Multivariate Copula-based SUR Tobit Models: A Modified Inference Function for Margins and Interval Estimation

Paulo Henrique Ferreira da Silva

Advisor: Prof. Dr. Francisco Louzada Neto

São Carlos, August, 2015

Multivariate Copula-based SUR Tobit Models: A Modified Inference Function for Margins and Interval Estimation

Paulo Henrique Ferreira da Silva

Advisor: Prof. Dr. Francisco Louzada Neto

Thesis submitted to Department of Statistics at Federal University of São Carlos for the award of degree of Doctor of Philosophy.

São Carlos, August, 2015

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária/UFSCar

S586mc

Silva, Paulo Henrique Ferreira da.
Multivariate Copula-based SUR Tobit Models : a modified inference function for margins and interval estimation / Paulo Henrique Ferreira da Silva. -- São Carlos : UFSCar, 2015.
154 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2015.

1. Probabilidades. 2. Intervalos de confiança. 3. *Bootstrap* (Estatística). 4. Cópula. 5. Ampliação de dados. 6. Método de inferência para marginais modificado. I. Título.

CDD: 519.2 (20ª)

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Paulo Henrique Ferreira da Silva, realizada em 30/09/2015:

Prof. Dr. Francisco Louzada Neto
USP

Adriano Kamimura Suzuki

Prof. Dr. Adriano Kamimura Suzuki
USP

Prof. Dr. Heleno Bolfarine
USP

Prof. Dr. Jorge Luis Bazán Guzmán
USP

Prof. Dr. Vicente Garibay Cancho
USP

Acknowledgements

I would like to express my special appreciation and thanks to my advisor Prof. Dr. Francisco Louzada Neto, for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been invaluable.

I take this opportunity to express gratitude to all of the members of the Department of Statistics at Federal University of São Carlos, for their help and support.

A special thanks to my lovely parents, Alzira and Paulo Sérgio. Words can not express how grateful I am to you for all of the sacrifices that you have made on my behalf. I would also like to thank to my fiancée, Ana Paula, for her love, patience and ongoing faith on me.

Finally, I thank God for letting me through all the difficulties. Thank you, Lord.

Abstract

In this thesis, we extend the analysis of multivariate Seemingly Unrelated Regression (SUR) Tobit models by modeling their nonlinear dependence structures through copulas. The capability in coupling together the different - and possibly non-normal - marginal distributions allows the flexible modeling for the SUR Tobit models. In addition, the ability to capture the tail dependence of the SUR Tobit models where some data are censored (e.g., in econometric analysis, clinical essays, wide range of political and social phenomena, among others, data are commonly left-censored at zero point, or right-censored at a point $d > 0$) is another useful feature of copulas. Our study proposes a modified version of the (classical) Inference Function for Margins (IFM) method by Joe & Xu (1996), which we refer to as MIFM method, to obtain the (point) estimates of the marginal and copula association parameters. More specifically, we use a (frequentist) data augmentation technique at the second stage of the IFM method (the first stage of the MIFM method is equivalent to the first stage of the IFM method) to generate the censored observations and then estimate the copula parameter. This procedure (data augmentation and copula parameter estimation) is repeated until convergence. Such modification at the second stage of the usual method is justified in order to obtain continuous marginal distributions, which ensures the uniqueness of the resulting copula, as stated by Sklar (1959)'s theorem; and also to provide an unbiased estimate of the copula association parameter (the IFM method provides a biased estimate of the copula parameter in the presence of censored observations in the margins). Since the usual asymptotic approach, that is the computation of the asymptotic covariance matrix of the parameter estimates, is troublesome in this case, we also propose the use of resampling procedures (bootstrap methods, like standard normal and percentile by Efron & Tibshirani (1993), and basic bootstrap by Davison & Hinkley (1997)) to obtain confidence intervals for the copula-based SUR Tobit model parameters.

Resumo

Nesta tese de doutorado, consideramos os chamados modelos SUR (da expressão Seemingly Unrelated Regression) Tobit multivariados e estendemos a análise de tais modelos ao empregar funções de cópula para modelar estruturas com dependência não linear. As cópulas, dentre outras características, possuem a importante habilidade (vantagem) de capturar/modelar a dependência na(s) cauda(s) do modelo SUR Tobit em que alguns dados são censurados (por exemplo, em análise econométrica, ensaios clínicos e em ampla gama de fenômenos políticos e sociais, dentre outros, os dados são geralmente censurados à esquerda no ponto zero, ou à direita em um ponto $d > 0$ qualquer). Neste trabalho, propomos uma versão modificada do método clássico da Inferência para as Marginais (IFM, da expressão Inference Function for Margins), originalmente proposto por Joe & Xu (1996), a qual chamamos de MIFM, para estimação (pontual) dos parâmetros do modelo SUR Tobit multivariado baseado em cópula. Mais especificamente, empregamos uma técnica (frequentista) de ampliação de dados no segundo estágio do método IFM (o primeiro estágio do método MIFM é igual ao primeiro estágio do método IFM) para gerar as observações censuradas e, então, estimamos o parâmetro de dependência da cópula. Repetimos tal procedimento (ampliação de dados e estimação do parâmetro da cópula) até obter convergência. As razões para esta modificação no segundo estágio do método usual, são as seguintes: primeiro, construir/obter distribuições marginais contínuas, atendendo, então, ao teorema de unicidade da cópula resultante de Sklar (Sklar, 1959); e segundo, fornecer uma estimativa não viesada para o parâmetro da cópula (uma vez que o método IFM produz estimativas viesadas do parâmetro da cópula na presença de observações censuradas nas marginais). Tendo em vista a dificuldade adicional em calcular/obter a matriz de covariâncias assintótica das estimativas dos parâmetros, também propomos o uso de procedimentos de reamostragem (métodos bootstrap, tais como normal padrão e percentil, propostos por Efron & Tibshirani (1993), e básico, proposto por Davison

& Hinkley (1997)) para a construção de intervalos de confiança para os parâmetros do modelo SUR Tobit baseado em cópula.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Tobit model refers to a class of regression models whose range of the dependent variable (or response variable) is somehow constrained. It was first proposed in 1958 by the 1981 Nobel Prize winner in Economic Sciences, James Tobin, to describe the relationship between a non-negative dependent variable y (the ratio of total durable goods expenditure to total disposable income, per household) and a vector of independent variables $\boldsymbol{x}$ (the age of the household head, and the ratio of liquid asset holdings to total disposable income) (see Tobin, 1958). Tobin called his model the *limited dependent variable model.* However, it and its various generalizations are popularly known among economists as *Tobit models*, a phrase coined by Goldberger (1964) because of similarities to *probit models* (the term *Tobit* aims to synthesize in one word the concept "Tobin's probit"). Tobit models are also known as *censored* or *truncated* regression models.

Particularly, the presence of censoring (left-censoring, right-censoring or both) occurs when data on the dependent variable is limited or lost. Examples are:

1. **Left-censoring.** Antibody concentration values in Haitian 12-month-old infants vaccinated against measles are determined through neutralization antibody assays with the lower detection limit of 0.1 IU (Moulton & Halsey, 1995). Thus, concentration values under or equal to 0.1 are reported as 0.1.

2. **Right-censoring.** People of all income levels are included in the sample, but for some reason high-income people have their income coded as R$ 100,000 (Bolfarine, Santos, Correia, Martínez, Goméz & Bazán, 2013).

3. **Both left- and right-censoring.** Scores of students on an academic aptitude test can be any value between 200 and 800, and it is not rare to observe students

answering all questions in the test correctly, thus receiving a score of 800 (even though it is likely that these students are not “truly” equal in aptitude); or students answering all of the questions incorrectly, thus receiving a score of 200 (although they may not all be of equal aptitude).

The Tobit specification is appropriate for the situation in which the sample proportion of censored observations is roughly equivalent to the remaining tail area of the assumed parametric distribution. The Cragg (1971) model, which in the classical literature is known as the *two-part model*, is an alternative to Tobit when the data rate below or above the threshold is quite different from the probability of the tail obtained with the assumed parametric model.

The censoring problem also arises in situations with the presence of multiple dependent variables. For example, Chen & Zhou (2011) consider the joint problem of censoring and simultaneity when working with multivariate microeconomic data.

The next section describes two real datasets that show such characteristics (i.e. censoring and multiple correlated dependent variables).

1.1 The data

This section introduces the datasets that will be used to illustrate the approaches proposed in this thesis.

1.1.1 U.S. salad dressing, tomato and lettuce consumption data

The United States is the second largest tomato and lettuce-producing country after China. In terms of consumption, tomatoes are the United States’ fourth most popular fresh-market vegetable after potatoes, lettuce and onions. Over the past few decades, per capita use of tomatoes has been on the rise (in 2011, 86.3 pounds per person of tomatoes were available for Americans to eat) as a result of the enduring popularity of salads, salad bars, and bacon-lettuce-tomato (BLT) and submarine (sub) sandwiches. On the other hand, the total lettuce consumption, i.e. consumption of all lettuce varieties by Americans reached a high record of 34.5 pounds per capita in 2004. However, as discussed in Mintel’s ”Bagged Salad and Salad Dressings - U.S., July 2008”, salad dressing sales have declined since 2005. Among other reasons, it is due to the fact that health-oriented consumers who eat large amounts of tomatoes, lettuce and other vegetables are curtailing consumption of salad dressings perceived as high in fat, calories and sodium.

Our study aims at establishing some factors (age, region/location and income, among others) that influence the consumption of tomatoes (including raw and cooked tomatoes, tomato juices, tomato sauces and mixtures having tomatoes as a main ingredient), lettuce (including all plain, Boston and Romaine lettuce reported separately or as part of a mixed salad or sandwich) and salad dressing products (including mayonnaise-type salad dressing reported separately or as part of a sandwich, and pourable salad dressings reported separately or as part of a mixture such as a salad) by U.S. individuals. This study is based on part of a dataset extracted from the 1994-1996 Continuing Survey of Food Intakes by Individuals (CSFII) (USDA, 2000). In the CSFII, two non-consecutive days of dietary data for individuals of all ages residing in the United States were collected through in-person interviews using 24-hour recall. Each sample person reported the amount of each food item consumed. Where two days were reported there is also a third record containing daily averages. Socioeconomic and demographic data for the sample households and their members were also collected in the CSFII. The size of the extracted sample here is $n = 400$ adults age 20 or older. We only consider one member per household.

Table 1.1 provides the definitions and sample statistics for all considered variables, where we observe the proportions of consuming individuals in the dataset to range from 85.00% for salad dressings, to 63.25% for tomatoes and 67.25% for lettuce. Among those consuming, an individual on average consumes 32.84 g of salad dressings, 66.56 g of tomatoes and 60.52 g of lettuce per day.

In Figure 1.1, the histograms, and in Figures 1.2 and 1.3, the three-dimensional (3D) and two-dimensional (2D) scatter plots, respectively, show some features of the data and model we work on: all three dependent variables (salad dressing, tomato and lettuce consumption) are limited (left-censored or lower-bounded by zero, since there are some individuals in the extracted sample who did not consume tomatoes, lettuce and/or salad dressings during the survey period) and there is a considerable positive association among salad dressing, tomato and lettuce consumption data (the Kendall tau rank correlation coefficient between salad dressing and tomato, salad dressing and lettuce, and tomato and lettuce consumption is 0.3522, 0.5572 and 0.3437, respectively). These features, as well as the presence of covariates (age, region and income), suggest that the relationship among

Table 1.1: Variable definitions and sample statistics ($n = 400$).

Variable	Definition	Mean	Standard Deviation
Dependent variables: amount consumed			
Salad dressing (in 100 g)	Quantity of salad dressings consumed	0.2791	0.2371
	Among the consuming ($n = 340$; 85.00%)	0.3284	0.2235
Tomato (in 400 g)	Quantity of tomatoes consumed	0.1052	0.1526
	Among the consuming ($n = 253$; 63.25%)	0.1664	0.1633
Lettuce (in 200 g)	Quantity of lettuce consumed	0.2035	0.2348
	Among the consuming ($n = 269$; 67.25%)	0.3026	0.2280
Continuous explanatory variable			
Income	Household income as the proportion of poverty threshold	2.3160	0.8404
Binary explanatory variables (yes = 1; no = 0)			
Age 20-30	Age is 20-30	0.1375	
Age 31-40	Age is 31-40	0.1600	
Age 41-50	Age is 41-50	0.1900	
Age 51-60	Age is 51-60	0.1725	
Age > 60	Age > 60 (reference)	0.3400	
Northeast	Resides in the Northeastern states	0.1850	
Midwest	Resides in the Midwestern states	0.2450	
South	Resides in the Southern states (reference)	0.3500	
West	Resides in the Western states	0.2200	

Source: Compiled from the CSFII, USDA, 1994-1996.

the reported salad dressing, tomato and lettuce consumption could be modeled through a trivariate regression model with limited (left-censored at zero) dependent variables.

1.1.2 Brazilian commercial bank customer churn data

Customer churn, also known as customer attrition or customer defection, has become a major issue for most banks in terms of representing the loss of clients or customers as they stop using certain products or services. According to Wang, Liu, Peng, Nie, Kou & Shi (2010), an important reason for customer churn analysis is that the cost of acquiring/developing a new customer is much higher than that of retaining an existing one. Generally, it costs up to five times as much to make a new sale to a new customer as it does to make an additional sale to an existing customer (Dixon, 1999; Slater & Narver, 2000). Reichheld & Sasser (1990) found that a bank can increase its profits by 85% by enhancing the customer retention rate by 5%.

Our study aims at establishing a few factors (age and income, among others) that influence the time (in years) to churn/cancel three credit products (hereafter, Products A, B and C for reasons of confidentiality) for 927 customers of a Brazilian commercial bank. These customers started a relationship with this financial institution almost at the same time (i.e. the same month) and about 10 years before the financial institution was acquired by a bank holding company. This process is popularly known as merger and

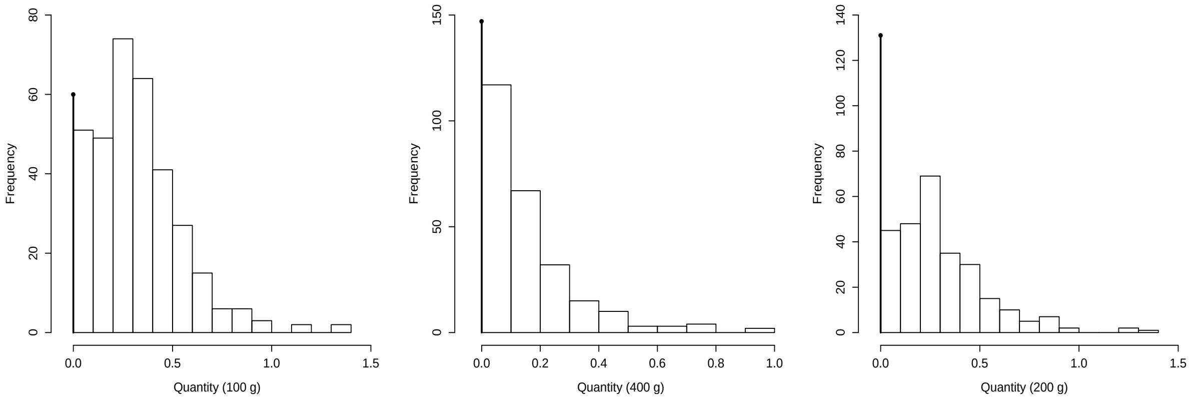


Figure 1.1: Distributions of the salad dressing (left panel), tomato (middle panel) and lettuce (right panel) consumption. The vertical line at zero on x axis represents individuals that did not consume salad dressings, tomatoes or lettuce during the survey period.

Table 1.2: Variable definitions and sample statistics ($n = 927$).

Variable	Definition	Mean	Standard Deviation
Dependent variables: in log of years			
Product A	Log of time to churn Product A	1.1200	0.9118
	Among the uncensored ($n = 777$; 83.82%)	0.8925	0.8192
Product B	Log of time to churn Product B	1.2610	0.8325
	Among the uncensored ($n = 745$; 80.37%)	1.0070	0.7306
Product C	Log of time to churn Product C	1.1500	0.8987
	Among the uncensored ($n = 765$; 82.52%)	0.9069	0.7996
Continuous explanatory variable			
Age	Age in completed years	43.2000	15.0241
Income	Monthly income in Brazilian reais (BRL)	1,524.0000	2,385.3710

acquisition (M&A) or takeover (Hildebrandt, 2007). Thus, the range of each dependent variable (time to churn Product A, time to churn Product B and time to churn Product C) is bounded by the interval zero year (i.e. customers close their accounts before completing the first year of the relationship) to ten years (i.e. customers still with the bank at the acquisition date).

Table 1.2 provides the definitions and sample statistics for all considered variables, where we observe the proportions of uncensored observations (i.e. customers whose log of time to churn is less than 2.3 or 10 years) in the dataset to range from 83.82% for Product A, to 80.37% for Product B and 82.52% for Product C. Among those uncensored, a customer on average churns Product A in 0.8925 log of years or 2.44 years; Product B in 1.0070 log of years or 2.74 years; and Product C in 0.9069 log of years or 2.48 years.

In Figure 1.4, the histograms, and in Figures 1.5 and 1.6, the 3D and 2D scatter plots, respectively, show some features of the data and model we work on: all three dependent variables (log-transformed) are limited (upper-bounded or right-censored at point $d = 2.3$

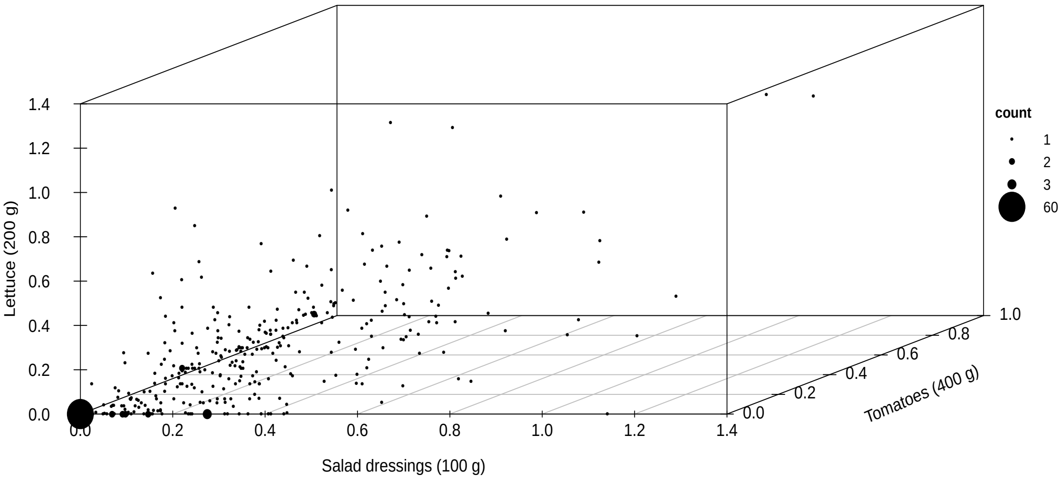


Figure 1.2: 3D scatter plot of salad dressing versus tomato versus lettuce. The bold ball sizes are related to the number of pair of data with the same dependent variable values.

or approximately 10 years) and there is a considerable positive association among the log of times to churn Products A, B and C (the Kendall tau rank correlation coefficient between the log of times to churn Products A and B, Products A and C, and Products B and C is 0.6386, 0.5389 and 0.5928, respectively). These features, as well as the presence of covariates (age and income), suggest that the relationship among the reported log of times to churn Products A, B and C could be modeled through a trivariate regression model with limited (right-censored at point $d = 2.3$) dependent variables.

1.2 Literature review

The multivariate Tobit models, which generalize univariate Tobit ones to systems of equations, is a class of models able to address the above-mentioned issues in Sections 1.1.1 and 1.1.2. There are several generalizations available in the literature, each designed to uniquely capture features of each particular application. See, e.g., Lee (1993) for a survey. Our thesis considers the Seemingly Unrelated Regression (SUR) Tobit model, which is a SUR-type model, i.e. a set of linear regression equations where all dependent variables are partially observed or censored. In the SUR models, each equation is a valid linear regression on its own and can be estimated separately, which is the reason why the system is called *seemingly unrelated* (Greene, 2003). However, some authors, like Davidson & MacKinnon (2003), suggest that the term *seemingly related* would be more appropriate, since the error terms are assumed to be correlated across the equations. See, e.g., Zellner (1962), Greene (2003, Chapter 14), Davidson & MacKinnon (2003, Chapter 12) and Zellner & Ando (2010) for more details on the SUR models; and Amemiya (1984) for a thorough review of various types of Tobit models.

Several estimation techniques have been proposed to implement the SUR Tobit model. See, e.g., Wales & Woodland (1983), Brown & Lankford (1992) and Kamakura & Wedel (2001) for the maximum likelihood (ML) estimation; Huang, Sloan & Adamache (1987) for the expectation-maximization; Meng & Rubin (1996) for the expectation-conditional maximization (ECM); and Huang (1999) for the Monte Carlo ECM (MCECM). Moreover, Huang (2001), Baranchuk & Chib (2008) and Taylor & Phaneuf (2009) implement the SUR Tobit model through the Bayesian approach using Gibbs samplers, while Chen & Zhou (2011) estimate the model parameters in the semiparametric context. However, all these estimation methods are cumbersome (i.e. computationally demanding and difficult to implement), especially for high dimensions. Trivedi & Zimmer (2005) suggest this as a reason why the SUR Tobit model is not well applied. These methods also assume normal marginal error distributions, which may be inappropriate in many real applications. In addition, modeling the dependence structure of the SUR Tobit model through the multivariate normal distribution is restricted to the linear relationship among marginal distributions through the correlation coefficients.

In order to relax the assumptions on the same normally-distributed margins and their linear dependence structure, we can use copulas to analyze the SUR Tobit model (Wichitaksorn, Choy & Gerlach, 2012). According to Sklar's theorem (Sklar, 1959), copulas are used to model the nonlinear dependence structure of the margins that can follow any arbitrary distributions. See, e.g., Joe (1997), McNeil, Frey & Embrechts (2005, Chapter 5) and Nelsen (2006) for further details on copulas. The copulas have been successfully applied in many financial and economic applications with continuous and discrete margins (Pitt, Chan & Kohn, 2006; Smith & Khaled, 2012; Panagiotelis, Czado & Joe, 2012). Nevertheless, the case of censored (or semi-continuous) margins has not been widely studied and applied, as pointed out by Wichitaksorn *et al.* (2012). Moreover, the tail coefficients from some copulas can reveal the dependence at the tails where some data are censored. Trivedi & Zimmer (2005) implement the bivariate SUR Tobit model through a few copulas (Clayton, Frank, Gaussian and Farlie-Gumbel-Morgenstern) to model the

U.S. out-of-pocket and non-out-of-pocket medical expenses data, finding that the two-stage ML/Inference Function for Margins (IFM) estimation results are unstable. This is not surprising considering the previous findings about the inconsistency of ML estimators of the parameters of the Tobit model with non-normal errors (Cameron & Trivedi, 2005). Yen & Lin (2008) estimate the copula-based censored equation system (a system of four meat products - beef, pork, poultry and fish - consumed by U.S. individuals) via the quasi-ML estimation method, yet considering the Frank copula with generalized log-Burr margins (the generalized log-Burr distribution nests the logistic distribution, which is kin to the normal distribution) exclusively. Finally, Wichitaksorn *et al.* (2012) apply and combine the data augmentation techniques by Geweke (1991), Chib (1992), Chib & Greenberg (1998), Pitt *et al.* (2006) and Smith & Khaled (2012) to simulate the unobserved marginal dependent variables and proceed with the bivariate copula-based SUR Tobit model implementation through Bayesian Markov Chain Monte Carlo methods as in other copula models with continuous margins. In their work, the relationship between the self-reported out-of-pocket and non-out-of-pocket medical expenses of elderly Americans, as well as the relationship between the wage earnings income of household head and members living in the rural households in Thailand, are described by bivariate SUR Tobit models with Student-t margins through four different copulas (Gaussian, Student-t, Frank and Clayton).

1.3 Objectives

In this thesis, inspired by the (Bayesian) work of Wichitaksorn *et al.* (2012), we propose/develop a modified version of the (classical) IFM method by Joe & Xu (1996), hereafter Modified Inference Function for Margins (MIFM) method, to implement the SUR Tobit model with arbitrary margins through copulas. The MIFM method consists of the most significant contribution of this thesis. For now, we consider only the (one-parameter) Clayton copula and its survival (or reflected) copula, as well as symmetric (normal), asymmetric (power-normal) and heavy-tailed (logistic) distributions for the marginal errors. The copula-based SUR Tobit models with asymmetric (power-normal) marginal errors is another major contribution of this thesis. These error choices were directed mainly by the dataset features detected in Sections 1.1.1 and 1.1.2. Regarding the first dataset, its features indicate that the relationship among the reported salad dressing, tomato and lettuce consumption, in the presence of covariates (age, region and income), could be modeled through the trivariate SUR Tobit model with left-censored (at zero point) normally-, power-normally- or logistically-distributed dependent variables based on the one-parameter Clayton copula. Note from Figure 1.1 that the assumption of normality of marginal errors, or equivalently, the assumption of left-censored normal distribution of the observed dependent variables does not seem to be a reasonable one to make (all distributions seem to have a right-tail heavier than the normal tail). From Figure 1.2, we see that there is a high number of 3-tuple zero ($n = 60$ observations); this seems to indicate the strongest relationship among the three dependent variables/margins in their lower regions (i.e. for low or no consumption of salad dressings, tomatoes and lettuce), where data are most concentrated. Therefore, the use of the Clayton copula with only one parameter is justified in order to accommodate the possible existence of lower tail dependence, as well as positive nonlinear dependence of the same magnitude (since the Kendall tau values for each pair of dependent variables are not so different; see Section 1.1.1). Furthermore, Figures 1.1 and 1.3 have indications that each pair of dependent variables could be modeled through the bivariate SUR Tobit model with left-censored (at zero point) normally-, power-normally- or logistically-distributed dependent variables based on the Clayton copula. On the other hand, the second dataset has indications that the relationship among the reported log of times to churn Products A, B and C, in the presence of covariates (age and income), could be modeled through the trivariate SUR Tobit model with right-censored (at point $d = 2.3$) normally-, power-normally- or logistically-distributed dependent variables based on the one-parameter Clayton survival copula. Note from Figure 1.4 that the assumption of normality of marginal errors, or equivalently, the assumption of right-censored normal distribution of the observed dependent variables may be doubtful. From Figure 1.5, we observe that there is a high number of 3-tuple 2.3 ($n = 95$ observations); which seems to indicate the strongest relationship among the three dependent variables in their upper regions (i.e. for high times or log of times to churn Products A, B and C). Thus, the use of the Clayton survival copula with just a single parameter is justified in order to accommodate the possible existence of upper tail dependence, as well as positive nonlinear dependence of the same magnitude (provided that the Kendall tau values for each pair of dependent variables are similar; see Section 1.1.2). Moreover, Figures 1.4 and 1.6 have indications that each pair of dependent variables could be modeled through the bivariate SUR Tobit model with right-censored (at point $d = 2.3$) normally-, power-normally- or logistically-distributed dependent variables based on the Clayton survival copula. In this work, we also decided for the Clayton and Clayton survival copulas guided by the literature, which states that these copula families have a remarkable and useful (as will be seen in Sections 2.1.1.1, 2.2.1.1, 3.1.1.1, 3.2.1.1, 4.1.1.1 and 4.2.1.1) invariance property under truncation.

In short, the MIFM method proposed in this thesis uses a (frequentist) data augmentation technique at the second stage of the IFM method (the IFM method provides biased estimates of the Clayton and Clayton survival copulas' association parameter, as will be seen in Sections 2.1.2.2, 2.2.2.2, 3.1.2.2 and 3.2.2.2) to generate the censored observations/margins and thus obtain a better (unbiased) estimate of the copula dependence parameter. This modification also aims to satisfy the Sklar's theorem, which states that marginal distributions should be continuous to ensure the uniqueness of the resulting copula. Since the usual asymptotic approximation, that is the computation of the asymptotic covariance matrix of the parameter estimates, is cumbersome in this case, we consider resampling procedures (a parametric resampling plan) to obtain confidence intervals for the copula-based SUR Tobit model parameters. More specifically, we use the standard normal and percentile methods by Efron & Tibshirani (1993), and the basic method by Davison & Hinkley (1997), to build bootstrap confidence intervals.

1.4 Overview

The thesis has the following organization. In Chapter 2, we present the bivariate copula-based SUR Tobit models (i.e. the bivariate Clayton copula-based SUR Tobit model and the bivariate Clayton survival copula-based SUR Tobit right-censored model, both with normal, power-normal and logistic distribution assumption for the marginal errors), discuss inference for the models' parameters, showing the models' implementations through the MIFM method and the confidence intervals construction using the bootstrap approach; present the simulation studies used to evaluate our proposed models and methods; and provide applications of our procedures to real datasets. In Chapter 3, we extend the bivariate ideas, i.e. the bivariate models and methods to the trivariate case. Chapter 3 also presents the simulation studies conducted and the empirical applications. In Chapter 4, we present a straightforward generalization of the models and methods proposed in this thesis for the m-variate ($m \geq 2$) case. Finally, Chapter 5 concludes the thesis with final remarks and a few indications for further studies.

It is useful to note that this thesis is organized as a series of papers. More advanced readers may skip ahead to Chapter 4 concerning multivariate models and methods after reading Chapter 1, and then proceed to Chapters 2 and 3 as they provide the simulation studies and empirical applications for particular cases of the multivariate approach, i.e. bivariate and trivariate models and methods, respectively.

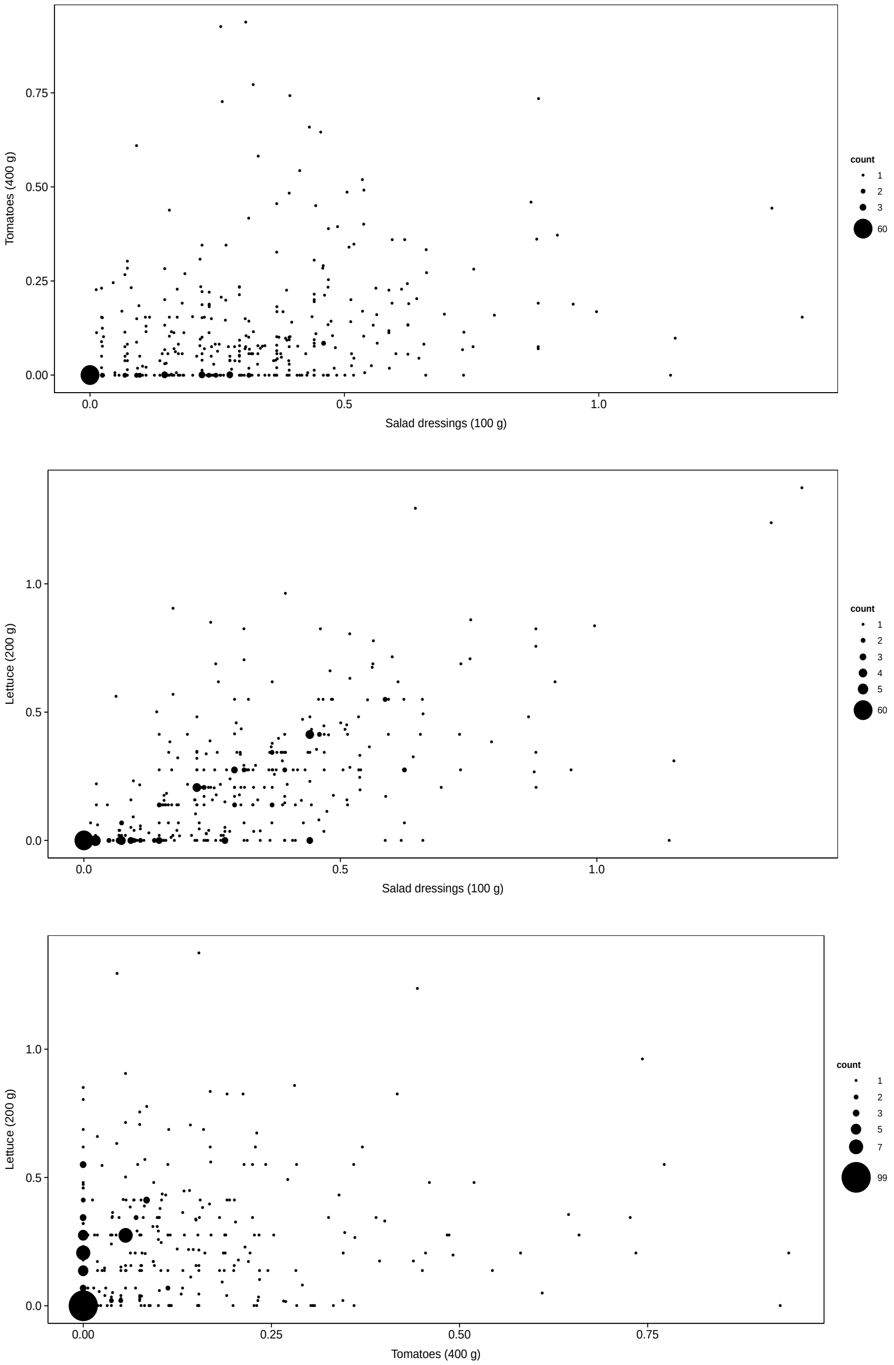


Figure 1.3: 2D scatter plots of salad dressing versus tomato (upper panel), salad dressing versus lettuce (middle panel) and tomato versus lettuce (lower panel). The bold ball sizes are related to the number of pair of data with the same dependent variable values.

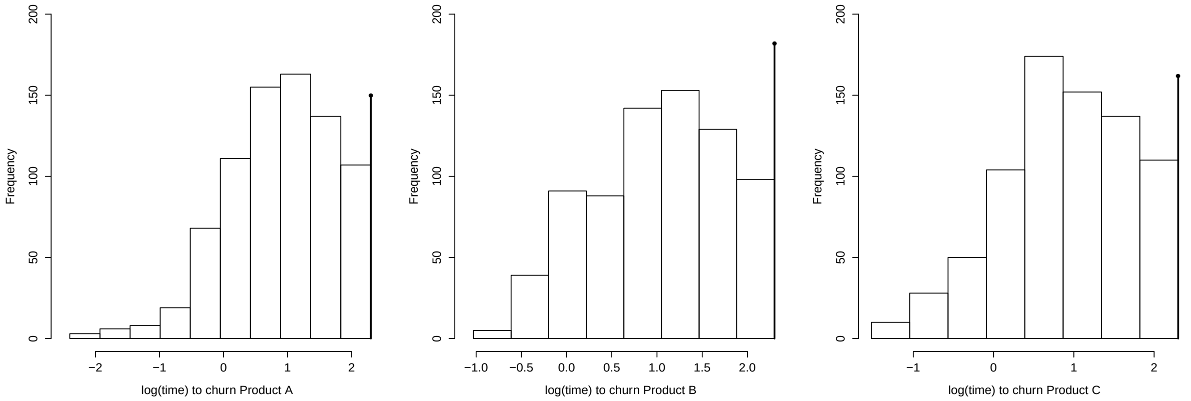


Figure 1.4: Distributions of the log(time) to churn Product A (left panel), log(time) to churn Product B (middle panel) and log(time) to churn Product C (right panel) variables. The vertical line at 2.3 on x axis represents customers still with the bank at the acquisition date.

