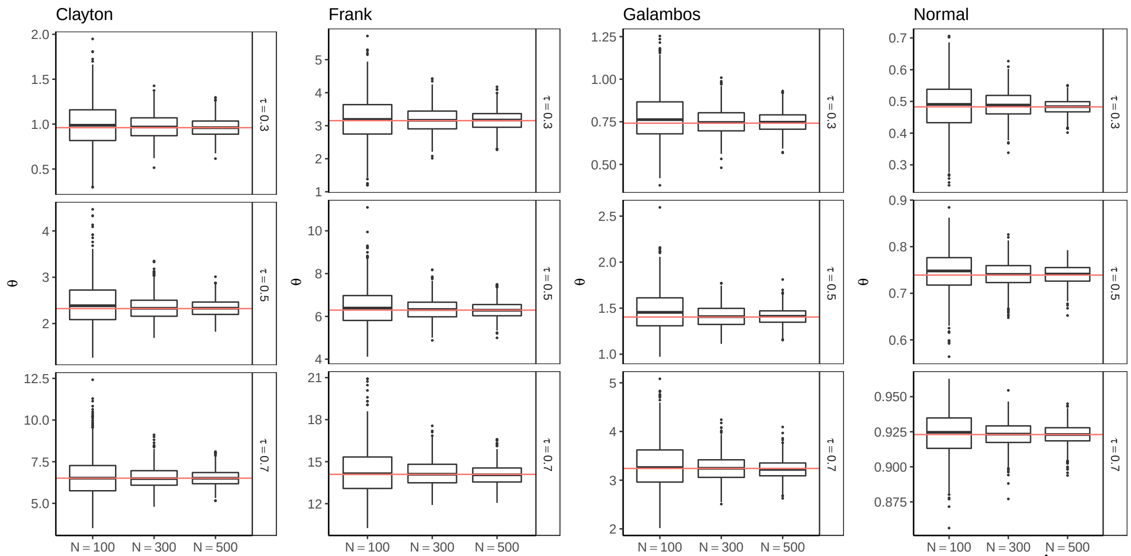

Figure 3–24: Boxplots of the estimates of the copula parameter estimator $\hat{\theta}$ when data comes from the copula in the title based on 1,000 simulations. Horizontal line: true copula parameter value.

Likelihood-based criteria performance results

Tables 3–3 and 3–4 and Figures 3–25 and 3–26 summarize the percentage of times the true copula model was selected in each scenario, when the marginal distributions of the counts were different. As in Study 1, selection of the true copula more than 99% of the times was achieved for high degrees of association ($\tau = 0.7$) and samples size $N \geq 300$ for all copula models. For sample size equal to 100 and small degree of association ($\tau = 0.3$), BIC failed to select the true copula between 12% up to 51% of the times depending on the true copula model. The Normal and Frank copulas were the hardest to be correctly selected by the BIC, in particular when the degree of association was low. Performance of the BIC was slightly better in Study 1, where the marginal distributions were the same. As presented in Table 3–4 and Figure 3–26, BIC correctly selected the Independent model more than 94% of the times. The performance of likelihood-based criteria was better when the degree of association and sample size increased.

Table 3–3: Percentage of times the true copula is selected using likelihood based criteria based on 1,000 simulations.

True Copula	$N = 100$			$N = 300$			$N = 500$		
	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$	$\tau = 0.3$	$\tau = 0.5$	$\tau = 0.7$
Clayton	87.9	96.8	98.9	98.5	99.3	100.0	99.9	99.3	100.0
Frank	62.2	80.2	91.6	84.8	96.8	100.0	93.6	98.0	99.9
Galambos	80.1	89.7	91.9	94.2	98.6	99.0	98.7	99.8	99.9
Normal	49.1	77.5	87.4	80.7	96.5	99.7	99.0	99.6	99.9

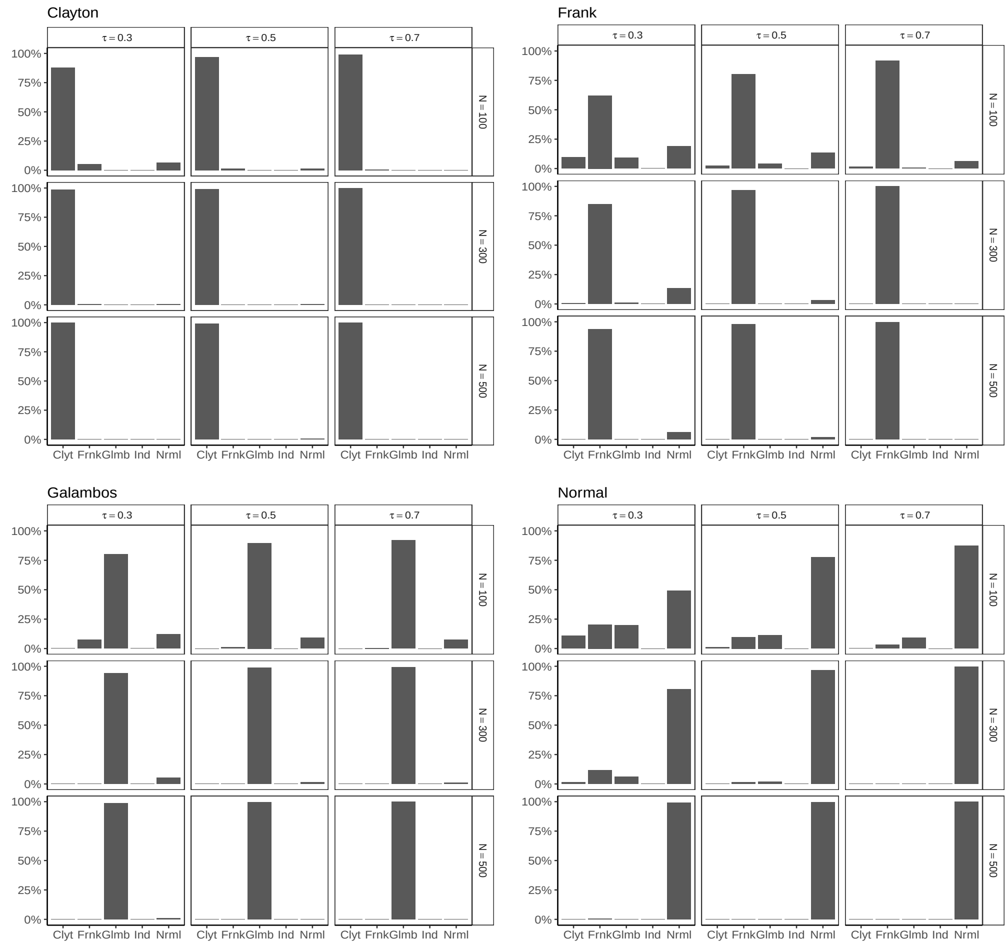

Figure 3–25: Percentage of times each copula is selected when data comes from the copula in the title using BIC based on 1,000 simulations.

Table 3–4: Percentage of times the true copula is selected using BIC based on 1,000 simulations.

True model	$N = 100$	$N = 300$	$N = 500$
Independent	94.2	97.0	97.2

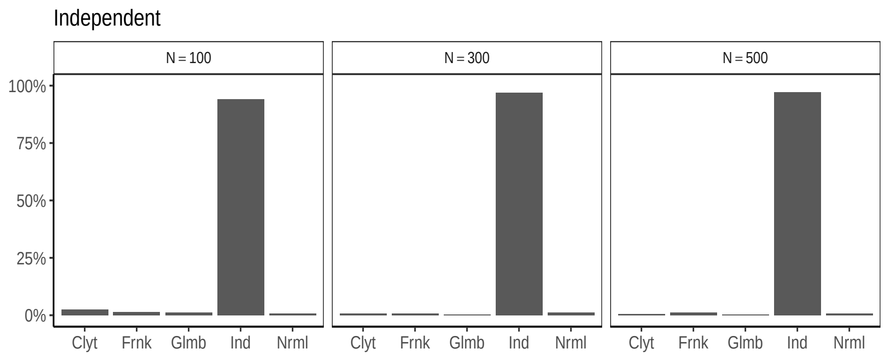

Figure 3–26: Percentage of times each copula is selected when data comes from the independent model using BIC based on 1,000 simulations.

CHAPTER 4
APPLICATION: FATAL AND NON-FATAL VEHICLE CRASHES IN HIGHWAYS OF PUERTO RICO

Detailed crash count data is obtained mostly for safety issues in different severity levels, e.g., Fatal, Injury, Property damage. Even though crash data by severity are innately multivariate, they are often analyzed separately without taking into account the dependence and shared variables that exist between severities, which can lead to inefficient estimates and incorrect interpretations of the results, as discussed in the simulation studies in Section 3.3. Efforts have been dedicated to taking into account the dependence and shared characteristics between collision types by simultaneously modeling crashes by severity. In this regard, road safety analysts have focused on the study and application of the multivariate Poisson Lognormal distribution allowing for the overdispersion that characterizes crash counts for jointly modeling crash data by collision type [2, 10, 23, 24, 31]. As mentioned in Section 2.3, this approach has some shortcomings and requires numerical methods for estimation. Moreover, this and other multivariate models for count data, as described in Section 2.3, are limited in the dependence structure and the marginal distributions they allow. In this Chapter, an application to fatal and non-fatal crashes and their relationship to the annual average daily traffic (AADT) is presented by jointly modeling fatal and non-fatal crashes using a copula-based regression model.

4.1 Data Description

The data used refer to the number of fatal and non-fatal crashes in segments of highways in Puerto Rico for a period of six years ($2004 - 2009$) collected by the Highways and Transportation Authority of Puerto Rico as described in [14]. For the application presented here three data points were identified as influential outliers using the outlierTest function from the car package and the influence.measures function in R, and were removed from the dataset to proceed with the statistical analysis. The marginal association between fatal and non-fatal crashes in the 141 segments of highways has a Kendall's τ of 0.537 (0.340 when accounting for the length of the segment). Figure 4–1 shows the marginal distribution of the counts of fatal and non-fatal crashes and the relationship between them. The covariate included in this analysis was the natural logarithm of the mean average annual daily traffic over the six years period (logMAADT), which measures how busy these highway segments are. The length of each segment was used as an offset. In Table 4–1 a summary of the variables is presented. Notably, the marginal means of the response variables are very different.

Table 4–1: Summary statistics of the crash data for the 6 year period.

Variable	Mean	SD	Min	Max
Fatal Crashes	1.84	2.11	0.00	10.00
Non-fatal Crashes	118.65	107.83	0.00	684.00
logMAADT	10.97	0.72	9.04	12.57
Length (km)	2.48	1.91	0.15	12.30

4.2 Modeling the Marginals

The first step when jointly modeling bivariate variable using copulas is to estimate suitable univariate marginal distributions for each count. Appropriate marginal fit to the data is needed to get a good estimation of the copula parameter θ [33]. A scatterplot of the rate of crashes versus logMAADT suggested a quadratic relationship, therefore, a polynomial regression model was fitted to both response

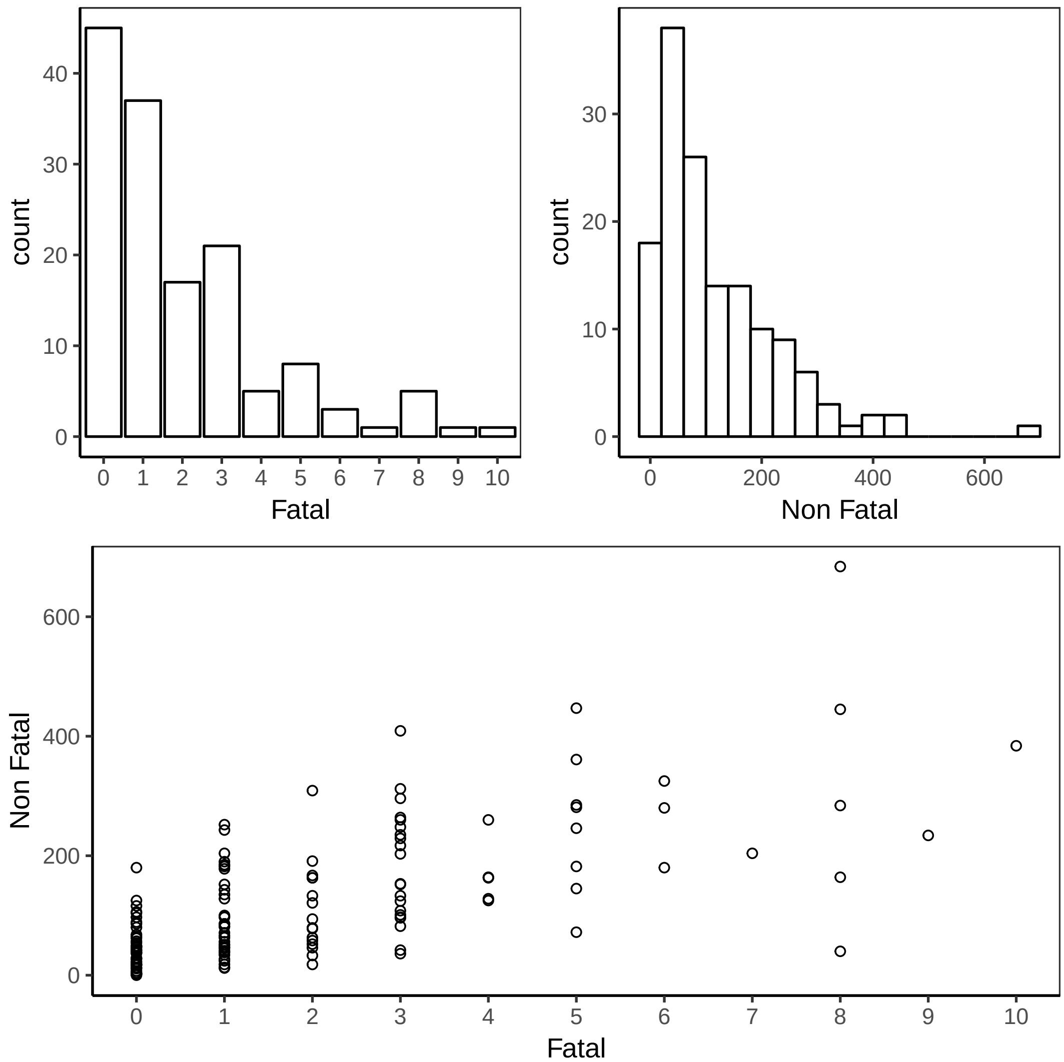

Figure 4–1: The vehicle crashes data in 141 segments of highways in Puerto Rico. ($\tau = 0.537$)

variables. After fitting Poisson models, the Pearson dispersion statistic and other tests, such as the Score and Lagrange multiplier tests, indicated that Poisson overdispersion was present in the counts of non-fatal crashes. On this account, a negative binomial regression model was fitted to the non-fatal crashes data to account for overdispersion, and negative binomial was chosen over Poisson as the marginal distribution based on AIC and BIC for this response. Table 4–2 includes the estimates of regression parameters obtained for the univariate models under the Independent column. Precisely, one of the advantages of using copula-based regression models is the possibility of working with the marginal distributions.