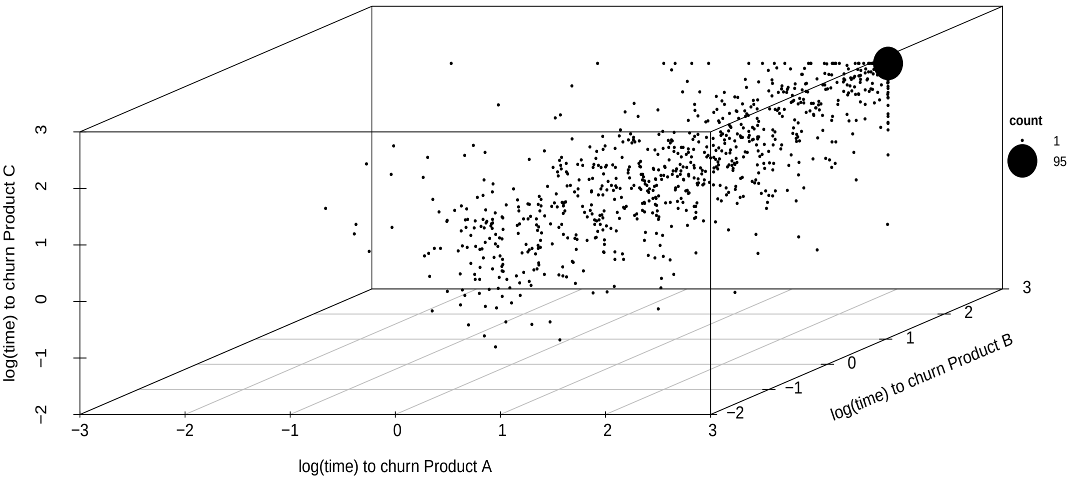


Figure 1.5: 3D scatter plot of log(time) to churn Product A versus log(time) to churn Product B versus log(time) to churn Product C. The bold ball sizes are related to the number of pair of data with the same dependent variable values.

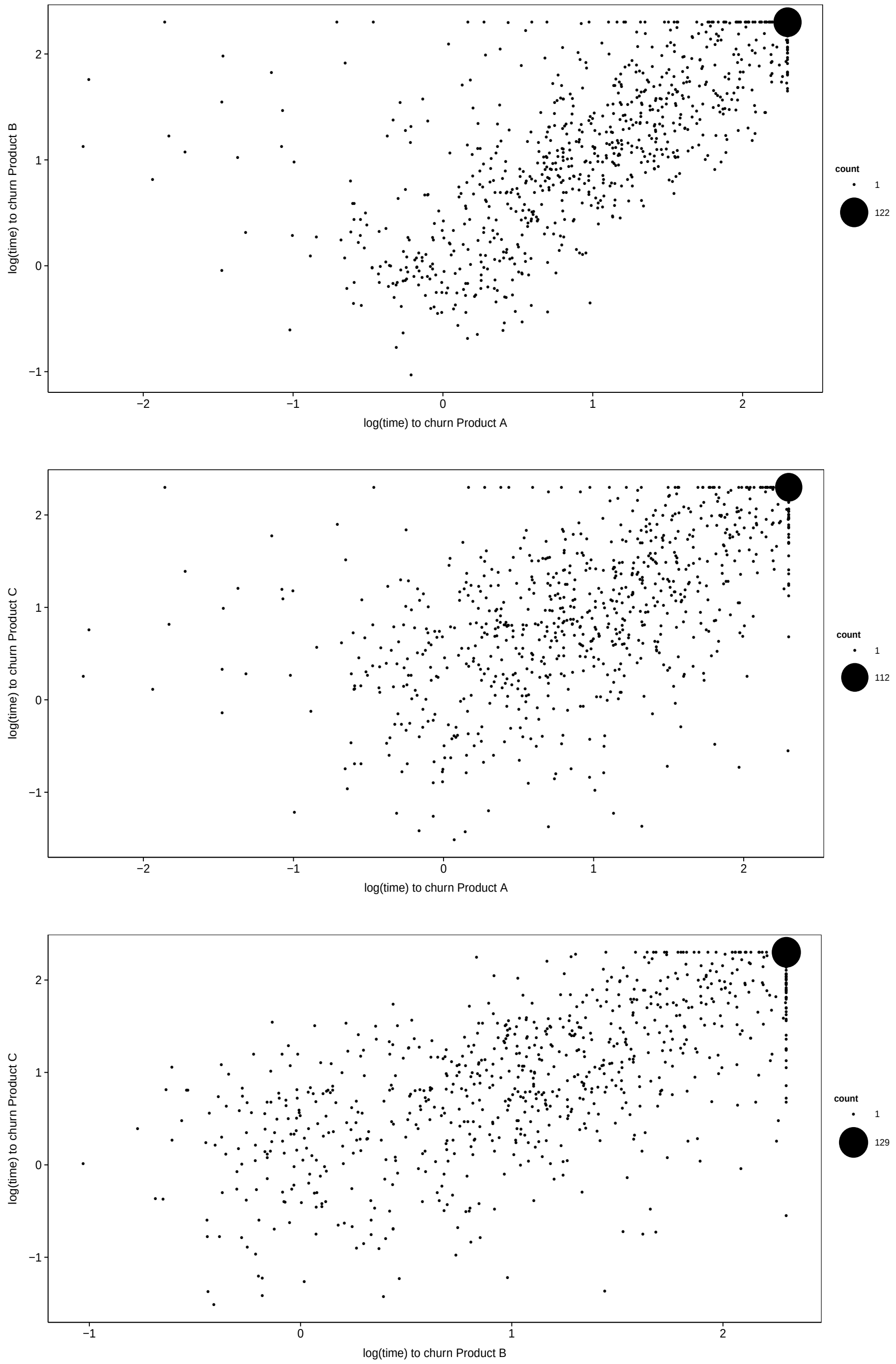


Figure 1.6: 2D scatter plots of log(time) to churn Product A versus log(time) to churn Product B (upper panel), log(time) to churn Product A versus log(time) to churn Product C (middle panel), and log(time) to churn Product B versus log(time) to churn Product C (lower panel). The bold ball sizes are related to the number of pair of data with the same dependent variable values.

Chapter 2

Bivariate Copula-based SUR Tobit Models

In this chapter, we present the bivariate copula-based SUR Tobit models proposed in this thesis. We first present the bivariate Clayton copula-based SUR Tobit model, i.e. the SUR Tobit model with two left-censored (at zero point) dependent variables whose dependence between them is modeled through the Clayton copula. Then, we present the bivariate Clayton survival copula-based SUR Tobit right-censored model, which is the SUR Tobit model with two right-censored (at point $d_j > 0$, $j = 1, 2$) dependent variables whose dependence structure between them is modeled by the Clayton survival copula. In both cases, we assume symmetric, asymmetric and heavy-tailed distributions for the marginal error terms. Discussions concerning the model implementation through the proposed MIFM method, as well as the confidence intervals construction from the bootstrap distribution of model parameters, are made for each proposed model. Simulation studies and applications to real datasets are also provided in this chapter.

2.1 Bivariate Clayton copula-based SUR Tobit model formulation

The SUR Tobit model with two left-censored (at zero point) dependent variables, or simply bivariate SUR Tobit model, is expressed as

$$y^*_{ij} = \boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j + \epsilon_{ij},$$

$$y_{ij} = \begin{cases} y^*_{ij} & \text{if } y^*_{ij} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, ..., n$ and $j = 1, 2$, where n is the number of observations, y_{ij}^{*} is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^{*} whenever y_{ij}^{*} is above zero and zero otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients and ϵ_{ij} is the margin j's error that follows some zero mean distribution.

Suppose that the marginal errors are no longer normal, but they are assumed to be distributed according to the power-normal (Gupta & Gupta, 2008) and logistic models, thus providing asymmetric and heavy-tailed alternatives to Tobin's model (Tobin, 1958). These choices of error distribution consist of expressing the density function of y_{ij} in the following forms.

- Normal marginal errors (i.e. $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$):

$$f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j\right) = \begin{cases} 1 - \Phi\left(\frac{\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{\sigma_j}\right) & \text{if } y_{ij} = 0, \\ \frac{1}{\sigma_j}\phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{\sigma_j}\right) & \text{if } y_{ij} > 0, \end{cases} \tag{2.1}$$

  (Trivedi & Zimmer, 2005), where $\phi(.)$ and $\Phi(.)$ are the standard normal probability density function (p.d.f.) and cumulative distribution function (c.d.f.), respectively. Note that if $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$, then we have marginal standard Tobit models or Type I Tobit models (Amemiya, 1984). The corresponding distribution function of y_{ij} is denoted by $F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j\right)$ and is obtained by replacing $\phi(.)$ with $\Phi(.)$ and removing $1 \,/\, \sigma_j$ from the second part of (2.1) (i.e. where $y_{ij} > 0$).

- Power-normal marginal errors (i.e. $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$):

$$f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j, \alpha_j\right) = \begin{cases} \left[\Phi\left(-\frac{\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{\sigma_j}\right)\right]^{\alpha_j} & \text{if } y_{ij} = 0, \\ \frac{\alpha_j}{\sigma_j}\phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{\sigma_j}\right)\left[\Phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{\sigma_j}\right)\right]^{\alpha_j - 1} & \text{if } y_{ij} > 0, \end{cases} \tag{2.2}$$

  where $\alpha_j > 0$ is a shape parameter that controls the amount of asymmetry in the distribution, as well as the distribution kurtosis (for $\alpha_j > 1$, the kurtosis is greater than that of the normal distribution, and for $0 < \alpha_j < 1$ the opposite is observed). Note that (2.1) is recovered when $\alpha_j = 1$. For further details on the power-normal distributions, see Gupta & Gupta (2008). If we assume $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$, then we have marginal power-normal Tobit models (Martínez-Floréz, Bolfarine & Gómez, 2013). The corresponding distribution function of y_{ij} is denoted by $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,\sigma_j,\alpha_j\right)$ and is obtained by replacing $\phi\left(.\right)$ with $\Phi\left(.\right)$ and removing $\alpha_j\,/\,\sigma_j$ from the second part of (2.2) (i.e. where $y_{ij}>0$).

- Logistic marginal errors (i.e. $\epsilon_{ij}\sim L\left(0,s_j\right)$):

$$f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,s_j\right)=\begin{cases}1-G\left(\frac{\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{s_j}\right) & \text{if } y_{ij}=0,\\ \frac{1}{s_j}g\left(\frac{y_{ij}-\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j}{s_j}\right) & \text{if } y_{ij}>0,\end{cases} \tag{2.3}$$

where $g\left(z\right)=e^z\,/\left(1+e^z\right)^2$ and $G\left(z\right)=1\,/\left(1+e^{-z}\right)$ are the $L\left(0,1\right)$ p.d.f. and c.d.f., respectively. The corresponding distribution function of y_{ij}, $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,s_j\right)$, is obtained by replacing $g\left(.\right)$ with $G\left(.\right)$ and removing $1\,/\,s_j$ from the second part of (2.3) (i.e. where $y_{ij}>0$).

Usually, the dependence between the error terms ϵ_{i1} and ϵ_{i2} is modeled through a bivariate distribution, especially the bivariate normal distribution (such specification characterizes the basic bivariate SUR Tobit model; see, e.g., Huang *et al.* (1987) for more details on this model). However, as commented before (in Section 1.2), a restriction in applying a bivariate distribution to the bivariate SUR Tobit model is the linear relationship between marginal distributions through the correlation coefficient. One way to overcome this restriction is to use a copula function to capture/model the nonlinear dependence structure in the bivariate SUR Tobit model.

Thus, for the censored outcomes y_{i1} and y_{i2}, the bivariate copula-based SUR Tobit distribution is given by

$$F\left(y_{i1},y_{i2}\right)=C\left(u_{i1},u_{i2}|\theta\right),$$

where, e.g., $u_{ij}=F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,\sigma_j\right)$ if $\epsilon_{ij}\sim N\left(0,\sigma_j^2\right)$, $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,\sigma_j,\alpha_j\right)$ if $\epsilon_{ij}\sim PN\left(0,\sigma_j,\alpha_j\right)$, and $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\beta}_j,s_j\right)$ if $\epsilon_{ij}\sim L\left(0,s_j\right)$, for $j=1,2$, and θ is the association parameter (or parameter vector) of the copula, which is assumed to be scalar.

Suppose C is the bidimensional Clayton (1978) copula, also referred to as the Cook & Johnson (1981) copula, originally studied by Kimeldorf & Sampson (1975). It takes the form

$$C\left(u_{i1},u_{i2}|\theta\right)=\left(u_{i1}^{-\theta}+u_{i2}^{-\theta}-1\right)^{-\frac{1}{\theta}}, \tag{2.4}$$

with θ restricted to the region $(0,\infty)$. The dependence between the margins increases with the value of θ, with $\theta\to0^+$ implying independence and $\theta\to\infty$ implying perfect positive dependence. The Clayton copula does not allow for negative dependence. In the survival analysis framework, there is an equivalence between the Clayton copula and the shared gamma frailty model (see, e.g., Goethals, Janssen & Duchateau, 2008). Trivedi & Zimmer (2005) point out that the Clayton copula is widely used to study correlated risks because it shows strong left tail dependence and relatively weak right tail dependence. Indeed, when correlation between two events is stronger in the left tail of the joint distribution, Clayton is usually an appropriate modeling choice.

2.1.1 Inference

In this subsection, we discuss inference (point and interval estimation) for the parameters of the bivariate Clayton copula-based SUR Tobit model. Particularly, by considering/assuming normal, power-normal and logistic distributions for the marginal errors.

2.1.1.1 Estimation through the MIFM method

According to Trivedi & Zimmer (2005), the log-likelihood function for the bivariate Clayton copula-based SUR Tobit model can be written in the following form [1]

$$\ell\left(\boldsymbol{\eta}\right)=\sum_{i=1}^{n}\log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1},\boldsymbol{\upsilon}_1\right),F_2\left(y_{i2}|\boldsymbol{x}_{i2},\boldsymbol{\upsilon}_2\right)|\theta\right)+\sum_{i=1}^{n}\sum_{j=1}^{2}\log f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right),\qquad(2.5)$$

where $\boldsymbol{\eta}=(\boldsymbol{\upsilon}_1,\boldsymbol{\upsilon}_2,\theta)$ is the vector of model parameters, $\boldsymbol{\upsilon}_j$ is the margin j's parameter vector, $f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$ is the c.d.f. of y_{ij}, and $c\left(u_{i1},u_{i2}|\theta\right)$, with $u_{ij}=F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton copula, which is calculated from (2.4) as

$$c\left(u_{i1},u_{i2}|\theta\right)=\frac{\partial^2 C\left(u_{i1},u_{i2}|\theta\right)}{\partial u_{i1}\partial u_{i2}}=\left(\theta+1\right)\left(u_{i1}u_{i2}\right)^{-\theta-1}\left(u_{i1}^{-\theta}+u_{i2}^{-\theta}-1\right)^{-\frac{1}{\theta}-2}.$$

For model estimation, the use of copula methods, as well as the log-likelihood function form given by (2.5), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}=\arg\max_{\boldsymbol{\upsilon}_j}\sum_{i=1}^{n}\log f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right),\qquad(2.6)$$

for $j=1,2$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$ by

$$\widehat{\theta}_{\mathrm{IFM}}=\arg\max_{\theta}\sum_{i=1}^{n}\log c\big(F_1\left(y_{i1}|\boldsymbol{x}_{i1},\widehat{\boldsymbol{\upsilon}}_{1,\mathrm{IFM}}\right),F_2\left(y_{i2}|\boldsymbol{x}_{i2},\widehat{\boldsymbol{\upsilon}}_{2,\mathrm{IFM}}\right)|\theta\big).\qquad(2.7)$$

[1] This is the same form as in the case of continuous margins.

Note that each maximization task (step) has a small number of parameters, which reduces the computational difficulty. However, the IFM method provides a biased estimate for the parameter θ in the presence of censored observations for both margins (as will be seen in Section 2.1.2.2). Since we are interested in the bivariate Clayton copula-based SUR Tobit model where both marginal distributions are censored/semi-continuous, we are dealing with the case where there is not a one-to-one relationship between the marginal distributions and the copula, i.e. there is more than one copula to join the marginal distributions. This constitutes a violation of the Sklar's theorem (Sklar, 1959). When it occurs, researchers often face problems in the copula model fitting and validation.

In order to facilitate the implementation of copula models with semi-continuous margins, the semi-continuous marginal distributions could be augmented to achieve continuity. More specifically, we can use a (frequentist) data augmentation technique to simulate the latent (i.e. unobserved) dependent variables in the censored margins, that is, we generate the unobserved data with all properties, e.g., mean, variance and dependence structure that match with the observed ones, and obtain the continuous marginal distributions (Wichitaksorn *et al.*, 2012). Thus, in order to obtain an unbiased estimate for the association parameter θ, we replace y_{ij} by the augmented data y^{a}_{ij}, or equivalently and more simply (thus, preferred by us), we can replace u_{ij} by the augmented uniform data u^{a}_{ij} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the continuous margin cases. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence occurs (MIFM method). The (frequentist) data augmentation technique we employ here is partially based on Algorithm A2 presented in Wichitaksorn *et al.* (2012). For alternative ways of implementing copula models with censored observations in the margins, but in a survival analysis framework, see, e.g., the classical work of Shih & Louis (1995), as well as its Bayesian counterpart developed by Romeo, Tanaka & Pedroso-de Lima (2006).

In the remaining part of this subsubsection, we discuss the MIFM method when using the Clayton copula to describe the nonlinear dependence structure of the bivariate SUR Tobit model with arbitrary margins (e.g., normal, power-normal and logistic distribution assumption for the marginal error terms). However, the proposed approach can be extended to other copula functions by applying different sampling algorithms. For the cases where only one of the dependent variables/margins is censored (i.e. when $y_{i1} > 0$ and

$y_{i2} = 0$, or $y_{i1} = 0$ and $y_{i2} > 0$), the uniform data augmentation is performed through the truncated conditional distribution of the Clayton copula. If the inverse conditional distribution of the copula used has a closed-form expression, which is the case of the Clayton copula (see, e.g., Armstrong, 2003), we can generate random numbers from its truncated version by applying the method by Devroye (1986, p. 38-39). Otherwise, numerical root-finding procedures are required. By observing the results in Oakes (2005), we see that the Clayton copula has a remarkable invariance property under truncation, such that the conditional distribution of u_{i1} and u_{i2} in a sub-region of a Clayton copula, with one corner at $(0, 0)$, can be written by means of a Clayton copula. That formulation enables a simple simulation scheme (see, e.g., the following online short note: `http://web.cecs.pdx.edu/~cgshirl/Documents/Research/Copula_Methods/Clayton%20Copula.pdf`) in the cases where both dependent variables/margins are censored (i.e. when $y_{i1} = y_{i2} = 0$). For copulas that do not have the truncation-invariance property, an iterative simulation scheme could be used.

The implementation of the bivariate Clayton copula-based SUR Tobit model with arbitrary margins through the proposed MIFM method can be described as follows. In particular, if the marginal error distributions are normal, then set $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)$; if marginal error distributions are power-normal, so $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \left[\Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)\right]^{\alpha_j}$; and if marginal error distributions are logistic, then $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, s_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = G\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right) = \left[1 + \exp\left\{-\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right\}\right]^{-1}$, for $j = 1, 2$ and $z \in \mathbb{R}$.

Stage 1. Estimate the marginal parameters using (2.6). Set $\hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$, for $j = 1, 2$.

Stage 2. Estimate the copula parameter using e.g., (2.7). Set $\hat{\theta}^{(1)}_{\mathrm{MIFM}} = \hat{\theta}_{\mathrm{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = y_{i2} = 0$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}|\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(0, b_{i1}) \times (0, b_{i2})$. This can be performed relatively easily using the following steps.

1. Draw (p,q) from $C\left(p,q|\hat{\theta}^{(\omega)}_{\text{MIFM}}\right) = \left(p^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + q^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} - 1\right)^{-1/\hat{\theta}^{(\omega)}_{\text{MIFM}}}$. See, e.g., Armstrong (2003) for the Clayton copula data generation.

2. Compute $b_{ij} = H_j\left(0|\boldsymbol{x}_{ij}, \hat{\boldsymbol{\upsilon}}_{j,\text{MIFM}}\right)$, for $j = 1, 2$.

3. Set $u^{\text{a}}_{i1} = \left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + b_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} - 1\right) p^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + 1 - b_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\text{MIFM}}}$.

4. Set $u^{\text{a}}_{i2} = \left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + b_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} - 1\right) q^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + 1 - b_{i1}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\text{MIFM}}}$.

If $y_{i1} = 0$ and $y_{i2} > 0$, then draw u^{a}_{i1} from $C\left(u^{\text{a}}_{i1}|u_{i2}, \hat{\theta}^{(\omega)}_{\text{MIFM}}\right)$ truncated to the interval $(0, b_{i1})$. This can be done according to the following steps.

1. Compute $u_{i2} = H_2\left(y_{i2}|\boldsymbol{x}_{i2}, \hat{\boldsymbol{\upsilon}}_{2,\text{MIFM}}\right)$.

2. Compute $b_{i1} = H_1\left(0|\boldsymbol{x}_{i1}, \hat{\boldsymbol{\upsilon}}_{1,\text{MIFM}}\right)$.

3. Draw t from $Uniform\,(0,1)$.

4. Compute $v_{i1} = t\left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + u_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} - 1\right)^{-1/\hat{\theta}^{(\omega)}_{\text{MIFM}}-1}\right] u_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}-1}$.

5. Set $u^{\text{a}}_{i1} = \left[\left(v_{i1}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}/\left(\hat{\theta}^{(\omega)}_{\text{MIFM}}+1\right)} - 1\right) u_{i2}^{-\hat{\theta}^{(\omega)}_{\text{MIFM}}} + 1\right]^{-1/\hat{\theta}^{(\omega)}_{\text{MIFM}}}$.

If $y_{i1} > 0$ and $y_{i2} = 0$, then draw u^{a}_{i2} from $C\left(u^{\text{a}}_{i2}|u_{i1}, \hat{\theta}^{(\omega)}_{\text{MIFM}}\right)$ truncated to the interval $(0, b_{i2})$. This can be done by following the five steps of the previous case (i.e. $y_{i1} = 0$ and $y_{i2} > 0$) by switching subscripts 1 and 2.

If $y_{i1} > 0$ and $y_{i2} > 0$, then set $u^{\text{a}}_{i1} = u_{i1} = H_1\left(y_{i1}|\boldsymbol{x}_{i1}, \hat{\boldsymbol{\upsilon}}_{1,\text{MIFM}}\right)$ and $u^{\text{a}}_{i2} = u_{i2} = H_2\left(y_{i2}|\boldsymbol{x}_{i2}, \hat{\boldsymbol{\upsilon}}_{2,\text{MIFM}}\right)$.

Given the generated/augmented marginal uniform data u^{a}_{ij}, we estimate the association parameter θ by [2]

$$\hat{\theta}^{(\omega+1)}_{\text{MIFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u^{\text{a}}_{i1}, u^{\text{a}}_{i2}|\theta\right).$$

The algorithm stops if a termination criterion is fulfilled, e.g. if $|\hat{\theta}^{(\omega+1)}_{\text{MIFM}} - \hat{\theta}^{(\omega)}_{\text{MIFM}}| < \xi$, where ξ is the tolerance parameter (e.g., $\xi = 10^{-3}$).

[2] The generated/augmented marginal uniform data u^{a}_{ij} should carry (ω) as a superscript, i.e. $u^{\text{a}(\omega)}_{ij}$, but we omit it so as not to clutter the notation.

2.1.1.2 Interval estimation

Joe & Xu (1996) suggest the use of the jackknife method for the estimation of the standard errors of the multivariate model parameter estimates when using the IFM approach. It makes the analytic derivatives no longer required to compute the inverse Godambe information matrix, which is the asymptotic covariance matrix associated with the vector of parameter estimates under some regularity conditions. See Joe (1997, p. 301-302) for the form of this matrix. However, we carried out a pilot simulation study whose results revealed that the jackknife is not valid to obtain standard errors of parameter estimates when using the MIFM approach, i.e. in the context of copula-based models with censored/semi-continuous margins (the jackknife method produces an overestimate of the standard error of the association parameter estimate). This implies that confidence intervals for the parameters of the bivariate Clayton copula-based SUR Tobit model cannot be constructed using this resampling technique. To overcome this problem, we propose the use of bootstrap methods to build confidence intervals.

Our bootstrap approach can be described as follows. Let η_h, $h = 1, ..., k$, be any component of the parameter vector $\boldsymbol{\eta}$ of the bivariate Clayton copula-based SUR Tobit model (see Section 2.1.1.1). By using a parametric resampling plan, we obtain the bootstrap estimates $\hat{\eta}^*_{h1}, \hat{\eta}^*_{h2}, ..., \hat{\eta}^*_{hB}$ of η_h through the MIFM method, where B is the number of bootstrap samples. Hinkley (1988) suggests that the minimum value of B depends on the parameter being estimated, but that it is often 100 or more. Then, we can derive confidence intervals from the bootstrap distribution through the following two methods, for instance.

- **Percentile bootstrap** (Efron & Tibshirani, 1993, p. 171). The $100\,(1-2\alpha)\,\%$ percentile confidence interval is defined by the $100\,(\alpha)$th and $100\,(1-\alpha)$th percentiles of the bootstrap distribution of $\hat{\eta}^*_h$:

$$\left[\hat{\eta}_h^{*(\alpha)}, \hat{\eta}_h^{*(1-\alpha)}\right].$$

  For Carpenter & Bithell (2000), simplicity is the attractive feature of this method. Moreover, no invalid parameter values can be included in the interval.

- **Standard normal interval** (Efron & Tibshirani, 1993, p. 154). Since most statistics are asymptotically normally distributed, in large samples we can use the standard error estimate, $\widehat{se}_h$, as well as the normal distribution, to yield a $100\,(1-2\alpha)\,\%$

confidence interval for η_h based on the original estimate (i.e. from the original data/sample) $\hat{\eta}_h$:

$$\left[\hat{\eta}_h - z^{(1-\alpha)}\widehat{se}_h, \hat{\eta}_h - z^{(\alpha)}\widehat{se}_h\right],$$

where $z^{(\alpha)}$ represents the $100\,(\alpha)$th percentile point of a standard normal distribution, and $\widehat{se}_h$ is the hth entry on the diagonal of the bootstrap-based covariance matrix estimate of the parameter vector estimate $\hat{\boldsymbol{\eta}}$, which is given by

$$\widehat{\mathbf{\Sigma}}_{\text{boot}} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)', \tag{2.8}$$

where $\hat{\boldsymbol{\eta}}_b^*$, $b = 1, ..., B$, is the bootstrap estimate of $\boldsymbol{\eta}$ and

$$\overline{\hat{\boldsymbol{\eta}}}^* = \left(\frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{1b}^*, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{2b}^*, \ldots, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{kb}^*\right).$$

2.1.2 Simulation study

A simulation study was performed to investigate the behavior of the MIFM estimates (focusing on the copula association parameter estimate) and check the coverage probabilities of bootstrap confidence intervals (constructed using the two methods described in Section 2.1.1.2) for the bivariate Clayton copula-based SUR Tobit model parameters. Here, we considered some circumstances that might arise in the development of bivariate copula-based SUR Tobit models, involving the sample size, the censoring percentage (i.e. the percentage of zero observations) in the dependent variables/margins and their interdependence degree. We also considered/assumed different distributions for the marginal error terms.

2.1.2.1 General specifications

In the simulation study, we applied the Clayton copula to model the nonlinear dependence structure of the bivariate SUR Tobit model. We set the true value for the association parameter θ at 0.67, 2 and 6, corresponding to a Kendall's tau association measure [3] of 0.25, 0.50 and 0.75, respectively. For the Clayton copula data generation, see, e.g., Armstrong (2003).

For $i = 1, ..., n$, the covariates for margin 1, $\boldsymbol{x}_{i1} = (x_{i1,0}, x_{i1,1})'$, were $x_{i1,0} = 1$ and $x_{i1,1}$ was randomly simulated from a standard normal distribution. While the covariates for

[3] The Kendall's tau for Clayton copula is given by $\tau_2 = \theta\,/\,(\theta + 2)$; see, e.g., Joe (1997, p. 78) and McNeil *et al.* (2005, p. 222).

margin 2, $\boldsymbol{x}_{i2} = \left(x_{i2,0}, x_{i2,1}\right)^{'}$, were generated as $x_{i2,0} = 1$ and $x_{i2,1}$ was randomly simulated from $N\left(1, 2^2\right)$. The model errors ϵ_{i1} and ϵ_{i2} were assumed to follow the distributions shown below:

- **Normal**: i.e. $\epsilon_{i1} \sim N\left(0, \sigma_1^2\right)$ and $\epsilon_{i2} \sim N\left(0, \sigma_2^2\right)$, where $\sigma_1 = 1$ and $\sigma_2 = 2$ are the standard deviations (scale parameters) for margins 1 and 2, respectively. To ensure a percentage of censoring (i.e. of zero observations) for both margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$ and $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$:

  - $\boldsymbol{\beta}_1 = (2.3, 1)$ and $\boldsymbol{\beta}_2 = (4, -0.5)$;
  - $\boldsymbol{\beta}_1 = (1.5, 1)$ and $\boldsymbol{\beta}_2 = (2.75, -0.5)$;
  - $\boldsymbol{\beta}_1 = (1, 1)$ and $\boldsymbol{\beta}_2 = (2, -0.5)$;
  - $\boldsymbol{\beta}_1 = (0.5, 1)$ and $\boldsymbol{\beta}_2 = (1.3, -0.5)$;
  - $\boldsymbol{\beta}_1 = (-0.02, 1)$ and $\boldsymbol{\beta}_2 = (0.5, -0.5)$;

  respectively. For $j = 1, 2$, the latent dependent variable of margin j, y_{ij}^*, was randomly simulated from $N\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, \sigma_j^2\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\max\left\{0, y_{ij}^*\right\}$.

- **Power-normal**: i.e. $\epsilon_{i1} \sim PN\left(0, \sigma_1, \alpha_1\right)$ and $\epsilon_{i2} \sim PN\left(0, \sigma_2, \alpha_2\right)$, where $\sigma_1 = 1$ and $\sigma_2 = 2$ are the scale parameters for margins 1 and 2, respectively; and $\alpha_1 = \alpha_2 = 1.75$ are the shape parameters for margins 1 and 2. To ensure a percentage of censoring for both margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$ and $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$:

  - $\boldsymbol{\beta}_1 = (1.7, 1)$ and $\boldsymbol{\beta}_2 = (2.8, -0.5)$;
  - $\boldsymbol{\beta}_1 = (0.9, 1)$ and $\boldsymbol{\beta}_2 = (1.6, -0.5)$;
  - $\boldsymbol{\beta}_1 = (0.4, 1)$ and $\boldsymbol{\beta}_2 = (0.9, -0.5)$;
  - $\boldsymbol{\beta}_1 = (0.05, 1)$ and $\boldsymbol{\beta}_2 = (0.4, -0.5)$;
  - $\boldsymbol{\beta}_1 = (-0.5, 1)$ and $\boldsymbol{\beta}_2 = (-0.4, -0.5)$;

  respectively. For $j = 1, 2$, the latent dependent variable of margin j, y_{ij}^*, was randomly simulated from $PN\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$; therefore, the observed dependent variable of margin j, y_{ij}, was obtained from $\max\left\{0, y_{ij}^*\right\}$.

- **Logistic**: i.e. $\epsilon_{i1} \sim L(0, s_1)$ and $\epsilon_{i2} \sim L(0, s_2)$, where $s_1 = 1$ and $s_2 = 2$ are the scale parameters for margins 1 and 2, respectively. To ensure a percentage of censoring for both margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$ and $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$:

    - $\boldsymbol{\beta}_1 = (3.3, 1)$ and $\boldsymbol{\beta}_2 = (5.8, 1)$;
    - $\boldsymbol{\beta}_1 = (2.1, 1)$ and $\boldsymbol{\beta}_2 = (3.1, 1)$;
    - $\boldsymbol{\beta}_1 = (1.3, 1)$ and $\boldsymbol{\beta}_2 = (1.7, 1)$;
    - $\boldsymbol{\beta}_1 = (0.8, 1)$ and $\boldsymbol{\beta}_2 = (0.5, 1)$;
    - $\boldsymbol{\beta}_1 = (-0.05, 1)$ and $\boldsymbol{\beta}_2 = (-0.9, 1)$;

  respectively. For $j = 1, 2$, the latent dependent variable of margin j, y^*_{ij}, was randomly simulated from $L\left(\boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j, s_j\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from max $\left\{0, y^*_{ij}\right\}$.

For each error distribution assumption (normal, power-normal and logistic), censoring percentage in the margins (5%, 15%, 25%, 35% and 50% of zero observations) and degree of dependence between them (low: $\theta = 0.67$, moderate: $\theta = 2$ and high: $\theta = 6$), we generated 100 datasets of sizes $n = 200$, 800 and 2000. These choices of sample sizes were based on some authors' indication (e.g., Joe, 2014) that large sample sizes are commonly required when working with copulas. Then, for each dataset (original sample), we obtained 500 bootstrap samples through a parametric resampling plan (parametric bootstrap approach), i.e. we fitted a bivariate Clayton copula-based SUR Tobit model with the corresponding error distributions to each dataset using the MIFM approach, and then generated a set of 500 new datasets (the same size as the original dataset/sample) from the estimated parametric model. The computing language was written in R statistical programming environment (R Core Team, 2014) and ran on a virtual machine of the Cloud-USP at ICMC, with Intel Xeon processor E5500 series, 8 core (virtual CPUs), 32 GB RAM.

We assessed the performance of the proposed models and methods through the coverage probabilities of the nominally 90% standard normal and percentile bootstrap confidence intervals, the Bias and the Mean Squared Error (MSE), in which the Bias and the MSE of each parameter η_h, $h = 1, ..., k$, are given by Bias $= M^{-1}\sum_{r=1}^{M}(\hat{\eta}^r_h - \eta_h)$ and

$\text{MSE} = M^{-1} \sum_{r=1}^{M} (\hat{\eta}_h^r - \eta_h)^2$, respectively, where $M = 100$ is the number of replications (original datasets/samples) and $\hat{\eta}_h^r$ is the estimated value of η_h at the rth replication.

2.1.2.2 Simulation results

In this subsubsection, we present the main results obtained from the simulation study performed with samples (datasets) of different sizes, percentages of censoring in the margins and degrees of dependence between them, regarding the bivariate Clayton copula-based SUR Tobit model parameters estimated using the MIFM approach. Since both the MIFM and IFM methods provide the same marginal parameter estimates (the first stage of the proposed method is similar to the first stage of the usual one, as seen in Section 2.1.1.1), we focus here on the Clayton copula parameter estimate. Some asymptotic results (such as asymptotic normality) associated with the IFM method appear in Joe & Xu (1996). We also show the results related to the estimated coverage probabilities of the 90% confidence intervals for θ, obtained by bootstrap methods (standard normal and percentile intervals).

Figures 2.1, 2.2 and 2.3 show the Bias and MSE of the observed MIFM estimates of θ for normal, power-normal and logistic marginal errors, respectively. From these figures, we observe that, regardless of the error distribution assumption, the percentage of censoring in the margins and their interdependence degree, the Bias and MSE of the MIFM estimator of θ are relatively low and tend to zero for large n, i.e. the MIFM estimator is asymptotically unbiased and consistent for the Clayton copula parameter.

Figures 2.4, 2.5 and 2.6 show the estimated coverage probabilities of the bootstrap confidence intervals for θ for normal, power-normal and logistic marginal errors, respectively. Observe that the estimated coverage probabilities are sufficiently high and close to the nominal value of 0.90, except for a few cases in which n is small to moderate ($n = 200$ and 800), the degree of dependence between the margins is high ($\theta = 6$) and the marginal errors follow non-normal (i.e. power-normal and logistic) distributions (see Figures 2.5(c) and 2.6(c)).

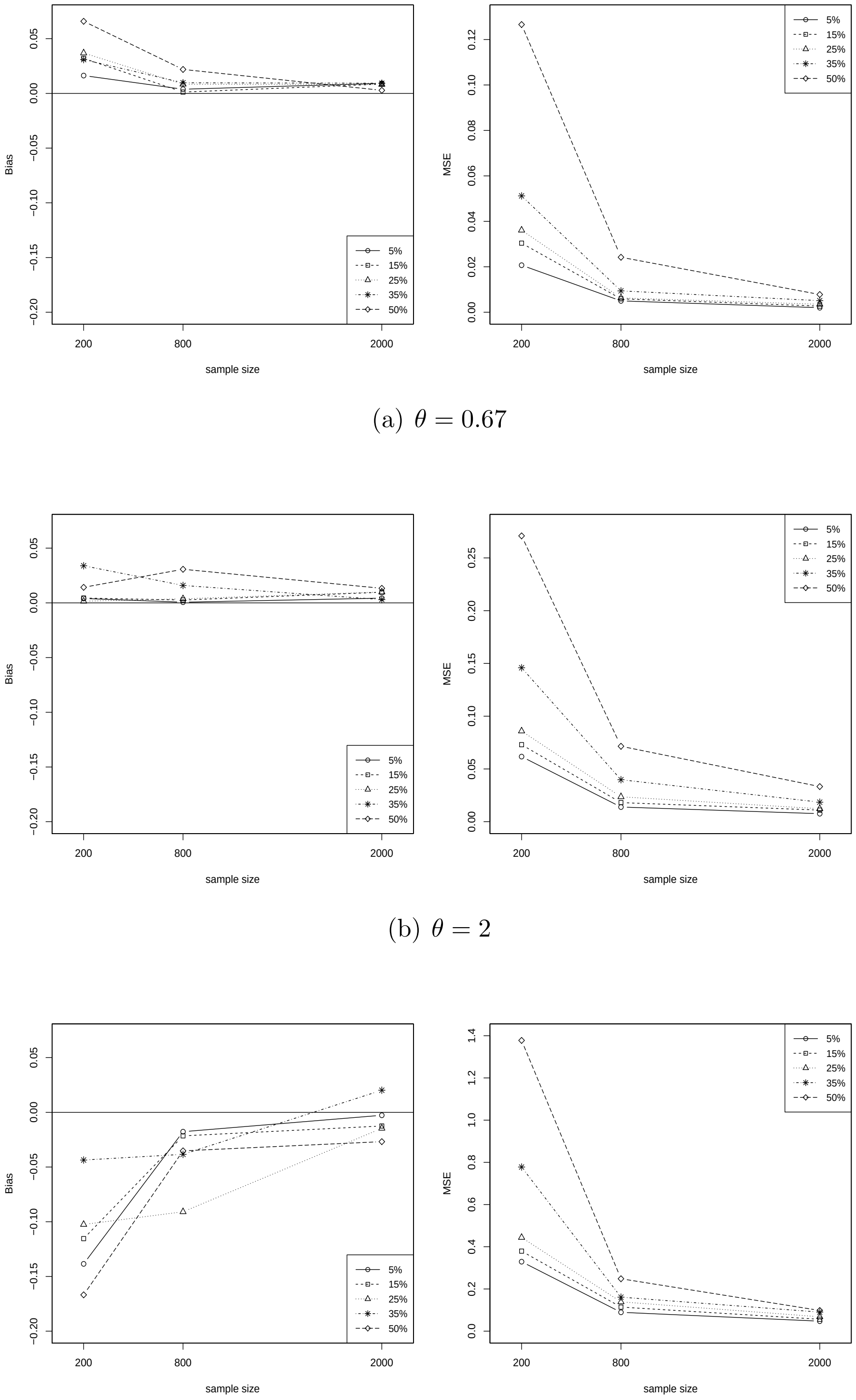
Finally, Figures 2.7, 2.8 and 2.9 compare, via boxplots, the observed MIFM estimates of θ with its estimates obtained through the IFM method for normal, power-normal and logistic marginal errors, respectively, and for $n = 2000$. It can be seen from Figure 2.7 that there is a certain equivalence between the two estimation methods (with a slight advantage for the MIFM method over the IFM method, in terms of bias) when the degree of dependence between the margins is relatively low, that is $\theta = 0.67$ (Figure 2.7(a)). However, the IFM method underestimates θ for dependence at a higher level, that is $\theta = 2$ and $\theta = 6$ (Figures 2.7(b) and 2.7(c), respectively). From Figure 2.8, we observe that the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 2.8(a)), and underestimates θ for dependence at a higher level, that is $\theta = 2$ and $\theta = 6$ (Figures 2.8(b) and 2.8(c), respectively). In Figure 2.9, we see that there is a certain equivalence between the two estimation methods (with a slight advantage for the MIFM method over the IFM method, in terms of bias) when the degree of dependence between the margins is moderate, that is $\theta = 2$ (Figure 2.9(b)). Nevertheless, the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 2.9(a)), and underestimates θ for dependence at a higher level, that is $\theta = 6$ (Figure 2.9(c)). Note also from Figures 2.7, 2.8 and 2.9 that the difference (distance) between the distributions of the IFM and MIFM estimates often increases as the percentage of censoring in the margins increases.

2.1.3 Application

Consider the consumption dataset described in Section 1.1.1. For the sake of illustration of our proposed bivariate models and methods, we assume that there are only two dependent variables: salad dressing and lettuce consumption (which show the highest Kendall tau correlation; see Section 1.1.1).

In this application, the relationship between the reported salad dressing (amount consumed in 100 grams) and lettuce (amount consumed in 200 grams) consumption by 400 U.S. adults is modeled by the bivariate SUR Tobit model with normal, power-normal and logistic marginal errors through the Clayton copula (see Section 1.3 for the reasons for this copula model choice). We include age, location (region) and income as the covariates and use them for both margins in all three candidate models.

Tables 2.1, 2.2 and 2.3 show the MIFM estimates for the parameters of the bivariate Clayton copula-based SUR Tobit model with normal, power-normal and logistic marginal errors, respectively, as well as the 90% confidence intervals obtained through the standard normal and percentile bootstrap methods. These tables also present the log-likelihood values for the three fitted models. We can then compare the bivariate Clayton copula-based SUR Tobit models by using some information criterion, e.g. the Akaike Information Crite-

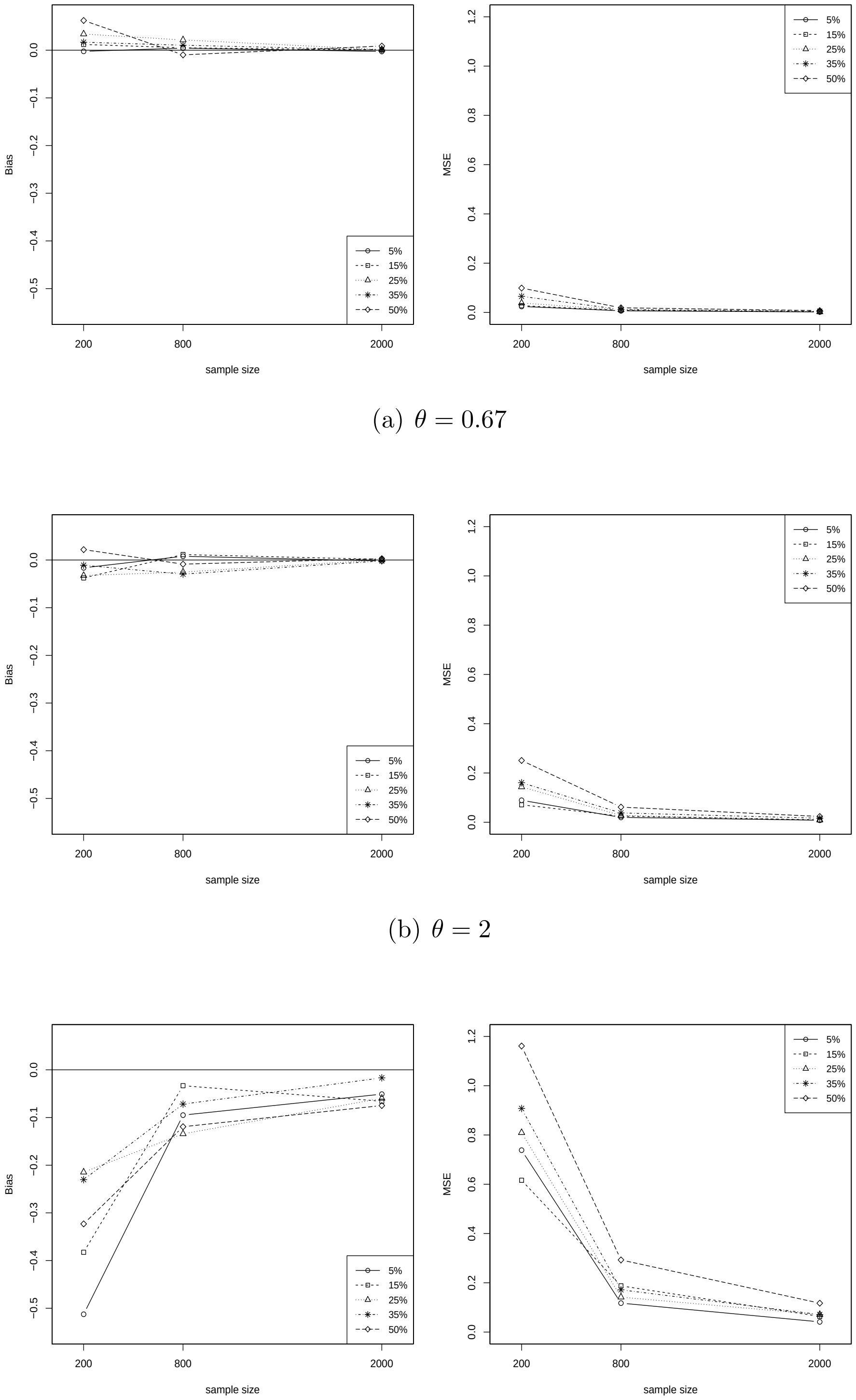


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.1: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (normal marginal errors).

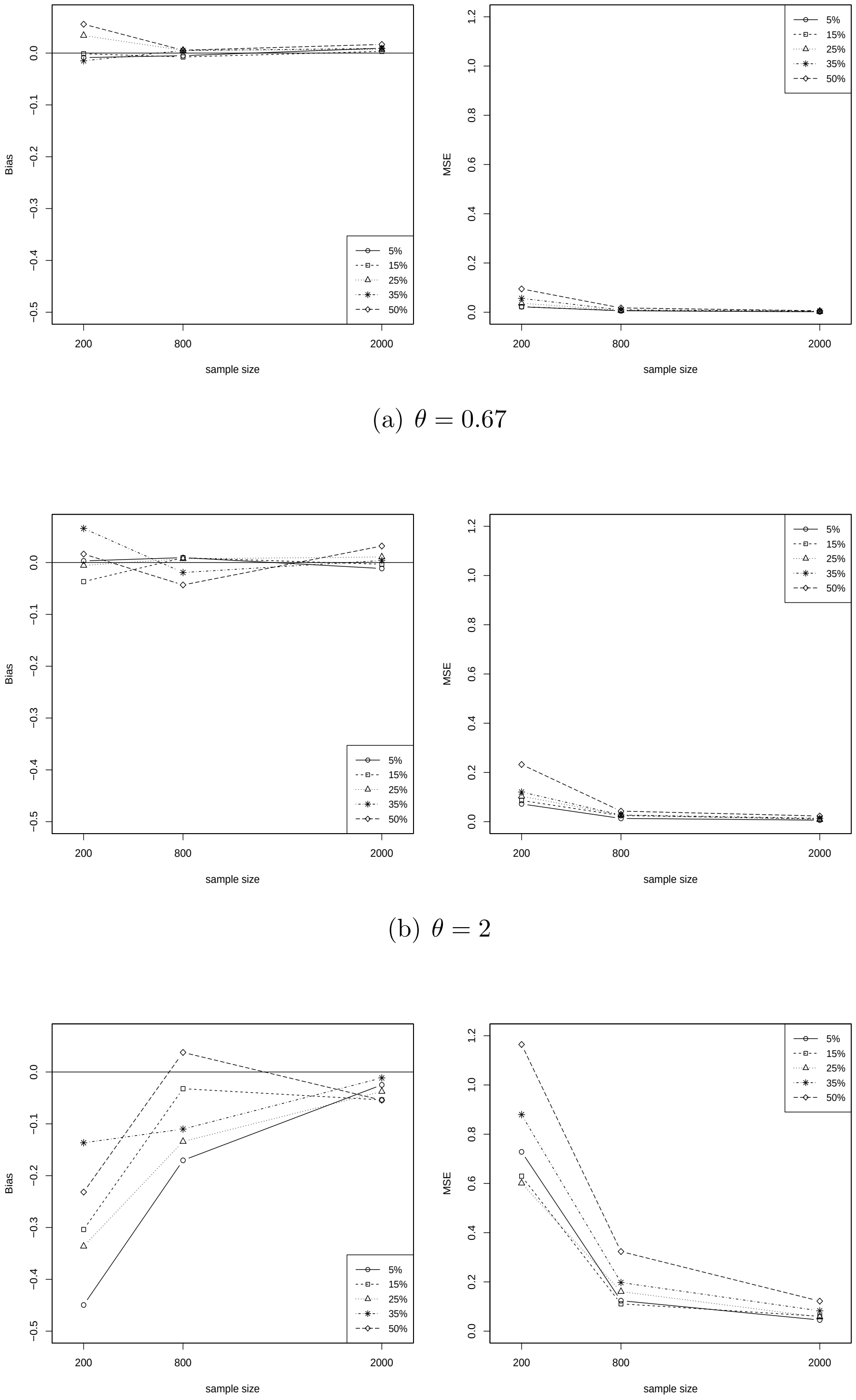


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.2: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (power-normal marginal errors).

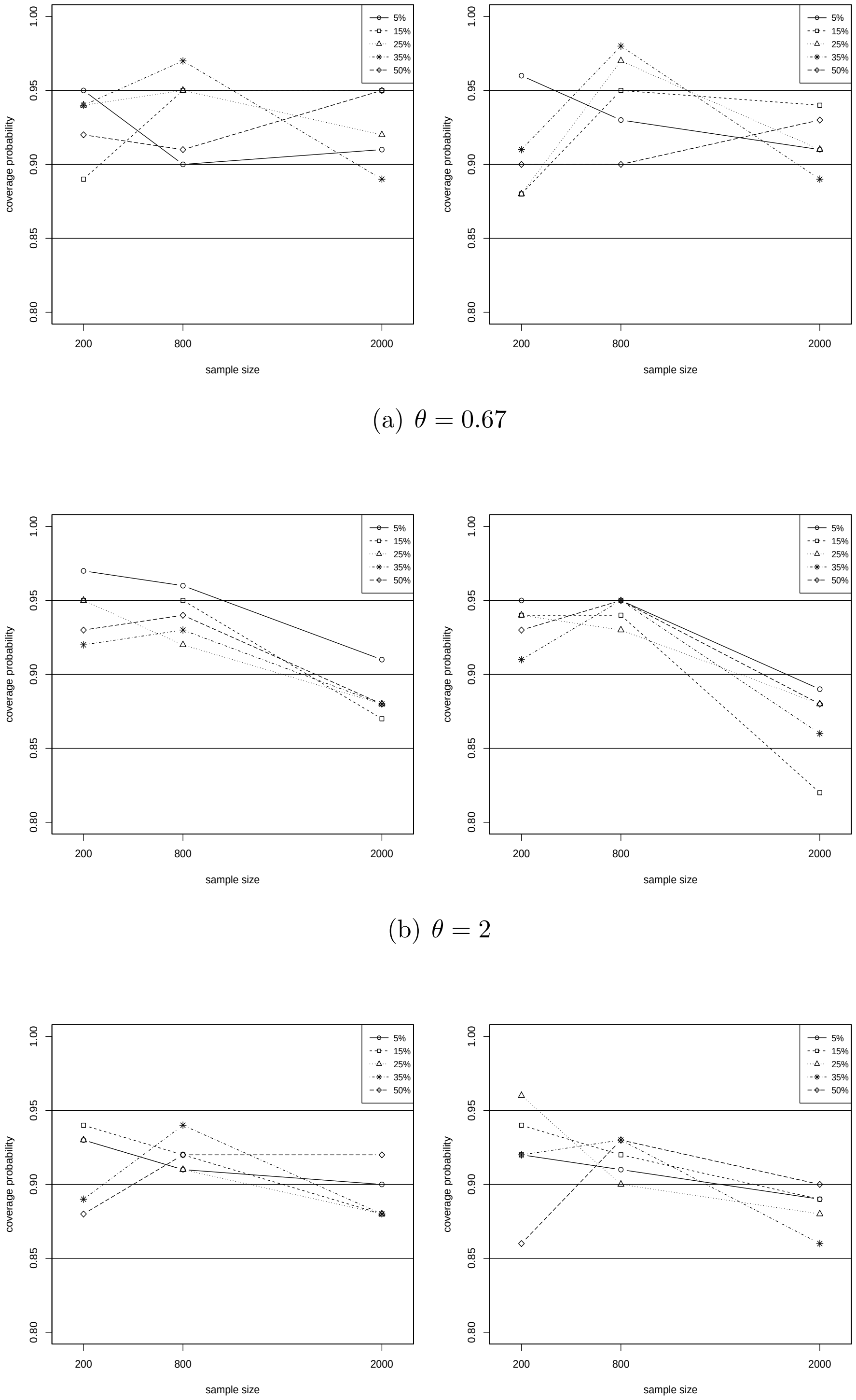


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.3: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (logistic marginal errors).

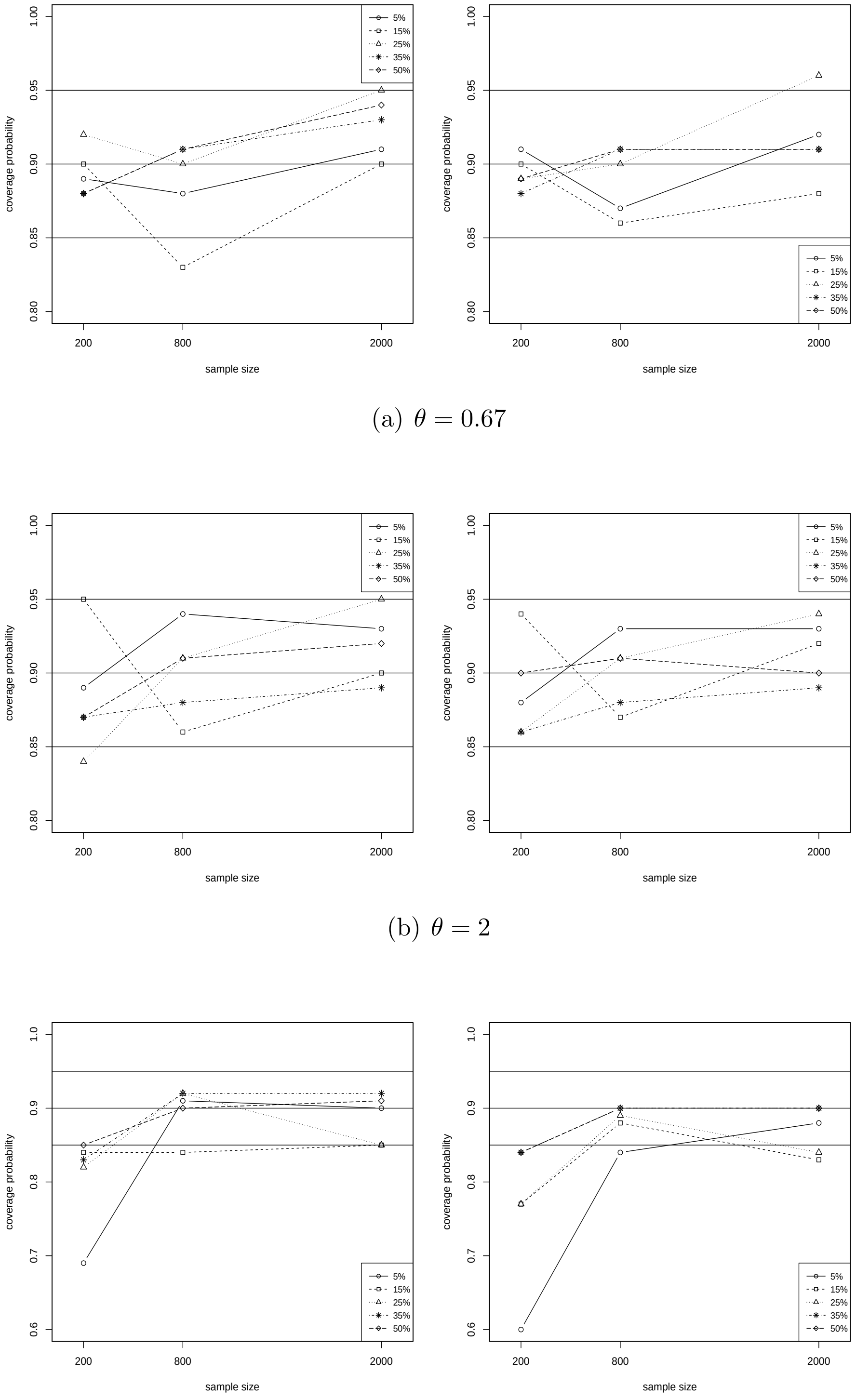


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.4: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

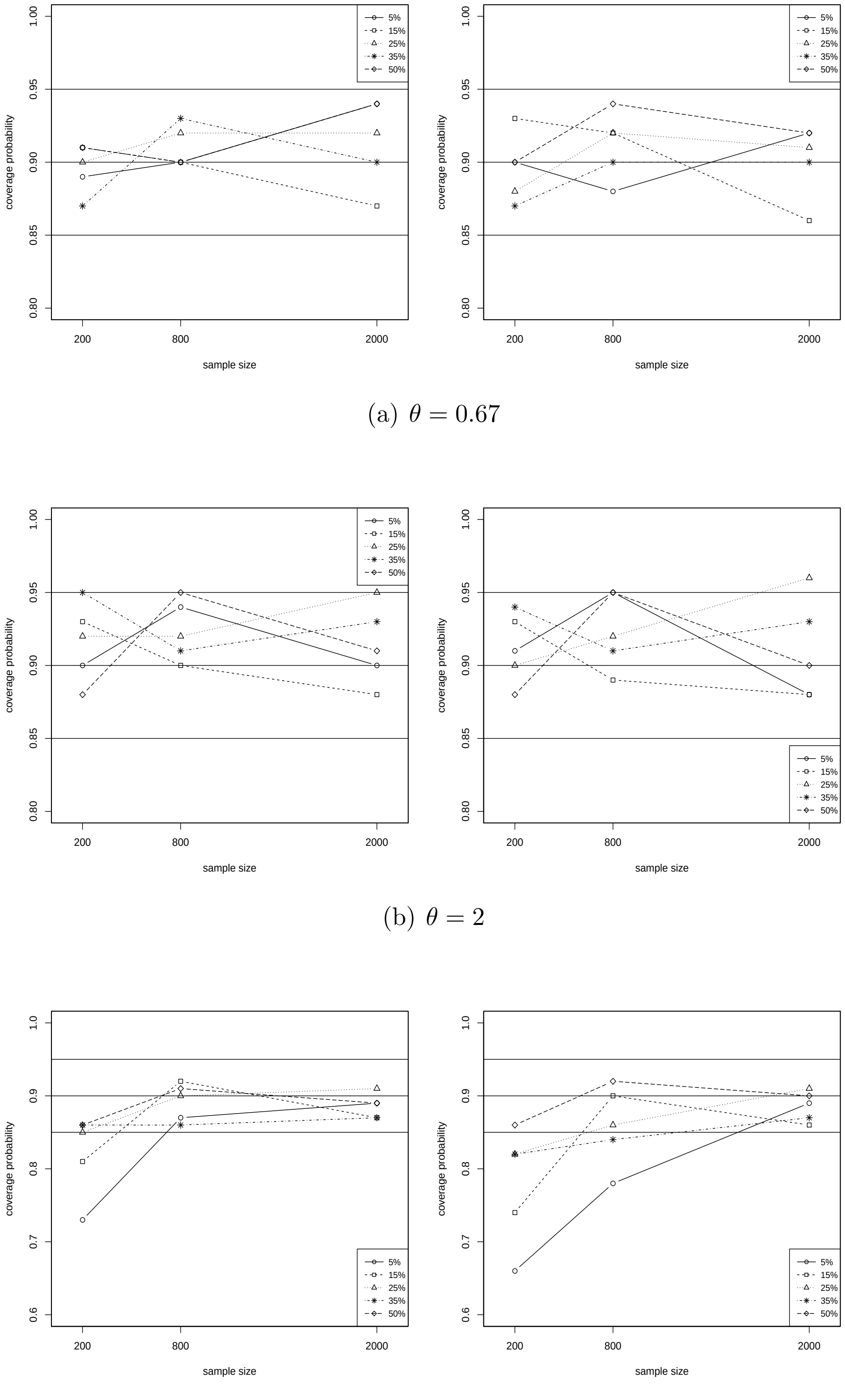


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.5: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (power-normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

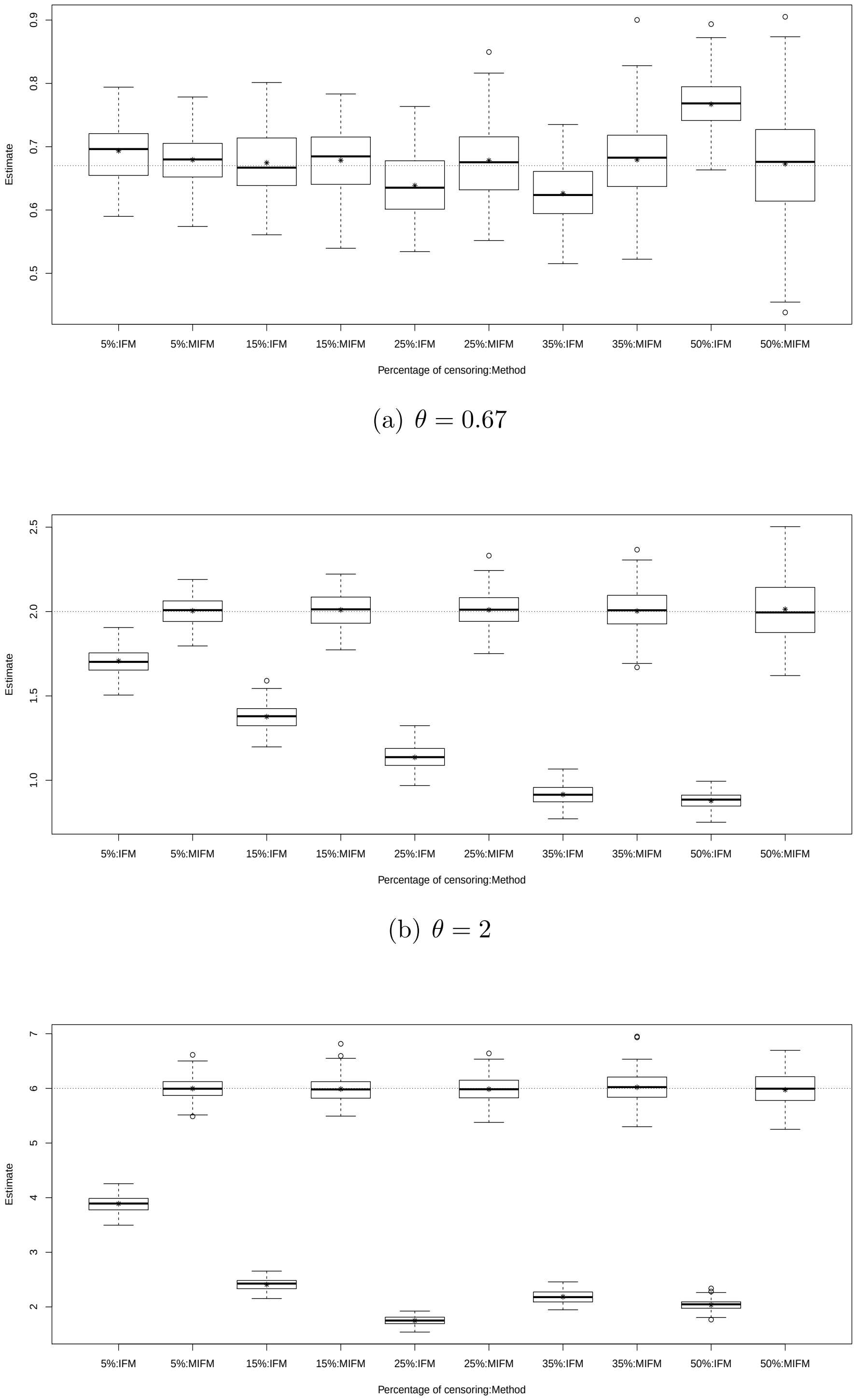


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.6: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (logistic marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

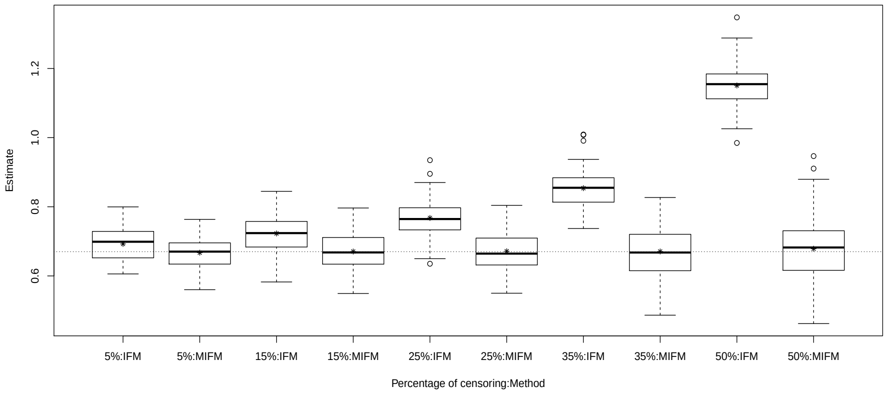


(a) $\theta = 0.67$

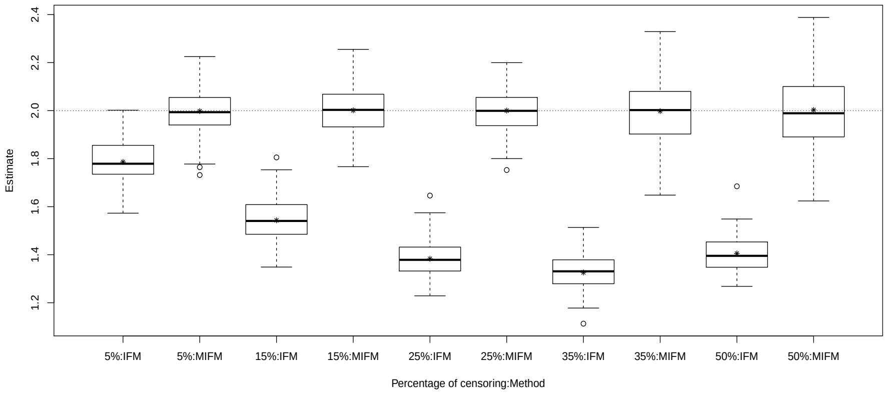
(b) $\theta = 2$

(c) $\theta = 6$

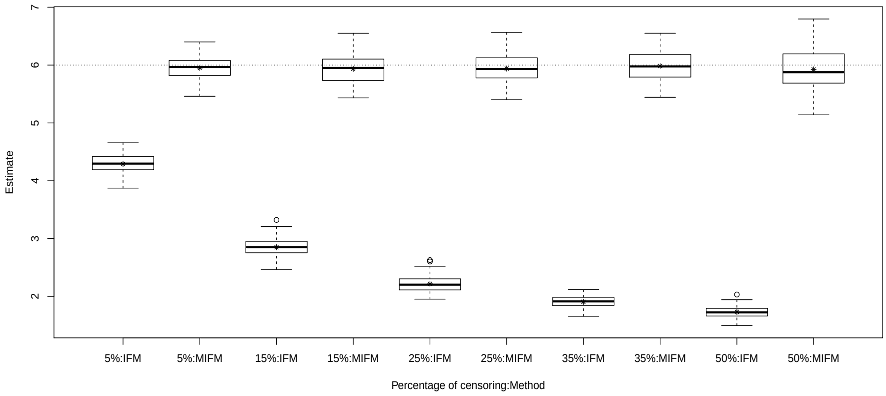
Figure 2.7: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.



(a) $\theta = 0.67$

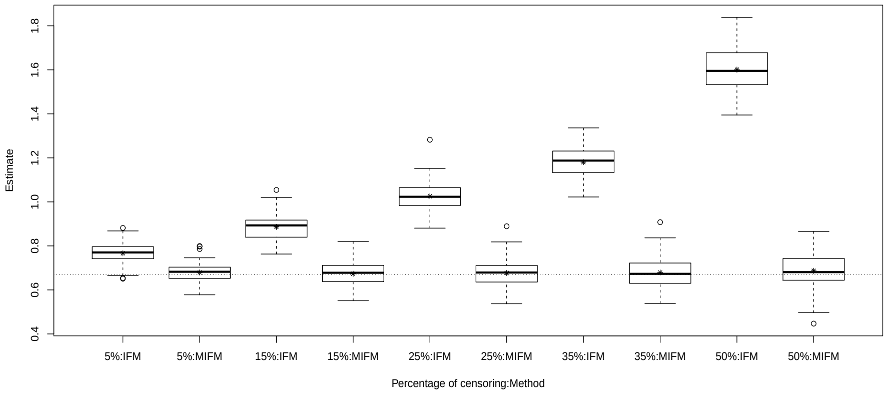


(b) $\theta = 2$

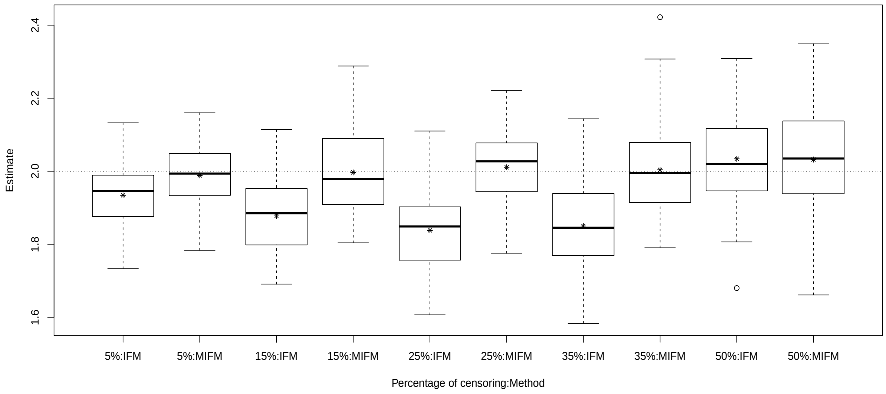


(c) $\theta = 6$

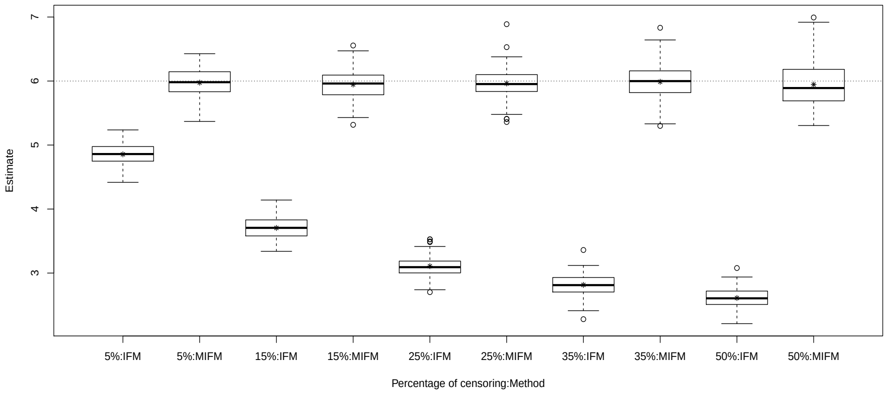
Figure 2.8: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (power-normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.