4.3 Joint Modeling Using Copulas

To jointly model the number of fatal and non-fatal crashes, families from the Archimedean, elliptical and extreme value classes of copulas were fitted to allow for different dependence structures and tail behaviors. Briefly, the copula families considered were: Frank and Normal copulas for their flexibility, Clayton to account for the apparent lower tail dependence (Figure 4–1), and Galambos for its usefulness in modeling the dependence of rare events. The copula representation of the joint distribution is given by

$$C(F_1(y_1; x), F_2(y_2; x); \theta), \tag{4.1}$$

where, respectively, y_1 and y_2 are the number of fatal and non-fatal crashes, $F_1(\cdot)$ and $F_1(\cdot)$ are the univariate marginals chosen to be Poisson and negative binomial, x is the covariate logMAADT, and θ is the copula parameter. The log-likelihood as defined in Eq. 3.2 was maximized via full maximum likelihood estimation to obtain an estimate of the regression and copula parameters for each candidate copula. The Nelder and Mead method, implemented in the `optim` function in the statistical software `R` was used to obtain the estimates of the parameters using the estimates provided by the univariate models as the initial values.

The parameters estimates obtained via full maximum likelihood based on the Independent and the candidate copula models are presented in Table 4–2. The Normal copula model is not included since optimization was not achieved after several unsuccessful attempts. According to AIC and BIC criteria, the Clayton copula regression model was selected as the one with the best fit with a BIC of 1909.08, followed by the Frank copula with a BIC of 1912.47, among the set of candidate copulas and over the Independent model. On the other hand, the Galambos copula had the worst fit among the set of candidate copula models probably because of the upper tail dependence this copula admits. To assess the goodness-of-fit of the chosen model (Clayton copula), a plot of the observed and fitted curve for different segment lengths is presented in Figure 4–2. Jointly modeling the crash data using the Clayton copula gives similar estimates to those of the univariate models and smaller standard errors (SE) for the estimates of the regression parameters when compared to the univariate approach. In other words, the estimates of the regression parameters are more precise when accounting for the dependence using the Clayton copula model. For example, the percentage decrease in SE for the three marginal means parameters was 70% and 23% for the fatal and non-fatal data, respectively. Consequently, smaller SE's give narrower confidence intervals for the regression coefficients that could result in the declaration of covariates or factors as significant. To put in another way, accounting for the dependence between counts provides additional information that might be sufficient to conclude that an independent variable is significant when, in contrast, by not taking the dependence into account, larger SE's would have led to incorrectly declaring a variable as non-significant. For example, under the univariate model, the Wald confidence interval of the multiplicative effect on the number of fatal crashes for a one unit increase in the squared logMAADT was $(0.55, 0.95)$ compared to $(0.67, 0.78)$ under the bivariate model using the Clayton copula.

Table 4–2: Estimated parameters (SE), AIC, BIC, and loglikelihood, under the Independent and the four copula models for the crash data ($N = 141$).

		Independent	Clayton	Frank	Galambos
μ_{iFatal}	Intercept	-49.96 (17.35)	-49.88 (5.3)	-50.15 (15.37)	-50.28 (20.76)
	logMAADT	8.14 (3.1)	8.14 (0.92)	8.19 (2.75)	8.19 (3.69)
	logMAADT^2	-0.33 (0.14)	-0.33 (0.04)	-0.33 (0.12)	-0.33 (0.16)
$\mu_{iNon\text{-}Fatal}$	Intercept	-46.03 (9.29)	-46.01 (7.23)	-45.86 (18.12)	-46.19 (11.15)
	logMAADT	8.02 (1.69)	8.02 (1.31)	7.98 (3.26)	8.04 (1.99)
	logMAADT^2	-0.32 (0.08)	-0.32 (0.06)	-0.31 (0.15)	-0.32 (0.09)
	ϕ_2	3.47 (0.44)	3.51 (0.49)	3.45 (0.99)	3.34 (0.63)
	$s(\theta)$	—	-0.69 (0.16)	1.77 (1.2)	-1.06 (0.35)
	AIC	1895.62	1885.49	1888.53	1891.25
	BIC	1916.26	1909.08	1912.12	1914.84
	loglik	-940.81	-934.74	-936.26	-937.63

With regard to the interpretation of the copula model's output, the coefficients are fairly alike for both counts, indicating that logMAADT had a similar marginal effect for both types of crashes. The model indicates that higher values of logMAADT are associated with a higher number of fatal and non-fatal crashes for the range of values of logMAADT observed in this dataset. In other words, busier highway segments tend to have more fatal and non-fatal crashes. The estimate for the copula parameter was $\theta = 0.50$. Evaluated at the mean value of logMAADT and logMAADT^2 (10.97 and 120.44, respectively), this estimate of the copula parameter transformed to the Kendall's τ measure of dependence as defined in 2.14 was $\tau = 0.16$. Recall that the true range of τ is dependent on the marginal means and the bounds do not attain ± 1 (see Section 2.4.4). The true range at this value of the logMAADT and given the estimates of the regression parameters is $[0, 0.64]$, since the Clayton copula only admits a positive dependence structure. Therefore, a $\tau = 0.16$ is one forth of the highest dependence that can be reached under this value of logMAADT.

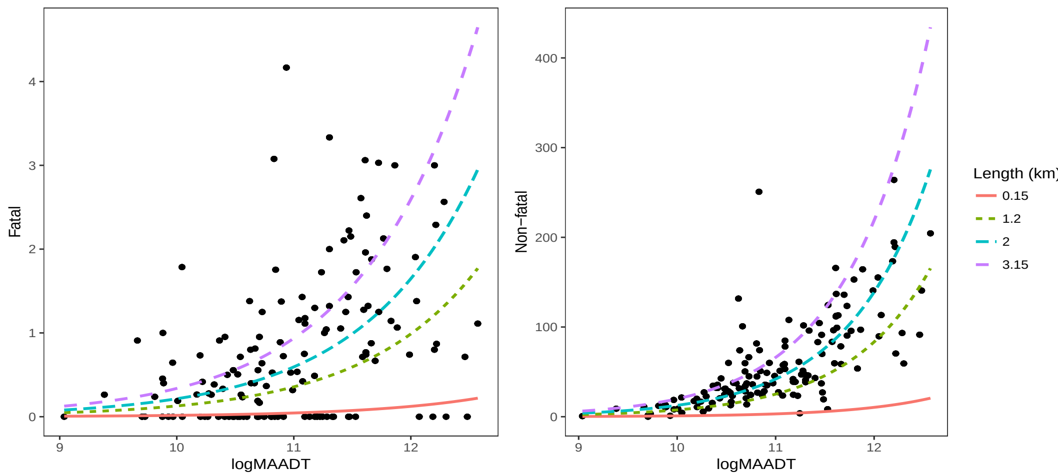

Figure 4–2: The vehicle crashes data in 141 segments of highways in Puerto Rico by logMAADT. Each curve represent the fitted values assuming different lengths of highway segments.

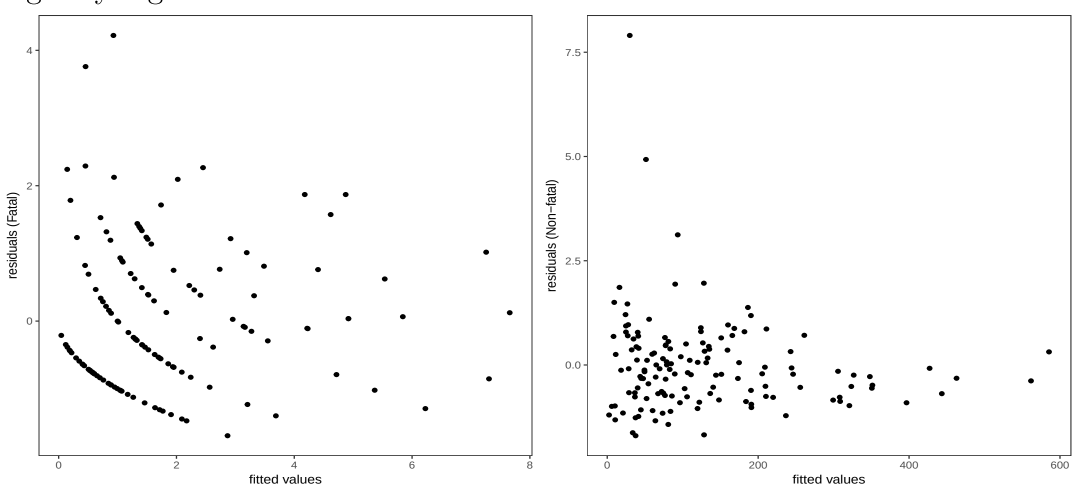

Figure 4–3: Residuals by the Clayton copula model for the number of fatal and non-fatal crashes. $(N = 141)$

The Clayton copula-based regression model fitted to the 141 highway segments provides the curve for different segment lengths of the observed data given the covariate logMAADT (Figure 4–2). The residual plots in Figure 4–3 indicates that there are some outliers in this dataset. In that regard, we tried fitting the models dropping them but more outliers appeared according to different criteria. This suggests that more exploration of these data points is required and more covariates, such as speed limit and lighting, are needed to obtain a better fit. Unfortunately, we did not have additional significant information to include in this application.

Table 4–3: Estimated parameters (SE), AIC, BIC, and loglikelihood, under the Poisson lognormal model for the crash data.

		Poisson lognormal
μ_{iFatal}	Intercept	-30.08 (31.33)
	logMAADT	4.58 (5.67)
	logMAADT^2	-0.17 (0.25)
$\mu_{iNon\text{-}Fatal}$	Intercept	-40.02 (9.78)
	logMAADT	6.87 (1.78)
	logMAADT^2	-0.26 (0.08)
Σ	$\sigma^2_{b_{Fatal}}$	0.06 (0.34)
	$\rho\sigma_{b_{Fatal}}\sigma_{b_{Non-Fatal}}$	0.13 (0.05)
	$\sigma^2_{b_{Non-Fatal}}$	0.30 (0.04)
	AIC	1892.3
	BIC	1918.9
	loglik	-937.15

For comparison purposes, the estimates obtained by fitting a Poisson lognormal model as described in 2.3 are presented in Table 4–3. According to AIC/BIC, this model gave a slightly better fit than the univariate (Independent) approach. One can see that the signs of the coefficients agree. However, the parameters are not comparable since this is a subject-specific effects model and so are their interpretation. For example, $\exp(\boldsymbol{x}^t\boldsymbol{\beta})$ can be interpreted as the incidence rate ratio for a change of one unit in the logMAADT given the same segment. In that sense, another advantage of copula-based regression models is that the parameters are already given as marginal effects and their interpretation is based on the population average.

4.4 Regression to the copula parameter

As proposed by the authors in [30] we introduced a regression part for the copula parameter θ with logMAADT as a covariate using the Clayton and Frank copula models to get a better understanding of how the logMAADT affects the association between fatal and non-fatal crashes. The results obtained by the Clayton and Frank copula are presented in Table 4–4. According to both models, the dependence between fatal and non-fatal crashes was weaker for higher logMAADT; although the covariate logMAADT was not statistically significant for the copula parameter in the Frank copula model. For instance, the point estimate for the copula parameter θ obtained from the best model (Clayton copula model) was 13.24 ($\tau = 0.72$) for the minimum observed logMAADT and 0.02 ($\tau = 0$) for the maximum, which suggest that the dependence between the number of fatal and non-fatal crashes on highway segments is lower on busier highways. Moreover, note that the AIC/BIC was slightly lower for the models when introducing the covariate in the copula parameter.

While we recognize the limitations of this dataset, it is a simple and good example to comply with the objective of illustrating the flexibility and ease of use of copula-based regression models for bivariate count data.

Table 4–4: Estimated parameters (SE), AIC, BIC, and loglikelihood, under the Clayton and Frank copula models for the crash data including a regression part for the copula parameter θ.

		Clayton	Frank
	Intercept	-49.85 (13.25)	-50.16 (4.3)
μ_{iFatal}	logMAADT	8.09 (2.37)	8.18 (0.78)
	logMAADT^2	-0.33 (0.11)	-0.33 (0.04)
	Intercept	-46.43 (7.58)	-46.25 (21.61)
$\mu_{iNon\text{-}Fatal}$	logMAADT	8.07 (1.38)	8.05 (3.89)
	logMAADT^2	-0.32 (0.06)	-0.32 (0.17)
	ϕ_2	3.54 (0.48)	3.48 (1.23)
$s(\theta)$	Intercept	19.79 (6.13)	17.73 (21.1)
	logMAADT	-1.9 (0.57)	-1.45 (1.92)
	AIC	1882.23	1887.66
	BIC	1908.77	1914.2
	loglik	-932.11	-934.83

CHAPTER 5
CONCLUSIONS AND FUTURE WORK

The main purpose of this work was to provide researchers a guide on when copula-based regression models could be considered to model bivariate count data. Also we wanted to provide an overview of the benefits of accounting for dependence between counts using copulas instead of modeling counts independently for different scenarios. In this context, a scenario where the counts had different means does not appear in the literature. Simulation studies and an application to fatal and non-fatal vehicle crashes were included to show the performance of our suggested approach when bivariate counts have different means.

Overall, copula-based regression models are very flexible by admitting different dependence structures and allowing different marginals for each response variable. This method of modeling correlated bivariate counts provides more efficient estimators for the regression parameters in comparison to the independent approach when adequate copula models are fitted, even when the data exhibits a low degree of association. However, caution should be taken in the selection of the set of candidate copulas since biased estimates can be produced by copulas with a dependence structure far from the true association. Previous exploration of the data is crucial in that regard. Also, traditional penalized likelihood-based criteria, such as AIC and BIC, have a fairly good performance in selecting the true model among a set of candidate copula models, yet slightly better when both marginal distributions had large means. Under the scenarios considered in this work, the estimates of the copula parameter seemed to be reasonable except when $N = 100$ where bias was observed. Simulation studies with sample sizes between $N = 100$ and $N = 300$ are needed to establish if copula models should be used for sample sizes less than 300 if one is interested in the interpretation of the copula parameter. In that sense, whether the estimate of the copula parameter in our application is biased or not is unclear since $N = 141$. We suggest a sample size of $N = 300$ or higher for the use of the copula approach for count data. As a last note, interpretation of the copula parameter about the dependence structure should be made carefully since the range of its transformation to a dependence measure may be narrower than $[-1, 1]$ depending on the copula, even for large marginal means. As shown in the application, another advantage of copula-based regression models is the marginal interpretation of the parameters compared to other parametric models where the interpretation is subject specific.

A long list of copula families exist in the literature, including nonparametric copulas. Therefore, future work related to this research should include other copula families and extensions to three or more counts. Also, a Bayesian approach to copula modeling for count data could be explored given the complexity of some copula models such as the Normal copula, in particular when covariates are introduced into the copula parameter. Inclusion of other goodness-of-fit and selection criteria available in the literature [12] for copula-based models besides likelihood-based criteria and MSE should be considered for comparison purposes.

APPENDICES

APPENDIX A
STUDY 1 AVERAGE BIAS OF ESTIMATES AND MONTE CARLO STANDARD ERROR (MCSE)

Table A–1: Average bias of estimates and Monte Carlo Standard Error (MCSE) based on 1,000 simulations.

		True Copula: Clayton				
		Clayton	Frank	Galambos	Independent	Normal
	$\tau = 0.3$					
	a_1	**0.024 (0.039)**	0.100 (0.041)	0.193 (0.041)	-0.006 (0.050)	0.191 (0.040)
	b_1	**-0.003 (0.004)**	-0.013 (0.004)	-0.018 (0.004)	-0.000 (0.005)	-0.020 (0.004)
	ϕ_1	**0.082 (0.010)**	0.006 (0.011)	-0.009 (0.011)	0.092 (0.010)	0.068 (0.011)
	a_2	**0.021 (0.039)**	0.087 (0.044)	0.259 (0.042)	0.016 (0.049)	0.206 (0.038)
	b_2	**-0.003 (0.004)**	-0.012 (0.004)	-0.024 (0.004)	-0.002 (0.005)	-0.021 (0.004)
	ϕ_2	**0.089 (0.011)**	0.011 (0.011)	-0.003 (0.011)	0.099 (0.011)	0.067 (0.011)
	$\tau = 0.5$					
$N = 100$	a_1	**-0.015 (0.031)**	-0.082 (0.030)	0.164 (0.038)	-0.032 (0.043)	0.106 (0.033)
	b_1	**0.001 (0.003)**	0.001 (0.003)	-0.012 (0.004)	0.003 (0.004)	-0.010 (0.003)
	ϕ_1	**0.070 (0.009)**	-0.114 (0.010)	-0.132 (0.010)	0.079 (0.010)	0.017 (0.010)
	a_2	**0.017 (0.031)**	-0.049 (0.031)	0.143 (0.040)	-0.043 (0.043)	0.073 (0.033)
	b_2	**-0.002 (0.003)**	-0.002 (0.003)	-0.009 (0.004)	0.004 (0.004)	-0.007 (0.003)
	ϕ_2	**0.059 (0.009)**	-0.120 (0.009)	-0.141 (0.009)	0.069 (0.009)	0.004 (0.009)
	$\tau = 0.7$					
	a_1	**0.020 (0.027)**	-0.079 (0.030)	0.179 (0.041)	0.082 (0.048)	0.075 (0.035)
	b_1	**-0.002 (0.003)**	-0.006 (0.003)	-0.010 (0.004)	-0.008 (0.005)	-0.006 (0.003)
	ϕ_1	**0.122 (0.010)**	-0.222 (0.010)	-0.246 (0.010)	0.137 (0.011)	0.023 (0.011)
	a_2	**0.028 (0.027)**	-0.083 (0.029)	0.177 (0.040)	0.096 (0.046)	0.096 (0.034)
	b_2	**-0.003 (0.003)**	-0.006 (0.003)	-0.010 (0.004)	-0.010 (0.005)	-0.008 (0.003)
	ϕ_2	**0.122 (0.010)**	-0.222 (0.010)	-0.256 (0.010)	0.132 (0.011)	0.019 (0.011)
	$\tau = 0.3$					
	a_1	**0.033 (0.021)**	0.093 (0.026)	0.254 (0.024)	0.052 (0.026)	0.281 (0.022)
	b_1	**-0.003 (0.002)**	-0.012 (0.003)	-0.023 (0.002)	-0.005 (0.003)	-0.028 (0.002)
	ϕ_1	**0.025 (0.005)**	-0.051 (0.006)	-0.071 (0.006)	0.027 (0.006)	0.014 (0.006)
	a_2	**0.021 (0.021)**	0.070 (0.026)	0.236 (0.024)	0.052 (0.026)	0.289 (0.022)
	b_2	**-0.002 (0.002)**	-0.009 (0.003)	-0.021 (0.002)	-0.005 (0.003)	-0.028 (0.002)
	ϕ_2	**0.025 (0.005)**	-0.051 (0.006)	-0.070 (0.006)	0.027 (0.005)	0.013 (0.006)
	$\tau = 0.5$					
$N = 300$	a_1	**-0.020 (0.019)**	-0.032 (0.019)	0.210 (0.023)	-0.003 (0.026)	0.149 (0.021)
	b_1	**0.002 (0.002)**	-0.004 (0.002)	-0.015 (0.002)	0.000 (0.003)	-0.014 (0.002)
	ϕ_1	**0.029 (0.005)**	-0.159 (0.006)	-0.183 (0.005)	0.028 (0.005)	-0.030 (0.005)
	a_2	**-0.019 (0.019)**	-0.038 (0.018)	0.212 (0.022)	-0.012 (0.025)	0.156 (0.021)
	b_2	**0.002 (0.002)**	-0.003 (0.002)	-0.016 (0.002)	0.001 (0.003)	-0.015 (0.002)
	ϕ_2	**0.035 (0.005)**	-0.153 (0.006)	-0.174 (0.005)	0.037 (0.006)	-0.023 (0.006)
	$\tau = 0.7$					
	a_1	**0.009 (0.015)**	-0.104 (0.014)	0.173 (0.021)	0.024 (0.024)	0.045 (0.015)
	b_1	**-0.001 (0.001)**	-0.004 (0.001)	-0.008 (0.002)	-0.002 (0.002)	-0.003 (0.001)
	ϕ_1	**0.042 (0.005)**	-0.306 (0.005)	-0.335 (0.005)	0.045 (0.006)	-0.060 (0.006)
	a_2	**0.021 (0.015)**	-0.095 (0.014)	0.187 (0.022)	0.021 (0.025)	0.045 (0.016)
	b_2	**-0.002 (0.002)**	-0.005 (0.001)	-0.010 (0.002)	-0.002 (0.002)	-0.003 (0.002)
	ϕ_2	**0.042 (0.005)**	-0.309 (0.005)	-0.338 (0.005)	0.041 (0.006)	-0.065 (0.006)
	$\tau = 0.3$					
	a_1	**0.020 (0.017)**	0.066 (0.020)	0.194 (0.020)	0.023 (0.020)	0.289 (0.019)
	b_1	**-0.002 (0.002)**	-0.009 (0.002)	-0.017 (0.002)	-0.003 (0.002)	-0.029 (0.002)
	ϕ_1	**0.009 (0.004)**	-0.064 (0.004)	-0.083 (0.004)	0.011 (0.004)	0.002 (0.004)
	a_2	**0.002 (0.017)**	0.045 (0.021)	0.167 (0.021)	-0.003 (0.020)	0.311 (0.019)
	b_2	**-0.000 (0.002)**	-0.007 (0.002)	-0.014 (0.002)	0.000 (0.002)	-0.031 (0.002)
	ϕ_2	**0.007 (0.004)**	-0.067 (0.004)	-0.086 (0.004)	0.007 (0.004)	-0.008 (0.004)
	$\tau = 0.5$					
$N = 500$	a_1	**0.004 (0.015)**	-0.004 (0.013)	0.264 (0.017)	0.025 (0.019)	0.183 (0.015)
	b_1	**-0.001 (0.001)**	-0.007 (0.001)	-0.021 (0.002)	-0.003 (0.002)	-0.018 (0.002)
	ϕ_1	**0.017 (0.004)**	-0.170 (0.004)	-0.192 (0.004)	0.017 (0.004)	-0.041 (0.004)
	a_2	**0.015 (0.014)**	-0.026 (0.013)	0.240 (0.016)	-0.008 (0.018)	0.173 (0.015)
	b_2	**-0.002 (0.001)**	-0.004 (0.001)	-0.019 (0.002)	0.001 (0.002)	-0.017 (0.001)
	ϕ_2	**0.019 (0.004)**	-0.169 (0.004)	-0.188 (0.004)	0.020 (0.004)	-0.039 (0.004)
	$\tau = 0.7$					
	a_1	**0.008 (0.014)**	-0.066 (0.013)	0.175 (0.019)	0.016 (0.021)	0.059 (0.014)
	b_1	**-0.001 (0.001)**	-0.008 (0.001)	-0.009 (0.002)	-0.002 (0.002)	-0.005 (0.001)
	ϕ_1	**0.019 (0.004)**	-0.330 (0.004)	-0.357 (0.004)	0.021 (0.004)	-0.084 (0.004)
	a_2	**0.003 (0.013)**	-0.081 (0.013)	0.162 (0.019)	0.006 (0.020)	0.051 (0.015)
	b_2	**-0.001 (0.001)**	-0.007 (0.001)	-0.008 (0.002)	-0.001 (0.002)	-0.004 (0.001)
	ϕ_2	**0.022 (0.004)**	-0.329 (0.004)	-0.358 (0.004)	0.021 (0.004)	-0.083 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table A–2: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Frank				
		Clayton	**Frank**	Galambos	Independent	Normal
$N = 100$	$\tau = 0.3$					
	a_1	0.250 (0.034)	**-0.001 (0.032)**	0.081 (0.035)	0.012 (0.046)	0.107 (0.032)
	b_1	-0.021 (0.003)	**-0.000 (0.003)**	-0.006 (0.004)	-0.002 (0.005)	-0.011 (0.003)
	ϕ_1	0.061 (0.010)	**0.086 (0.009)**	-0.078 (0.010)	0.094 (0.010)	0.121 (0.010)
	a_2	0.265 (0.036)	**0.008 (0.032)**	0.121 (0.037)	0.034 (0.046)	0.084 (0.033)
	b_2	-0.023 (0.004)	**-0.001 (0.003)**	-0.009 (0.004)	-0.004 (0.005)	-0.009 (0.003)
	ϕ_2	0.075 (0.010)	**0.096 (0.010)**	-0.073 (0.010)	0.102 (0.010)	0.131 (0.010)
	$\tau = 0.5$					
	a_1	0.261 (0.033)	**0.024 (0.034)**	0.116 (0.035)	-0.004 (0.043)	0.103 (0.033)
	b_1	-0.024 (0.003)	**-0.003 (0.003)**	-0.009 (0.004)	-0.000 (0.004)	-0.011 (0.003)
	ϕ_1	0.061 (0.010)	**0.068 (0.010)**	-0.057 (0.009)	0.072 (0.010)	0.089 (0.010)
	a_2	0.315 (0.032)	**0.025 (0.033)**	0.127 (0.036)	-0.003 (0.043)	0.121 (0.032)
	b_2	-0.029 (0.003)	**-0.003 (0.003)**	-0.010 (0.004)	0.000 (0.004)	-0.012 (0.003)
	ϕ_2	0.053 (0.010)	**0.067 (0.009)**	-0.056 (0.009)	0.070 (0.009)	0.085 (0.009)
	$\tau = 0.7$					
	a_1	0.207 (0.031)	**0.058 (0.026)**	0.089 (0.030)	0.080 (0.043)	0.046 (0.026)
	b_1	-0.015 (0.003)	**-0.006 (0.003)**	-0.005 (0.003)	-0.008 (0.004)	-0.005 (0.003)
	ϕ_1	0.080 (0.011)	**0.095 (0.010)**	-0.129 (0.009)	0.105 (0.011)	0.141 (0.010)
	a_2	0.205 (0.031)	**0.073 (0.025)**	0.091 (0.028)	0.094 (0.042)	0.046 (0.025)
	b_2	-0.015 (0.003)	**-0.008 (0.002)**	-0.005 (0.003)	-0.010 (0.004)	-0.005 (0.002)
	ϕ_2	0.071 (0.010)	**0.098 (0.009)**	-0.140 (0.009)	0.099 (0.010)	0.136 (0.010)
$N = 300$	$\tau = 0.3$					
	a_1	0.392 (0.024)	**0.036 (0.022)**	0.163 (0.023)	0.035 (0.026)	0.242 (0.022)
	b_1	-0.038 (0.002)	**-0.004 (0.002)**	-0.015 (0.002)	-0.004 (0.003)	-0.024 (0.002)
	ϕ_1	0.023 (0.006)	**0.026 (0.006)**	-0.036 (0.006)	0.027 (0.006)	0.042 (0.006)
	a_2	0.454 (0.024)	**0.042 (0.021)**	0.179 (0.023)	0.058 (0.026)	0.279 (0.023)
	b_2	-0.044 (0.002)	**-0.004 (0.002)**	-0.016 (0.002)	-0.006 (0.003)	-0.028 (0.002)
	ϕ_2	0.023 (0.006)	**0.024 (0.005)**	-0.039 (0.005)	0.024 (0.005)	0.039 (0.006)
	$\tau = 0.5$					
	a_1	0.363 (0.020)	**0.046 (0.019)**	0.158 (0.019)	0.040 (0.025)	0.151 (0.020)
	b_1	-0.033 (0.002)	**-0.005 (0.002)**	-0.013 (0.002)	-0.004 (0.003)	-0.015 (0.002)
	ϕ_1	0.025 (0.006)	**0.028 (0.005)**	-0.103 (0.005)	0.028 (0.006)	0.045 (0.006)
	a_2	0.356 (0.019)	**0.016 (0.019)**	0.149 (0.020)	0.019 (0.025)	0.127 (0.019)
	b_2	-0.033 (0.002)	**-0.002 (0.002)**	-0.012 (0.002)	-0.002 (0.002)	-0.013 (0.002)
	ϕ_2	0.028 (0.006)	**0.034 (0.006)**	-0.096 (0.005)	0.035 (0.006)	0.052 (0.006)
	$\tau = 0.7$					
	a_1	0.162 (0.017)	**0.003 (0.012)**	0.060 (0.013)	0.014 (0.022)	0.010 (0.012)
	b_1	-0.010 (0.002)	**-0.000 (0.001)**	-0.002 (0.001)	-0.001 (0.002)	-0.001 (0.001)
	ϕ_1	-0.004 (0.006)	**0.020 (0.005)**	-0.214 (0.005)	0.022 (0.006)	0.067 (0.006)
	a_2	0.191 (0.016)	**0.016 (0.012)**	0.074 (0.014)	0.029 (0.022)	0.019 (0.012)
	b_2	-0.013 (0.002)	**-0.002 (0.001)**	-0.003 (0.001)	-0.003 (0.002)	-0.002 (0.001)
	ϕ_2	-0.003 (0.006)	**0.019 (0.005)**	-0.216 (0.005)	0.018 (0.005)	0.063 (0.005)
$N = 500$	$\tau = 0.3$					
	a_1	0.418 (0.020)	**0.027 (0.017)**	0.201 (0.018)	0.019 (0.020)	0.232 (0.019)
	b_1	-0.041 (0.002)	**-0.003 (0.002)**	-0.019 (0.002)	-0.002 (0.002)	-0.023 (0.002)
	ϕ_1	0.003 (0.004)	**0.009 (0.004)**	-0.053 (0.004)	0.008 (0.004)	0.028 (0.004)
	a_2	0.421 (0.021)	**0.012 (0.017)**	0.161 (0.017)	-0.007 (0.020)	0.290 (0.019)
	b_2	-0.041 (0.002)	**-0.001 (0.002)**	-0.015 (0.002)	0.001 (0.002)	-0.029 (0.002)
	ϕ_2	0.005 (0.004)	**0.009 (0.004)**	-0.051 (0.004)	0.009 (0.004)	0.029 (0.004)
	$\tau = 0.5$					
	a_1	0.343 (0.015)	**0.018 (0.014)**	0.170 (0.014)	0.019 (0.018)	0.178 (0.014)
	b_1	-0.032 (0.001)	**-0.002 (0.001)**	-0.014 (0.001)	-0.002 (0.002)	-0.018 (0.001)
	ϕ_1	0.014 (0.004)	**0.016 (0.004)**	-0.113 (0.004)	0.016 (0.004)	0.037 (0.004)
	a_2	0.348 (0.015)	**0.005 (0.014)**	0.135 (0.015)	-0.010 (0.018)	0.159 (0.014)
	b_2	-0.032 (0.002)	**-0.001 (0.001)**	-0.011 (0.001)	0.001 (0.002)	-0.016 (0.001)
	ϕ_2	0.012 (0.004)	**0.016 (0.004)**	-0.110 (0.004)	0.017 (0.004)	0.039 (0.004)
	$\tau = 0.7$					
	a_1	0.199 (0.014)	**0.008 (0.009)**	0.088 (0.013)	0.022 (0.018)	0.018 (0.010)
	b_1	-0.014 (0.001)	**-0.001 (0.001)**	-0.004 (0.001)	-0.002 (0.002)	-0.002 (0.001)