(a) $\theta = 0.67$



(b) $\theta = 2$



(c) $\theta = 6$

Figure 2.9: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (logistic marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.

rion (AIC) (Akaike, 1973, 1974) and the Bayesian Information Criterion (BIC) (Schwarz, 1978), which are defined by $-2\ell\left(\hat{\boldsymbol{\eta}}\right)+2k$ and $-2\ell\left(\hat{\boldsymbol{\eta}}\right)+k\log\left(n\right)$, respectively. The preferred model is the one with the smaller value on each criterion. The AIC and BIC criterion values for the three fitted models are shown in Tables 2.1, 2.2 and 2.3. Observe that the bivariate Clayton copula-based SUR Tobit model with logistic marginal errors has the smallest AIC and BIC criterion values and therefore provides the best fit to the salad dressing and lettuce consumption data. Appendix B provides the R codes for fitting this favorite model using the MIFM approach, as well as for building standard normal and percentile bootstrap confidence intervals for its parameters. From the Kolmogorov-Smirnov goodness-of-fit tests (see, e.g., Conover, 1971, p. 295-301) of augmented marginal residuals [4], we obtain p-values equal to 0.9020 and 0.9356 for the salad dressings and lettuce models, respectively. Thus, the logistic distribution assumption for the marginal errors is valid. The results reported in Table 2.3 reveal that individuals aged 20-40 years consume more salad dressings than those over 60 years of age. Su & Arab (2006) found a similar effect of age on salad dressing consumption. According to the 90% percentile interval, individuals aged 41-50 years consume more lettuce than those over 60 years of age. Regional effects are also notable, as individuals from the Northeast and Midwest consume more salad dressings, and individuals from the Midwest and West consume more lettuce than those residing in the South. The household income has a positive effect on the consumption of both salad dressings and lettuce. The MIFM estimate of the Clayton copula parameter $\left(\hat{\theta}_{\text{MIFM}}=2.3853, \text{obtained after 7 iterations}\right)$ and its 90% bootstrap-based confidence intervals show us that the relationship between salad dressing and lettuce consumption is positive (the estimated Kendall's tau is $\hat{\tau}_2=\hat{\theta}_{\text{MIFM}}/\left(\hat{\theta}_{\text{MIFM}}+2\right)=0.5439$, which is close to the value of the nonparametric association measure presented in Section 1.1.1) and significant at the 10% level (the lower limits of the 90% bootstrap-based confidence intervals for θ are greater than and far above zero), justifying joint estimation of the censored equations through the Clayton copula to improve statistical efficiency. Furthermore, the estimated coefficient of tail dependence for Clayton copula, $\hat{\lambda}_{\text{L}}=0.7478$, obtained from $2^{-1/\hat{\theta}_{\text{MIFM}}}$ (McNeil *et al.*, 2005, p. 209), shows the positive dependence at the lower tail, i.e. for low or no consumption of salad dressings and lettuce.

[4] The augmented residuals are the differences between the augmented observed and predicted responses, i.e. $e_{ij}^{\text{a}}=y_{ij}^{\text{a}}-\boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}$, for $i=1,...,n$ and $j=1,2$, where $y_{ij}^{\text{a}}=\boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}+\hat{s}_{j,\text{MIFM}}G^{-1}\left(u_{ij}^{\text{a}}\right)$, with $G^{-1}\left(.\right)$ being the inverse function of the $L\left(0,1\right)$ c.d.f.; or simply, $e_{ij}^{\text{a}}=\hat{s}_{j,\text{MIFM}}G^{-1}\left(u_{ij}^{\text{a}}\right)$.

Table 2.1: Estimation results of bivariate Clayton copula-based SUR Tobit model with normal marginal errors for salad dressing and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals	
Salad dressing	Estimate	Standard Normal	Percentile
Intercept	0.1130	[0.0391; 0.1868]	[0.0407; 0.1854]
Age 20-30	0.1106	[0.0396; 0.1817]	[0.0412; 0.1830]
Age 31-40	0.1011	[0.0347; 0.1675]	[0.0369; 0.1668]
Age 41-50	0.0633	[0.0029; 0.1238]	[0.0029; 0.1210]
Age 51-60	-0.0030	[-0.0682; 0.0622]	[-0.0646; 0.0678]
Northeast	0.0784	[0.0140; 0.1429]	[0.0107; 0.1393]
Midwest	0.0521	[-0.0089; 0.1130]	[-0.0078; 0.1113]
West	0.0544	[-0.0080; 0.1167]	[-0.0071; 0.1158]
Income	0.0277	[0.0041; 0.0512]	[0.0045; 0.0514]
σ_1	0.2636	[0.2464; 0.2809]	[0.2448; 0.2773]
		90% Confidence Intervals	
Lettuce	Estimate	Standard Normal	Percentile
Intercept	-0.1084	[-0.2048; -0.0121]	[-0.1989; -0.0102]
Age 20-30	0.1051	[0.0179; 0.1923]	[0.0122; 0.1903]
Age 31-40	0.0786	[-0.0046; 0.1617]	[-0.0082; 0.1646]
Age 41-50	0.0908	[0.0188; 0.1628]	[0.0134; 0.1590]
Age 51-60	0.0232	[-0.0575; 0.1038]	[-0.0605; 0.1040]
Northeast	0.0588	[-0.0160; 0.1336]	[-0.0236; 0.1273]
Midwest	0.1065	[0.0391; 0.1739]	[0.0324; 0.1690]
West	0.0946	[0.0204; 0.1688]	[0.0205; 0.1640]
Income	0.0604	[0.0300; 0.0908]	[0.0290; 0.0898]
σ_2	0.3101	[0.2853; 0.3349]	[0.2824; 0.3315]
θ	2.7704	[2.2748; 3.2660]	[2.2736; 3.2302]
Log-likelihood	-140.3680		
AIC	322.7360		
BIC	406.5567		

Table 2.2: Estimation results of bivariate Clayton copula-based SUR Tobit model with power-normal marginal errors for salad dressing and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals	
Salad dressing	Estimate	Standard Normal	Percentile
Intercept	0.1291	[0.0044; 0.2537]	[0.0521; 0.2781]
Age 20-30	0.1080	[0.0396; 0.1764]	[0.0371; 0.1755]
Age 31-40	0.0908	[0.0264; 0.1552]	[0.0293; 0.1528]
Age 41-50	0.0457	[-0.0186; 0.1099]	[-0.0217; 0.1082]
Age 51-60	-0.0199	[-0.0840; 0.0442]	[-0.0829; 0.0453]
Northeast	0.0675	[0.0036; 0.1315]	[-0.0002; 0.1304]
Midwest	0.0287	[-0.0305; 0.0878]	[-0.0341; 0.0855]
West	0.0424	[-0.0129; 0.0977]	[-0.0148; 0.0957]
Income	0.0249	[0.0006; 0.0492]	[0.0013; 0.0502]
σ_1	0.2686	[0.2290; 0.3082]	[0.2149; 0.2911]
α_1	1.0475	[0.6013; 1.4937]	[0.5524; 1.3666]
		90% Confidence Intervals	
Lettuce	Estimate	Standard Normal	Percentile
Intercept	-0.0674	[-0.2308; 0.0961]	[-0.2057; 0.1082]
Age 20-30	0.1006	[0.0181; 0.1831]	[0.0123; 0.1751]
Age 31-40	0.0769	[-0.0055; 0.1593]	[-0.0042; 0.1578]
Age 41-50	0.0833	[0.0030; 0.1636]	[-0.0009; 0.1641]
Age 51-60	0.0199	[-0.0555; 0.0952]	[-0.0557; 0.0963]
Northeast	0.0462	[-0.0341; 0.1266]	[-0.0387; 0.1198]
Midwest	0.0961	[0.0238; 0.1684]	[0.0253; 0.1605]
West	0.0797	[0.0128; 0.1466]	[0.0110; 0.1464]
Income	0.0609	[0.0309; 0.0909]	[0.0316; 0.0899]
σ_2	0.3021	[0.2482; 0.3559]	[0.2365; 0.3396]
α_2	0.8939	[0.4311; 1.3566]	[0.4460; 1.3239]
θ	2.7864	[2.2828; 3.2899]	[2.2925; 3.2591]
Log-likelihood	-144.0309		
AIC	334.0618		
BIC	425.8655		

Table 2.3: Estimation results of bivariate Clayton copula-based SUR Tobit model with logistic marginal errors for salad dressing and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals	
Salad dressing	Estimate	Standard Normal	Percentile
Intercept	0.1124	[0.0372; 0.1876]	[0.0406; 0.1852]
Age 20-30	0.0968	[0.0304; 0.1633]	[0.0341; 0.1701]
Age 31-40	0.0977	[0.0370; 0.1583]	[0.0379; 0.1599]
Age 41-50	0.0480	[-0.0148; 0.1108]	[-0.0119; 0.1119]
Age 51-60	0.0024	[-0.0600; 0.0647]	[-0.0630; 0.0651]
Northeast	0.0744	[0.0124; 0.1365]	[0.0133; 0.1386]
Midwest	0.0559	[0.0011; 0.1108]	[0.0016; 0.1078]
West	0.0570	[-0.0017; 0.1157]	[-0.0015; 0.1152]
Income	0.0275	[0.0046; 0.0503]	[0.0051; 0.0496]
s_1	0.1459	[0.1351; 0.1567]	[0.1347; 0.1550]
		90% Confidence Intervals	
Lettuce	Estimate	Standard Normal	Percentile
Intercept	-0.0837	[-0.1745; 0.0072]	[-0.1729; 0.0031]
Age 20-30	0.0804	[-0.0021; 0.1629]	[-0.0030; 0.1658]
Age 31-40	0.0718	[-0.0033; 0.1469]	[-0.0046; 0.1452]
Age 41-50	0.0721	[-0.0064; 0.1506]	[0.0002; 0.1518]
Age 51-60	0.0133	[-0.0634; 0.0901]	[-0.0700; 0.0915]
Northeast	0.0662	[-0.0111; 0.1436]	[-0.0069; 0.1438]
Midwest	0.0936	[0.0255; 0.1617]	[0.0270; 0.1601]
West	0.0850	[0.0135; 0.1565]	[0.0168; 0.1562]
Income	0.0559	[0.0273; 0.0845]	[0.0263; 0.0843]
s_2	0.1743	[0.1592; 0.1893]	[0.1583; 0.1874]
θ	2.3853	[1.9555; 2.8151]	[1.9695; 2.7993]
Log-likelihood	-129.9304		
AIC	301.8608		
BIC	385.6815		

For purposes of comparison, we also fit, via the MCECM algorithm of Huang (1999) adapted to bivariate logistic distribution, what we call here the basic bivariate SUR Tobit model with logistic marginal errors, that is the bivariate SUR Tobit model whose dependence between the error terms ϵ_{i1} and ϵ_{i2}, $i = 1, ..., n$, is modeled through the classical bivariate logistic distribution as defined by Gumbel (1961). The estimation results, obtained after 3 iterations (i.e. in fewer iterations than required by the MIFM method, but the adapted MCECM algorithm is much more time consuming), are presented in Table 2.4. The standard errors in Table 2.4 were derived from the bootstrap-based covariance matrix estimate given by (2.8) (bootstrap standard errors) [5]. It can be seen from Tables 2.3 and 2.4 that the marginal parameter estimates obtained through the adapted MCECM and MIFM methods are similar. However, the bivariate Clayton copula-based SUR Tobit model with logistic marginal errors overcomes the basic bivariate SUR Tobit model with logistic marginal errors in both AIC and BIC criterion. This indicates that the gain for introducing the Clayton copula to model the nonlinear dependence structure of the bivariate SUR Tobit model with logistic marginal errors, was substantial for this dataset.

2.2 Bivariate Clayton survival copula-based SUR Tobit right-censored model formulation

The SUR Tobit model with two right-censored dependent variables, or simply bivariate SUR Tobit right-censored model, is expressed as

$$y_{ij}^* = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j + \epsilon_{ij},$$

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* < d_j, \\ d_j & \text{otherwise,} \end{cases}$$

for $i = 1, ..., n$ and $j = 1, 2$, where n is the number of observations, d_j is the censoring point/threshold of margin j (which is assumed to be known and constant [6], here), y_{ij}^* is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^*

[5] But now with $\boldsymbol{\eta}$ denoting the parameter vector of the basic bivariate SUR Tobit model with logistic marginal errors.

[6] See, e.g., Omori & Miyawaki (2010) for examples of Tobit models with unknown and covariate dependent thresholds.

Table 2.4: Estimation results of basic bivariate SUR Tobit model with logistic marginal errors for salad dressing and lettuce consumption in the U.S. in 1994-1996.

Salad dressing	Estimate	Standard Error
Intercept	0.1288 *	0.0390
Age 20-30	0.0965 *	0.0378
Age 31-40	0.0751 *	0.0331
Age 41-50	0.0531 *	0.0313
Age 51-60	-0.0126	0.0323
Northeast	0.0587 *	0.0335
Midwest	0.0610 *	0.0292
West	0.0537 *	0.0293
Income	0.0328 *	0.0123
s_1	0.1340 *	0.0056
Lettuce	Estimate	Standard Error
Intercept	-0.0699	0.0485
Age 20-30	0.0876 *	0.0456
Age 31-40	0.0758 *	0.0413
Age 41-50	0.0750 *	0.0403
Age 51-60	-0.0049	0.0408
Northeast	0.0590	0.0409
Midwest	0.1003 *	0.0357
West	0.0740 *	0.0353
Income	0.0628 *	0.0156
s_2	0.1621 *	0.0080
Log-likelihood	-142.5471	
AIC	325.0942	
BIC	404.9234	

* Denotes significant at the 10% level.

whenever y^*_{ij} is below d_j and d_j otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients and ϵ_{ij} is the margin j's error that follows some zero mean distribution.

Suppose that the marginal errors are no longer normal, but they are assumed to be distributed according to the power-normal (Gupta & Gupta, 2008) and logistic models. Then, the density function of y_{ij} takes the following forms.

- Normal marginal errors (i.e. $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$):

$$f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j\right) = \begin{cases} \frac{1}{\sigma_j}\phi\left(\frac{y_{ij}-\boldsymbol{x}'_{ij}\boldsymbol{\beta}_j}{\sigma_j}\right) & \text{if } y_{ij} < d_j, \\ 1 - \Phi\left(\frac{y_{ij}-\boldsymbol{x}'_{ij}\boldsymbol{\beta}_j}{\sigma_j}\right) & \text{if } y_{ij} = d_j, \end{cases} \tag{2.9}$$

and the corresponding distribution function of y_{ij} is obtained by

$$F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j\right) = \begin{cases} \Phi\left(\frac{y_{ij}-\boldsymbol{x}'_{ij}\boldsymbol{\beta}_j}{\sigma_j}\right) & \text{if } y_{ij} < d_j, \\ 1 & \text{if } y_{ij} \geq d_j. \end{cases} \tag{2.10}$$

- Power-normal marginal errors (i.e. $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$):

$$f_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j, \alpha_j\right) = \begin{cases} \frac{\alpha_j}{\sigma_j} \phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{\sigma_j}\right) \left[\Phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{\sigma_j}\right)\right]^{\alpha_j - 1} & \text{if } y_{ij} < d_j, \\ 1 - \left[\Phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{\sigma_j}\right)\right]^{\alpha_j} & \text{if } y_{ij} = d_j, \end{cases} \tag{2.11}$$

and the corresponding distribution function of y_{ij} is given by

$$F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j, \alpha_j\right) = \begin{cases} \left[\Phi\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{\sigma_j}\right)\right]^{\alpha_j} & \text{if } y_{ij} < d_j, \\ 1 & \text{if } y_{ij} \geq d_j. \end{cases} \tag{2.12}$$

- Logistic marginal errors (i.e. $\epsilon_{ij} \sim L\left(0, s_j\right)$):

$$f_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, s_j\right) = \begin{cases} \frac{1}{s_j} g\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{s_j}\right) & \text{if } y_{ij} < d_j, \\ 1 - G\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{s_j}\right) & \text{if } y_{ij} = d_j, \end{cases} \tag{2.13}$$

where $g\left(z\right) = e^{z} \,/\, \left(1 + e^{z}\right)^{2}$ and $G\left(z\right) = 1 \,/\, \left(1 + e^{-z}\right)$ are the $L\left(0, 1\right)$ p.d.f. and c.d.f., respectively. The corresponding distribution function of y_{ij} is obtained by

$$F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, s_j\right) = \begin{cases} G\left(\frac{y_{ij} - \boldsymbol{x}_{ij}^{'} \boldsymbol{\beta}_j}{s_j}\right) & \text{if } y_{ij} < d_j, \\ 1 & \text{if } y_{ij} \geq d_j. \end{cases} \tag{2.14}$$

Generally, the dependence between the error terms ϵ_{i1} and ϵ_{i2} is modeled via a bivariate distribution, especially the bivariate normal distribution (this specification characterizes the basic bivariate SUR Tobit right-censored model). Nevertheless, as commented before (in Section 1.2), one of the restrictions in applying a bivariate distribution to the bivariate SUR Tobit right-censored model is the linear relationship between marginal distributions through the correlation coefficient. To overcome this restriction, we can use a copula function to model the nonlinear dependence structure in the bivariate SUR Tobit right-censored model.

Therefore, for the censored outcomes y_{i1} and y_{i2}, the bivariate copula-based SUR Tobit right-censored distribution is given by

$$F\left(y_{i1}, y_{i2}\right) = C\left(u_{i1}, u_{i2} | \theta\right),$$

where, e.g., u_{ij} is given by (2.10) if $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$, (2.12) if $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$, or (2.14) if $\epsilon_{ij} \sim L\left(0, s_j\right)$, for $j = 1, 2$, and θ is the association parameter (or parameter vector) of the copula, which is assumed to be scalar.

Suppose C is the bidimensional Clayton survival copula, which is also referred to as the reflected or rotated by 180 degrees version of the Clayton (1978) copula. It takes the form of

$$C\left(u_{i1}, u_{i2}|\theta\right) = u_{i1} + u_{i2} - 1 + \left[\left(1-u_{i1}\right)^{-\theta} + \left(1-u_{i2}\right)^{-\theta} - 1\right]^{-\frac{1}{\theta}} \tag{2.15}$$

(Georges, Lamy, Nicolas, Quibel & Roncalli, 2001), with θ restricted to the region $(0, \infty)$. The dependence between the margins increases with the value of θ, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect positive dependence. Unlike the Clayton copula, the Clayton survival copula is not Archimedean and is usually an appropriate modeling choice when the correlation between two events is stronger in the upper tail of the joint distribution.

2.2.1 Inference

In this subsection, we discuss inference (point and interval estimation) for the parameters of the bivariate Clayton survival copula-based SUR Tobit right-censored model. Particularly, by considering/assuming normal, power-normal and logistic distributions for the marginal error terms in the model.

2.2.1.1 Estimation through the MIFM method

Following Trivedi & Zimmer (2005), the log-likelihood function for the bivariate Clayton survival copula-based SUR Tobit right-censored model can be written in the following form

$$\ell\left(\boldsymbol{\eta}\right) = \sum_{i=1}^{n} \log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \boldsymbol{\upsilon}_1\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \boldsymbol{\upsilon}_2\right)|\theta\right) + \sum_{i=1}^{n}\sum_{j=1}^{2} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right), \tag{2.16}$$

where $\boldsymbol{\eta} = (\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2, \theta)$ is the vector of model parameters, $\boldsymbol{\upsilon}_j$ is the margin j's parameter vector, $f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, $F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the c.d.f. of y_{ij}, and $c\left(u_{i1}, u_{i2}|\theta\right)$, with $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton survival copula, which is calculated from (2.15) as

$$\begin{aligned} c\left(u_{i1}, u_{i2}|\theta\right) = \frac{\partial^2 C\left(u_{i1}, u_{i2}|\theta\right)}{\partial u_{i1}\partial u_{i2}} &= (\theta+1)\left[\left(1-u_{i1}\right)\left(1-u_{i2}\right)\right]^{-\theta-1} \times \\ &\times \left[\left(1-u_{i1}\right)^{-\theta} + \left(1-u_{i2}\right)^{-\theta} - 1\right]^{-\frac{1}{\theta}-2}. \end{aligned}$$

Using copula methods, as well as the log-likelihood function form given by (2.16), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}} = \arg\max_{\boldsymbol{\upsilon}_j} \sum_{i=1}^{n} \log f_j \left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j \right), \tag{2.17}$$

for $j = 1, 2$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$ by

$$\widehat{\theta}_{\mathrm{IFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\big(F_1 \left(y_{i1} | \boldsymbol{x}_{i1}, \widehat{\boldsymbol{\upsilon}}_{1,\mathrm{IFM}} \right), F_2 \left(y_{i2} | \boldsymbol{x}_{i2}, \widehat{\boldsymbol{\upsilon}}_{2,\mathrm{IFM}} \right) | \theta \big). \tag{2.18}$$

However, the IFM method provides a biased estimate for the parameter θ in the presence of censored observations for both margins (as will be seen in Section 2.2.2.2), which occurs because there is a violation of Sklar's theorem in this case (see discussion in Section 2.1.1.1). In order to obtain an unbiased estimate for the association parameter θ, we can augment the semi-continuous/censored marginal distributions to achieve continuity. More specifically, we can replace y_{ij} by the augmented data y_{ij}^{a}, or equivalently and more simply (thus, preferred by us), we can replace u_{ij} by the augmented uniform data u_{ij}^{a} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the cases of continuous margins. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence is achieved (MIFM method). The (frequentist) data augmentation technique we employ here is partially based on Algorithm A2 presented in Wichitaksorn *et al.* (2012).

In the remaining part of this subsubsection, we discuss the MIFM method when using the Clayton survival copula to describe the nonlinear dependence structure of the bivariate SUR Tobit right-censored model with arbitrary margins (e.g., normal, power-normal and logistic distribution assumption for the marginal error terms). Nevertheless, the proposed approach can be extended to other copula functions by applying different sampling algorithms. For the cases where just a single dependent variable/margin is censored (i.e. when $y_{i1} < d_1$ and $y_{i2} = d_2$, or $y_{i1} = d_1$ and $y_{i2} < d_2$), the uniform data augmentation is performed through the truncated conditional distribution of the Clayton survival copula. Since the inverse conditional distribution of the Clayton survival copula has a closed-form expression (see Appendix A), we can generate random numbers from its truncated version by applying the method by Devroye (1986, p. 38-39). Otherwise, numerical root-finding procedures would be required. As the Clayton survival copula, as well as the Clayton copula has a remarkable invariance property under truncation, the conditional distribution of

u_{i1} and u_{i2} in a sub-region of a Clayton survival copula, with one corner at $(1,1)$, can be written by means of a Clayton survival copula. This enables a simple simulation scheme in the cases where both dependent variables/margins are censored (i.e. when $y_{i1} = d_1$ and $y_{i2} = d_2$). For copulas that do not have the truncation-invariance property, an iterative simulation scheme could be adopted.

The implementation of the bivariate Clayton survival copula-based SUR Tobit right-censored model with arbitrary margins through the proposed MIFM method can be described as follows. In particular, if the marginal error distributions are normal, then set $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)$; if marginal error distributions are power-normal, so $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \left[\Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)\right]^{\alpha_j}$; and if marginal error distributions are logistic, then $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, s_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = G\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right) = \left[1 + \exp\left\{-\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right\}\right]^{-1}$, for $j = 1, 2$ and $z \in \mathbb{R}$.

Stage 1. Estimate the marginal parameters using (2.17). Set $\hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$, for $j = 1, 2$.

Stage 2. Estimate the copula parameter using e.g., (2.18). Set $\hat{\theta}^{(1)}_{\mathrm{MIFM}} = \hat{\theta}_{\mathrm{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = d_1$ and $y_{i2} = d_2$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}|\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(a_{i1}, 1) \times (a_{i2}, 1)$. This can be performed relatively easily using the following steps.

1. Draw (p, q) from $C\left(p, q|\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right) = p + q - 1 + \left[(1-p)^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-q)^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$. See Appendix A for the Clayton survival copula data generation.
2. Compute $a_{ij} = H_j\left(d_j|\boldsymbol{x}_{ij}, \hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2$.
3. Set $u^{\mathrm{a}}_{i1} = 1 - \left\{\left[(1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-a_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right](1-p)^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 1 - (1-a_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right\}^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$
4. Set $u^{\mathrm{a}}_{i2} = 1 - \left\{\left[(1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-a_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right](1-q)^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 1 - (1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right\}^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$

If $y_{i1} = d_1$ and $y_{i2} < d_2$, then draw u^{a}_{i1} from $C\left(u^{\mathrm{a}}_{i1}|u_{i2}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i1}, 1)$. This can be done by following the next steps.

1. Compute $u_{i2} = H_2\left(y_{i2}|\boldsymbol{x}_{i2}, \hat{\boldsymbol{\upsilon}}_{2,\mathrm{MIFM}}\right)$.

2. Compute $a_{i1} = H_1\left(d_1|\boldsymbol{x}_{i1}, \hat{\boldsymbol{\upsilon}}_{1,\mathrm{MIFM}}\right)$.

3. Draw t from $Uniform\,(0, 1)$.

4. Compute $v_{i1} = t + (1-t)\left\{1 - (1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}-1}\left[(1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}-1}\right\}$.

5. Set $u^{\mathrm{a}}_{i1} = 1 - \left\{1 + (1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\left[(1-v_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}/\left(\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}+1\right)} - 1\right]\right\}^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

If $y_{i1} < d_1$ and $y_{i2} = d_2$, then draw u^{a}_{i2} from $C\left(u^{\mathrm{a}}_{i2}|u_{i1}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i2}, 1)$. This can be done through the five steps of the previous case (i.e. when $y_{i1} = d_1$ and $y_{i2} < d_2$) by switching subscripts 1 and 2.

If $y_{i1} < d_1$ and $y_{i2} < d_2$, then set $u^{\mathrm{a}}_{i1} = u_{i1} = H_1\left(y_{i1}|\boldsymbol{x}_{i1}, \hat{\boldsymbol{\upsilon}}_{1,\mathrm{MIFM}}\right)$ and $u^{\mathrm{a}}_{i2} = u_{i2} = H_2\left(y_{i2}|\boldsymbol{x}_{i2}, \hat{\boldsymbol{\upsilon}}_{2,\mathrm{MIFM}}\right)$.

Given the generated/augmented marginal uniform data u^{a}_{ij}, we estimate the association parameter θ by [7]

$$\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}|\theta\right).$$

The algorithm terminates when it satisfies the stopping/convergence criterion: $|\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} - \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}| < \xi$, where ξ is the tolerance parameter (e.g., $\xi = 10^{-3}$).

2.2.1.2 Interval estimation

Joe & Xu (1996) combine the IFM method by using the jackknife method for the estimation of the standard errors of the multivariate model parameter estimates. It makes the analytic derivatives no longer required to compute the inverse Godambe information matrix, which is the asymptotic covariance matrix associated with the vector of parameter estimates under certain regularity conditions. See Joe (1997, p. 301-302) for this matrix form. Nevertheless, we carried out a pilot simulation study with results indicating that

[7] The generated/augmented marginal uniform data u^{a}_{ij} should carry (ω) as a superscript $\left(\text{i.e. } u^{\mathrm{a}(\omega)}_{ij}\right)$, but we omit it so as not to clutter the notation.

the jackknife is not valid to obtain standard errors of parameter estimates when using the MIFM approach (the jackknife method produces an overestimate of the standard error of the association parameter estimate). This implies that confidence intervals for the parameters of the bivariate Clayton survival copula-based SUR Tobit right-censored model cannot be constructed using this resampling technique. To overcome this problem, we propose a bootstrap approach for deriving confidence intervals. For more details about our bootstrap approach, we refer to Section 2.1.1.2.

2.2.2 Simulation study

A simulation study was performed to examine the behavior of the MIFM estimates, focusing on the copula association parameter estimate; and check the coverage probabilities of different confidence intervals (derived using the bootstrap approach mentioned in Section 2.2.1.2 and described in Section 2.1.1.2) for the bivariate Clayton survival copula-based SUR Tobit right-censored model parameters. Here, we considered some circumstances that might arise in the development of bivariate copula-based SUR Tobit right-censored models, involving the sample size, the censoring percentage (i.e. the percentage of d_1 and d_2 observations in the margins 1 and 2, respectively) in the dependent variables/margins and their interdependence degree. We also considered/assumed different distributions for the marginal error terms.

2.2.2.1 General specifications

In the simulation study, we applied the Clayton survival copula to model the nonlinear dependence structure of the bivariate SUR Tobit right-censored model. We set the true value for the association parameter θ at 0.67, 2 and 6, corresponding to a Kendall's tau association measure [8] of 0.25, 0.50 and 0.75, respectively. See Appendix A for the Clayton survival copula data generation.

For $i = 1, ..., n$, the covariates for margin 1, $\boldsymbol{x}_{i1} = (x_{i1,0}, x_{i1,1})^{'}$, were $x_{i1,0} = 1$ and $x_{i1,1}$ was randomly simulated from $N(2, 1^2)$. While the covariates for margin 2, $\boldsymbol{x}_{i2} = (x_{i2,0}, x_{i2,1})^{'}$, were generated as $x_{i2,0} = 1$ and $x_{i2,1}$ was randomly simulated from $N(1, 2^2)$. The model errors ϵ_{i1} and ϵ_{i2} were assumed to follow the following distributions:

- **Normal**: i.e. $\epsilon_{i1} \sim N(0, \sigma_1^2)$ and $\epsilon_{i2} \sim N(0, \sigma_2^2)$, where $\sigma_1 = 1$ and $\sigma_2 = 2$ are the

[8] The Kendall's tau for the Clayton survival copula is $\tau_2 = \theta/(\theta + 2)$, which is the same for the Clayton copula.

standard deviations (scale parameters) for margins 1 and 2, respectively. To ensure a percentage of censoring for both margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = 5$ and assume the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$ and $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$:

- $\boldsymbol{\beta}_1 = (0.7, 1)$ and $\boldsymbol{\beta}_2 = (-0.6, 1)$;
- $\boldsymbol{\beta}_1 = (1.5, 1)$ and $\boldsymbol{\beta}_2 = (1.1, 1)$;
- $\boldsymbol{\beta}_1 = (2, 1)$ and $\boldsymbol{\beta}_2 = (2.1, 1)$;
- $\boldsymbol{\beta}_1 = (2.5, 1)$ and $\boldsymbol{\beta}_2 = (3, 1)$;
- $\boldsymbol{\beta}_1 = (3, 1)$ and $\boldsymbol{\beta}_2 = (4, 1)$;

respectively. For $j = 1, 2$, the latent dependent variable of margin j, y^*_{ij}, was randomly simulated from $N\left(\boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j, \sigma^2_j\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y^*_{ij}, d_j\right\}$.

- **Power-normal**: i.e. $\epsilon_{i1} \sim PN\,(0, \sigma_1, \alpha_1)$ and $\epsilon_{i2} \sim PN\,(0, \sigma_2, \alpha_2)$, where $\sigma_1 = 1$ and $\sigma_2 = 2$ are the scale parameters for margins 1 and 2, respectively; and $\alpha_1 = \alpha_2 = 0.5$ are the shape parameters for margins 1 and 2. To ensure a percentage of censoring for both margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = 5$ and assume the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$ and $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$:

  - $\boldsymbol{\beta}_1 = (1.1, 1)$ and $\boldsymbol{\beta}_2 = (0.3, 1)$;
  - $\boldsymbol{\beta}_1 = (2.1, 1)$ and $\boldsymbol{\beta}_2 = (2.1, 1)$;
  - $\boldsymbol{\beta}_1 = (2.6, 1)$ and $\boldsymbol{\beta}_2 = (3.2, 1)$;
  - $\boldsymbol{\beta}_1 = (3.1, 1)$ and $\boldsymbol{\beta}_2 = (4.2, 1)$;
  - $\boldsymbol{\beta}_1 = (3.7, 1)$ and $\boldsymbol{\beta}_2 = (5.4, 1)$;

  respectively. For $j = 1, 2$, the latent dependent variable of margin j, y^*_{ij}, was randomly simulated from $PN\left(\boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$; therefore, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y^*_{ij}, d_j\right\}$.

- **Logistic**: i.e. $\epsilon_{i1} \sim L\,(0, s_1)$ and $\epsilon_{i2} \sim L\,(0, s_2)$, where $s_1 = 1$ and $s_2 = 2$ are the scale parameters for margins 1 and 2, respectively. To ensure a percentage of censoring for both margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = 5$ and assume the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$ and $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$:

- $\boldsymbol{\beta}_1 = (-0.3, 1)$ and $\boldsymbol{\beta}_2 = (-2.5, 1)$;
- $\boldsymbol{\beta}_1 = (0.9, 1)$ and $\boldsymbol{\beta}_2 = (-0.2, 1)$;
- $\boldsymbol{\beta}_1 = (1.7, 1)$ and $\boldsymbol{\beta}_2 = (1.5, 1)$;
- $\boldsymbol{\beta}_1 = (2.3, 1)$ and $\boldsymbol{\beta}_2 = (2.5, 1)$;
- $\boldsymbol{\beta}_1 = (3, 1)$ and $\boldsymbol{\beta}_2 = (4, 1)$;

respectively. For $j = 1, 2$, the latent dependent variable of margin j, y_{ij}^{*}, was randomly simulated from $L\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, s_j\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y_{ij}^{*}, d_j\right\}$.