	ϕ_1	-0.012 (0.004)	**0.013 (0.004)**	-0.218 (0.004)	0.014 (0.004)	0.060 (0.004)
	a_2	0.196 (0.013)	**0.002 (0.010)**	0.078 (0.012)	0.008 (0.018)	0.005 (0.010)
	b_2	-0.013 (0.001)	**-0.000 (0.001)**	-0.003 (0.001)	-0.001 (0.002)	-0.001 (0.001)
	ϕ_2	-0.005 (0.004)	**0.016 (0.004)**	-0.218 (0.004)	0.016 (0.004)	0.062 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table A–3: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Galambos				
		Clayton	Frank	**Galambos**	Independent	Normal
N = 100	$\tau = 0.3$					
	a_1	0.052 (0.041)	0.059 (0.040)	**0.035 (0.035)**	0.036 (0.046)	0.049 (0.038)
	b_1	-0.004 (0.004)	-0.004 (0.004)	**-0.004 (0.004)**	-0.004 (0.005)	-0.005 (0.004)
	ϕ_1	0.059 (0.010)	0.078 (0.010)	**0.085 (0.010)**	0.089 (0.010)	0.094 (0.010)
	a_2	0.048 (0.039)	0.045 (0.038)	**-0.011 (0.036)**	0.020 (0.045)	0.009 (0.036)
	b_2	-0.004 (0.004)	-0.003 (0.004)	**0.001 (0.004)**	-0.002 (0.004)	-0.001 (0.004)
	ϕ_2	0.041 (0.010)	0.057 (0.010)	**0.063 (0.010)**	0.071 (0.010)	0.076 (0.010)
	$\tau = 0.5$					
	a_1	0.010 (0.036)	-0.011 (0.035)	**0.012 (0.027)**	-0.034 (0.042)	-0.054 (0.028)
	b_1	0.002 (0.004)	0.005 (0.003)	**-0.002 (0.003)**	0.003 (0.004)	0.005 (0.003)
	ϕ_1	0.025 (0.010)	0.038 (0.010)	**0.061 (0.009)**	0.070 (0.010)	0.096 (0.010)
	a_2	0.006 (0.038)	-0.024 (0.036)	**0.015 (0.027)**	-0.023 (0.042)	-0.063 (0.029)
	b_2	0.002 (0.004)	0.006 (0.004)	**-0.002 (0.003)**	0.002 (0.004)	0.006 (0.003)
	ϕ_2	0.022 (0.010)	0.036 (0.010)	**0.067 (0.010)**	0.068 (0.010)	0.093 (0.010)
	$\tau = 0.7$					
	a_1	0.005 (0.035)	-0.088 (0.040)	**-0.019 (0.018)**	-0.084 (0.043)	-0.105 (0.025)
	b_1	0.004 (0.004)	0.015 (0.004)	**0.000 (0.002)**	0.007 (0.004)	0.008 (0.003)
	ϕ_1	0.027 (0.011)	0.015 (0.011)	**0.101 (0.010)**	0.111 (0.011)	0.164 (0.011)
	a_2	-0.034 (0.035)	-0.131 (0.040)	**-0.036 (0.018)**	-0.105 (0.043)	-0.126 (0.025)
	b_2	0.008 (0.003)	0.019 (0.004)	**0.002 (0.002)**	0.009 (0.004)	0.011 (0.003)
	ϕ_2	0.029 (0.011)	0.012 (0.010)	**0.105 (0.010)**	0.115 (0.011)	0.167 (0.011)
N = 300	$\tau = 0.3$					
	a_1	0.066 (0.022)	0.023 (0.022)	**0.008 (0.019)**	-0.000 (0.026)	0.029 (0.021)
	b_1	-0.006 (0.002)	-0.001 (0.002)	**-0.001 (0.002)**	-0.000 (0.003)	-0.003 (0.002)
	ϕ_1	0.010 (0.006)	0.023 (0.006)	**0.030 (0.006)**	0.033 (0.006)	0.043 (0.006)
	a_2	0.025 (0.021)	0.004 (0.022)	**-0.021 (0.020)**	-0.030 (0.025)	0.015 (0.021)
	b_2	-0.001 (0.002)	0.001 (0.002)	**0.002 (0.002)**	0.003 (0.003)	-0.002 (0.002)
	ϕ_2	0.003 (0.006)	0.020 (0.005)	**0.026 (0.005)**	0.028 (0.005)	0.039 (0.005)
	$\tau = 0.5$					
	a_1	0.012 (0.021)	-0.022 (0.021)	**-0.013 (0.016)**	-0.025 (0.024)	-0.027 (0.017)
	b_1	0.002 (0.002)	0.006 (0.002)	**0.001 (0.002)**	0.003 (0.002)	0.002 (0.002)
	ϕ_1	-0.011 (0.006)	-0.002 (0.005)	**0.027 (0.005)**	0.028 (0.005)	0.060 (0.005)
	a_2	0.042 (0.020)	-0.005 (0.021)	**0.001 (0.014)**	-0.009 (0.023)	-0.010 (0.017)
	b_2	-0.001 (0.002)	0.005 (0.002)	**-0.000 (0.001)**	0.001 (0.002)	0.001 (0.002)
	ϕ_2	-0.025 (0.006)	-0.014 (0.005)	**0.015 (0.005)**	0.015 (0.005)	0.047 (0.005)
	$\tau = 0.7$					
	a_1	0.013 (0.019)	-0.087 (0.027)	**0.008 (0.008)**	0.001 (0.023)	-0.115 (0.014)
	b_1	0.005 (0.002)	0.016 (0.003)	**-0.001 (0.001)**	-0.001 (0.002)	0.010 (0.001)
	ϕ_1	-0.050 (0.006)	-0.063 (0.006)	**0.033 (0.005)**	0.036 (0.006)	0.095 (0.006)
	a_2	0.000 (0.019)	-0.094 (0.029)	**-0.000 (0.009)**	-0.010 (0.023)	-0.123 (0.015)
	b_2	0.006 (0.002)	0.017 (0.003)	**-0.000 (0.001)**	0.000 (0.002)	0.011 (0.001)
	ϕ_2	-0.054 (0.006)	-0.066 (0.006)	**0.029 (0.005)**	0.033 (0.006)	0.091 (0.006)
N = 500	$\tau = 0.3$					
	a_1	0.041 (0.018)	0.035 (0.018)	**-0.003 (0.015)**	-0.008 (0.020)	0.005 (0.016)
	b_1	-0.003 (0.002)	-0.002 (0.002)	**0.000 (0.002)**	0.001 (0.002)	-0.001 (0.002)
	ϕ_1	-0.012 (0.004)	0.004 (0.004)	**0.012 (0.004)**	0.012 (0.004)	0.023 (0.004)
	a_2	0.052 (0.016)	0.046 (0.017)	**0.003 (0.015)**	-0.002 (0.019)	0.005 (0.015)
	b_2	-0.004 (0.002)	-0.003 (0.002)	**-0.000 (0.002)**	0.000 (0.002)	-0.001 (0.002)
	ϕ_2	-0.014 (0.004)	0.001 (0.004)	**0.009 (0.004)**	0.010 (0.004)	0.021 (0.004)
	$\tau = 0.5$					
	a_1	0.044 (0.016)	0.007 (0.016)	**0.005 (0.012)**	0.006 (0.018)	-0.024 (0.013)
	b_1	-0.001 (0.002)	0.003 (0.002)	**-0.001 (0.001)**	-0.001 (0.002)	0.002 (0.001)
	ϕ_1	-0.016 (0.004)	-0.008 (0.004)	**0.018 (0.004)**	0.021 (0.004)	0.053 (0.004)
	a_2	0.048 (0.016)	0.028 (0.016)	**0.008 (0.011)**	0.013 (0.018)	-0.017 (0.012)
	b_2	-0.002 (0.002)	0.001 (0.002)	**-0.001 (0.001)**	-0.001 (0.002)	0.001 (0.001)
	ϕ_2	-0.027 (0.004)	-0.015 (0.004)	**0.012 (0.004)**	0.012 (0.004)	0.044 (0.004)
	$\tau = 0.7$					
	a_1	0.012 (0.015)	-0.097 (0.027)	**0.005 (0.006)**	-0.000 (0.019)	-0.103 (0.014)
	b_1	0.005 (0.001)	0.017 (0.003)	**-0.001 (0.001)**	-0.000 (0.002)	0.009 (0.001)
	ϕ_1	-0.060 (0.005)	-0.074 (0.004)	**0.023 (0.004)**	0.027 (0.005)	0.084 (0.005)
	a_2	0.017 (0.016)	-0.108 (0.027)	**0.014 (0.006)**	0.004 (0.019)	-0.099 (0.014)
	b_2	0.005 (0.002)	0.018 (0.003)	**-0.002 (0.001)**	-0.001 (0.002)	0.009 (0.001)
	ϕ_2	-0.055 (0.004)	-0.071 (0.004)	**0.028 (0.004)**	0.032 (0.004)	0.090 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table A–4: [STUDY 1: Average bias of estimates and Monte Carlo Standard Error (MCSE) for Normal copula.]Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Normal				
		Clayton	Frank	Galambos	Independent	**Normal**
$N = 100$	$\tau = 0.3$					
	a_1	0.325 (0.036)	0.278 (0.042)	0.172 (0.034)	0.017 (0.043)	**0.009 (0.036)**
	b_1	-0.032 (0.004)	-0.028 (0.004)	-0.016 (0.003)	-0.002 (0.004)	**-0.001 (0.004)**
	ϕ_1	0.075 (0.010)	0.052 (0.010)	0.031 (0.010)	0.090 (0.010)	**0.084 (0.010)**
	a_2	0.345 (0.036)	0.270 (0.044)	0.198 (0.035)	0.056 (0.043)	**0.019 (0.034)**
	b_2	-0.034 (0.004)	-0.027 (0.004)	-0.019 (0.003)	-0.006 (0.004)	**-0.002 (0.003)**
	ϕ_2	0.057 (0.010)	0.037 (0.010)	0.021 (0.010)	0.075 (0.010)	**0.069 (0.010)**
	$\tau = 0.5$					
	a_1	0.146 (0.034)	-0.004 (0.033)	0.065 (0.032)	-0.074 (0.043)	**-0.077 (0.033)**
	b_1	-0.012 (0.003)	0.001 (0.003)	-0.005 (0.003)	0.007 (0.004)	**0.007 (0.003)**
	ϕ_1	0.043 (0.010)	0.004 (0.010)	-0.005 (0.010)	0.089 (0.010)	**0.081 (0.010)**
	a_2	0.229 (0.035)	0.081 (0.031)	0.184 (0.032)	0.028 (0.044)	**0.001 (0.032)**
	b_2	-0.021 (0.004)	-0.008 (0.003)	-0.017 (0.003)	-0.003 (0.004)	**-0.001 (0.003)**
	ϕ_2	0.059 (0.010)	-0.009 (0.009)	-0.025 (0.009)	0.081 (0.009)	**0.074 (0.009)**
	$\tau = 0.7$					
	a_1	0.121 (0.033)	0.044 (0.026)	0.035 (0.025)	0.051 (0.042)	**0.009 (0.023)**
	b_1	-0.007 (0.003)	-0.004 (0.003)	-0.001 (0.003)	-0.006 (0.004)	**-0.002 (0.002)**
	ϕ_1	0.042 (0.011)	-0.067 (0.010)	-0.038 (0.010)	0.136 (0.011)	**0.123 (0.011)**
	a_2	0.128 (0.034)	0.062 (0.026)	0.076 (0.025)	0.085 (0.042)	**0.024 (0.024)**
	b_2	-0.008 (0.003)	-0.005 (0.003)	-0.005 (0.002)	-0.009 (0.004)	**-0.003 (0.002)**
	ϕ_2	0.033 (0.011)	-0.070 (0.011)	-0.043 (0.010)	0.129 (0.011)	**0.115 (0.011)**
$N = 300$	$\tau = 0.3$					
	a_1	0.369 (0.023)	0.269 (0.031)	0.160 (0.022)	0.016 (0.026)	**0.031 (0.022)**
	b_1	-0.036 (0.002)	-0.027 (0.003)	-0.015 (0.002)	-0.002 (0.003)	**-0.003 (0.002)**
	ϕ_1	0.008 (0.006)	-0.008 (0.006)	-0.027 (0.006)	0.024 (0.006)	**0.022 (0.006)**
	a_2	0.358 (0.023)	0.257 (0.031)	0.173 (0.023)	0.022 (0.026)	**0.038 (0.021)**
	b_2	-0.035 (0.002)	-0.026 (0.003)	-0.016 (0.002)	-0.002 (0.003)	**-0.004 (0.002)**
	ϕ_2	0.012 (0.006)	-0.008 (0.005)	-0.024 (0.005)	0.026 (0.005)	**0.024 (0.005)**
	$\tau = 0.5$					
	a_1	0.294 (0.018)	0.059 (0.022)	0.185 (0.018)	0.059 (0.024)	**0.051 (0.018)**
	b_1	-0.027 (0.002)	-0.005 (0.002)	-0.016 (0.002)	-0.006 (0.002)	**-0.005 (0.002)**
	ϕ_1	-0.000 (0.006)	-0.058 (0.006)	-0.070 (0.005)	0.027 (0.006)	**0.024 (0.006)**
	a_2	0.272 (0.018)	0.055 (0.022)	0.168 (0.018)	0.040 (0.025)	**0.036 (0.019)**
	b_2	-0.025 (0.002)	-0.005 (0.002)	-0.015 (0.002)	-0.004 (0.002)	**-0.004 (0.002)**
	ϕ_2	-0.001 (0.006)	-0.059 (0.006)	-0.075 (0.006)	0.025 (0.006)	**0.022 (0.006)**
	$\tau = 0.7$					
	a_1	0.103 (0.019)	0.032 (0.012)	0.054 (0.012)	0.035 (0.023)	**0.008 (0.012)**
	b_1	-0.004 (0.002)	-0.001 (0.001)	-0.002 (0.001)	-0.004 (0.002)	**-0.001 (0.001)**
	ϕ_1	-0.059 (0.006)	-0.166 (0.006)	-0.135 (0.006)	0.032 (0.006)	**0.028 (0.006)**
	a_2	0.087 (0.019)	0.018 (0.013)	0.051 (0.012)	0.034 (0.023)	**0.002 (0.012)**
	b_2	-0.003 (0.002)	-0.000 (0.001)	-0.002 (0.001)	-0.003 (0.002)	**-0.000 (0.001)**
	ϕ_2	-0.060 (0.006)	-0.166 (0.006)	-0.138 (0.005)	0.031 (0.006)	**0.027 (0.006)**
$N = 500$	$\tau = 0.3$					
	a_1	0.330 (0.019)	0.227 (0.030)	0.174 (0.019)	0.012 (0.020)	**0.018 (0.017)**
	b_1	-0.032 (0.002)	-0.023 (0.003)	-0.016 (0.002)	-0.001 (0.002)	**-0.002 (0.002)**
	ϕ_1	-0.003 (0.004)	-0.021 (0.004)	-0.041 (0.004)	0.011 (0.004)	**0.010 (0.004)**
	a_2	0.359 (0.019)	0.266 (0.030)	0.209 (0.019)	0.042 (0.021)	**0.023 (0.018)**
	b_2	-0.035 (0.002)	-0.026 (0.003)	-0.020 (0.002)	-0.005 (0.002)	**-0.003 (0.002)**
	ϕ_2	0.001 (0.004)	-0.015 (0.004)	-0.036 (0.004)	0.016 (0.004)	**0.015 (0.004)**
	$\tau = 0.5$					
	a_1	0.256 (0.015)	0.019 (0.020)	0.177 (0.014)	0.055 (0.019)	**0.036 (0.014)**
	b_1	-0.023 (0.001)	-0.001 (0.002)	-0.016 (0.001)	-0.006 (0.002)	**-0.004 (0.001)**
	ϕ_1	-0.010 (0.004)	-0.067 (0.004)	-0.079 (0.004)	0.017 (0.004)	**0.016 (0.004)**
	a_2	0.252 (0.015)	0.021 (0.020)	0.179 (0.014)	0.054 (0.019)	**0.030 (0.015)**
	b_2	-0.023 (0.001)	-0.002 (0.002)	-0.016 (0.001)	-0.006 (0.002)	**-0.003 (0.001)**
	ϕ_2	-0.007 (0.004)	-0.065 (0.004)	-0.081 (0.004)	0.018 (0.004)	**0.017 (0.004)**
	$\tau = 0.7$					
	a_1	0.111 (0.018)	-0.004 (0.013)	0.051 (0.012)	0.022 (0.021)	**0.010 (0.011)**
	b_1	-0.006 (0.002)	0.002 (0.001)	-0.002 (0.001)	-0.003 (0.002)	**-0.001 (0.001)**
	ϕ_1	-0.065 (0.004)	-0.173 (0.004)	-0.143 (0.004)	0.026 (0.004)	**0.023 (0.004)**
	a_2	0.103 (0.018)	-0.004 (0.013)	0.055 (0.012)	0.026 (0.021)	**0.010 (0.011)**
	b_2	-0.005 (0.002)	0.002 (0.001)	-0.003 (0.001)	-0.003 (0.002)	**-0.001 (0.001)**
	ϕ_2	-0.065 (0.004)	-0.169 (0.004)	-0.138 (0.004)	0.029 (0.004)	**0.026 (0.004)**

Note: **Bold** numbers indicate results from the true copula.

Table A–5: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Model: Independent				
		Clayton	Frank	Galambos	**Independent**	Normal
$N = 100$	a_1	0.268 (0.043)	-0.043 (0.039)	0.069 (0.039)	**-0.082 (0.044)**	-0.056 (0.038)
	b_1	-0.027 (0.004)	0.004 (0.004)	-0.008 (0.004)	**0.007 (0.004)**	0.005 (0.004)
	ϕ_1	0.077 (0.010)	0.078 (0.010)	0.078 (0.010)	**0.081 (0.010)**	0.078 (0.010)
	a_2	0.394 (0.042)	0.013 (0.039)	0.178 (0.040)	**-0.005 (0.045)**	0.012 (0.039)
	b_2	-0.039 (0.004)	-0.001 (0.004)	-0.018 (0.004)	**0.000 (0.004)**	-0.001 (0.004)
	ϕ_2	0.070 (0.010)	0.070 (0.010)	0.081 (0.011)	**0.075 (0.010)**	0.070 (0.010)
$N = 300$	a_1	0.418 (0.029)	0.003 (0.022)	0.107 (0.024)	**0.000 (0.026)**	0.005 (0.022)
	b_1	-0.042 (0.003)	-0.000 (0.002)	-0.011 (0.002)	**-0.000 (0.003)**	-0.000 (0.002)
	ϕ_1	0.027 (0.006)	0.028 (0.006)	0.033 (0.006)	**0.030 (0.006)**	0.029 (0.006)
	a_2	0.360 (0.028)	-0.015 (0.022)	0.115 (0.024)	**-0.037 (0.025)**	-0.014 (0.022)
	b_2	-0.036 (0.003)	0.002 (0.002)	-0.011 (0.002)	**0.004 (0.003)**	0.001 (0.002)
	ϕ_2	0.017 (0.006)	0.022 (0.006)	0.027 (0.006)	**0.023 (0.006)**	0.022 (0.006)
$N = 500$	a_1	0.421 (0.023)	0.032 (0.018)	0.098 (0.020)	**0.030 (0.020)**	0.045 (0.017)
	b_1	-0.042 (0.002)	-0.003 (0.002)	-0.010 (0.002)	**-0.003 (0.002)**	-0.004 (0.002)
	ϕ_1	0.024 (0.005)	0.027 (0.004)	0.032 (0.005)	**0.027 (0.004)**	0.026 (0.004)
	a_2	0.430 (0.023)	0.039 (0.017)	0.115 (0.020)	**0.038 (0.020)**	0.045 (0.017)
	b_2	-0.043 (0.002)	-0.004 (0.002)	-0.011 (0.002)	**-0.004 (0.002)**	-0.005 (0.002)
	ϕ_2	0.006 (0.004)	0.010 (0.004)	0.011 (0.004)	**0.010 (0.004)**	0.009 (0.004)

Note: **Bold** numbers indicate results from the true model.

APPENDIX B
STUDY 2 AVERAGE BIAS OF ESTIMATES AND MONTE CARLO STANDARD ERROR (MCSE)

Table B–1: Average bias of estimates and Monte Carlo Standard Error (MCSE) based on 1,000 simulations.