For each error distribution assumption (normal, power-normal and logistic), censoring percentage in the margins (5%, 15%, 25%, 35% and 50%) and degree of dependence between them (low: $\theta = 0.67$, moderate: $\theta = 2$ and high: $\theta = 6$), we generated 100 datasets of sizes $n = 200$, 800 and 2000. Then, for each dataset (original sample), we obtained 500 bootstrap samples through a parametric resampling plan (parametric bootstrap approach), i.e. we fitted a bivariate Clayton survival copula-based SUR Tobit right-censored model with the corresponding error distributions to each dataset through the MIFM approach, and then generated a set of 500 new datasets (the same size as the original dataset/sample) from the estimated parametric model. The computing language was written in R statistical programming environment (R Core Team, 2014) and ran on a virtual machine of the Cloud-USP at ICMC, with Intel Xeon processor E5500 series, 8 core (virtual CPUs), 32 GB RAM.

We assessed the performance of the proposed models and methods through the coverage probabilities of the standard normal and percentile bootstrap confidence intervals (the nominal value is 0.90 or 90%), the Bias and the Mean Squared Error (MSE), in which the Bias and the MSE of each parameter η_h, $h = 1, ..., k$, are given by Bias $= M^{-1}\sum_{r=1}^{M}\left(\hat{\eta}_h^r - \eta_h\right)$ and MSE $= M^{-1}\sum_{r=1}^{M}\left(\hat{\eta}_h^r - \eta_h\right)^2$, respectively, where $M = 100$ is the number of replications (original datasets/samples) and $\hat{\eta}_h^r$ is the estimated value of η_h at the rth replication.

2.2.2.2 Simulation results

In this subsubsection, we present the main results obtained from the simulation study performed with samples (datasets) of different sizes, percentages of censoring in the margins and degrees of dependence between them, regarding the bivariate Clayton survival copula-based SUR Tobit right-censored model parameters estimated using the MIFM method. Since both the MIFM and IFM methods provide the same marginal parameter estimates (the first stage of the proposed method is similar to the first stage of the usual one, as seen in Section 2.2.1.1), we focus here on the Clayton survival copula parameter estimate. Some asymptotic results (such as asymptotic normality) associated with the IFM method appear in Joe & Xu (1996). We also show the results related to the estimated coverage probabilities of the 90% confidence intervals for θ, obtained through bootstrap methods (standard normal and percentile intervals).

Figures 2.10, 2.11 and 2.12 show the Bias and MSE of the observed MIFM estimates of θ for normal, power-normal and logistic marginal errors, respectively. From these figures, we observe that, regardless of the error distribution assumption, the percentage of censoring in the margins and their interdependence degree, the Bias and MSE of the MIFM estimator of θ are relatively low and tend to zero for large n, i.e. the MIFM estimator is asymptotically unbiased (despite some random fluctuations of Bias, mainly for $n \geq 800$) and consistent for the Clayton survival copula parameter.

Figures 2.13, 2.14 and 2.15 show the estimated coverage probabilities of the bootstrap confidence intervals for θ for normal, power-normal and logistic marginal errors, respectively. Observe that the estimated coverage probabilities are sufficiently high and close to the nominal value of 0.90, except for a few cases in which n is small to moderate ($n = 200$ and 800), the degree of dependence between the margins is high ($\theta = 6$) and the marginal errors follow non-normal (i.e. power-normal and logistic) distributions (see Figures 2.14(c) and 2.15(c)).

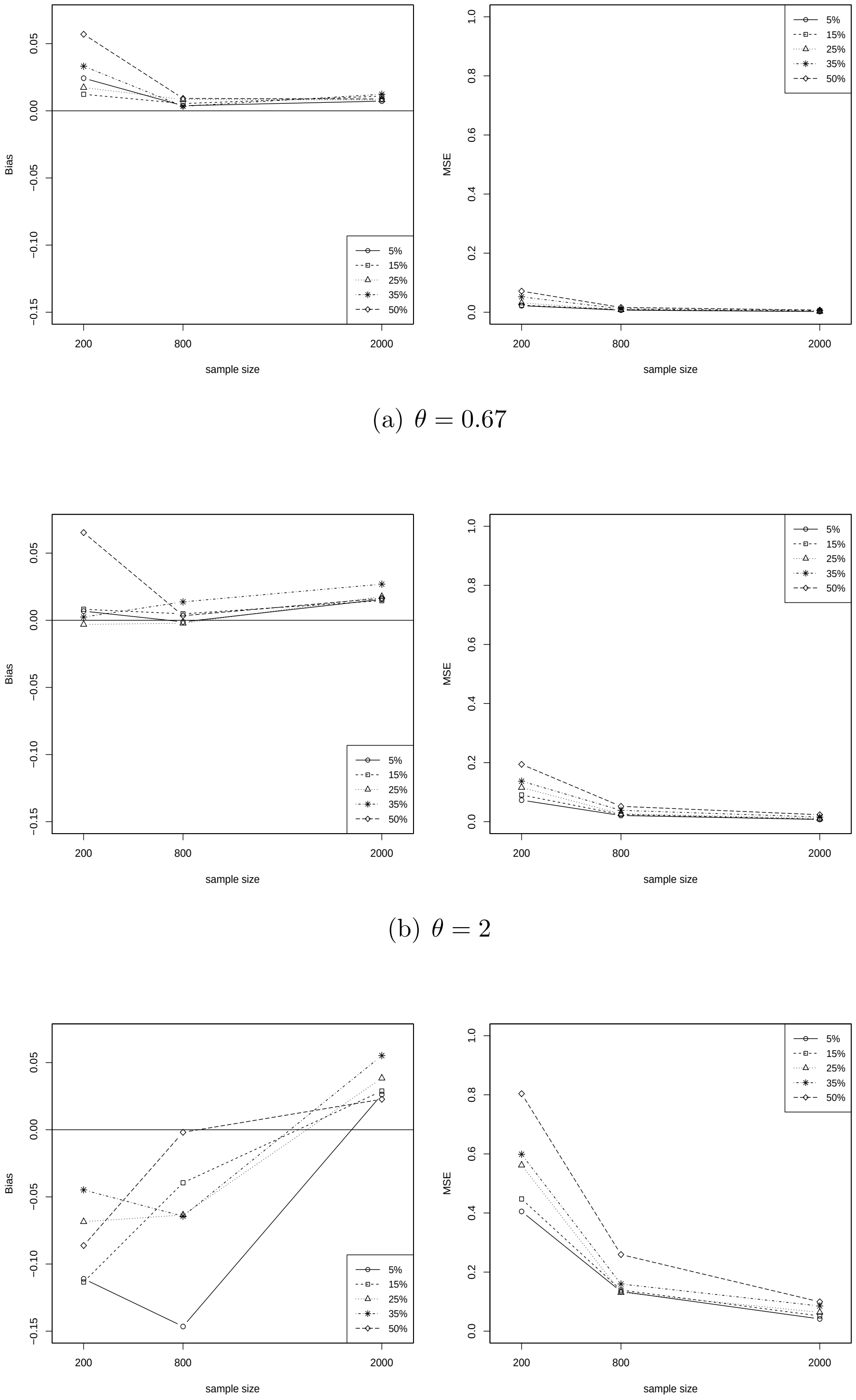
Finally, Figures 2.16, 2.17 and 2.18 compare, via boxplots, the observed MIFM estimates of θ with its estimates obtained through the IFM method for normal, power-normal and logistic marginal errors, respectively, and for $n = 2000$. It can be seen from Figure 2.16 that the MIFM method outperforms the IFM method, which overestimates θ for dependence and censoring at any level. From Figure 2.17, we observe that the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 2.17(a)), and underestimates θ for dependence at a higher level, that is $\theta = 2$ and $\theta = 6$ (Figures 2.17(b) and 2.17(c), respectively). In Figure 2.18, we see that there is a certain equivalence between the two estimation methods (with a slight advantage for the MIFM method over the IFM method, in terms of bias) when the degree of dependence between the margins is moderate, that is $\theta = 2$ (Figure 2.18(b)). However, the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 2.18(a)), and underestimates θ for dependence at a higher level, that is $\theta = 6$ (Figure 2.18(c)). Note also from Figures 2.16, 2.17 and 2.18 that the difference (distance) between the distributions of the IFM and MIFM estimates often increases with the percentage of censoring in the margins.

2.2.3 Application

Consider the customer churn dataset described in Section 1.1.2. For the sake of illustration of the bivariate models and methods proposed throughout Section 2.2, we assume that there are only two dependent variables: log(time) to churn Product A and log(time) to churn Product B (which show the highest Kendall tau correlation; see Section 1.1.2).

In this application, the relationship between the reported log(time) to churn Product A and log(time) to churn Product B (right-censored at $d_1 = d_2 = 2.3$, or approximately 10 years) of 927 customers of a Brazilian commercial bank is modeled by the bivariate SUR Tobit right-censored model with normal, power-normal and logistic marginal errors through the Clayton survival copula (see Section 1.3 for the reasons for this copula model choice). We include age and income as the covariates and use them for both margins in all three candidate models.

Tables 2.5, 2.6 and 2.7 show the MIFM estimates for the parameters of the bivariate Clayton survival copula-based SUR Tobit right-censored model with normal, power-normal and logistic marginal errors, respectively, as well as the 90% confidence intervals obtained through the standard normal and percentile bootstrap methods. Tables 2.5, 2.6 and 2.7 also present the log-likelihood, AIC and BIC criterion values for the three fitted models. Note that the bivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors has the smallest AIC and BIC criterion values and therefore provides the best fit to the customer churn data. The R codes for fitting this preferred model using the MIFM method, as well as for building standard normal and percentile bootstrap confidence intervals for its parameters, are available in Appendix B.

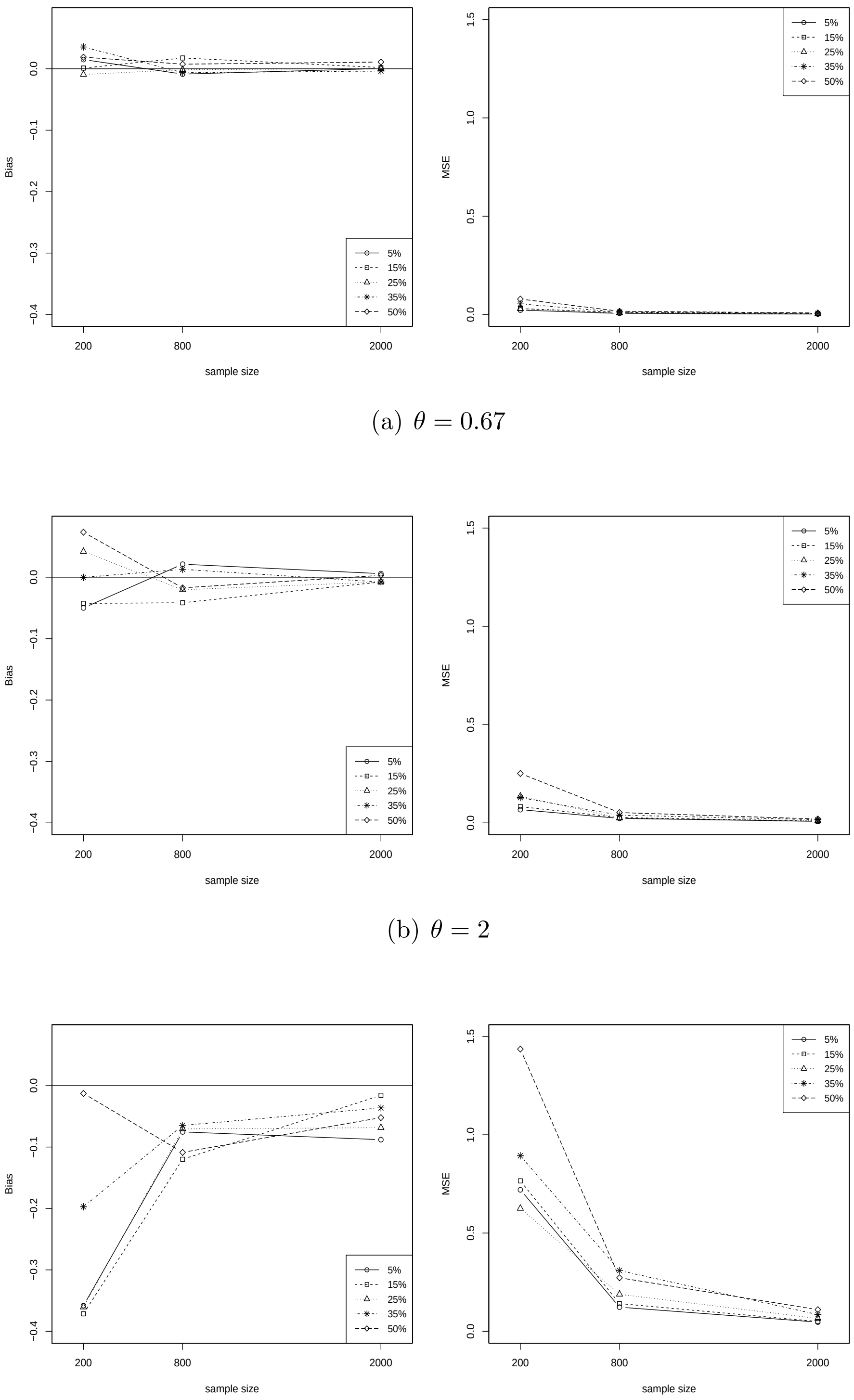


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.10: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (normal marginal errors).

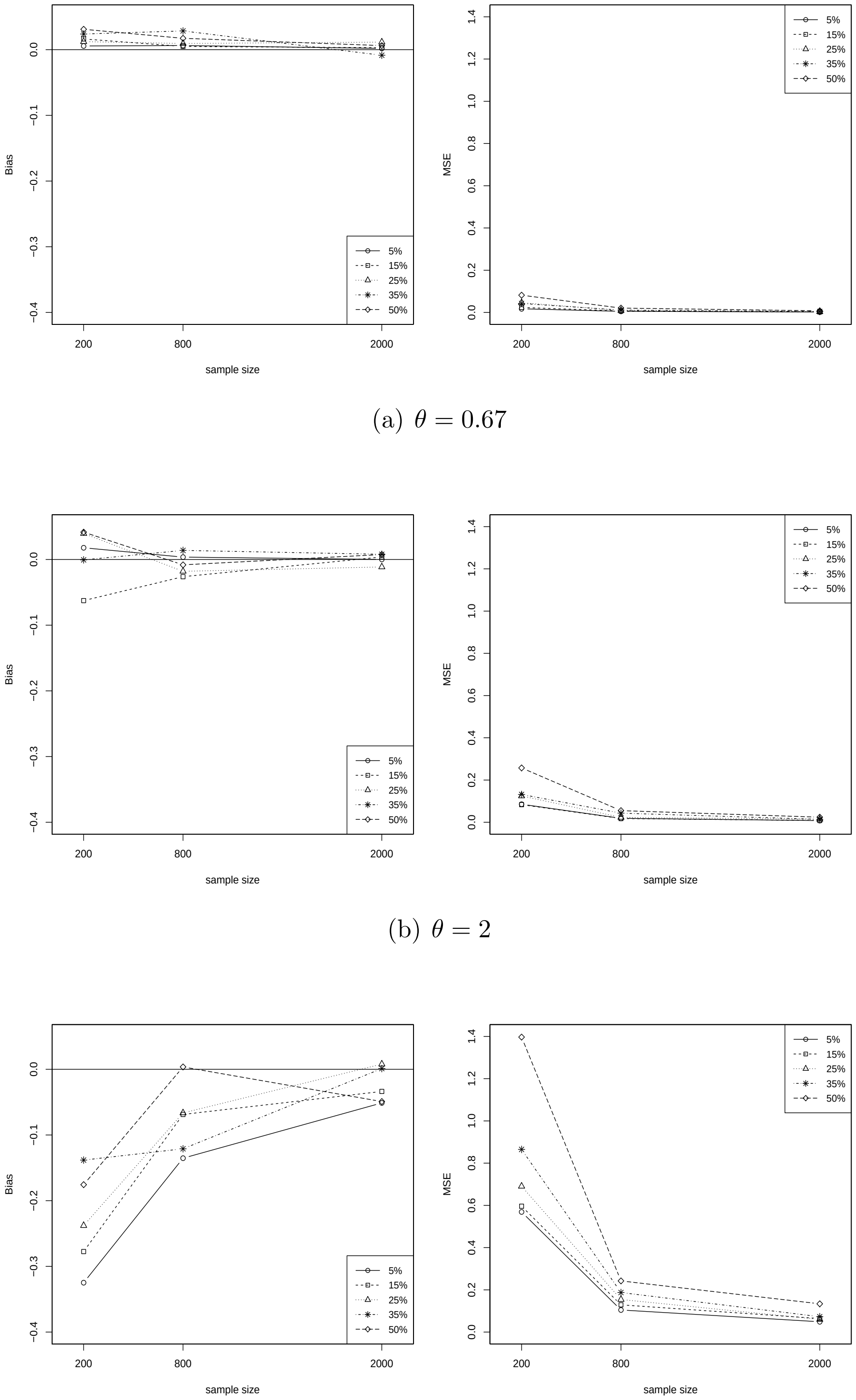


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.11: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (power-normal marginal errors).

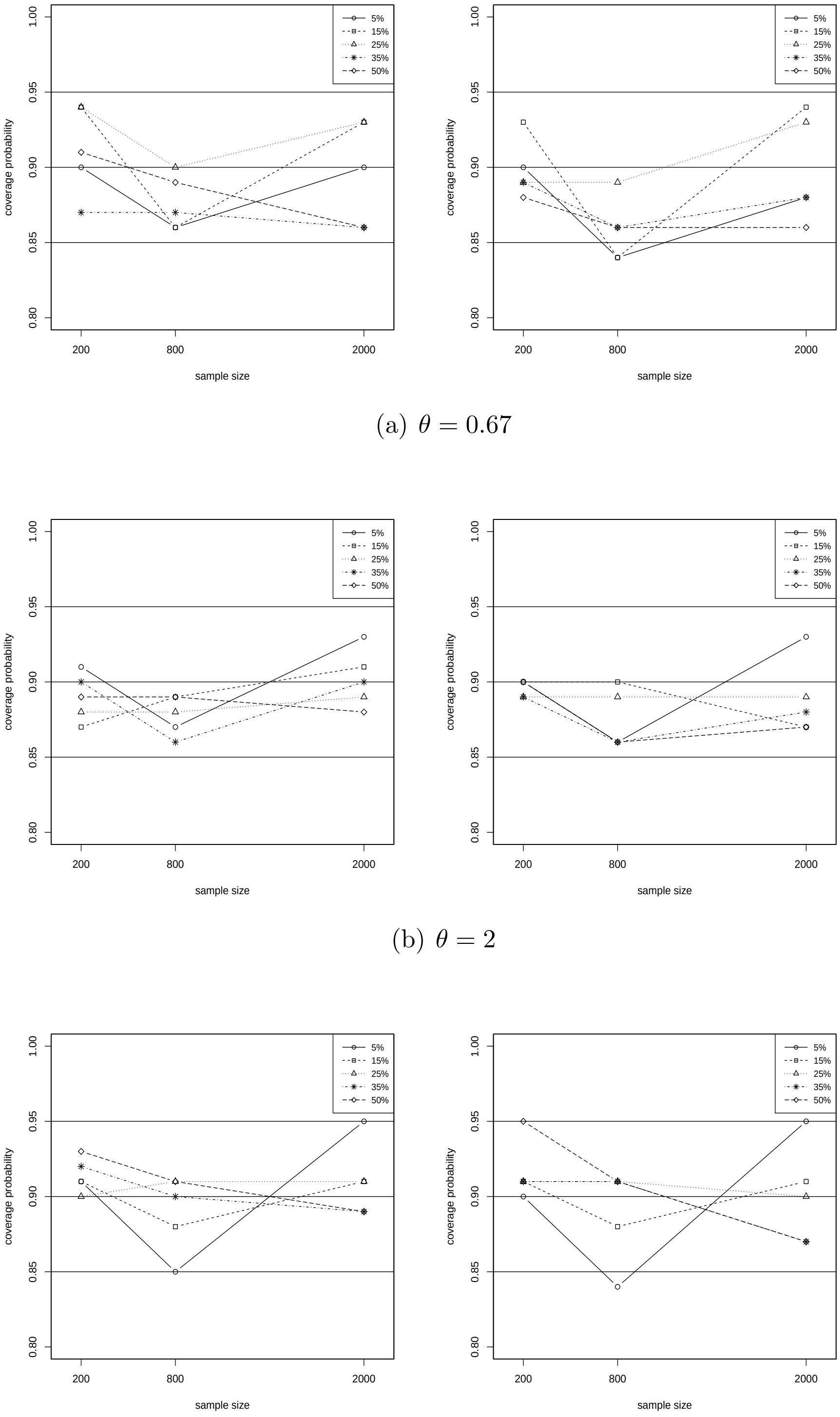


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.12: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (logistic marginal errors).

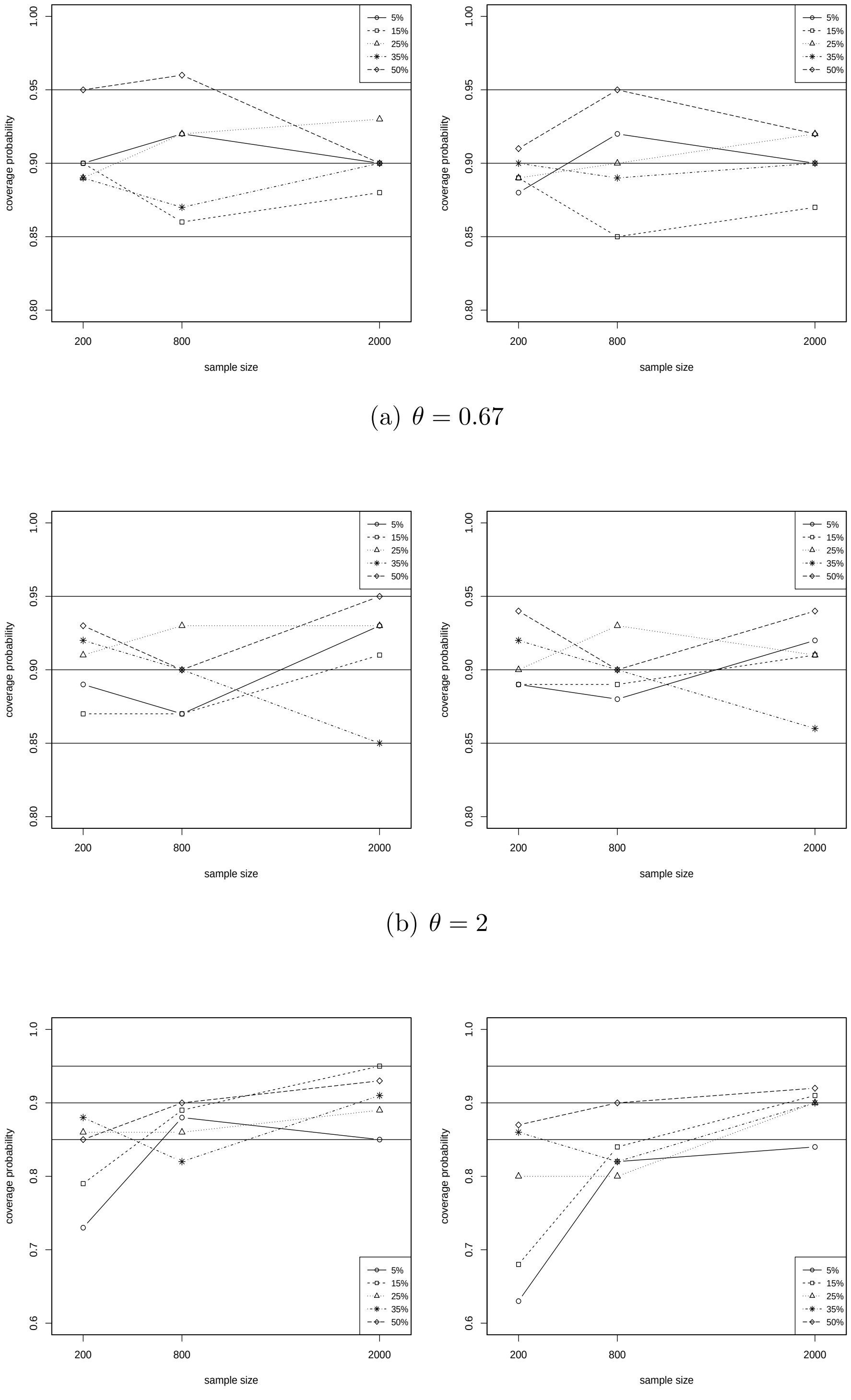


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.13: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

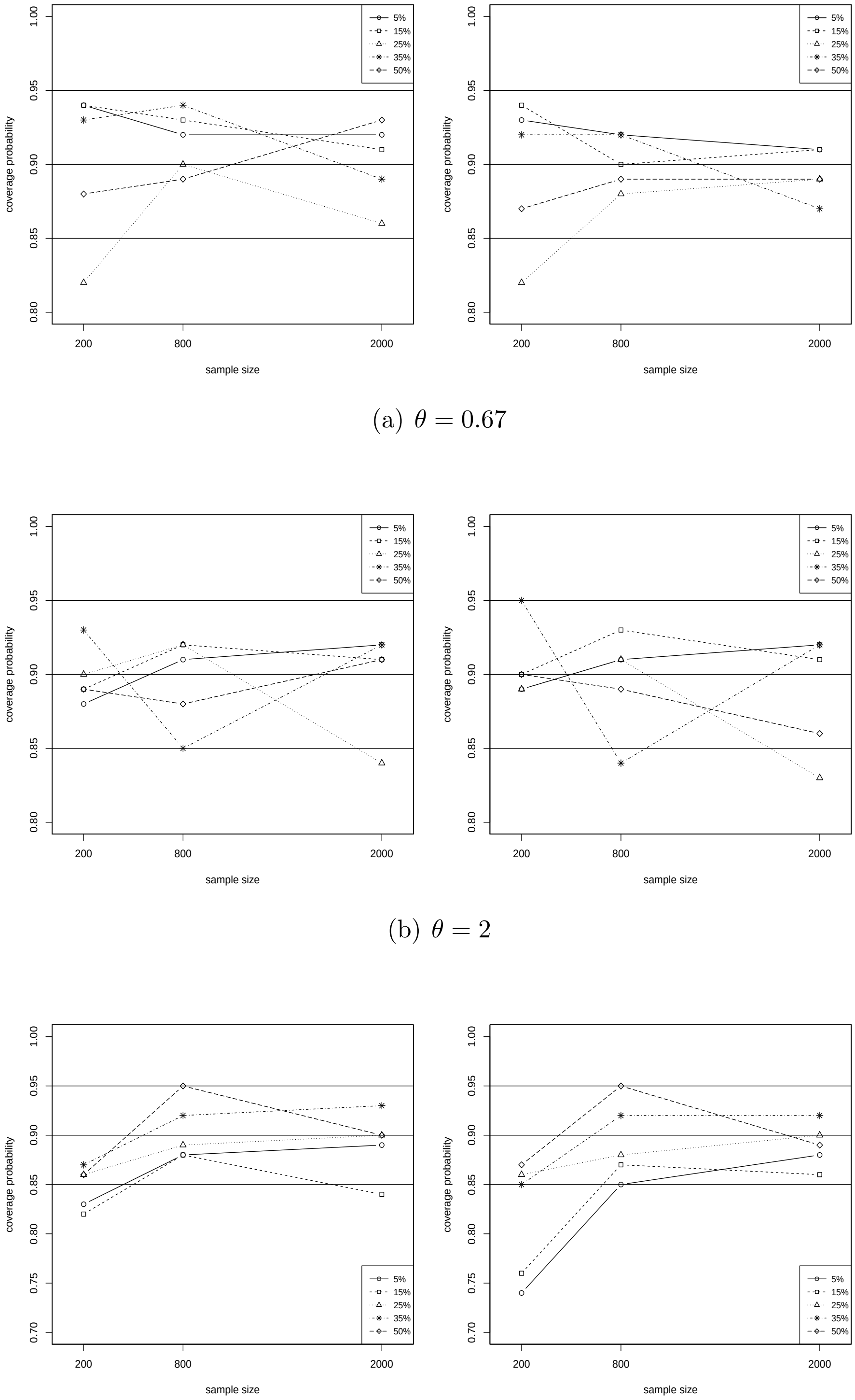


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.14: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (power-normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

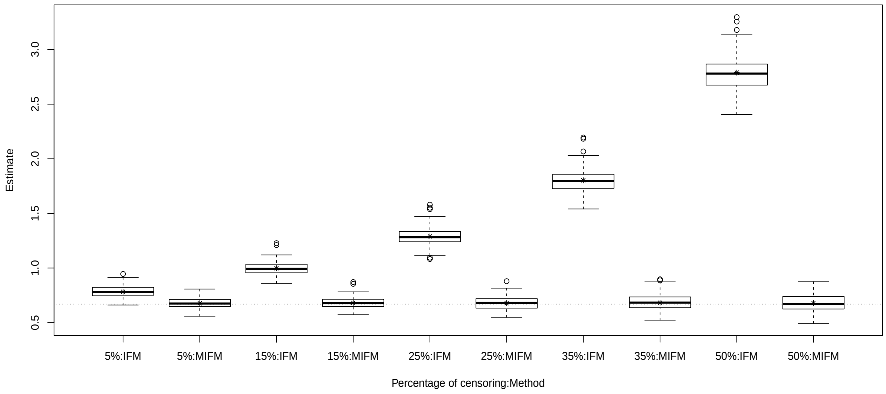


(a) $\theta = 0.67$

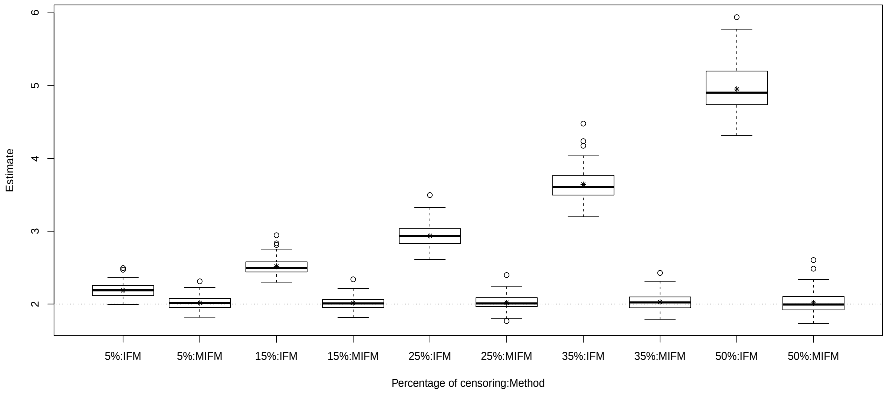
(b) $\theta = 2$

(c) $\theta = 6$

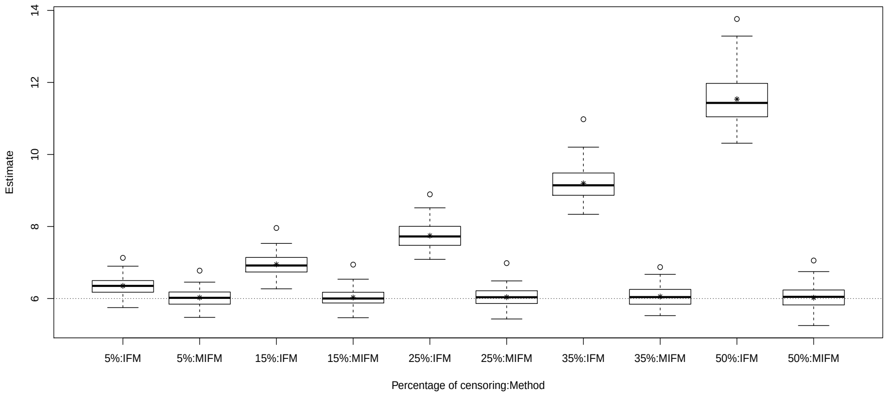
Figure 2.15: Coverage probabilities (CPs) of the 90% standard normal (panels on the left) and percentile (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence between them (logistic marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.



(a) $\theta = 0.67$

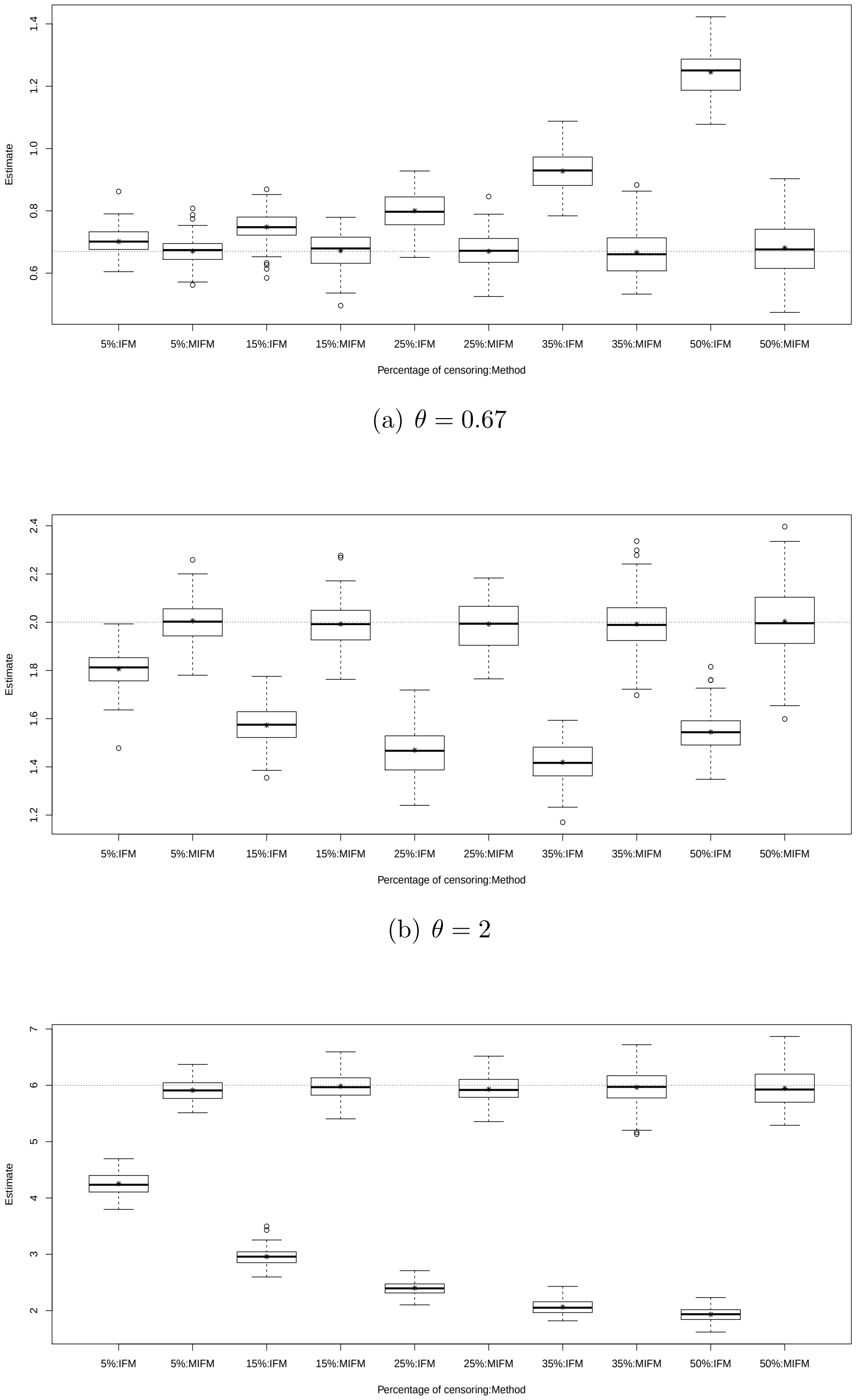


(b) $\theta = 2$



(c) $\theta = 6$

Figure 2.16: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.