		True Copula: Clayton				
		Clayton	Frank	Galambos	Independent	Normal
$N=100$	$\tau=0.3$					
	a_1	**0.076 (0.039)**	0.037 (0.039)	0.112 (0.042)	0.052 (0.049)	0.065 (0.042)
	b_1	**-0.008 (0.004)**	-0.007 (0.004)	-0.009 (0.004)	-0.006 (0.005)	-0.007 (0.004)
	ϕ_1	**0.912 (0.088)**	0.407 (0.087)	-0.281 (0.069)	1.152 (0.105)	0.322 (0.077)
	a_2	**0.085 (0.040)**	0.086 (0.039)	0.156 (0.042)	0.091 (0.046)	0.087 (0.040)
	b_2	**-0.009 (0.004)**	-0.012 (0.004)	-0.014 (0.004)	-0.010 (0.005)	-0.009 (0.004)
	ϕ_2	**0.072 (0.010)**	-0.002 (0.010)	-0.041 (0.010)	0.077 (0.010)	0.047 (0.010)
	$\tau=0.5$					
	a_1	**-0.014 (0.033)**	0.017 (0.031)	0.291 (0.037)	-0.001 (0.044)	0.093 (0.034)
	b_1	**0.001 (0.003)**	-0.009 (0.003)	-0.025 (0.004)	-0.000 (0.004)	-0.009 (0.003)
	ϕ_1	**0.466 (0.063)**	-0.738 (0.045)	-1.434 (0.037)	0.488 (0.064)	-0.549 (0.047)
	a_2	**0.004 (0.034)**	0.070 (0.034)	0.177 (0.038)	-0.020 (0.042)	0.099 (0.035)
	b_2	**-0.001 (0.003)**	-0.014 (0.003)	-0.013 (0.004)	0.002 (0.004)	-0.010 (0.003)
	ϕ_2	**0.037 (0.009)**	-0.128 (0.009)	-0.188 (0.009)	0.046 (0.009)	-0.015 (0.009)
	$\tau=0.7$					
	a_1	**0.083 (0.030)**	-0.042 (0.029)	0.282 (0.040)	0.070 (0.049)	0.122 (0.036)
	b_1	**-0.008 (0.003)**	-0.007 (0.003)	-0.021 (0.004)	-0.007 (0.005)	-0.011 (0.004)
	ϕ_1	**1.378 (0.106)**	-0.362 (0.066)	-1.783 (0.040)	1.719 (0.138)	-0.368 (0.064)
	a_2	**0.110 (0.031)**	0.112 (0.030)	0.213 (0.039)	0.101 (0.047)	0.149 (0.035)
	b_2	**-0.011 (0.003)**	-0.022 (0.003)	-0.014 (0.004)	-0.010 (0.005)	-0.014 (0.004)
	ϕ_2	**0.089 (0.009)**	-0.120 (0.009)	-0.312 (0.008)	0.111 (0.011)	-0.027 (0.009)
$N=300$	$\tau=0.3$					
	a_1	**0.017 (0.020)**	0.025 (0.019)	0.104 (0.022)	0.029 (0.024)	0.038 (0.020)
	b_1	**-0.002 (0.002)**	-0.005 (0.002)	-0.008 (0.002)	-0.003 (0.002)	-0.004 (0.002)
	ϕ_1	**0.152 (0.034)**	-0.329 (0.032)	-0.874 (0.027)	0.163 (0.035)	-0.315 (0.031)
	a_2	**0.042 (0.021)**	0.037 (0.021)	0.089 (0.020)	0.031 (0.024)	0.065 (0.021)
	b_2	**-0.004 (0.002)**	-0.006 (0.002)	-0.006 (0.002)	-0.003 (0.002)	-0.006 (0.002)
	ϕ_2	**0.014 (0.005)**	-0.060 (0.005)	-0.098 (0.005)	0.015 (0.005)	-0.010 (0.005)
	$\tau=0.5$					
	a_1	**-0.011 (0.020)**	0.012 (0.018)	0.327 (0.027)	0.005 (0.025)	0.155 (0.020)
	b_1	**0.001 (0.002)**	-0.008 (0.002)	-0.028 (0.003)	-0.000 (0.003)	-0.015 (0.002)
	ϕ_1	**0.149 (0.033)**	-0.917 (0.025)	-1.588 (0.021)	0.141 (0.034)	-0.729 (0.026)
	a_2	**0.007 (0.020)**	0.098 (0.019)	0.251 (0.026)	0.004 (0.024)	0.174 (0.020)
	b_2	**-0.001 (0.002)**	-0.017 (0.002)	-0.020 (0.003)	-0.001 (0.002)	-0.017 (0.002)
	ϕ_2	**0.020 (0.005)**	-0.150 (0.005)	-0.211 (0.005)	0.023 (0.005)	-0.036 (0.005)
	$\tau=0.7$					
	a_1	**0.013 (0.015)**	-0.083 (0.012)	0.157 (0.019)	-0.014 (0.025)	0.060 (0.016)
	b_1	**-0.001 (0.001)**	-0.003 (0.001)	-0.009 (0.002)	0.002 (0.002)	-0.005 (0.002)
	ϕ_1	**0.302 (0.038)**	-0.991 (0.025)	-2.173 (0.016)	0.310 (0.040)	-0.963 (0.026)
	a_2	**0.002 (0.016)**	0.058 (0.014)	0.060 (0.019)	-0.020 (0.024)	0.057 (0.017)
	b_2	**-0.000 (0.002)**	-0.017 (0.001)	0.001 (0.002)	0.002 (0.002)	-0.005 (0.002)
	ϕ_2	**0.022 (0.005)**	-0.193 (0.004)	-0.387 (0.004)	0.025 (0.005)	-0.097 (0.005)
$N=500$	$\tau=0.3$					
	a_1	**0.022 (0.016)**	0.001 (0.016)	0.078 (0.018)	0.018 (0.020)	0.032 (0.017)
	b_1	**-0.002 (0.002)**	-0.003 (0.002)	-0.005 (0.002)	-0.002 (0.002)	-0.003 (0.002)
	ϕ_1	**0.198 (0.028)**	-0.302 (0.025)	-0.851 (0.021)	0.193 (0.028)	-0.291 (0.025)
	a_2	**0.045 (0.016)**	0.039 (0.016)	0.074 (0.015)	0.051 (0.018)	0.051 (0.016)
	b_2	**-0.004 (0.002)**	-0.006 (0.002)	-0.005 (0.002)	-0.005 (0.002)	-0.005 (0.002)
	ϕ_2	**0.012 (0.004)**	-0.062 (0.004)	-0.102 (0.004)	0.012 (0.004)	-0.013 (0.004)
	$\tau=0.5$					
	a_1	**0.010 (0.015)**	-0.001 (0.014)	0.291 (0.024)	0.019 (0.019)	0.149 (0.015)
	b_1	**-0.001 (0.001)**	-0.007 (0.001)	-0.025 (0.002)	-0.002 (0.002)	-0.015 (0.001)
	ϕ_1	**0.137 (0.027)**	-0.926 (0.021)	-1.601 (0.017)	0.144 (0.028)	-0.702 (0.022)
	a_2	**0.018 (0.015)**	0.079 (0.016)	0.181 (0.022)	0.010 (0.018)	0.153 (0.016)
	b_2	**-0.002 (0.001)**	-0.015 (0.002)	-0.013 (0.002)	-0.001 (0.002)	-0.015 (0.002)
	ϕ_2	**0.010 (0.004)**	-0.160 (0.004)	-0.224 (0.004)	0.012 (0.004)	-0.043 (0.004)
	$\tau=0.7$					
	a_1	**0.000 (0.013)**	-0.081 (0.010)	0.191 (0.017)	-0.028 (0.021)	0.051 (0.014)
	b_1	**-0.000 (0.001)**	-0.003 (0.001)	-0.013 (0.002)	0.003 (0.002)	-0.004 (0.001)
	ϕ_1	**0.180 (0.029)**	-1.057 (0.019)	-2.213 (0.013)	0.203 (0.031)	-1.008 (0.021)
	a_2	**-0.004 (0.013)**	0.068 (0.011)	0.102 (0.016)	-0.020 (0.020)	0.065 (0.014)
	b_2	**0.000 (0.001)**	-0.018 (0.001)	-0.003 (0.002)	0.002 (0.002)	-0.006 (0.001)
	ϕ_2	**0.013 (0.004)**	-0.202 (0.003)	-0.398 (0.003)	0.015 (0.004)	-0.106 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table B–2: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Frank				
		Clayton	**Frank**	Galambos	Independent	Normal
	$\tau = 0.3$					
	a_1	0.046 (0.041)	**-0.017 (0.039)**	0.070 (0.042)	-0.001 (0.048)	0.046 (0.040)
	b_1	-0.004 (0.004)	**0.001 (0.004)**	-0.006 (0.004)	-0.001 (0.005)	-0.005 (0.004)
	ϕ_1	1.302 (0.121)	**1.064 (0.103)**	0.219 (0.085)	1.319 (0.132)	1.076 (0.102)
	a_2	0.015 (0.042)	**-0.001 (0.039)**	0.094 (0.041)	0.009 (0.047)	0.022 (0.038)
	b_2	-0.001 (0.004)	**-0.000 (0.004)**	-0.008 (0.004)	-0.002 (0.005)	-0.003 (0.004)
	ϕ_2	0.061 (0.010)	**0.082 (0.010)**	0.007 (0.010)	0.086 (0.011)	0.084 (0.010)
	$\tau = 0.5$					
$N = 100$	a_1	0.148 (0.033)	**0.032 (0.030)**	0.180 (0.034)	0.030 (0.043)	0.034 (0.033)
	b_1	-0.012 (0.003)	**-0.003 (0.003)**	-0.016 (0.003)	-0.003 (0.004)	-0.003 (0.003)
	ϕ_1	0.801 (0.067)	**0.433 (0.063)**	-0.820 (0.044)	0.482 (0.065)	0.376 (0.058)
	a_2	0.137 (0.036)	**0.005 (0.030)**	0.119 (0.036)	0.018 (0.041)	0.018 (0.033)
	b_2	-0.010 (0.004)	**-0.000 (0.003)**	-0.009 (0.004)	-0.002 (0.004)	-0.002 (0.003)
	ϕ_2	0.002 (0.009)	**0.049 (0.009)**	-0.099 (0.009)	0.051 (0.009)	0.048 (0.009)
	$\tau = 0.7$					
	a_1	0.144 (0.034)	**-0.074 (0.028)**	0.117 (0.032)	-0.046 (0.047)	0.024 (0.030)
	b_1	-0.010 (0.003)	**0.007 (0.003)**	-0.008 (0.003)	0.004 (0.005)	-0.003 (0.003)
	ϕ_1	0.728 (0.077)	**0.618 (0.068)**	-1.298 (0.044)	0.848 (0.083)	0.517 (0.076)
	a_2	0.087 (0.037)	**-0.070 (0.029)**	0.083 (0.032)	-0.033 (0.047)	0.041 (0.031)
	b_2	-0.003 (0.004)	**0.006 (0.003)**	-0.005 (0.003)	0.003 (0.005)	-0.005 (0.003)
	ϕ_2	-0.037 (0.010)	**0.052 (0.009)**	-0.225 (0.009)	0.067 (0.010)	0.032 (0.010)
	$\tau = 0.3$					
	a_1	0.046 (0.021)	**0.010 (0.019)**	0.063 (0.021)	0.008 (0.024)	0.004 (0.020)
	b_1	-0.004 (0.002)	**-0.001 (0.002)**	-0.005 (0.002)	-0.001 (0.002)	-0.000 (0.002)
	ϕ_1	0.368 (0.037)	**0.187 (0.035)**	-0.516 (0.029)	0.205 (0.035)	0.195 (0.035)
	a_2	0.067 (0.020)	**0.039 (0.020)**	0.077 (0.020)	0.040 (0.023)	0.011 (0.020)
	b_2	-0.005 (0.002)	**-0.004 (0.002)**	-0.006 (0.002)	-0.004 (0.002)	-0.001 (0.002)
	ϕ_2	-0.004 (0.005)	**0.015 (0.005)**	-0.060 (0.005)	0.016 (0.005)	0.018 (0.005)
	$\tau = 0.5$					
$N = 300$	a_1	0.189 (0.020)	**0.010 (0.018)**	0.107 (0.020)	0.024 (0.024)	0.076 (0.018)
	b_1	-0.016 (0.002)	**-0.001 (0.002)**	-0.008 (0.002)	-0.002 (0.002)	-0.007 (0.002)
	ϕ_1	0.596 (0.035)	**0.189 (0.034)**	-1.023 (0.025)	0.184 (0.035)	0.159 (0.032)
	a_2	0.160 (0.020)	**0.024 (0.018)**	0.085 (0.019)	0.016 (0.023)	0.087 (0.017)
	b_2	-0.013 (0.002)	**-0.002 (0.002)**	-0.006 (0.002)	-0.002 (0.002)	-0.009 (0.002)
	ϕ_2	-0.015 (0.005)	**0.026 (0.005)**	-0.131 (0.005)	0.027 (0.006)	0.027 (0.005)
	$\tau = 0.7$					
	a_1	0.177 (0.017)	**0.015 (0.012)**	0.109 (0.015)	0.012 (0.024)	0.009 (0.014)
	b_1	-0.012 (0.002)	**-0.001 (0.001)**	-0.007 (0.001)	-0.001 (0.002)	-0.001 (0.001)
	ϕ_1	0.191 (0.039)	**0.219 (0.035)**	-1.624 (0.020)	0.231 (0.038)	-0.005 (0.035)
	a_2	0.128 (0.017)	**0.022 (0.012)**	0.033 (0.015)	0.016 (0.023)	0.018 (0.015)
	b_2	-0.005 (0.002)	**-0.002 (0.001)**	0.001 (0.001)	-0.001 (0.002)	-0.001 (0.001)
	ϕ_2	-0.086 (0.005)	**0.015 (0.005)**	-0.282 (0.005)	0.009 (0.005)	-0.019 (0.005)
	$\tau = 0.3$					
	a_1	0.031 (0.016)	**0.002 (0.015)**	0.065 (0.016)	0.013 (0.019)	0.022 (0.016)
	b_1	-0.002 (0.002)	**-0.000 (0.001)**	-0.005 (0.002)	-0.002 (0.002)	-0.002 (0.002)
	ϕ_1	0.256 (0.029)	**0.116 (0.028)**	-0.580 (0.023)	0.115 (0.028)	0.114 (0.028)
	a_2	0.014 (0.016)	**-0.006 (0.015)**	0.041 (0.015)	-0.009 (0.018)	0.009 (0.016)
	b_2	-0.000 (0.002)	**0.000 (0.002)**	-0.003 (0.001)	0.001 (0.002)	-0.001 (0.002)
	ϕ_2	-0.013 (0.004)	**0.008 (0.004)**	-0.069 (0.004)	0.007 (0.004)	0.009 (0.004)
	$\tau = 0.5$					
$N = 500$	a_1	0.257 (0.015)	**0.005 (0.014)**	0.130 (0.016)	0.031 (0.018)	0.107 (0.013)
	b_1	-0.023 (0.001)	**-0.001 (0.001)**	-0.010 (0.002)	-0.003 (0.002)	-0.011 (0.001)
	ϕ_1	0.515 (0.027)	**0.104 (0.027)**	-1.072 (0.019)	0.108 (0.027)	0.130 (0.026)
	a_2	0.218 (0.015)	**-0.007 (0.013)**	0.069 (0.015)	0.009 (0.017)	0.090 (0.013)
	b_2	-0.018 (0.001)	**0.001 (0.001)**	-0.004 (0.002)	-0.001 (0.002)	-0.009 (0.001)
	ϕ_2	-0.022 (0.004)	**0.016 (0.004)**	-0.139 (0.004)	0.016 (0.004)	0.022 (0.004)
	$\tau = 0.7$					
	a_1	0.170 (0.016)	**0.002 (0.011)**	0.152 (0.015)	0.013 (0.021)	0.014 (0.013)
	b_1	-0.012 (0.002)	**-0.000 (0.001)**	-0.012 (0.001)	-0.002 (0.002)	-0.001 (0.001)
	ϕ_1	0.121 (0.031)	**0.191 (0.028)**	-1.634 (0.017)	0.201 (0.032)	-0.025 (0.030)
	a_2	0.104 (0.016)	**-0.006 (0.011)**	0.061 (0.014)	-0.005 (0.020)	-0.003 (0.013)
	b_2	-0.003 (0.002)	**0.000 (0.001)**	-0.002 (0.001)	0.000 (0.002)	0.000 (0.001)
	ϕ_2	-0.084 (0.004)	**0.017 (0.004)**	-0.279 (0.004)	0.017 (0.004)	-0.013 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table B–3: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Galambos				
		Clayton	Frank	**Galambos**	Independent	Normal
$N = 100$	$\tau = 0.3$					
	a_1	0.099 (0.039)	0.088 (0.035)	**0.049 (0.036)**	0.074 (0.044)	0.028 (0.036)
	b_1	-0.009 (0.004)	-0.007 (0.004)	**-0.005 (0.004)**	-0.007 (0.004)	-0.003 (0.004)
	ϕ_1	0.282 (0.058)	0.229 (0.057)	**0.221 (0.056)**	0.249 (0.057)	0.567 (0.061)
	a_2	0.095 (0.038)	0.043 (0.036)	**0.067 (0.033)**	0.070 (0.042)	0.051 (0.033)
	b_2	-0.008 (0.004)	-0.002 (0.004)	**-0.007 (0.003)**	-0.007 (0.004)	-0.005 (0.003)
	ϕ_2	0.005 (0.010)	0.034 (0.009)	**0.041 (0.009)**	0.043 (0.009)	0.056 (0.010)
	$\tau = 0.5$					
	a_1	0.170 (0.039)	0.207 (0.034)	**0.068 (0.031)**	0.059 (0.043)	0.027 (0.030)
	b_1	-0.013 (0.004)	-0.015 (0.003)	**-0.006 (0.003)**	-0.005 (0.004)	-0.002 (0.003)
	ϕ_1	0.455 (0.066)	0.306 (0.063)	**0.348 (0.062)**	0.347 (0.063)	0.895 (0.069)
	a_2	0.073 (0.041)	0.132 (0.034)	**0.047 (0.028)**	0.019 (0.042)	-0.003 (0.030)
	b_2	-0.003 (0.004)	-0.008 (0.003)	**-0.004 (0.003)**	-0.001 (0.004)	0.001 (0.003)
	ϕ_2	-0.039 (0.009)	0.011 (0.009)	**0.035 (0.009)**	0.035 (0.009)	0.068 (0.009)
	$\tau = 0.7$					
	a_1	0.130 (0.039)	0.126 (0.030)	**-0.036 (0.023)**	-0.047 (0.047)	-0.021 (0.028)
	b_1	-0.007 (0.004)	-0.005 (0.003)	**0.003 (0.002)**	0.004 (0.005)	0.001 (0.003)
	ϕ_1	0.387 (0.082)	0.186 (0.069)	**0.712 (0.078)**	1.067 (0.101)	1.483 (0.094)
	a_2	0.102 (0.042)	0.079 (0.031)	**0.006 (0.021)**	-0.016 (0.047)	0.025 (0.027)
	b_2	-0.003 (0.004)	0.000 (0.003)	**-0.001 (0.002)**	0.001 (0.005)	-0.003 (0.003)
	ϕ_2	-0.090 (0.010)	-0.030 (0.009)	**0.052 (0.009)**	0.057 (0.010)	0.106 (0.010)
$N = 300$	$\tau = 0.3$					
	a_1	0.083 (0.022)	0.079 (0.020)	**0.040 (0.020)**	0.039 (0.024)	0.057 (0.020)
	b_1	-0.007 (0.002)	-0.006 (0.002)	**-0.004 (0.002)**	-0.004 (0.002)	-0.006 (0.002)
	ϕ_1	0.244 (0.035)	0.179 (0.034)	**0.150 (0.032)**	0.179 (0.034)	0.491 (0.036)
	a_2	0.066 (0.022)	0.037 (0.020)	**0.028 (0.019)**	0.028 (0.024)	0.021 (0.019)
	b_2	-0.005 (0.002)	-0.002 (0.002)	**-0.003 (0.002)**	-0.003 (0.002)	-0.002 (0.002)
	ϕ_2	-0.022 (0.005)	0.002 (0.005)	**0.009 (0.005)**	0.010 (0.005)	0.025 (0.005)
	$\tau = 0.5$					
	a_1	0.161 (0.022)	0.167 (0.021)	**0.040 (0.018)**	0.033 (0.024)	0.043 (0.016)
	b_1	-0.013 (0.002)	-0.012 (0.002)	**-0.004 (0.002)**	-0.004 (0.002)	-0.005 (0.002)
	ϕ_1	0.270 (0.036)	0.107 (0.035)	**0.107 (0.031)**	0.146 (0.035)	0.757 (0.038)
	a_2	0.077 (0.022)	0.079 (0.020)	**0.029 (0.016)**	0.010 (0.023)	0.012 (0.016)
	b_2	-0.004 (0.002)	-0.003 (0.002)	**-0.003 (0.002)**	-0.001 (0.002)	-0.002 (0.002)
	ϕ_2	-0.055 (0.005)	-0.012 (0.005)	**0.011 (0.005)**	0.012 (0.005)	0.053 (0.005)
	$\tau = 0.7$					
	a_1	0.188 (0.022)	0.136 (0.014)	**-0.002 (0.011)**	-0.014 (0.025)	-0.013 (0.014)
	b_1	-0.013 (0.002)	-0.005 (0.001)	**0.000 (0.001)**	0.001 (0.003)	0.001 (0.001)
	ϕ_1	-0.095 (0.039)	-0.244 (0.034)	**0.280 (0.035)**	0.396 (0.042)	1.059 (0.046)
	a_2	0.125 (0.023)	0.041 (0.014)	**-0.003 (0.009)**	-0.016 (0.025)	-0.015 (0.014)
	b_2	-0.005 (0.002)	0.004 (0.001)	**0.000 (0.001)**	0.001 (0.002)	0.001 (0.001)
	ϕ_2	-0.120 (0.006)	-0.060 (0.005)	**0.026 (0.005)**	0.028 (0.006)	0.089 (0.006)
$N = 500$	$\tau = 0.3$					
	a_1	0.085 (0.016)	0.064 (0.015)	**0.040 (0.015)**	0.038 (0.018)	0.031 (0.015)
	b_1	-0.007 (0.002)	-0.004 (0.001)	**-0.004 (0.001)**	-0.004 (0.002)	-0.003 (0.001)
	ϕ_1	0.186 (0.029)	0.107 (0.028)	**0.083 (0.026)**	0.115 (0.028)	0.438 (0.030)
	a_2	0.053 (0.016)	0.027 (0.015)	**0.017 (0.014)**	0.022 (0.017)	0.015 (0.015)
	b_2	-0.004 (0.002)	-0.001 (0.001)	**-0.002 (0.001)**	-0.002 (0.002)	-0.002 (0.001)
	ϕ_2	-0.023 (0.004)	0.002 (0.004)	**0.009 (0.004)**	0.009 (0.004)	0.025 (0.004)
	$\tau = 0.5$					
	a_1	0.145 (0.016)	0.162 (0.017)	**0.036 (0.014)**	0.029 (0.019)	0.048 (0.012)
	b_1	-0.012 (0.002)	-0.012 (0.002)	**-0.004 (0.001)**	-0.003 (0.002)	-0.005 (0.001)
	ϕ_1	0.226 (0.029)	0.093 (0.029)	**0.087 (0.026)**	0.101 (0.028)	0.716 (0.030)
	a_2	0.079 (0.018)	0.055 (0.017)	**0.026 (0.012)**	0.014 (0.018)	0.021 (0.012)
	b_2	-0.004 (0.002)	-0.001 (0.002)	**-0.003 (0.001)**	-0.001 (0.002)	-0.002 (0.001)
	ϕ_2	-0.059 (0.004)	-0.011 (0.004)	**0.009 (0.004)**	0.009 (0.004)	0.050 (0.004)
	$\tau = 0.7$					
	a_1	0.196 (0.018)	0.126 (0.012)	**-0.004 (0.009)**	0.005 (0.020)	-0.010 (0.012)
	b_1	-0.014 (0.002)	-0.004 (0.001)	**0.000 (0.001)**	-0.001 (0.002)	0.000 (0.001)
	ϕ_1	-0.116 (0.032)	-0.250 (0.028)	**0.253 (0.028)**	0.352 (0.034)	1.066 (0.038)
	a_2	0.131 (0.018)	0.040 (0.012)	**0.005 (0.008)**	0.008 (0.020)	-0.007 (0.011)
	b_2	-0.005 (0.002)	0.005 (0.001)	**-0.001 (0.001)**	-0.001 (0.002)	0.000 (0.001)
	ϕ_2	-0.119 (0.004)	-0.056 (0.004)	**0.027 (0.004)**	0.030 (0.004)	0.092 (0.004)

Note: **Bold** numbers indicate results from the true copula.

Table B–4: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Copula: Normal				
		Clayton	Frank	Galambos	Independent	**Normal**
$N = 100$	$\tau = 0.3$					
	a_1	-0.006 (0.040)	-0.027 (0.037)	0.021 (0.041)	-0.060 (0.046)	**0.001 (0.038)**
	b_1	0.002 (0.004)	0.003 (0.004)	-0.001 (0.004)	0.006 (0.005)	**-0.000 (0.004)**
	ϕ_1	1.368 (0.117)	1.108 (0.113)	0.460 (0.094)	1.511 (0.141)	**1.221 (0.108)**
	a_2	-0.037 (0.041)	0.008 (0.039)	0.024 (0.039)	-0.033 (0.046)	**0.026 (0.039)**
	b_2	0.004 (0.004)	-0.001 (0.004)	-0.002 (0.004)	0.003 (0.005)	**-0.003 (0.004)**
	ϕ_2	0.025 (0.010)	0.025 (0.010)	-0.011 (0.010)	0.052 (0.010)	**0.047 (0.010)**
	$\tau = 0.5$					
	a_1	0.244 (0.034)	0.160 (0.033)	0.175 (0.031)	0.105 (0.044)	**0.088 (0.033)**
	b_1	-0.022 (0.003)	-0.015 (0.003)	-0.016 (0.003)	-0.011 (0.004)	**-0.009 (0.003)**
	ϕ_1	0.344 (0.060)	-0.081 (0.056)	-0.603 (0.047)	0.380 (0.061)	**0.266 (0.059)**
	a_2	0.188 (0.037)	0.144 (0.033)	0.178 (0.032)	0.116 (0.044)	**0.078 (0.032)**
	b_2	-0.016 (0.004)	-0.014 (0.003)	-0.016 (0.003)	-0.012 (0.004)	**-0.008 (0.003)**
	ϕ_2	-0.017 (0.009)	-0.027 (0.010)	-0.066 (0.009)	0.052 (0.010)	**0.041 (0.009)**
	$\tau = 0.7$					
	a_1	0.174 (0.033)	0.020 (0.028)	0.038 (0.023)	0.036 (0.044)	**0.023 (0.023)**
	b_1	-0.013 (0.003)	0.000 (0.003)	-0.002 (0.002)	-0.004 (0.004)	**-0.003 (0.002)**
	ϕ_1	0.201 (0.068)	-0.353 (0.054)	-0.607 (0.047)	0.923 (0.082)	**0.537 (0.065)**
	a_2	0.170 (0.035)	0.050 (0.027)	0.065 (0.021)	0.089 (0.041)	**0.068 (0.022)**
	b_2	-0.011 (0.003)	-0.003 (0.003)	-0.005 (0.002)	-0.009 (0.004)	**-0.007 (0.002)**
	ϕ_2	-0.090 (0.009)	-0.118 (0.009)	-0.120 (0.008)	0.047 (0.009)	**0.030 (0.009)**
$N = 300$	$\tau = 0.3$					
	a_1	0.056 (0.021)	0.033 (0.019)	0.060 (0.020)	0.006 (0.024)	**0.001 (0.020)**
	b_1	-0.005 (0.002)	-0.003 (0.002)	-0.005 (0.002)	-0.001 (0.002)	**-0.000 (0.002)**
	ϕ_1	0.214 (0.036)	-0.004 (0.034)	-0.428 (0.031)	0.153 (0.036)	**0.143 (0.036)**
	a_2	0.065 (0.021)	0.032 (0.020)	0.056 (0.019)	0.020 (0.023)	**0.018 (0.020)**
	b_2	-0.006 (0.002)	-0.003 (0.002)	-0.004 (0.002)	-0.002 (0.002)	**-0.002 (0.002)**
	ϕ_2	-0.009 (0.005)	-0.010 (0.005)	-0.049 (0.005)	0.016 (0.005)	**0.015 (0.005)**
	$\tau = 0.5$					
	a_1	0.194 (0.019)	0.096 (0.019)	0.164 (0.017)	0.055 (0.024)	**0.056 (0.018)**
	b_1	-0.017 (0.002)	-0.009 (0.002)	-0.015 (0.002)	-0.006 (0.002)	**-0.006 (0.002)**
	ϕ_1	0.316 (0.037)	-0.201 (0.033)	-0.670 (0.028)	0.252 (0.037)	**0.224 (0.036)**
	a_2	0.107 (0.021)	0.071 (0.019)	0.105 (0.016)	0.030 (0.023)	**0.028 (0.016)**
	b_2	-0.008 (0.002)	-0.006 (0.002)	-0.009 (0.002)	-0.003 (0.002)	**-0.003 (0.002)**
	ϕ_2	-0.038 (0.005)	-0.056 (0.005)	-0.097 (0.005)	0.019 (0.005)	**0.018 (0.005)**
	$\tau = 0.7$					
	a_1	0.135 (0.020)	0.072 (0.014)	0.045 (0.013)	-0.044 (0.025)	**-0.007 (0.013)**
	b_1	-0.008 (0.002)	-0.004 (0.001)	-0.002 (0.001)	0.004 (0.002)	**0.001 (0.001)**
	ϕ_1	-0.200 (0.039)	-0.636 (0.030)	-0.797 (0.028)	0.421 (0.044)	**0.331 (0.039)**
	a_2	0.084 (0.021)	0.045 (0.015)	0.009 (0.011)	-0.030 (0.024)	**0.002 (0.012)**
	b_2	-0.002 (0.002)	-0.001 (0.002)	0.002 (0.001)	0.003 (0.002)	**-0.000 (0.001)**
	ϕ_2	-0.100 (0.005)	-0.135 (0.005)	-0.139 (0.005)	0.031 (0.005)	**0.029 (0.005)**
$N = 500$	$\tau = 0.3$					
	a_1	0.034 (0.012)	0.014 (0.011)	0.042 (0.012)	-0.006 (0.014)	**0.007 (0.012)**
	b_1	-0.003 (0.001)	-0.001 (0.001)	-0.003 (0.001)	0.000 (0.001)	**-0.001 (0.001)**
	ϕ_1	0.114 (0.020)	-0.108 (0.019)	-0.527 (0.017)	0.034 (0.020)	**0.036 (0.020)**
	a_2	0.017 (0.012)	0.016 (0.011)	0.016 (0.010)	-0.002 (0.013)	**0.011 (0.012)**
	b_2	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.001)	**-0.001 (0.001)**
	ϕ_2	-0.023 (0.003)	-0.024 (0.003)	-0.061 (0.003)	0.002 (0.003)	**0.002 (0.003)**
	$\tau = 0.5$					
	a_1	0.189 (0.014)	0.120 (0.014)	0.135 (0.013)	0.029 (0.018)	**0.036 (0.014)**
	b_1	-0.016 (0.001)	-0.011 (0.001)	-0.012 (0.001)	-0.003 (0.002)	**-0.004 (0.001)**
	ϕ_1	0.240 (0.027)	-0.246 (0.025)	-0.711 (0.022)	0.196 (0.028)	**0.175 (0.027)**
	a_2	0.129 (0.015)	0.113 (0.013)	0.094 (0.013)	0.033 (0.017)	**0.030 (0.012)**
	b_2	-0.010 (0.002)	-0.010 (0.001)	-0.007 (0.001)	-0.003 (0.002)	**-0.003 (0.001)**
	ϕ_2	-0.043 (0.004)	-0.062 (0.004)	-0.105 (0.004)	0.013 (0.004)	**0.012 (0.004)**
	$\tau = 0.7$					
	a_1	0.169 (0.017)	0.076 (0.012)	0.035 (0.010)	0.009 (0.020)	**0.014 (0.010)**
	b_1	-0.012 (0.002)	-0.005 (0.001)	-0.002 (0.001)	-0.001 (0.002)	**-0.001 (0.001)**
	ϕ_1	-0.303 (0.029)	-0.750 (0.023)	-0.899 (0.021)	0.275 (0.033)	**0.186 (0.029)**
	a_2	0.120 (0.017)	0.061 (0.012)	0.010 (0.010)	0.030 (0.020)	**0.031 (0.010)**
	b_2	-0.006 (0.002)	-0.003 (0.001)	0.001 (0.001)	-0.003 (0.002)	**-0.003 (0.001)**
	ϕ_2	-0.113 (0.004)	-0.146 (0.004)	-0.148 (0.004)	0.023 (0.004)	**0.018 (0.004)**

Note: **Bold** numbers indicate results from the true copula.

Table B–5: Average bias of estimates and Monte Carlo Standard Error(MCSE) based on 1,000 simulations.

		True Model: Independent				
		Clayton	Frank	Galambos	**Independent**	Normal
$N = 100$	a_1	0.117 (0.042)	-0.018 (0.038)	0.071 (0.040)	**-0.025 (0.044)**	0.037 (0.039)
	b_1	-0.012 (0.004)	0.002 (0.004)	-0.007 (0.004)	**0.002 (0.004)**	-0.004 (0.004)
	ϕ_1	0.555 (0.068)	0.456 (0.066)	0.643 (0.067)	**0.476 (0.065)**	0.448 (0.063)
	a_2	0.166 (0.041)	-0.017 (0.038)	0.020 (0.039)	**-0.018 (0.043)**	0.036 (0.038)
	b_2	-0.017 (0.004)	0.002 (0.004)	-0.002 (0.004)	**0.002 (0.004)**	-0.004 (0.004)
	ϕ_2	0.072 (0.010)	0.069 (0.010)	0.069 (0.010)	**0.072 (0.010)**	0.069 (0.010)
$N = 300$	a_1	0.139 (0.026)	0.016 (0.023)	0.055 (0.024)	**0.016 (0.026)**	0.033 (0.022)
	b_1	-0.014 (0.003)	-0.001 (0.002)	-0.005 (0.002)	**-0.002 (0.003)**	-0.003 (0.002)
	ϕ_1	0.292 (0.041)	0.207 (0.038)	0.439 (0.043)	**0.223 (0.038)**	0.213 (0.038)
	a_2	0.146 (0.026)	-0.000 (0.023)	0.073 (0.025)	**-0.000 (0.026)**	0.020 (0.023)
	b_2	-0.015 (0.003)	0.000 (0.002)	-0.007 (0.003)	**-0.000 (0.003)**	-0.002 (0.002)
	ϕ_2	0.023 (0.006)	0.024 (0.006)	0.026 (0.006)	**0.026 (0.006)**	0.025 (0.006)
$N = 500$	a_1	0.112 (0.020)	0.031 (0.017)	0.056 (0.020)	**0.036 (0.019)**	0.027 (0.016)
	b_1	-0.011 (0.002)	-0.003 (0.002)	-0.005 (0.002)	**-0.004 (0.002)**	-0.003 (0.002)
	ϕ_1	0.200 (0.031)	0.133 (0.029)	0.328 (0.032)	**0.140 (0.029)**	0.131 (0.029)
	a_2	0.130 (0.021)	-0.009 (0.017)	0.079 (0.020)	**-0.020 (0.019)**	-0.006 (0.017)
	b_2	-0.013 (0.002)	0.001 (0.002)	-0.008 (0.002)	**0.002 (0.002)**	0.001 (0.002)
	ϕ_2	0.013 (0.004)	0.014 (0.004)	0.015 (0.004)	**0.014 (0.004)**	0.013 (0.004)

Note: **Bold** numbers indicate results from the true model.