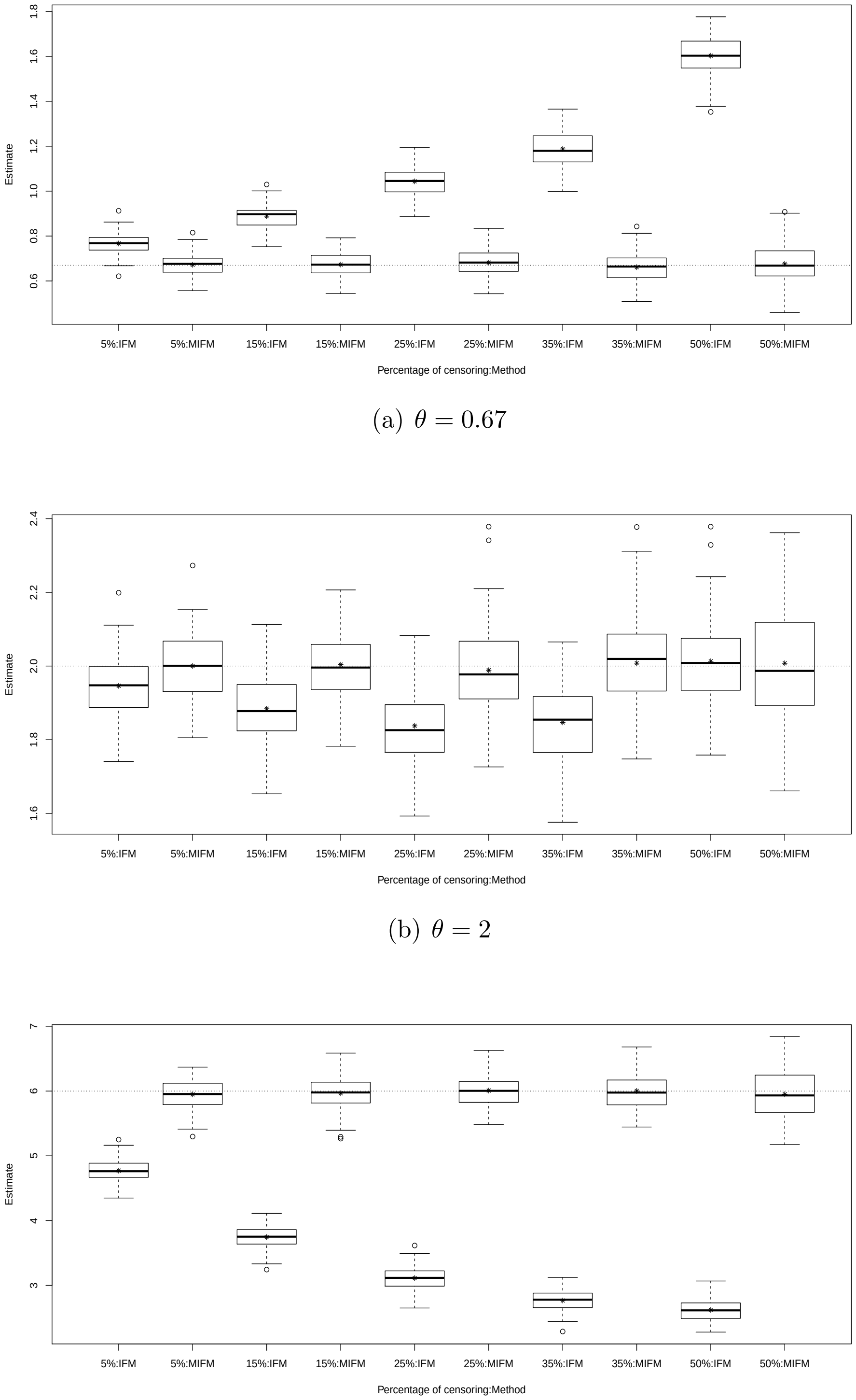


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.17: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (power-normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.



(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 2.18: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (logistic marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.

From the Lilliefors (Kolmogorov-Smirnov) normality tests (see, e.g., Thode, 2002, Section 5.1.1) of augmented marginal residuals [9], we obtain p-values equal to 0.8252 and 0.1369 for Product A and Product B models, respectively. Hence, the normality assumption for the marginal errors is valid. See, e.g., Holden (2004) and Caudill & Mixon (2009) for other approaches to testing normality of Tobit model residuals. The results reported in Table 2.5 reveal significant positive effects of age and income on log of time to churn Products A and B. The MIFM estimate of the Clayton survival copula parameter $\big(\hat{\theta}_{\text{MIFM}} = 2.7809$, obtained after 22 iterations$\big)$ and its 90% bootstrap-based confidence intervals reveal that the relationship between the log(time) to churn Product A and log(time) to churn Product B is positive (the estimated Kendall's tau is $\hat{\tau}_2 = \hat{\theta}_{\text{MIFM}} / \big(\hat{\theta}_{\text{MIFM}} + 2\big) = 0.5817$, which is not distant from the value of the nonparametric association measure presented in Section 1.1.2) and significant at the 10% level (the lower limits of the 90% bootstrap-based confidence intervals for θ are greater than and far above zero), justifying joint estimation of the censored equations through the Clayton survival copula to improve statistical efficiency. Furthermore, the estimated coefficient of tail dependence for the Clayton survival copula, $\hat{\lambda}_{\text{U}} = 0.7794$, obtained from $2^{-1/\hat{\theta}_{\text{MIFM}}}$ (the upper tail dependence coefficient for Clayton survival copula is equal to the lower tail dependence coefficient for Clayton copula), shows positive dependence at the upper tail, i.e. for high times or log of times to churn Products A and B.

For comparison purposes, we also fit the basic bivariate SUR Tobit right-censored model (that is the bivariate SUR Tobit right-censored model whose dependence between the marginal error terms ϵ_{i1} and ϵ_{i2}, $i = 1, ..., n$, is modeled through the bivariate normal distribution) using the MCECM algorithm of Huang (1999) adapted for right-censored bivariate normal data. The estimation results (obtained after 14 iterations) are presented in Table 2.8. The standard errors were derived from the bootstrap estimate of the covariance matrix (bootstrap standard errors). Note that all of the parameter estimates are significant at the 10% level. Moreover, the marginal parameter estimates obtained through the (adapted) MCECM and MIFM methods are similar (see Tables 2.5 and 2.8). However, the bivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors overcomes the basic bivariate SUR Tobit right-censored model in

[9] The augmented residuals are the differences between the augmented observed and predicted responses, i.e. $e^{\text{a}}_{ij} = y^{\text{a}}_{ij} - \boldsymbol{x}'_{ij}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}$, for $i = 1, ..., n$ and $j = 1, 2$, where $y^{\text{a}}_{ij} = \boldsymbol{x}'_{ij}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}} + \hat{\sigma}_{j,\text{MIFM}}\Phi^{-1}\left(u^{\text{a}}_{ij}\right)$, with $\Phi^{-1}(.)$ being the inverse function of the $N(0,1)$ c.d.f.; or simply, $e^{\text{a}}_{ij} = \hat{\sigma}_{j,\text{MIFM}}\Phi^{-1}\left(u^{\text{a}}_{ij}\right)$.

Table 2.5: Estimation results of bivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors for the customer churn data (Products A and B).

		90% Confidence Intervals	
Product A	Estimate	Standard Normal	Percentile
Intercept	0.1775	[0.0121; 0.3429]	[0.0091; 0.3372]
Age	0.0226	[0.0189; 0.0262]	[0.0191; 0.0263]
Income	4×10^{-5}	[1×10^{-5}; 6×10^{-5}]	[2×10^{-5}; 7×10^{-5}]
σ_1	0.9928	[0.9519; 1.0337]	[0.9517; 1.0343]
		90% Confidence Intervals	
Product B	Estimate	Standard Normal	Percentile
Intercept	0.2233	[0.0706; 0.3759]	[0.0648; 0.3683]
Age	0.0238	[0.0203; 0.0272]	[0.0202; 0.0275]
Income	8×10^{-5}	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]
σ_2	0.9098	[0.8723; 0.9472]	[0.8706; 0.9441]
θ	2.7809	[2.5139; 3.0480]	[2.5197; 3.0640]
Log-likelihood	-1920.9360		
AIC	3859.8700		
BIC	3903.3600		

Table 2.6: Estimation results of bivariate Clayton survival copula-based SUR Tobit right-censored model with power-normal marginal errors for the customer churn data (Products A and B).

		90% Confidence Intervals	
Product A	Estimate	Standard Normal	Percentile
Intercept	0.5195	[-0.2169; 1.2560]	[-0.2362; 1.1766]
Age	0.0229	[0.0190; 0.0267]	[0.0188; 0.0264]
Income	4×10^{-5}	[1×10^{-5}; 6×10^{-5}]	[2×10^{-5}; 6×10^{-5}]
σ_1	0.8594	[0.5874; 1.1315]	[0.6030; 1.0901]
α_1	0.6481	[-0.1033; 1.3995]	[0.2615; 1.3443]
		90% Confidence Intervals	
Product B	Estimate	Standard Normal	Percentile
Intercept	0.2230	[-0.5526; 0.9986]	[-0.6708; 0.8783]
Age	0.0237	[0.0200; 0.0273]	[0.0202; 0.0273]
Income	8×10^{-5}	[5×10^{-5}; 1×10^{-4}]	[5×10^{-5}; 1×10^{-4}]
σ_2	0.9113	[0.6540; 1.1686]	[0.6579; 1.2068]
α_2	1.0060	[-0.4712; 2.4831]	[0.4085; 2.6101]
θ	2.7395	[2.4364; 3.0425]	[2.4582; 3.0555]
Log-likelihood	-1923.7740		
AIC	3869.5480		
BIC	3922.6990		

Table 2.7: Estimation results of bivariate Clayton survival copula-based SUR Tobit right-censored model with logistic marginal errors for the customer churn data (Products A and B).

		90% Confidence Intervals	
Product A	Estimate	Standard Normal	Percentile
Intercept	0.1566	[-0.0148; 0.3280]	[-0.0110; 0.3273]
Age	0.0231	[0.0194; 0.0268]	[0.0195; 0.0266]
Income	4×10^{-5}	$[1 \times 10^{-5}; 6 \times 10^{-5}]$	$[1 \times 10^{-5}; 7 \times 10^{-5}]$
s_1	0.5750	[0.5482; 0.6017]	[0.5466; 0.5989]
		90% Confidence Intervals	
Product B	Estimate	Standard Normal	Percentile
Intercept	0.1592	[-0.0065; 0.3248]	[0.0026; 0.3375]
Age	0.0252	[0.0216; 0.0289]	[0.0214; 0.0285]
Income	8×10^{-5}	$[5 \times 10^{-5}; 1.1 \times 10^{-4}]$	$[6 \times 10^{-5}; 1.2 \times 10^{-4}]$
s_2	0.5363	[0.5092; 0.5633]	[0.5070; 0.5633]
θ	2.6925	[2.4193; 2.9656]	[2.4261; 2.9793]
Log-likelihood	-1940.6240		
AIC	3899.2470		
BIC	3942.7350		

Table 2.8: Estimation results of basic bivariate SUR Tobit right-censored model for the customer churn data (Products A and B).

Product A	Estimate	Standard Error
Intercept	0.2241	0.0901
Age	0.0209	0.0018
Income	4×10^{-5}	1×10^{-5}
σ_1	0.9464	0.0213
Product B	Estimate	Standard Error
Intercept	0.2514	0.0927
Age	0.0233	0.0020
Income	7×10^{-5}	1×10^{-5}
σ_2	0.9019	0.0231
ρ †	0.7389	0.0158
Log-likelihood	-1948.5050	
AIC	3915.0100	
BIC	3958.5000	

† Denotes the linear correlation coefficient.

both AIC and BIC criterion. This indicates that the gain for introducing the Clayton survival copula to model the nonlinear dependence structure of the bivariate SUR Tobit right-censored model was substantial for this dataset.

2.3 Final remarks

In this chapter, we extended the analysis of the SUR Tobit model with two left-censored or right-censored dependent variables by modeling its nonlinear dependence structure through copulas and assuming non-normal marginal error distributions. Our decision for two parametric families of copula (Clayton copula for the bivariate SUR Tobit model, and Clayton survival copula for the bivariate SUR Tobit right-censored model), as well as non-normal (power-normal and logistic) distribution assumption for the marginal error terms, were mainly motivated by the real data at hand (U.S. consumption data and Brazilian bank customer churn data). Furthermore, some advantages arose from these copula choices, regarding the development of the MIFM method for obtaining the estimates of the bivariate models' parameters. First, the Clayton copula and its survival copula are known to be preserved under truncation, which enabled simple simulation schemes in the cases where both dependent variables/margins were censored. Second, the existence of closed-form expressions for the inverse of the conditional Clayton and Clayton survival copulas' distributions enabled simple simulation schemes when just a single dependent variable/-margin was censored, by applying the method by Devroye (1986, p. 38-39). These copulas also have the ability to capture/model the tail dependence, especially the lower (case of Clayton copula) or upper (case of Clayton survival copula) tail where some data are censored.

In the simulation studies, we assessed the performance of our proposed bivariate models and methods, obtaining satisfactory results (unbiased estimates of the copula parameter, high and near the nominal value coverage probabilities of the bootstrap-based confidence intervals) regardless of the error distribution assumption, the censoring percentage in the margins and their degree of interdependence.

We also constructed bootstrap confidence intervals using the Bias-Corrected and Accelerated (BCa) method by Efron (1987), but its simulation (coverage probabilities) and real application (lower and upper limits) results were similar to those of the standard normal and percentile methods. Thus, the BCa method, which adjusts for both bias and skewness in the bootstrap distribution, is practically useless here.

Finally, we pointed out the applicability of our proposed bivariate models and methods for real datasets, where we found that the gain for introducing the copulas was substantial for these datasets.

Although it is relatively rare to analyze the SUR Tobit model with over two dimensions, our proposed approach can be straightforwardly applied to high-dimensional SUR Tobit models. This will be the subject of the next chapters.

Chapter 3

Trivariate Copula-based SUR Tobit Models

In this chapter, we propose a straightforward trivariate extension of our previously proposed bivariate models and methods. We first present the trivariate Clayton copula-based SUR Tobit model, which is the SUR Tobit model with three left-censored (at zero point) dependent variables whose dependence among them is modeled through the (tridimensional) Clayton copula. Then, we present the trivariate Clayton survival copula-based SUR Tobit right-censored model, i.e. the SUR Tobit model with three right-censored (at point $d_j > 0$, $j = 1, 2, 3$) dependent variables whose dependence structure among them is modeled by the (tridimensional) Clayton survival copula. As in the previous chapter, we assume symmetric (normal), asymmetric (power-normal) and heavy-tailed (logistic) distributions for the marginal error terms. Discussions concerning the model implementation using the proposed (extended) MIFM method, as well as the confidence interval construction from the bootstrap distribution of model parameters, are made for each proposed model. Simulation studies and applications for real datasets are also provided in this chapter.

3.1 Trivariate Clayton copula-based SUR Tobit model formulation

The SUR Tobit model with three left-censored (at zero point) dependent variables, or simply trivariate SUR Tobit model, is expressed as

$$y_{ij}^{*} = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j + \epsilon_{ij},$$

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* > 0, \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, ..., n$ and $j = 1, 2, 3$, where n is the number of observations, y_{ij}^* is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^* whenever y_{ij}^* is above zero and zero otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients and ϵ_{ij} is the margin j's error that follows some zero mean distribution.

As in the previous chapter, we suppose that the marginal errors are no longer normal, but they are assumed to be distributed according to the power-normal (Gupta & Gupta, 2008) and logistic models, thus providing asymmetric and heavy-tailed alternatives to Tobins model (Tobin, 1958). These choices of error distribution consist of expressing the density function of y_{ij} in the forms given by (2.1), (2.2) and (2.3), respectively.

The dependence among the error terms ϵ_{i1}, ϵ_{i2} and ϵ_{i3} is modeled in the usual way through a trivariate distribution, especially the trivariate normal distribution (this specification characterizes the basic trivariate SUR Tobit model). However, applying a trivariate distribution to the trivariate SUR Tobit model is limited to the linear relationship among marginal distributions through the correlation coefficients. Moreover, estimation methods for high-dimensional SUR Tobit models are often computationally demanding and difficult to implement (see comments in Section 1.2). To overcome these problems, we can use copula functions to model the nonlinear dependence structure in the trivariate SUR Tobit model.

For the censored outcomes y_{i1}, y_{i2} and y_{i3}, the trivariate copula-based SUR Tobit distribution is given by

$$F\left(y_{i1}, y_{i2}, y_{i3}\right) = C\left(u_{i1}, u_{i2}, u_{i3} | \theta\right),$$

where, e.g., $u_{ij} = F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j\right)$ if $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$, $F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$ if $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$, or $F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{\beta}_j, s_j\right)$ if $\epsilon_{ij} \sim L\left(0, s_j\right)$, for $j = 1, 2, 3$ (see Section 2.1), and θ is the copula association parameter (or parameter vector), which is assumed to be scalar.

Let us suppose that C is the tridimensional Clayton copula, which takes the form

$$C\left(u_{i1}, u_{i2}, u_{i3} | \theta\right) = \left(u_{i1}^{-\theta} + u_{i2}^{-\theta} + u_{i3}^{-\theta} - 2\right)^{-\frac{1}{\theta}}, \tag{3.1}$$

with θ restricted to the region $(0, \infty)$. The dependence among the margins increases with the value of θ, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect positive dependence. This Archimedean copula shows lower tail dependence and is characterized by zero upper tail dependence (De Luca & Rivieccio, 2012; Di Bernardino & Rullière, 2014).

3.1.1 Inference

In this subsection, we discuss inference (point and interval estimation) for the parameters of the trivariate Clayton copula-based SUR Tobit model. Particularly, by considering/assuming normal, power-normal and logistic distributions for the marginal error terms in the model.

3.1.1.1 Estimation through the (extended) MIFM method

Following Trivedi & Zimmer (2005) and Anastasopoulos, Shankar, Haddock & Mannering (2012), we can write the log-likelihood function for the trivariate Clayton copula-based SUR Tobit model in the following form [1]

$$\begin{aligned} \ell(\boldsymbol{\eta}) = & \sum_{i=1}^{n} \log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \boldsymbol{\upsilon}_1\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \boldsymbol{\upsilon}_2\right), F_3\left(y_{i3}|\boldsymbol{x}_{i3}, \boldsymbol{\upsilon}_3\right)|\theta\right) + \\ & + \sum_{i=1}^{n}\sum_{j=1}^{3} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right), \end{aligned} \tag{3.2}$$

where $\boldsymbol{\eta} = (\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2, \boldsymbol{\upsilon}_3, \theta)$ is the vector of model parameters, $\boldsymbol{\upsilon}_j$ is the margin j's parameter vector, $f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, $F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the c.d.f. of y_{ij}, and $c\left(u_{i1}, u_{i2}, u_{i3}|\theta\right)$, with $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton copula, which is calculated from (3.1) as

$$\begin{aligned} c\left(u_{i1}, u_{i2}, u_{i3}|\theta\right) &= \frac{\partial^3 C\left(u_{i1}, u_{i2}, u_{i3}|\theta\right)}{\partial u_{i1}\partial u_{i2}\partial u_{i3}} = \\ &= (\theta+1)(2\theta+1)\left(u_{i1}u_{i2}u_{i3}\right)^{-\theta-1}\left(u_{i1}^{-\theta} + u_{i2}^{-\theta} + u_{i3}^{-\theta} - 2\right)^{-\frac{1}{\theta}-3}. \end{aligned}$$

For model estimation, the use of copula methods, as well as the log-likelihood function form given by (3.2), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}} = \arg\max_{\boldsymbol{\upsilon}_j} \sum_{i=1}^{n} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right), \tag{3.3}$$

[1] This is the same form as in the case of continuous margins.

for $j = 1, 2, 3$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$ by

$$\widehat{\theta}_{\mathrm{IFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\big(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \widehat{\boldsymbol{\upsilon}}_{1,\mathrm{IFM}}\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \widehat{\boldsymbol{\upsilon}}_{2,\mathrm{IFM}}\right), F_3\left(y_{i3}|\boldsymbol{x}_{i3}, \widehat{\boldsymbol{\upsilon}}_{3,\mathrm{IFM}}\right)|\theta\big). \tag{3.4}$$

Note that each maximization task (step) has a small number of parameters, which reduces the computational difficulty. However, the IFM method provides a biased estimate for the parameter θ in the presence of censored observations in the margins (as will be seen in Section 3.1.2.2). Since we are interested in the trivariate Clayton copula-based SUR Tobit model where all marginal distributions are censored/semi-continuous, we are dealing with the case where there is not a one-to-one relationship between the marginal distributions and the copula, i.e. there is more than one copula to join the marginal distributions. This constitutes a violation of Sklar's theorem (Sklar, 1959). When it occurs, researchers often face problems in the copula model fitting and validation.

In order to facilitate the implementation of copula models with semi-continuous margins, the semi-continuous marginal distributions could be augmented to achieve continuity. More specifically, we can use a (frequentist) data augmentation technique to simulate the latent (unobserved) dependent variables in the censored margins, i.e. we generate the unobserved data with all properties, e.g., mean, variance and dependence structure that match the observed ones, and obtain the continuous marginal distributions (Wichitaksorn *et al.*, 2012). Thus, in order to obtain an unbiased estimate for the association parameter θ, we replace y_{ij} by the augmented data y_{ij}^{a}, or equivalently and more simply (thus, preferred by us), we can replace u_{ij} by the augmented uniform data u_{ij}^{a} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the continuous margin cases. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence occurs. The (frequentist) data augmentation technique we use here is partially based on Algorithm A2 presented in Wichitaksorn *et al.* (2012).

In the remaining part of this subsubsection, we discuss the proposed estimation method (an extension of the MIFM method proposed in Section 2.1.1.1 for the trivariate case) when using the Clayton copula to describe the nonlinear dependence structure of the trivariate SUR Tobit model with arbitrary margins (e.g., normal, power-normal and logistic distribution assumption for the marginal error terms). However, the proposed approach can be extended to other copula functions by applying different sampling algorithms. For the cases where just a single dependent variable/margin is censored (i.e. when $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} > 0$, or $y_{i1} > 0$ and $y_{i2} = 0$ and $y_{i3} > 0$, or $y_{i1} > 0$ and $y_{i2} > 0$ and $y_{i3} = 0$), the uniform data augmentation is performed through the (univariate) truncated conditional distribution of the Clayton copula. For the cases where two of the dependent variables/margins are censored (i.e. when $y_{i1} = 0$ and $y_{i2} = 0$ and $y_{i3} > 0$, or $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} = 0$, or $y_{i1} > 0$ and $y_{i2} = 0$ and $y_{i3} = 0$), the uniform data augmentation is performed through the (bivariate) truncated conditional distribution of the Clayton copula, e.g., by iterative (i.e. successive) conditioning. If the inverse conditional distribution of the copula used has a closed-form expression, which is the case of the Clayton copula (see, e.g., Cherubini, Luciano & Vecchiato, 2004, p. 184-185), we can generate random numbers from its truncated version by applying the method by Devroye (1986, p. 38-39). Otherwise, numerical root-finding procedures are required. By observing the results in Sungur (1999, 2002), we see that the (tridimensional) Clayton copula has the truncation dependence invariance property, such that the conditional distribution of u_{i1}, u_{i2} and u_{i3} in a sub-region of a Clayton copula, with one corner at $(0, 0, 0)$, can be written by means of a Clayton copula. That formulation enables a simple simulation scheme in the cases where all dependent variables/margins are censored (i.e. when $y_{i1} = y_{i2} = y_{i3} = 0$). For copulas that do not have the truncation-invariance property, an iterative simulation scheme could be adopted.

The implementation of the trivariate Clayton copula-based SUR Tobit model with arbitrary margins through the proposed (extended) MIFM method can be described as follows. In particular, if the marginal error distributions are normal, then set $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)$; if marginal error distributions are power-normal, so $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \left[\Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)\right]^{\alpha_j}$; and if marginal error distributions are logistic, then $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, s_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = G\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right) = \left[1 + \exp\left\{-\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right\}\right]^{-1}$, for $j = 1, 2, 3$ and $z \in \mathbb{R}$.

Stage 1. Estimate the marginal parameters using (3.3). Set $\hat{\boldsymbol{\upsilon}}_{j,\text{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\text{IFM}}$, for $j = 1, 2, 3$.

Stage 2. Estimate the copula parameter using, e.g., (3.4). Set $\hat{\theta}_{\text{MIFM}}^{(1)} = \hat{\theta}_{\text{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = y_{i2} = y_{i3} = 0$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, u^{\mathrm{a}}_{i3})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, u^{\mathrm{a}}_{i3} | \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(0, b_{i1}) \times (0, b_{i2}) \times (0, b_{i3})$. This can be performed relatively easily using the following steps.

1. Draw (p, q, r) from $C\left(p, q, r | \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right) = \left(p^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + q^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + r^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2\right)^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$. See, e.g., Cherubini *et al.* (2004, p. 184-185) for the multidimensional Clayton copula data generation using a conditional approach (conditional sampling).

2. Compute $b_{ij} = H_j\left(0 | \boldsymbol{x}_{ij}, \hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, 3$.

3. Set $u^{\mathrm{a}}_{i1} = \left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2\right) p^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 2 - b_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - b_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

4. Set $u^{\mathrm{a}}_{i2} = \left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2\right) q^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 2 - b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - b_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

5. Set $u^{\mathrm{a}}_{i3} = \left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + b_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2\right) r^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 2 - b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - b_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

If $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} > 0$, then draw u^{a}_{i1} from $C\left(u^{\mathrm{a}}_{i1} | u_{i2}, u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(0, b_{i1})$. This can be done according to the following steps.

1. Compute $u_{ij} = H_j\left(y_{ij} | \boldsymbol{x}_{ij}, \hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}}\right)$, for $j = 2, 3$.

2. Compute $b_{i1} = H_1\left(0 | \boldsymbol{x}_{i1}, \hat{\boldsymbol{\upsilon}}_{1,\mathrm{MIFM}}\right)$.

3. Draw t from $Uniform\,(0, 1)$.

4. Compute $v_{i1} = t\left[\left(b_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + u_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + u_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2\right) \Big/ \left(u_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + u_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right)\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} - 2}$.

5. Set $u^{\mathrm{a}}_{i1} = \left[v_{i1}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} / \left(2\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} + 1\right)} \left(u_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + u_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1\right) + 2 - u_{i2}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - u_{i3}^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}\right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

If $y_{i1} > 0$ and $y_{i2} = 0$ and $y_{i3} > 0$, then draw u^{a}_{i2} from $C\left(u^{\mathrm{a}}_{i2} | u_{i1}, u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(0, b_{i2})$. This can be done by following the five steps of the previous case (i.e. $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} > 0$) by switching subscripts 1 and 2.

If $y_{i1} > 0$ and $y_{i2} > 0$ and $y_{i3} = 0$, then draw u_{i3}^{a} from $C\left(u_{i3}^{\mathrm{a}}|u_{i1}, u_{i2}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the interval $(0, b_{i3})$. This can be done through the five steps of the penultimate case (i.e. $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} > 0$) by switching subscripts 1 and 3.

If $y_{i1} = 0$ and $y_{i2} = 0$ and $y_{i3} > 0$, then draw $(u_{i1}^{\mathrm{a}}, u_{i2}^{\mathrm{a}})$ from $C\left(u_{i1}^{\mathrm{a}}, u_{i2}^{\mathrm{a}}|u_{i3}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the region $(0, b_{i1}) \times (0, b_{i2})$. This can be performed relatively easily using the following steps (iterative conditioning).

1. Draw u_{i2}^{a} from $C\left(u_{i2}^{\mathrm{a}}|u_{i3}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the interval $(0, b_{i2})$. This can be done in the same manner as in the case of just a single censored dependent variable/margin in Section 2.1.1.1 (note that here C is the bidimensional Clayton copula given by (2.4)).

2. Draw u_{i1}^{a} from $C\left(u_{i1}^{\mathrm{a}}|u_{i2}^{\mathrm{a}}, u_{i3}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the interval $(0, b_{i1})$. This can be done according to the five steps of the second case (i.e. $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} > 0$).

If $y_{i1} = 0$ and $y_{i2} > 0$ and $y_{i3} = 0$, then draw $(u_{i1}^{\mathrm{a}}, u_{i3}^{\mathrm{a}})$ from $C\left(u_{i1}^{\mathrm{a}}, u_{i3}^{\mathrm{a}}|u_{i2}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the region $(0, b_{i1}) \times (0, b_{i3})$. This can be done by following the steps of the previous case (i.e. $y_{i1} = 0$ and $y_{i2} = 0$ and $y_{i3} > 0$) by switching subscripts 2 and 3.

If $y_{i1} > 0$ and $y_{i2} = 0$ and $y_{i3} = 0$, then draw $(u_{i2}^{\mathrm{a}}, u_{i3}^{\mathrm{a}})$ from $C\left(u_{i2}^{\mathrm{a}}, u_{i3}^{\mathrm{a}}|u_{i1}, \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}\right)$ truncated to the region $(0, b_{i2}) \times (0, b_{i3})$. This can be done by following the steps of the penultimate case (i.e. $y_{i1} = 0$ and $y_{i2} = 0$ and $y_{i3} > 0$) by switching subscripts 1 and 3.

If $y_{i1} > 0$ and $y_{i2} > 0$ and $y_{i3} > 0$, then set $u_{ij}^{\mathrm{a}} = u_{ij} = H_j\left(y_{ij}|\boldsymbol{x}_{ij}, \hat{\boldsymbol{v}}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, 3$.

Given the generated/augmented marginal uniform data u_{ij}^{a}, we estimate the association parameter θ by [2]

$$\hat{\theta}_{\mathrm{MIFM}}^{(\omega+1)} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u_{i1}^{\mathrm{a}}, u_{i2}^{\mathrm{a}}, u_{i3}^{\mathrm{a}}|\theta\right).$$

The algorithm stops if a termination criterion is fulfilled, e.g. if $|\hat{\theta}_{\mathrm{MIFM}}^{(\omega+1)} - \hat{\theta}_{\mathrm{MIFM}}^{(\omega)}| < \xi$, where ξ is the tolerance parameter (e.g., $\xi = 10^{-3}$).

3.1.1.2 Interval estimation

In this subsubsection, we propose the use of bootstrap methods to build confidence intervals for the parameters of the trivariate Clayton copula-based SUR Tobit model. It

[2] The generated/augmented marginal uniform data u_{ij}^{a} should carry (ω) as a superscript, i.e. $u_{ij}^{\mathrm{a}(\omega)}$, but we omit it so as not to clutter the notation.

makes the analytic derivatives no longer required to compute the asymptotic covariance matrix associated with the vector of parameter estimates.

Our bootstrap approach can be described as follows. Let η_h, $h = 1, ..., k$, be any component of the parameter vector $\boldsymbol{\eta}$ of the trivariate Clayton copula-based SUR Tobit model (see Section 3.1.1.1). By using a parametric resampling plan, we obtain the bootstrap estimates $\hat{\eta}^*_{h1}, \hat{\eta}^*_{h2}, ..., \hat{\eta}^*_{hB}$ of η_h through the (extended) MIFM method, where B is the number of bootstrap samples. Hinkley (1988) suggests that the minimum value of B will depend on the parameter being estimated, but that it will often be 100 or more. Then, we can derive confidence intervals from the bootstrap distribution through the following three methods, for instance.

- **Percentile bootstrap** (Efron & Tibshirani, 1993, p. 171). The $100\,(1-2\alpha)\,\%$ percentile confidence interval is defined by the $100\,(\alpha)$th and $100\,(1-\alpha)$th percentiles of the bootstrap distribution of $\hat{\eta}^*_h$:

$$\left[\hat{\eta}_h^{*(\alpha)}, \hat{\eta}_h^{*(1-\alpha)}\right].$$

  For Carpenter & Bithell (2000), simplicity is the attraction of this method. Note that no estimates of the standard errors are required. Furthermore, no invalid parameter values can be included in the interval.

- **Basic bootstrap** (Davison & Hinkley, 1997, p. 194). The basic bootstrap is one of the simplest schemes to build confidence intervals. We proceed in a similar way to the percentile bootstrap, using the percentiles of the bootstrap distribution of $\hat{\eta}^*_h$, but with the following different formula (note the inversion of the left and right quantiles!):

$$\left[2\hat{\eta}_h - \hat{\eta}_h^{*(1-\alpha)}, 2\hat{\eta}_h - \hat{\eta}_h^{*(\alpha)}\right],$$

  where $\hat{\eta}_h$ is the original estimate (i.e. from the original data) of η_h, obtained through the proposed (extended) MIFM method. Note that if there is a parameter constraint, such as $\eta_h > 0$, the $100\,(1-2\alpha)\,\%$ basic confidence interval given above may include invalid parameter values.

- **Standard normal interval** (Efron & Tibshirani, 1993, p. 154). Since most statistics are asymptotically normally distributed, in large samples we can use the standard error estimate, $\widehat{se}_h$, as well as the normal distribution, to yield a $100\,(1-2\alpha)\,\%$

confidence interval for η_h based on the original estimate $\hat{\eta}_h$:

$$\left[\hat{\eta}_h - z^{(1-\alpha)}\widehat{se}_h, \hat{\eta}_h - z^{(\alpha)}\widehat{se}_h\right],$$

where $z^{(\alpha)}$ represents the $100\,(\alpha)$th percentile point of a standard normal distribution, and $\widehat{se}_h$ is the hth entry on the diagonal of the bootstrap-based covariance matrix estimate of the parameter vector estimate $\hat{\boldsymbol{\eta}}$, which is given by

$$\widehat{\boldsymbol{\Sigma}}_{\text{boot}} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)', \tag{3.5}$$

where $\hat{\boldsymbol{\eta}}_b^*$, $b = 1, ..., B$, is the bootstrap estimate of $\boldsymbol{\eta}$ and

$$\overline{\hat{\boldsymbol{\eta}}}^* = \left(\frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{1b}^*, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{2b}^*, \ldots, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{kb}^*\right).$$

3.1.2 Simulation study

In this subsection, we present the main results of the simulation study that we conducted to examine the behavior of the MIFM estimates (focusing on the copula association parameter estimate) and check the coverage probabilities of bootstrap confidence intervals (constructed using the three methods described in Section 3.1.1.2) for the trivariate Clayton copula-based SUR Tobit model parameters. Here, we considered some circumstances that might arise in the development of trivariate copula-based SUR Tobit models, involving the sample size, the censoring percentage (i.e. the percentage of zero observations) in the dependent variables/margins and their interdependence degree. We also considered/assumed different distributions for the marginal error terms.