APPENDIX C
LOGLIKELIHOODS FOR THE BIVARIATE REGRESSION COPULA-BASED MODELS WITH NEGATIVE BINOMIAL MARGINALS

```
#---------- Clayton ----------

NBC.loglik <- function(pars){
n = length(y1)
X = as.matrix(cbind( rep(1,n), x) ) # Design Matrix
npars = ncol(X)

pars <- as.list(pars)
names(pars) <- c(paste("b", 0:(npars-1),"1", sep = ""),
paste("b", 0:(npars-1),"2", sep = ""),
"phi1", "phi2", 'theta') # Dispersion Parameters
list2env(pars , envir = .GlobalEnv)

beta1 = as.numeric(pars[1:npars])
beta2 = as.numeric(pars[(npars+1):(2*npars)])
theta = exp(as.numeric(pars[length(pars)]))

mu1 = exp( X%*%beta1 ); mu2 = exp( X%*%beta2 ) # means

F1 = function(y){
pnbinom(y, mu = mu1[i], size = phi1)
} # Marginal CDF X1 - NB (mu = mu1, size = phi1)

F2 = function(y){
pnbinom(y, mu = mu2[i], size = phi2)
} # Marginal CDF X2 - NB (mu = mu2, size = phi2)

C = function(u1, u2){
(u1 ^ -theta + u2 ^ -theta - 1) ^ -(1/theta)
} # Joint Distribution function by the Clayton Copula
   representation

h = NULL # Joint pmf for discrete random variables
```

```
for (i in 1:n) {

if (y1[i] == 0 & y2[i] == 0) {
h[i] = C( F1(0), F2(0) )
}
if (y1[i] != 0 & y2[i] == 0) {
h[i] = C( F1(y1[i]), F2(0) ) - C( F1(y1[i]-1), F2(0) )
}
if (y1[i] == 0 & y2[i] != 0) {
h[i] = C( F1(0), F2(y2[i]) ) - C( F1(0), F2(y2[i] - 1) )
}
if (y1[i] != 0 & y2[i] != 0) {
h[i] = C( F1(y1[i]), F2(y2[i]) ) - C( F1(y1[i] - 1), F2(
   y2[i]) ) - C( F1(y1[i]), F2(y2[i] - 1) ) + C( F1(y1[i]
    - 1), F2(y2[i] - 1) )
}
} # Joint PMF for discrete random variables

sum(log(h))
}

#---------- Normal ----------

NBN.loglik <- function(pars){
n = length(y1)
X = as.matrix(cbind( rep(1,n), x) ) # Design Matrix
npars = ncol(X)

pars <- as.list(pars)
names(pars) <- c(paste("b", 0:(npars-1),"1", sep = ""),
paste("b", 0:(npars-1),"2", sep = ""),
"phi1", "phi2", 'theta') # Dispersion Parameters
list2env(pars , envir = .GlobalEnv)

beta1 = as.numeric(pars[1:npars])
beta2 = as.numeric(pars[(npars+1):(2*npars)])
theta = ( exp(as.numeric(pars[length(pars)])) - 1 ) / ( 1
    + exp(as.numeric(pars[length(pars)])) )

mu1 = exp( X%*%beta1 ); mu2 = exp( X%*%beta2 ) # means

F1 = function(y){
pnbinom(y, mu = mu1[i], size = phi1)
} # Marginal CDF X1 - NB (mu = mu1, size = phi1)

F2 = function(y){
pnbinom(y, mu = mu2[i], size = phi2)
} # Marginal CDF X2 - NB (mu = mu2, size = phi2)
```

```
C = function(u1, u2){
corr <- diag(2)
corr[lower.tri(corr)] <- theta
corr[upper.tri(corr)] <- theta
pmvnorm(lower = -Inf, upper = c(qnorm(u1), qnorm(u2)),
   corr = corr) #cor = cov
} # Joint Distribution function by the Normal Copula
   representation

h = NULL # Joint pmf for discrete random variables

for (i in 1:n) {

if (y1[i] == 0 & y2[i] == 0) {
h[i] = C( F1(0), F2(0) )
}
if (y1[i] != 0 & y2[i] == 0) {
h[i] = C( F1(y1[i]), F2(0) ) - C( F1(y1[i]-1), F2(0) )
}
if (y1[i] == 0 & y2[i] != 0) {
h[i] = C( F1(0), F2(y2[i]) ) - C( F1(0), F2(y2[i] - 1) )
}
if (y1[i] != 0 & y2[i] != 0) {
h[i] = C( F1(y1[i]), F2(y2[i]) ) - C( F1(y1[i] - 1), F2(
   y2[i]) ) - C( F1(y1[i]), F2(y2[i] - 1) ) + C( F1(y1[i]
    - 1), F2(y2[i] - 1) )
}
} # Joint PMF for discrete random variables

sum(log(h))

}

#---------- Galambos ----------

NBG.loglik <- function(pars){
n = length(y1)
X = as.matrix(cbind( rep(1,n), x) ) # Design Matrix
npars = ncol(X)

pars <- as.list(pars)
names(pars) <- c(paste("b", 0:(npars-1),"1", sep = ""),
paste("b", 0:(npars-1),"2", sep = ""),
"phi1", "phi2", 'theta') # Dispersion Parameters
list2env(pars , envir = .GlobalEnv)

beta1 = as.numeric(pars[1:npars])
beta2 = as.numeric(pars[(npars+1):(2*npars)])
theta = exp(as.numeric(pars[length(pars)]))
```

```r
mu1 = exp( X%*%beta1 ); mu2 = exp( X%*%beta2 ) # means

F1 = function(y){
pnbinom(y, mu = mu1[i], size = phi1)
} # Marginal CDF X1 - NB (mu = mu1, size = phi1)

F2 = function(y){
pnbinom(y, mu = mu2[i], size = phi2)
} # Marginal CDF X2 - NB (mu = mu2, size = phi2)

C = function(u1, u2){
u1 * u2 * exp( ( (-log(u1))^-theta + (-log(u2))^-theta )
   ^ (-1/theta) )
} # Joint Distribution function by the Galambos Copula
   representation

h = NULL # Joint pmf for discrete random variables

for (i in 1:n) {

if (y1[i] == 0 & y2[i] == 0) {
h[i] = C( F1(0), F2(0) )
}
if (y1[i] != 0 & y2[i] == 0) {
h[i] = C( F1(y1[i]), F2(0) ) - C( F1(y1[i]-1), F2(0) )
}
if (y1[i] == 0 & y2[i] != 0) {
h[i] = C( F1(0), F2(y2[i]) ) - C( F1(0), F2(y2[i] - 1) )
}
if (y1[i] != 0 & y2[i] != 0) {
h[i] = C( F1(y1[i]), F2(y2[i]) ) - C( F1(y1[i] - 1), F2(
   y2[i]) ) - C( F1(y1[i]), F2(y2[i] - 1) ) + C( F1(y1[i]
    - 1), F2(y2[i] - 1) )
}
} # Joint PMF for discrete random variables

sum(log(h))
}

#---------- Frank ----------

NBF.loglik <- function(pars){
n = length(y1)
X = cbind( rep(1,n), x ) # Design Matrix

npars = ncol(X)

pars <- as.list(pars)
```

```
names(pars) <- c(paste("b", 0:(npars-1),"1", sep = ""),
paste("b", 0:(npars-1),"2", sep = ""),
"phi1", "phi2") # Dispersion Parameters
list2env(pars , envir = .GlobalEnv)

beta1 = as.numeric(pars[1:npars])
beta2 = as.numeric(pars[(npars+1):(2*npars)])
theta = as.numeric(pars[length(pars)])

mu1 = exp( X%*%beta1 ); mu2 = exp( X%*%beta2 ) # means

F1 = function(y){
pnbinom(y, mu = mu1[i], size = phi1)
} # Marginal CDF X1 - NB (mu = mu1, size = phi1)

F2 = function(y){
pnbinom(y, mu = mu2[i], size = phi2)
} # Marginal CDF X2 - NB (mu = mu2, size = phi2)

C = function(u1, u2){
-theta^-1 * log( 1 + (exp( -theta*u1 ) - 1) * (exp( -
   theta*u2 ) - 1) / (exp( -theta) - 1) )
} # Joint Distribution function by the Frank Copula
   representation

h = NULL # Joint pmf for discrete random variables

for (i in 1:n) {

if (y1[i] == 0 & y2[i] == 0) {
h[i] = C( F1(0), F2(0) )
}
if (y1[i] != 0 & y2[i] == 0) {
h[i] = C( F1(y1[i]), F2(0) ) - C( F1(y1[i]-1), F2(0) )
}
if (y1[i] == 0 & y2[i] != 0) {
h[i] = C( F1(0), F2(y2[i]) ) - C( F1(0), F2(y2[i] - 1) )
}
if (y1[i] != 0 & y2[i] != 0) {
h[i] = C( F1(y1[i]), F2(y2[i]) ) - C( F1(y1[i] - 1), F2(
   y2[i]) ) - C( F1(y1[i]), F2(y2[i] - 1) ) + C( F1(y1[i]
    - 1), F2(y2[i] - 1) )
}
} # Joint PMF for discrete random variables

sum(log(h))
}
```

REFERENCE LIST

[1] Alan Agresti. *Categorical data analysis*. New York: John Wiley & Sons, 2002.

[2] Jonathan Aguero-Valverde and Paul Jovanis. Bayesian multivariate Poisson lognormal models for crash severity modeling and site ranking. *Transportation Research Record: Journal of the Transportation Research Board*, (2136):82–91, 2009.

[3] J. Aitchison and C.H. Ho. The multivariate Poisson-log normal distribution. *Biometrika*, 76(4):643–653, 1989.

[4] Taoufik Bouezmarni, Mhamed Mesfioui, and Abdelouahid Tajar. On concordance measures for discrete data and dependence properties of Poisson model. *Journal of Probability and Statistics*, 2009, 2010.

[5] A Colin Cameron, Tong Li, Pravin K Trivedi, and David M Zimmer. Modeling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, 7(2):566–584, 2004.

[6] A Colin Cameron and Pravin K Trivedi. *Regression analysis of count data*. Cambridge University Press, 2013.

[7] Siddhartha Chib and Rainer Winkelmann. Markov chain monte carlo analysis of correlated count data. *Journal of Business & Economic Statistics*, 2012.

[8] Nan-Ting Chou. Bivariate count data regression models-a SAS® macro program. *Statistics and Data Analysis, Paper*, 355, 2011.

[9] Michel Denuit and Philippe Lambert. Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis*, 93(1):40–57, 2005.

[10] Karim El-Basyouny and Tarek Sayed. Collision prediction models using Poisson-lognormal regression. *Accident Analysis and Prevention*, 41:820–828, 2009.

[11] Christian Genest and Johanna Neslehova. A primer on copulas for count data. *Astin Bulletin*, 37(9):475–515, 2007.

[12] Christian Genest, Bruno Rémillard, and David Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2):199–213, 2009.

[13] Alan Genz, Frank Bretz, Tetsuhisa Miwa, Xuefei Mi, Friedrich Leisch, Fabian Scheipl, and Torsten Hothorn. *mvtnorm: Multivariate Normal and t Distributions*, 2018. https://CRAN.R-project.org/package=mvtnorm, R package version 1.0-8.

[14] Erika Jaramillo Giraldo. *A methodology to incorporate highway safety into transportation planning using generalized linear mixed models*. PhD thesis, University of Puerto Rico at Mayagüez, 2014.

[15] Gordon Gudendorf and Johan Segers. *Extreme-Value Copulas*, pages 127–145. Springer, 2010.

[16] Shiferaw Gurmu and John Elder. A simple bivariate count data regression model. *Economics Bulletin*, 3(11):1–10, 2007.

[17] Shiferaw Gurmu and John Elder. Flexible bivariate count data regression models. *Journal of Business & Economic Statistics*, 30(2):265–274, 2012.

[18] Joseph M. Hilbe. *Modeling count data*. Cambridge University Press, 2014.

[19] Marius Hofert, Ivan Kojadinovic, Martin Maechler, and Jun Yan. *copula: Multivariate Dependence with Copulas*, 2017. https://CRAN.R-project.org/package=copula, R package version 0.999-18.

[20] Harry Joe. *Dependence modeling with copulas*. Monographs on Statistics and Applied Probability (Book 134). Chapman and Hall/CRC, 2015.

[21] Dimitris Karlis and Evdokia Xekalaki. Mixed Poisson distributions. *International Statistical Review*, 73(1):35–58, 2005.

[22] Subrahmaniam Kocherlakota and Kathleen Kocherlakota. Bivariate discrete distributions. *Encyclopedia of Statistical Sciences*, 1992.

[23] Zhibin Li, Wei Wang, Pan Liu, Lu Bai, and Muqing Du. Analysis of crash risks by collision type at freeway diverge area using multivariate modeling technique. *Journal of Transportation Engineering*, 141, 2015.

[24] Jianming Ma, Kara M. Kockelman, and Paul Damien. A multivariate Poisson lognormal regression model for prediction of crash counts by severity, using bayesian methods. *Accident Analysis and Prevention*, 40(3):964–975, 2008.

[25] Ian McHale and Phil Scarf. Modeling soccer matches using bivariate discrete distributions with general dependence structure. *Statistica Neerlandica*, 61(4):432–445, 2007.

[26] Mhamed Mesfioui and Jean-François Quessy. Concordance measures for multivariate non-continuous random vectors. *Journal of Multivariate Analysis*, 101(10):2398–2410, 2010.

[27] Roger B Nelsen. *An introduction to copulas.* Springer Science & Business Media, 2007.

[28] Johanna Nešlehová. On rank correlation measures for non-continuous random variables. *Journal of Multivariate Analysis*, 98(3):544–567, 2007.

[29] Aristidis K Nikoloulopoulos. A survey on multivariate copula-based models for multivariate discrete response data. In *Copulae in Mathematical and Quantitative Finance*, pages 231–249. Springer, 2013.

[30] Aristidis K. Nikoloulopoulos and Dimitris Karlis. Regression in a copula model for bivariate count data. *Journal of Applied Statistics*, 37(9):1555–1568, 2010.

[31] Eun Sug Park and Dominique Lord. Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record: Journal of the Transportation Research Board*, (2019):1–6, 2007.

[32] M Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959.

[33] Pravin K Trivedi and David M Zimmer. *Copula modeling: an introduction for practitioners*. Now Publishers Inc, 2007.

[34] Linda J. Willson, J. Leroy Folks, and J. H. Young. Multistage estimation compared with fixed-sample-size estimation of the negative binomial parameter k. *Biometrics*, 40(1):109–117, 1984.

[35] Yunlong Zhang, P E Assistant, Ye Zhirui, and Dominique Lord. Estimating dispersion parameter of negative binomial distribution for analysis of crash data: Bootstrapped maximum likelihood method. *Transportation Research Record*, 2019, 12 2006.