3.1.2.1 General specifications

In the simulation study, we applied the Clayton copula to model the nonlinear dependence structure of the trivariate SUR Tobit model. We set the true value for the association parameter θ at 0.67, 2 and 6, corresponding to a Kendall's tau association measure [3] of 0.25, 0.50 and 0.75, respectively. For the multidimensional Clayton copula data generation, see, e.g., Cherubini *et al.* (2004, p. 184-185) (conditional sampling).

[3] The Kendall's tau for the m-dimensional Clayton copula with parameter θ is given by $\tau_m = \left(2^{m-1}-1\right)^{-1}\left\{-1 + 2^m \prod_{p=0}^{m-1}\left(1+p\theta\right)/\left(2+p\theta\right)\right\}$ (Genest, Nešlehová & Ben Ghorbal, 2011). After some simple calculations, we find that for $m = 3$, $\tau_3 = \theta\,/\,(\theta+2)$.

For $i = 1, ..., n$, the covariates for margin 1, $\boldsymbol{x}_{i1} = (x_{i1,0}, x_{i1,1})^{'}$, were $x_{i1,0} = 1$ and $x_{i1,1}$ was randomly simulated from a standard normal distribution. The covariates for margin 2, $\boldsymbol{x}_{i2} = (x_{i2,0}, x_{i2,1})^{'}$, were generated as $x_{i2,0} = 1$ and $x_{i2,1}$ was randomly simulated from $N(1, 2^2)$. Finally, the covariates for margin 3, $\boldsymbol{x}_{i3} = (x_{i3,0}, x_{i3,1})^{'}$, were generated as $x_{i3,0} = 1$ and $x_{i3,1}$ was randomly simulated from $Uniform(0, 5)$. The model errors ϵ_{i1}, ϵ_{i2} and ϵ_{i3} were assumed to follow the following distributions:

- **Normal**: i.e. $\epsilon_{i1} \sim N(0, \sigma_1^2)$, $\epsilon_{i2} \sim N(0, \sigma_2^2)$ and $\epsilon_{i3} \sim N(0, \sigma_3^2)$, where $\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 2$ are the standard deviations (scale parameters) for margins 1, 2 and 3, respectively. To ensure a percentage of censoring (i.e. of zero observations) for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$, $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$ and $\boldsymbol{\beta}_3 = (\beta_{3,0}, \beta_{3,1})^{'}$:

  - $\boldsymbol{\beta}_1 = (2.3, 1)$, $\boldsymbol{\beta}_2 = (4, -0.5)$ and $\boldsymbol{\beta}_3 = (1.5, 1)$;
  - $\boldsymbol{\beta}_1 = (1.5, 1)$, $\boldsymbol{\beta}_2 = (2.75, -0.5)$ and $\boldsymbol{\beta}_3 = (0.1, 1)$;
  - $\boldsymbol{\beta}_1 = (1, 1)$, $\boldsymbol{\beta}_2 = (2, -0.5)$ and $\boldsymbol{\beta}_3 = (-0.8, 1)$;
  - $\boldsymbol{\beta}_1 = (0.5, 1)$, $\boldsymbol{\beta}_2 = (1.3, -0.5)$ and $\boldsymbol{\beta}_3 = (-1.5, 1)$;
  - $\boldsymbol{\beta}_1 = (-0.02, 1)$, $\boldsymbol{\beta}_2 = (0.5, -0.5)$ and $\boldsymbol{\beta}_3 = (-2.5, 1)$;

  respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y_{ij}^*, was randomly simulated from $N\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, \sigma_j^2\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\max\left\{0, y_{ij}^*\right\}$.

- **Power-normal**: i.e. $\epsilon_{i1} \sim PN(0, \sigma_1, \alpha_1)$, $\epsilon_{i2} \sim PN(0, \sigma_2, \alpha_2)$ and $\epsilon_{i3} \sim PN(0, \sigma_3, \alpha_3)$, where $\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 2$ are the scale parameters for margins 1, 2 and 3, respectively; and $\alpha_1 = \alpha_2 = \alpha_3 = 1.75$ are the shape parameters for margins 1, 2 and 3. To ensure a percentage of censoring for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$, $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$ and $\boldsymbol{\beta}_3 = (\beta_{3,0}, \beta_{3,1})^{'}$:

  - $\boldsymbol{\beta}_1 = (1.7, 1)$, $\boldsymbol{\beta}_2 = (2.8, -0.5)$ and $\boldsymbol{\beta}_3 = (0.2, 1)$;
  - $\boldsymbol{\beta}_1 = (0.9, 1)$, $\boldsymbol{\beta}_2 = (1.6, -0.5)$ and $\boldsymbol{\beta}_3 = (-1.1, 1)$;
  - $\boldsymbol{\beta}_1 = (0.4, 1)$, $\boldsymbol{\beta}_2 = (0.9, -0.5)$ and $\boldsymbol{\beta}_3 = (-1.9, 1)$;
  - $\boldsymbol{\beta}_1 = (0.05, 1)$, $\boldsymbol{\beta}_2 = (0.4, -0.5)$ and $\boldsymbol{\beta}_3 = (-2.5, 1)$;

  - $\boldsymbol{\beta}_1 = (-0.5, 1)$, $\boldsymbol{\beta}_2 = (-0.4, -0.5)$ and $\boldsymbol{\beta}_3 = (-3.4, 1)$;

  respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y_{ij}^*, was randomly simulated from $PN\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$; therefore, the observed dependent variable of margin j, y_{ij}, was obtained from $\max\left\{0, y_{ij}^*\right\}$.

- **Logistic**: i.e. $\epsilon_{i1} \sim L(0, s_1)$, $\epsilon_{i2} \sim L(0, s_2)$ and $\epsilon_{i3} \sim L(0, s_3)$, where $s_1 = 1$, $s_2 = 2$ and $s_3 = 1.5$ are the scale parameters for margins 1, 2 and 3, respectively. To ensure a percentage of censoring for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we assumed the following true values for $\boldsymbol{\beta}_1 = (\beta_{1,0}, \beta_{1,1})^{'}$, $\boldsymbol{\beta}_2 = (\beta_{2,0}, \beta_{2,1})^{'}$ and $\boldsymbol{\beta}_3 = (\beta_{3,0}, \beta_{3,1})^{'}$:

  - $\boldsymbol{\beta}_1 = (3.3, 1)$, $\boldsymbol{\beta}_2 = (5.8, 1)$ and $\boldsymbol{\beta}_3 = (3.4, 0.5)$;
  - $\boldsymbol{\beta}_1 = (2.1, 1)$, $\boldsymbol{\beta}_2 = (3.1, 1)$ and $\boldsymbol{\beta}_3 = (1.5, 0.5)$;
  - $\boldsymbol{\beta}_1 = (1.3, 1)$, $\boldsymbol{\beta}_2 = (1.7, 1)$ and $\boldsymbol{\beta}_3 = (0.5, 0.5)$;
  - $\boldsymbol{\beta}_1 = (0.8, 1)$, $\boldsymbol{\beta}_2 = (0.5, 1)$ and $\boldsymbol{\beta}_3 = (-0.3, 0.5)$;
  - $\boldsymbol{\beta}_1 = (-0.05, 1)$, $\boldsymbol{\beta}_2 = (-0.9, 1)$ and $\boldsymbol{\beta}_3 = (-1.2, 0.5)$;

  respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y_{ij}^*, was randomly simulated from $L\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, s_j\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\max\left\{0, y_{ij}^*\right\}$.

For each error distribution assumption (normal, power-normal and logistic), censoring percentage in the margins (5%, 15%, 25%, 35% and 50% of zero observations) and degree of dependence among them (low: $\theta = 0.67$, moderate: $\theta = 2$ and high: $\theta = 6$), we generated 100 datasets of sizes $n = 200$, 800 and 2000. These choices of sample sizes were based on some authors' indication (e.g., Joe, 2014) that large sample sizes are commonly required when working with copulas. Then, for each dataset (original sample), we obtained 500 bootstrap samples through a parametric resampling plan (parametric bootstrap approach), i.e. we fitted a trivariate Clayton copula-based SUR Tobit model with the corresponding error distributions to each dataset using the (extended) MIFM approach, and then generated a set of 500 new datasets (the same size as the original dataset/sample) from the estimated parametric model. The computing language was written in R statistical programming environment (R Core Team, 2014) and ran on a virtual machine of the Cloud-USP at ICMC, with Intel Xeon processor E5500 series, 8 core (virtual CPUs), 32 GB RAM.

We assessed the performance of the proposed models and methods through the coverage probabilities of the nominally 90% standard normal, percentile and basic bootstrap confidence intervals, the Bias and the Mean Squared Error (MSE), in which the Bias and the MSE of each parameter η_h, $h = 1, ..., k$, are given by Bias $= M^{-1} \sum_{r=1}^{M} (\hat{\eta}_h^r - \eta_h)$ and MSE $= M^{-1} \sum_{r=1}^{M} (\hat{\eta}_h^r - \eta_h)^2$, respectively, where $M = 100$ is the number of replications (original datasets/samples) and $\hat{\eta}_h^r$ is the estimated value of η_h at the rth replication.

3.1.2.2 Simulation results

In this subsubsection, we present the main results obtained from the simulation study performed with samples (datasets) of different sizes, percentages of censoring in the margins and degrees of dependence among them, regarding the trivariate Clayton copula-based SUR Tobit model parameters estimated using the (extended) MIFM approach. Since both the (extended) MIFM and IFM methods provide the same marginal parameter estimates (the first stage of the proposed method is similar to the first stage of the usual one, as seen in Section 3.1.1.1), we focus here on the Clayton copula parameter estimate. For some asymptotic results (such as asymptotic normality) associated with the IFM method, see, e.g., Joe & Xu (1996). We also show the results related to the estimated coverage probabilities of the 90% confidence intervals for θ, obtained by bootstrap methods (standard normal, percentile and basic intervals).

Figures 3.1, 3.2 and 3.3 show the Bias and MSE of the observed MIFM estimates of θ for normal, power-normal and logistic marginal errors, respectively. From these figures, we observe that, regardless of the error distribution assumption, the percentage of censoring in the margins and their interdependence degree, the Bias and MSE of the MIFM estimator of θ are relatively low and tend to zero for large n, i.e. the MIFM estimator is asymptotically unbiased and consistent for the Clayton copula parameter.

Figures 3.4, 3.5 and 3.6 show the estimated coverage probabilities of the bootstrap confidence intervals for θ for normal, power-normal and logistic marginal errors, respectively. Observe that the estimated coverage probabilities are sufficiently high and close to the nominal value of 0.90, except for the percentile intervals in general, and for a few cases in which n is mainly small to moderate ($n = 200$ and 800) and the degree of dependence among the margins is high ($\theta = 6$) (see Figures 3.4(c), 3.5(c) and 3.6(c)).

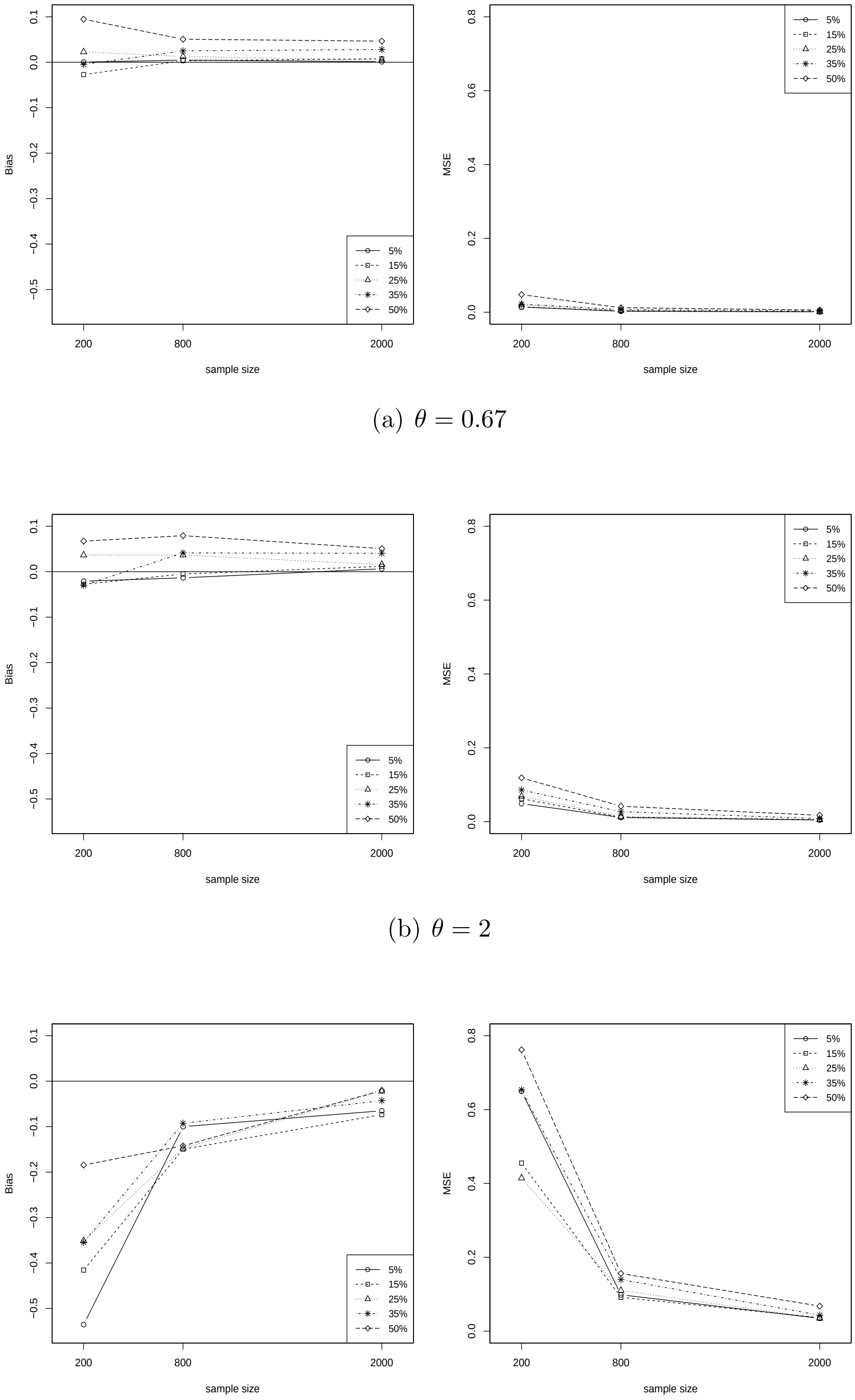
Finally, Figures 3.7, 3.8 and 3.9 compare, via boxplots, the observed MIFM estimates of θ with its estimates obtained through the IFM method for normal, power-normal and logistic marginal errors, respectively, and for $n = 2000$. It can be seen from Figure 3.7 that the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 3.7(a)), but underestimates θ for dependence at a higher level, that is $\theta = 2$ and $\theta = 6$ (Figures 3.7(b) and 3.7(c), respectively). Similar behavior is observed for the plots in Figure 3.8. In Figure 3.9, we see that there is a certain equivalence between the two estimation methods (with a slight advantage for the (extended) MIFM method over the IFM method, in terms of bias) when the degree of dependence among the margins is moderate, that is $\theta = 2$ (Figure 3.9(b)); however, the IFM method overestimates θ for dependence at a lower level, which is $\theta = 0.67$ (Figure 3.9(a)), and underestimates θ for dependence at a higher level, which is $\theta = 6$ (Figure 3.9(c)). Note also from Figures 3.7, 3.8 and 3.9 that the difference (distance) between the distributions of the IFM and MIFM estimates often increases as the percentage of censoring in the margins increases.

3.1.3 Application

In this subsection, we illustrate the applicability of our proposed trivariate models and methods for the salad dressing, tomato and lettuce data described in Section 1.1.1.

In this application, the relationship among the reported salad dressing (amount consumed in 100 grams), tomato (amount consumed in 400 grams) and lettuce (amount consumed in 200 grams) consumption by 400 U.S. adults is modeled by the trivariate SUR Tobit model with normal, power-normal and logistic marginal errors through the Clayton copula (see Sections 1.1.1 and 1.3 for the reasons for this choice of model). We include age, location (region) and income as the covariates and use them for all margins in all three candidate models.

Tables 3.1, 3.2 and 3.3 show the MIFM estimates for the parameters of the trivariate Clayton copula-based SUR Tobit model with normal, power-normal and logistic marginal errors, respectively, as well as the 90% confidence intervals obtained through the standard normal, percentile and basic bootstrap methods. Tables 3.1, 3.2 and 3.3 also present the log-likelihood, AIC and BIC criterion values for the three fitted models. Note that the trivariate Clayton copula-based SUR Tobit model with logistic marginal errors has the smallest AIC and BIC criterion values and therefore provides the best fit for the salad

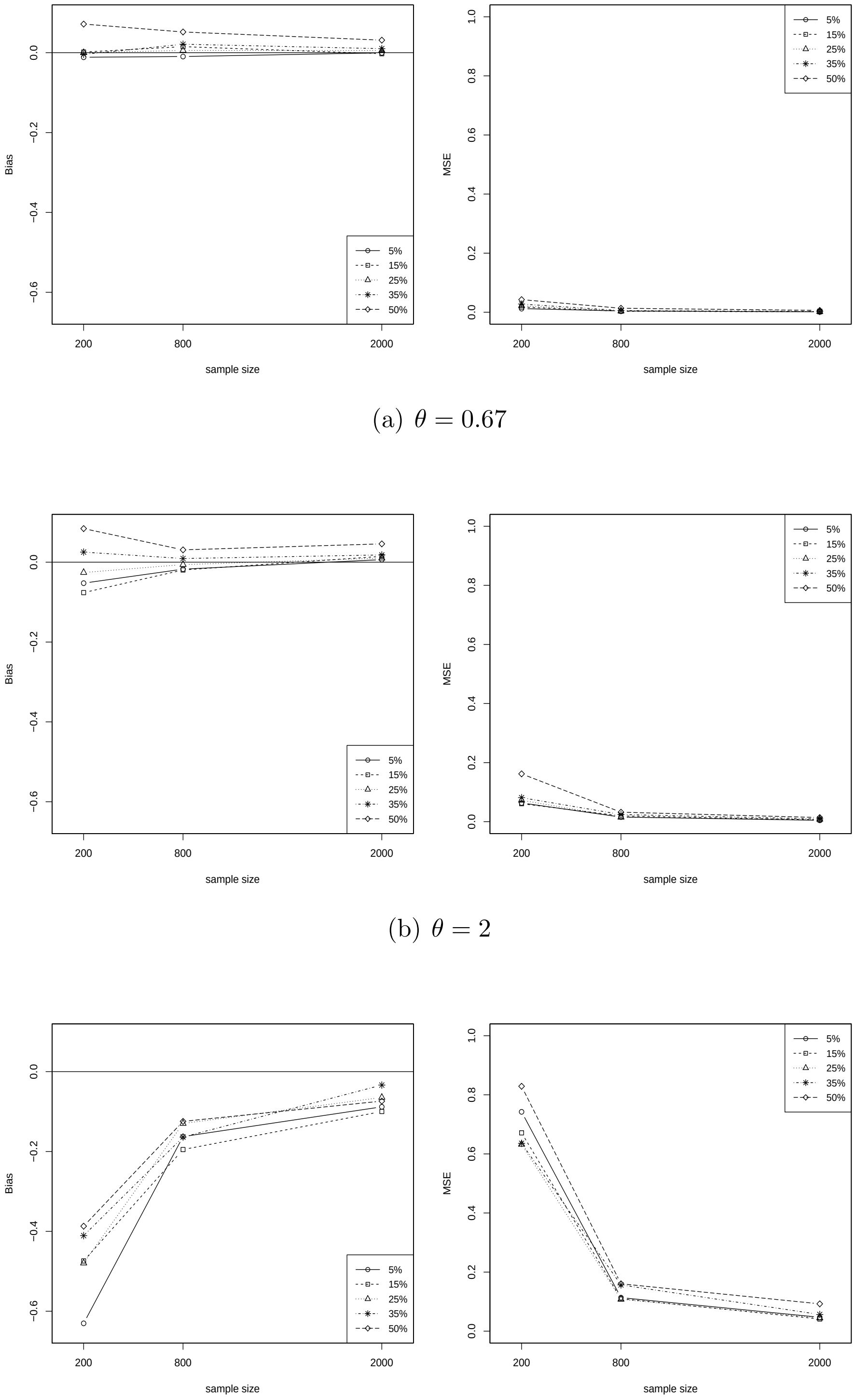


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.1: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (normal marginal errors).

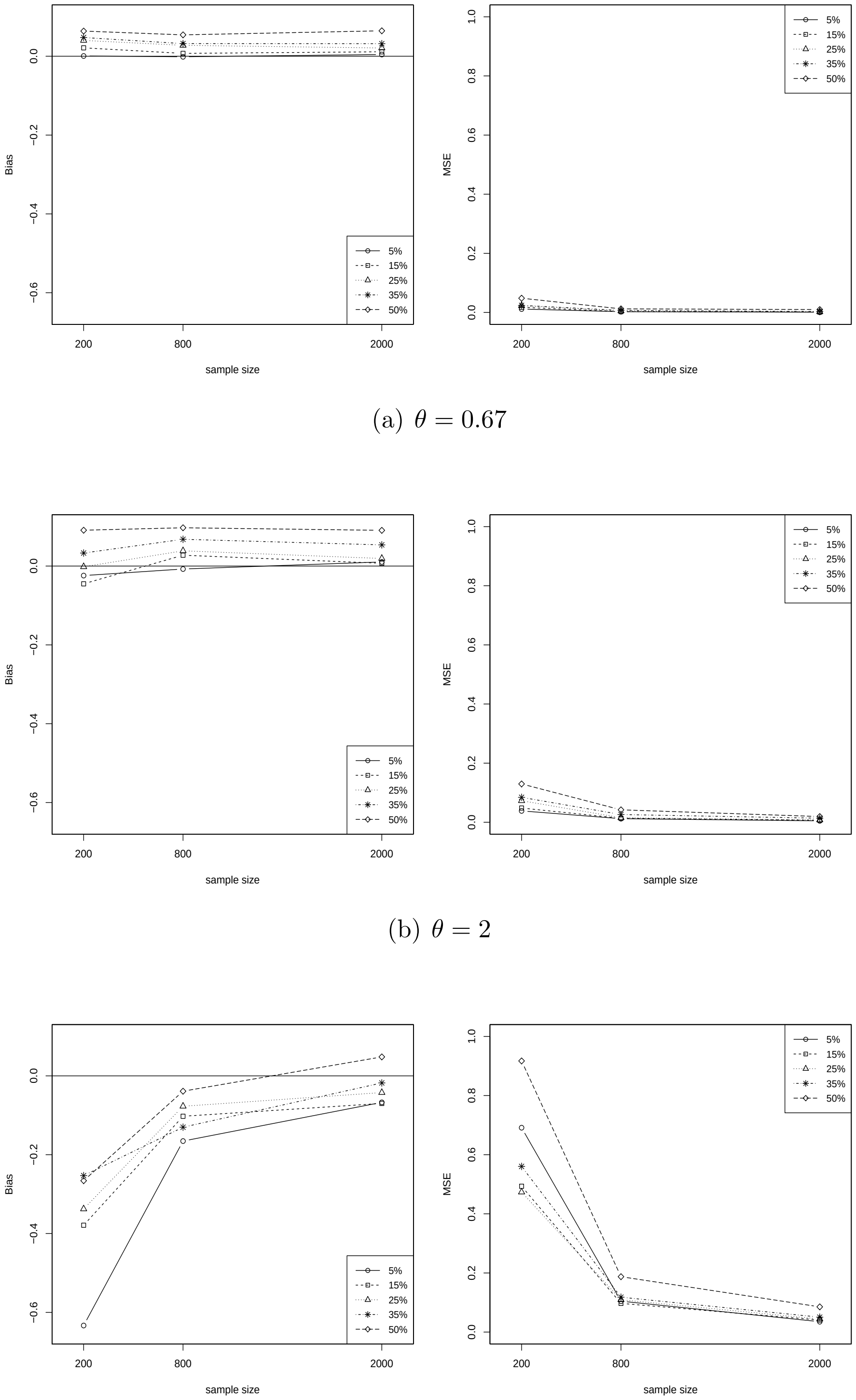


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.2: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (power-normal marginal errors).

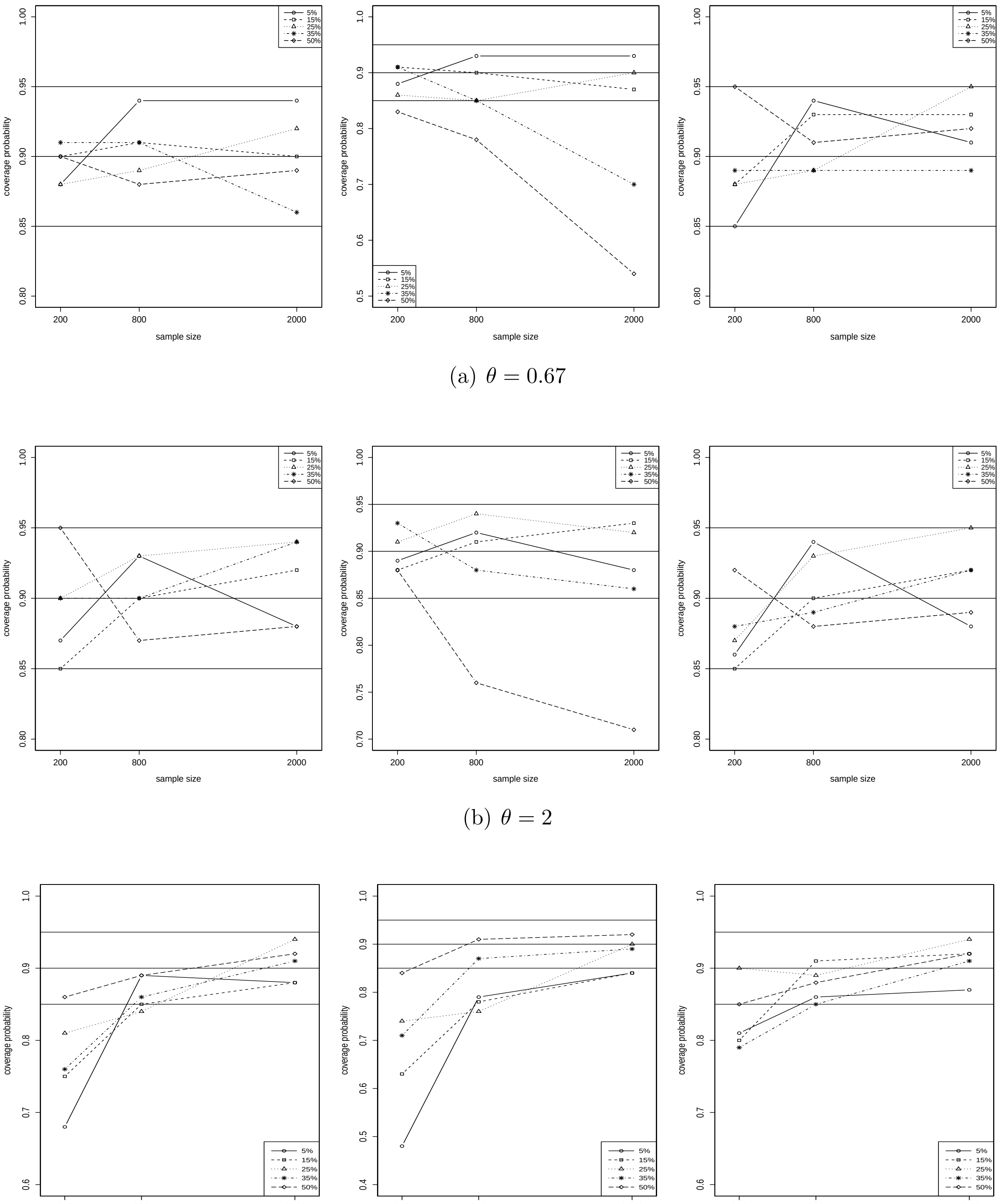


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.3: Bias and MSE of the MIFM estimate of the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (logistic marginal errors).

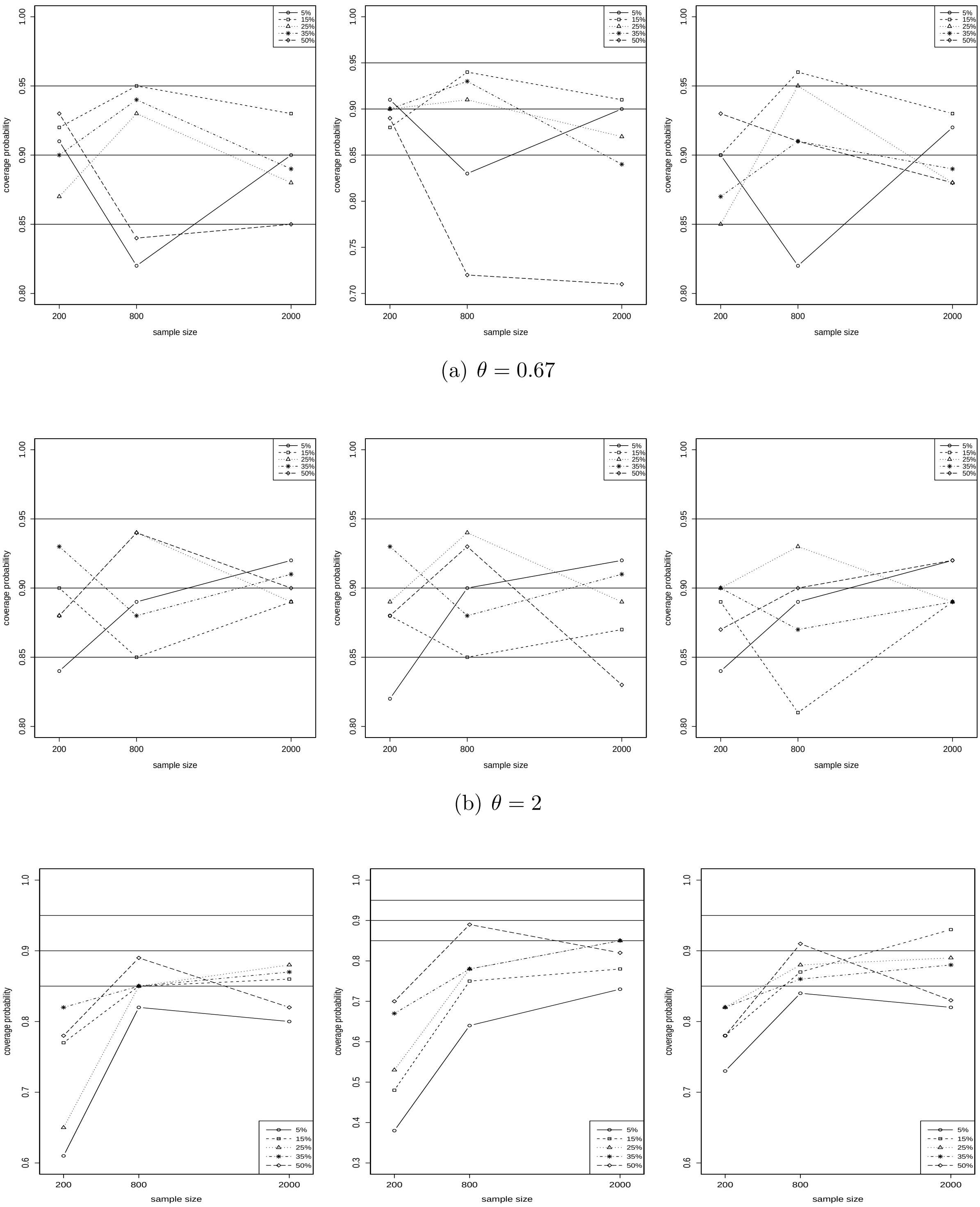


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.4: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

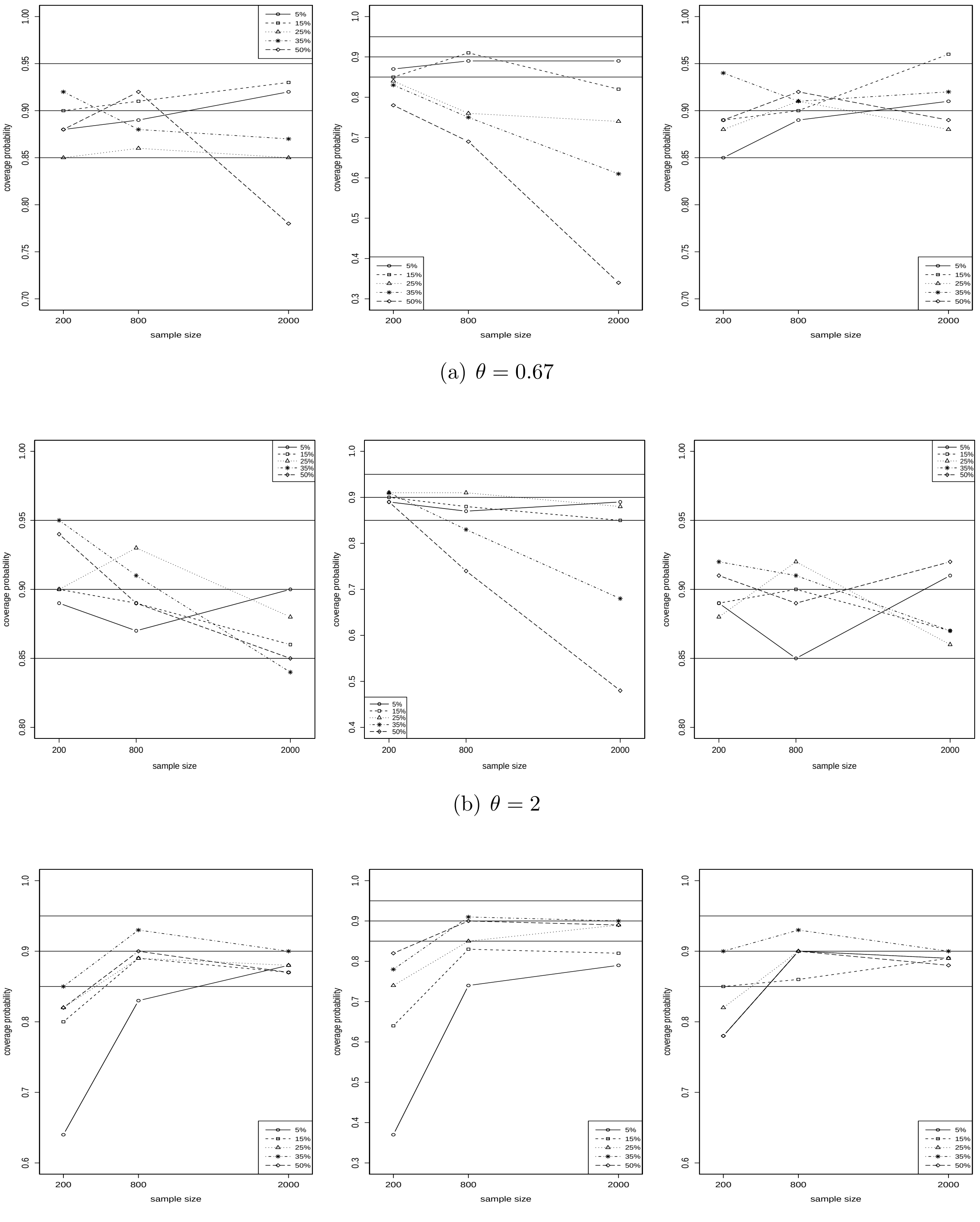


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.5: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (power-normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

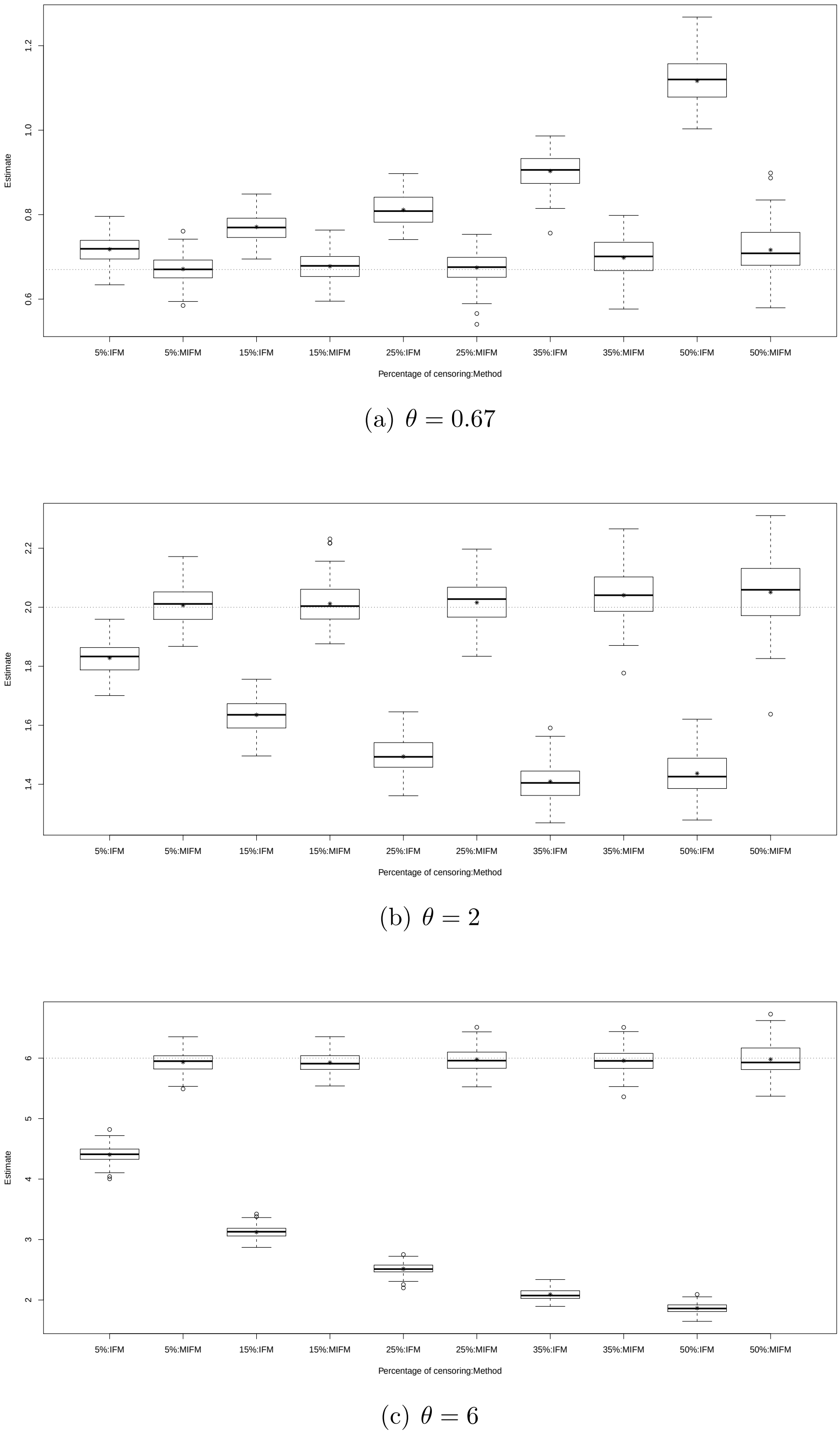


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.6: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (logistic marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

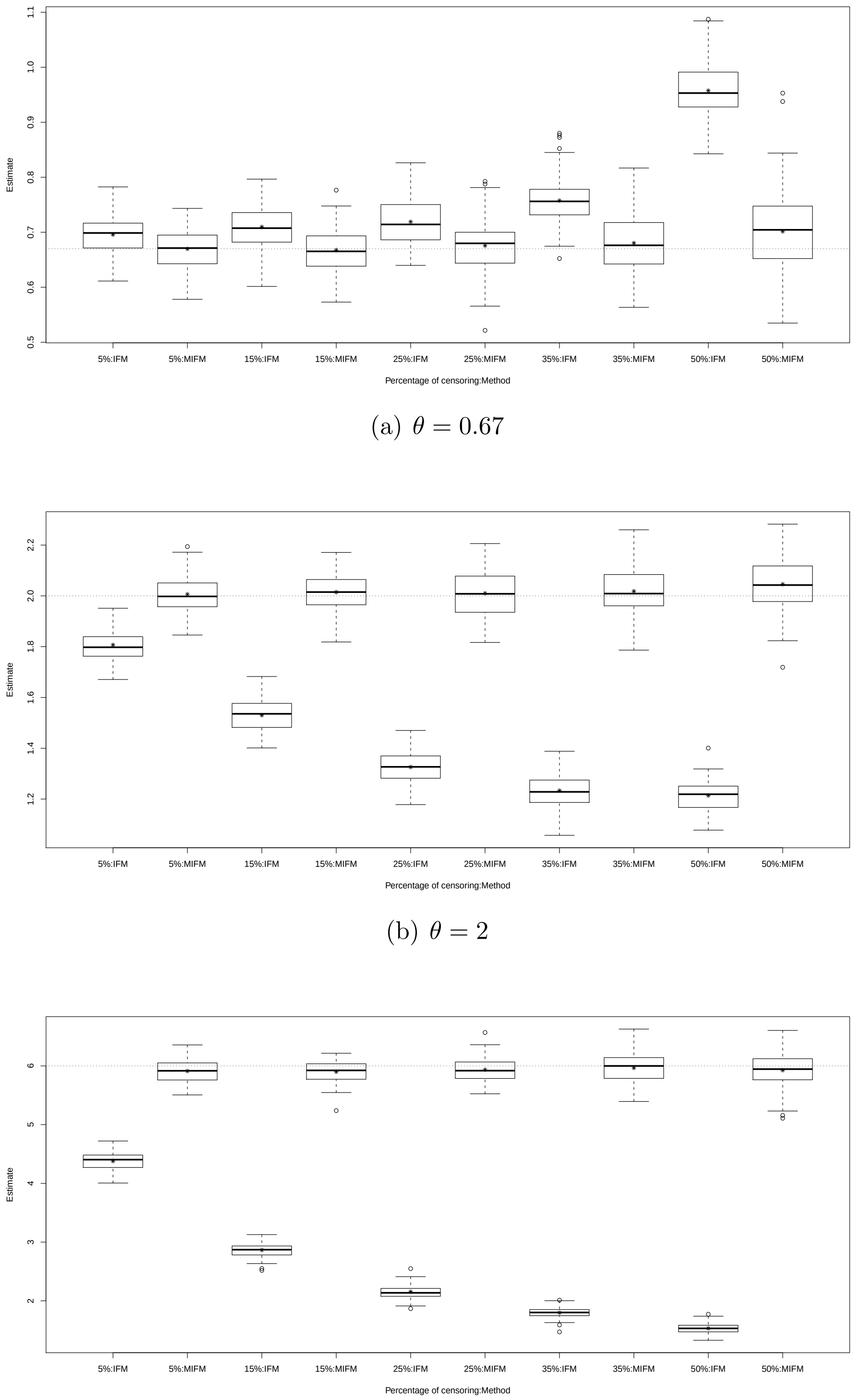


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.7: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.

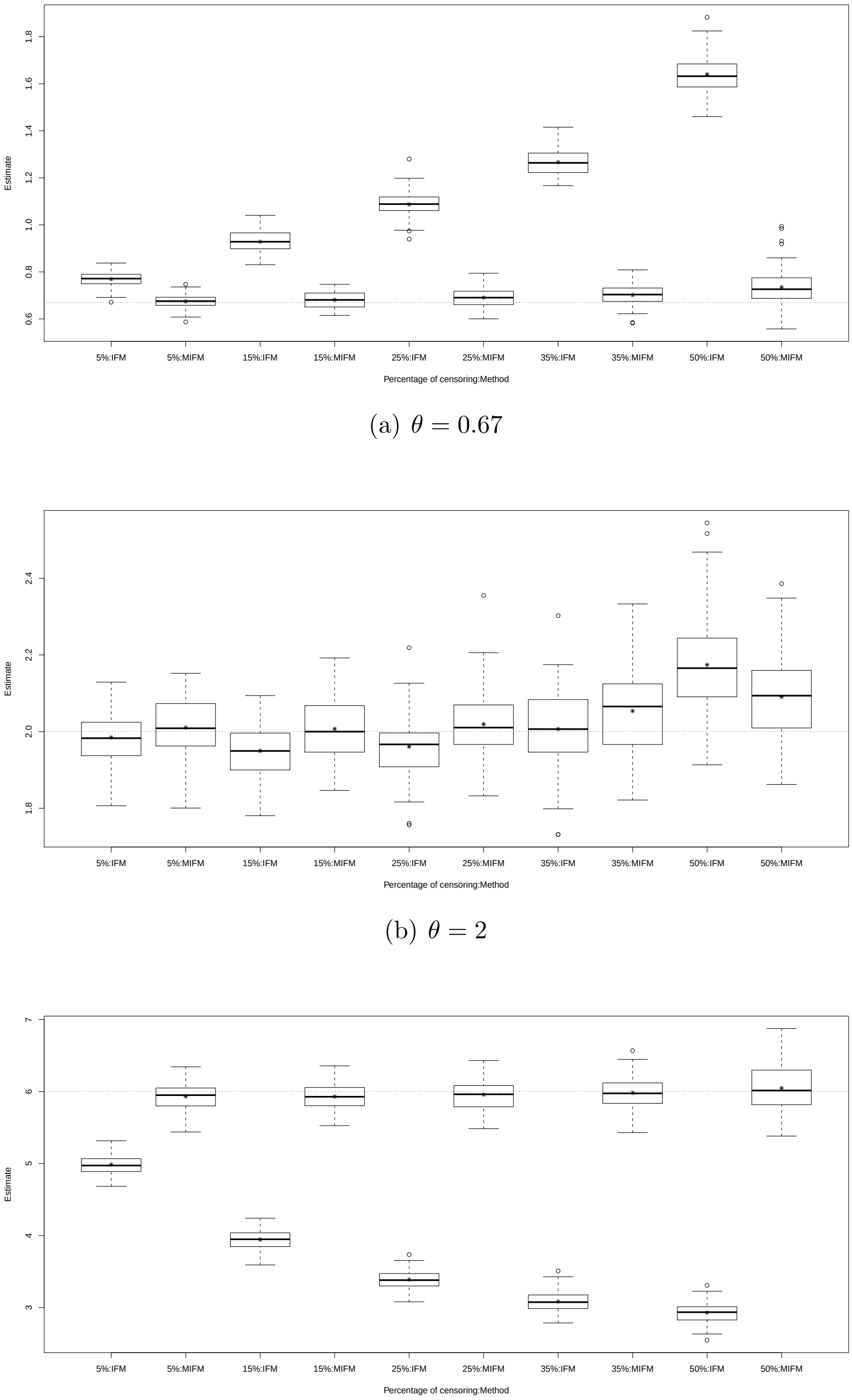


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.8: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (power-normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.



(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.9: Comparison between the IFM and MIFM estimates of the Clayton copula parameter, for $n = 2000$ (logistic marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton copula parameter.

dressing, tomato and lettuce data. From the Kolmogorov-Smirnov goodness-of-fit tests of augmented marginal residuals [4], we obtain p-values equal to 0.6599, 0.0995 and 0.8483 for the salad dressing, tomato and lettuce models, respectively. Thus, the logistic distribution assumption for the marginal errors is valid (at the 5% level). The results reported in Table 3.3 reveal that individuals aged 20-40 years consume more salad dressings, and individuals aged 20-30 years consume more lettuce than those over 60 years of age. Su & Arab (2006) found a similar effect of age on salad dressing consumption. Regional effects are also notable, as individuals from the Northeast and West (according to the 90% basic bootstrap confidence interval) consume more salad dressings, individuals from the Northeast consume more tomatoes, and individuals from the Midwest and West consume more lettuce than individuals residing in the South. The household income has a positive effect on the consumption of all these food items. The MIFM estimate of the Clayton copula parameter $\left(\hat{\theta}_{\text{MIFM}} = 1.6390, \text{obtained after 21 iterations}\right)$ and its 90% bootstrap-based confidence intervals show us that the relationship among salad dressing, tomato and lettuce consumption is positive (the estimated Kendall's tau is $\hat{\tau}_3 = \hat{\theta}_{\text{MIFM}} / \left(\hat{\theta}_{\text{MIFM}} + 2\right) = 0.4504$) and significant at the 10% level (the lower limits of the 90% bootstrap-based confidence intervals for θ are greater than and far above zero), justifying joint estimation of the censored equations through the Clayton copula to improve statistical efficiency. Furthermore, the estimated trivariate tail dependence coefficients for Clayton copula, $\hat{\lambda}_{\text{L}}^{1|23} = 0.7808$ and $\hat{\lambda}_{\text{L}}^{12|3} = 0.5116$, obtained from $(3\,/\,2)^{-1/\hat{\theta}_{\text{MIFM}}}$ and $3^{-1/\hat{\theta}_{\text{MIFM}}}$ (cf. De Luca & Rivieccio, 2012; Di Bernardino & Rullière, 2014), respectively, show the positive dependence at the lower tail of the joint distribution, i.e. for low or no consumption of salad dressings, tomatoes and lettuce.

For purposes of comparison, we also fit, via the MCECM algorithm of Huang (1999) adapted to trivariate logistic distribution, what we call here the basic trivariate SUR Tobit model with logistic marginal errors, that is the trivariate SUR Tobit model whose dependence among the marginal errors ϵ_{i1}, ϵ_{i2} and ϵ_{i3}, $i = 1, ..., n$, is modeled through the classical trivariate logistic distribution as proposed by Malik & Abraham (1973). The estimation results, obtained after 3 iterations (i.e. in much fewer iterations than required by the (extended) MIFM method, but the adapted MCECM algorithm is much more

[4] The augmented residuals are the differences between the augmented observed and predicted responses, i.e. $e_{ij}^{\text{a}} = y_{ij}^{\text{a}} - \boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}$, for $i = 1, ..., n$ and $j = 1, 2, 3$, where $y_{ij}^{\text{a}} = \boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}} + \hat{s}_{j,\text{MIFM}} G^{-1}\left(u_{ij}^{\text{a}}\right)$, with $G^{-1}(.)$ being the inverse function of the $L(0,1)$ c.d.f.; or simply, $e_{ij}^{\text{a}} = \hat{s}_{j,\text{MIFM}} G^{-1}\left(u_{ij}^{\text{a}}\right)$.

Table 3.1: Estimation results of trivariate Clayton copula-based SUR Tobit model with normal marginal errors for salad dressing, tomato and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals		
Salad dressing	Estimate	Standard Normal	Percentile	Basic
Intercept	0.1130	[0.0336; 0.1924]	[0.0422; 0.1969]	[0.0291; 0.1837]
Age 20-30	0.1106	[0.0353; 0.1860]	[0.0363; 0.1853]	[0.0360; 0.1850]
Age 31-40	0.1011	[0.0337; 0.1685]	[0.0295; 0.1650]	[0.0372; 0.1726]
Age 41-50	0.0633	[-0.0009; 0.1276]	[0.0014; 0.1266]	[0.00003; 0.12526]
Age 51-60	-0.0030	[-0.0729; 0.0669]	[-0.0745; 0.0665]	[-0.0725; 0.0685]
Northeast	0.0784	[0.0152; 0.1417]	[0.0115; 0.1383]	[0.0185; 0.1453]
Midwest	0.0521	[-0.0082; 0.1123]	[-0.0063; 0.1137]	[-0.0095; 0.1105]
West	0.0544	[-0.0027; 0.1114]	[-0.0051; 0.1100]	[-0.0013; 0.1138]
Income	0.0277	[0.0022; 0.0531]	[0.0035; 0.0504]	[0.0049; 0.0518]
σ_1	0.2636	[0.2461; 0.2812]	[0.2445; 0.2797]	[0.2476; 0.2828]
		90% Confidence Intervals		
Tomato	Estimate	Standard Normal	Percentile	Basic
Intercept	-0.0554	[-0.1183; 0.0076]	[-0.1184; 0.0070]	[-0.1177; 0.0077]
Age 20-30	0.0292	[-0.0311; 0.0895]	[-0.0305; 0.0903]	[-0.0318; 0.0890]
Age 31-40	0.0404	[-0.0171; 0.0980]	[-0.0203; 0.0914]	[-0.0105; 0.1012]
Age 41-50	0.0369	[-0.0166; 0.0905]	[-0.0138; 0.0907]	[-0.0168; 0.0877]
Age 51-60	-0.0351	[-0.0910; 0.0208]	[-0.0925; 0.0157]	[-0.0860; 0.0223]
Northeast	0.1000	[0.0457; 0.1543]	[0.0465; 0.1521]	[0.0479; 0.1535]
Midwest	0.0129	[-0.0373; 0.0632]	[-0.0413; 0.0591]	[-0.0332; 0.0671]
West	0.0177	[-0.0302; 0.0655]	[-0.0336; 0.0648]	[-0.0295; 0.0690]
Income	0.0295	[0.0092; 0.0498]	[0.0096; 0.0509]	[0.0082; 0.0494]
σ_2	0.2088	[0.1913; 0.2263]	[0.1887; 0.2241]	[0.1934; 0.2288]
		90% Confidence Intervals		
Lettuce	Estimate	Standard Normal	Percentile	Basic
Intercept	-0.1084	[-0.2056; -0.0112]	[-0.2052; -0.0081]	[-0.2088; -0.0117]
Age 20-30	0.1051	[0.0186; 0.1916]	[0.0243; 0.1898]	[0.0204; 0.1858]
Age 31-40	0.0786	[-0.0065; 0.1636]	[-0.0107; 0.1609]	[-0.0038; 0.1679]
Age 41-50	0.0908	[0.0126; 0.1690]	[0.0075; 0.1666]	[0.0149; 0.1741]
Age 51-60	0.0232	[-0.0561; 0.1024]	[-0.0574; 0.0989]	[-0.0526; 0.1037]
Northeast	0.0588	[-0.0194; 0.1370]	[-0.0178; 0.1359]	[-0.0183; 0.1354]
Midwest	0.1065	[0.0342; 0.1788]	[0.0325; 0.1771]	[0.0360; 0.1805]
West	0.0946	[0.0235; 0.1657]	[0.0209; 0.1622]	[0.0270; 0.1684]
Income	0.0604	[0.0288; 0.0919]	[0.0280; 0.0903]	[0.0304; 0.0928]
σ_3	0.3101	[0.2862; 0.3341]	[0.2826; 0.3295]	[0.2908; 0.3376]
θ	1.7323	[1.4244; 2.0401]	[1.4503; 2.0734]	[1.3911; 2.0143]
Log-likelihood	-150.8005			
AIC	363.6011			
BIC	487.3365			

Table 3.2: Estimation results of trivariate Clayton copula-based SUR Tobit model with power-normal marginal errors for salad dressing, tomato and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals		
Salad dressing	Estimate	Standard Normal	Percentile	Basic
Intercept	-1.6897	[-1.8411; -1.5384]	[-1.8082; -1.5142]	[-1.8653; -1.5713]
Age 20-30	0.0839	[0.0091; 0.1588]	[0.0250; 0.1734]	[-0.0056; 0.1428]
Age 31-40	0.0580	[-0.0122; 0.1282]	[0.0079; 0.1470]	[-0.0310; 0.1081]
Age 41-50	0.0553	[-0.0079; 0.1185]	[0.0093; 0.1278]	[-0.0173; 0.1013]
Age 51-60	0.0189	[-0.0498; 0.0877]	[-0.0373; 0.1020]	[-0.0641; 0.0751]
Northeast	0.0479	[-0.0157; 0.1116]	[-0.0060; 0.1235]	[-0.0276; 0.1018]
Midwest	0.0347	[-0.0252; 0.0946]	[-0.0182; 0.1030]	[-0.0336; 0.0876]
West	0.0446	[-0.0128; 0.1020]	[-0.0022; 0.1113]	[-0.0222; 0.0913]
Income	0.0218	[-0.0014; 0.0450]	[-0.0045; 0.0430]	[0.0005; 0.0481]
σ_1	0.6384	[0.5873; 0.6895]	[0.5739; 0.6746]	[0.6021; 0.7029]
α_1	302.8540	[292.3302; 313.3779]	[293.0191; 311.9069]	[293.8011; 312.6889]
		90% Confidence Intervals		
Tomato	Estimate	Standard Normal	Percentile	Basic
Intercept	-1.4305	[-1.5655; -1.2956]	[-1.5527; -1.2840]	[-1.5770; -1.3084]
Age 20-30	0.0212	[-0.0223; 0.0646]	[-0.0211; 0.0657]	[-0.0234; 0.0634]
Age 31-40	0.0332	[-0.0125; 0.0788]	[-0.0127; 0.0792]	[-0.0128; 0.0791]
Age 41-50	0.0296	[-0.0114; 0.0707]	[-0.0103; 0.0721]	[-0.0128; 0.0696]
Age 51-60	-0.0240	[-0.0696; 0.0215]	[-0.0694; 0.0239]	[-0.0720; 0.0213]
Northeast	0.0586	[0.0187; 0.0985]	[0.0232; 0.0992]	[0.0179; 0.0940]
Midwest	0.0105	[-0.0306; 0.0516]	[-0.0270; 0.0509]	[-0.0299; 0.0480]
West	0.0161	[-0.0232; 0.0554]	[-0.0206; 0.0571]	[-0.0248; 0.0529]
Income	0.0258	[0.0086; 0.0430]	[0.0073; 0.0426]	[0.0091; 0.0444]
σ_2	0.4631	[0.4237; 0.5024]	[0.4210; 0.4984]	[0.4278; 0.5052]
α_2	533.6174	[529.4562; 537.7787]	[531.3214; 537.8499]	[529.3850; 535.9135]
		90% Confidence Intervals		
Lettuce	Estimate	Standard Normal	Percentile	Basic
Intercept	-2.2898	[-2.5958; -1.9838]	[-2.4363; -1.7990]	[-2.7806; -2.1433]
Age 20-30	0.0397	[-0.0628; 0.1422]	[-0.0390; 0.1623]	[-0.0830; 0.1184]
Age 31-40	0.0384	[-0.0558; 0.1326]	[-0.0375; 0.1512]	[-0.0743; 0.1144]
Age 41-50	0.0629	[-0.0210; 0.1467]	[0.0004; 0.1616]	[-0.0359; 0.1254]
Age 51-60	-0.0122	[-0.1053; 0.0809]	[-0.0768; 0.1146]	[-0.1390; 0.0523]
Northeast	0.0883	[-0.0089; 0.1855]	[0.0168; 0.2132]	[-0.0367; 0.1597]
Midwest	0.1137	[0.0193; 0.2081]	[0.0427; 0.2404]	[-0.0129; 0.1848]
West	0.1063	[0.0155; 0.1972]	[0.0458; 0.2265]	[-0.0138; 0.1668]
Income	0.0561	[0.0204; 0.0917]	[0.0091; 0.0795]	[0.0326; 0.1031]
σ_3	0.7446	[0.6481; 0.8411]	[0.5857; 0.7819]	[0.7073; 0.9036]
α_3	422.7770	[403.3985; 442.1556]	[399.3004; 436.4070]	[409.1470; 446.2536]
θ	1.5470	[1.2514; 1.8426]	[1.1511; 1.7351]	[1.3589; 1.9429]
Log-likelihood	-152.0613			
AIC	372.1227			
BIC	507.8325			

Table 3.3: Estimation results of trivariate Clayton copula-based SUR Tobit model with logistic marginal errors for salad dressing, tomato and lettuce consumption in the U.S. in 1994-1996.

		90% Confidence Intervals		
Salad dressing	Estimate	Standard Normal	Percentile	Basic
Intercept	0.1124	[0.0375; 0.1873]	[0.0391; 0.1910]	[0.0338; 0.1857]
Age 20-30	0.0968	[0.0294; 0.1642]	[0.0297; 0.1592]	[0.0344; 0.1639]
Age 31-40	0.0977	[0.0326; 0.1627]	[0.0304; 0.1605]	[0.0348; 0.1649]
Age 41-50	0.0480	[-0.0147; 0.1107]	[-0.0142; 0.1076]	[-0.0116; 0.1102]
Age 51-60	0.0024	[-0.0597; 0.0644]	[-0.0608; 0.0627]	[-0.0579; 0.0655]
Northeast	0.0744	[0.0136; 0.1353]	[0.0122; 0.1299]	[0.0190; 0.1367]
Midwest	0.0559	[-0.0004; 0.1123]	[-0.0024; 0.1122]	[-0.0003; 0.1143]
West	0.0570	[-0.0010; 0.1150]	[-0.0048; 0.1115]	[0.0025; 0.1188]
Income	0.0275	[0.0039; 0.0510]	[0.0031; 0.0530]	[0.0019; 0.0518]
s_1	0.1459	[0.1352; 0.1566]	[0.1331; 0.1543]	[0.1375; 0.1588]
		90% Confidence Intervals		
Tomato	Estimate	Standard Normal	Percentile	Basic
Intercept	-0.0358	[-0.0873; 0.0158]	[-0.0910; 0.0137]	[-0.0852; 0.0195]
Age 20-30	0.0207	[-0.0287; 0.0700]	[-0.0272; 0.0685]	[-0.0271; 0.0685]
Age 31-40	0.0348	[-0.0121; 0.0817]	[-0.0145; 0.0820]	[-0.0123; 0.0841]
Age 41-50	0.0201	[-0.0276; 0.0677]	[-0.0294; 0.0679]	[-0.0278; 0.0695]
Age 51-60	-0.0251	[-0.0719; 0.0216]	[-0.0696; 0.0197]	[-0.0700; 0.0194]
Northeast	0.0677	[0.0221; 0.1132]	[0.0224; 0.1131]	[0.0222; 0.1129]
Midwest	0.0111	[-0.0312; 0.0535]	[-0.0323; 0.0551]	[-0.0329; 0.0546]
West	0.0191	[-0.0237; 0.0619]	[-0.0247; 0.0640]	[-0.0258; 0.0629]
Income	0.0249	[0.0077; 0.0421]	[0.0090; 0.0426]	[0.0073; 0.0409]
s_2	0.1069	[0.0969; 0.1168]	[0.0953; 0.1156]	[0.0981; 0.1185]
		90% Confidence Intervals		
Lettuce	Estimate	Standard Normal	Percentile	Basic
Intercept	-0.0837	[-0.1717; 0.0043]	[-0.1754; 0.0027]	[-0.1700; 0.0081]
Age 20-30	0.0804	[0.0029; 0.1579]	[0.0010; 0.1526]	[0.0082; 0.1598]
Age 31-40	0.0718	[-0.0078; 0.1514]	[-0.0120; 0.1483]	[-0.0047; 0.1556]
Age 41-50	0.0721	[-0.0057; 0.1499]	[-0.0082; 0.1521]	[-0.0079; 0.1523]
Age 51-60	0.0133	[-0.0629; 0.0895]	[-0.0615; 0.0878]	[-0.0611; 0.0881]
Northeast	0.0662	[-0.0096; 0.1420]	[-0.0148; 0.1350]	[-0.0026; 0.1472]
Midwest	0.0936	[0.0231; 0.1641]	[0.0221; 0.1694]	[0.0178; 0.1651]
West	0.0850	[0.0131; 0.1569]	[0.0096; 0.1526]	[0.0174; 0.1605]
Income	0.0559	[0.0268; 0.0850]	[0.0281; 0.0829]	[0.0289; 0.0837]
s_3	0.1743	[0.1599; 0.1886]	[0.1582; 0.1876]	[0.1609; 0.1903]
θ	1.6390	[1.3643; 1.9137]	[1.3985; 1.9346]	[1.3435; 1.8795]
Log-likelihood	-129.0396			
AIC	320.0792			
BIC	443.8146			

time consuming), are presented in Table 3.4. The standard errors were derived from the bootstrap-based covariance matrix estimate given by (3.5) (bootstrap standard errors) [5]. It can be seen from Tables 3.3 and 3.4 that the marginal parameter estimates obtained through the adapted MCECM and (extended) MIFM methods are similar. However, the trivariate Clayton copula-based SUR Tobit model with logistic marginal errors overcomes the basic trivariate SUR Tobit model with logistic marginal errors in both AIC and BIC criterion. This indicates that there was a gain by introducing the Clayton copula to model the nonlinear dependence structure of the trivariate SUR Tobit model with logistic marginal errors, for this dataset.

3.2 Trivariate Clayton survival copula-based SUR Tobit right-censored model formulation

The SUR Tobit model with three right-censored dependent variables, or simply trivariate SUR Tobit right-censored model, is expressed as

$$y_{ij}^* = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j + \epsilon_{ij},$$

$$y_{ij} = \begin{cases} y_{ij}^* & \text{if } y_{ij}^* < d_j, \\ d_j & \text{otherwise,} \end{cases}$$

for $i = 1, ..., n$ and $j = 1, 2, 3$, where n is the number of observations, d_j is the censoring point/threshold of margin j (which is assumed to be known and constant, here), y_{ij}^* is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^* whenever y_{ij}^* is below d_j and d_j otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients and ϵ_{ij} is the margin j's error that follows some zero mean distribution.

As in the previous chapter, we suppose that the marginal errors are no longer normal, but they are assumed to be distributed according to the power-normal (Gupta & Gupta, 2008) and logistic models. Then, the density function of y_{ij} takes the forms given by (2.9), (2.11) and (2.13), respectively; and the distribution function of y_{ij} is obtained by (2.10), (2.12) and (2.14), respectively.

[5] But now with $\boldsymbol{\eta}$ denoting the parameter vector of the basic trivariate SUR Tobit model with logistic marginal errors.

Table 3.4: Estimation results of basic trivariate SUR Tobit model with logistic marginal errors for salad dressing, tomato and lettuce consumption in the U.S. in 1994-1996.

Salad dressing	Estimate	Standard Error
Intercept	0.1304 *	0.0401
Age 20-30	0.0838 *	0.0370
Age 31-40	0.0812 *	0.0362
Age 41-50	0.0504	0.0329
Age 51-60	-0.0043	0.0337
Northeast	0.0639 *	0.0312
Midwest	0.0572 *	0.0303
West	0.0535 *	0.0316
Income	0.0294 *	0.0130
s_1	0.1388 *	0.0058
Tomato	Estimate	Standard Error
Intercept	-0.0269	0.0300
Age 20-30	0.0106	0.0291
Age 31-40	0.0337	0.0276
Age 41-50	0.0222	0.0252
Age 51-60	-0.0223	0.0275
Northeast	0.0589 *	0.0258
Midwest	0.0165	0.0234
West	0.0210	0.0250
Income	0.0226 *	0.0095
s_2	0.1030 *	0.0053
Lettuce	Estimate	Standard Error
Intercept	-0.0481	0.0481
Age 20-30	0.0777 *	0.0431
Age 31-40	0.0647	0.0409
Age 41-50	0.0654 *	0.0387
Age 51-60	0.0004	0.0375
Northeast	0.0561	0.0399
Midwest	0.0909 *	0.0368
West	0.0707 *	0.0361
Income	0.0529 *	0.0156
s_3	0.1646 *	0.0081
Log-likelihood	-136.3096	
AIC	332.6192	
BIC	452.3632	

* Denotes significant at the 10% level.

The dependence among the error terms ϵ_{i1}, ϵ_{i2} and ϵ_{i3} is modeled in the usual way through a trivariate distribution, especially the trivariate normal distribution (this specification characterizes the basic trivariate SUR Tobit right-censored model). Nevertheless, applying a trivariate distribution to the trivariate SUR Tobit right-censored model is limited to the linear relationship among marginal distributions through the correlation coefficients. Furthermore, estimation methods for high-dimensional SUR Tobit right-censored models are often computationally demanding and difficult to implement (see comments in Section 1.2). To overcome these problems, we can use copula functions to model the nonlinear dependence structure in the trivariate SUR Tobit right-censored model.

Thus, for the censored outcomes y_{i1}, y_{i2} and y_{i3}, the trivariate copula-based SUR Tobit right-censored distribution is given by

$$F\left(y_{i1}, y_{i2}, y_{i3}\right) = C\left(u_{i1}, u_{i2}, u_{i3} | \theta\right),$$

where, e.g., u_{ij} is given by (2.10) if $\epsilon_{ij} \sim N\left(0, \sigma_j^2\right)$, (2.12) if $\epsilon_{ij} \sim PN\left(0, \sigma_j, \alpha_j\right)$, or (2.14) if $\epsilon_{ij} \sim L\left(0, s_j\right)$, for $j = 1, 2, 3$ (see Section 2.2); and θ is the copula association parameter (or parameter vector), which is assumed to be scalar.

Let us suppose that C is the tridimensional Clayton survival copula, which, according to Joe (2014, p. 28), takes the form of

$$\begin{aligned} C\left(u_{i1}, u_{i2}, u_{i3} | \theta\right) = u_{i1} + u_{i2} + u_{i3} - 2 + \left[\left(1 - u_{i1}\right)^{-\theta} + \left(1 - u_{i2}\right)^{-\theta} - 1\right]^{-\frac{1}{\theta}} + \\ + \left[\left(1 - u_{i1}\right)^{-\theta} + \left(1 - u_{i3}\right)^{-\theta} - 1\right]^{-\frac{1}{\theta}} + \\ + \left[\left(1 - u_{i2}\right)^{-\theta} + \left(1 - u_{i3}\right)^{-\theta} - 1\right]^{-\frac{1}{\theta}} - \\ - \left[\left(1 - u_{i1}\right)^{-\theta} + \left(1 - u_{i2}\right)^{-\theta} + \left(1 - u_{i3}\right)^{-\theta} - 2\right]^{-\frac{1}{\theta}}, \end{aligned} \tag{3.6}$$

with θ restricted to the region $(0, \infty)$. The dependence among the margins increases with the value of θ, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect positive dependence. This copula shows upper tail dependence and is characterized by zero lower tail dependence.

3.2.1 Inference

In this subsection, we discuss inference (point and interval estimation) for the parameters of the trivariate Clayton survival copula-based SUR Tobit right-censored model. Particularly, by considering/assuming normal, power-normal and logistic distributions for the marginal error terms.

3.2.1.1 Estimation through the (extended) MIFM method

Following Trivedi & Zimmer (2005) and Anastasopoulos *et al.* (2012), we can write the log-likelihood function for the trivariate Clayton survival copula-based SUR Tobit right-censored model in the form

$$\begin{aligned}\ell\left(\boldsymbol{\eta}\right) = &\sum_{i=1}^{n} \log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1},\boldsymbol{\upsilon}_1\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2},\boldsymbol{\upsilon}_2\right), F_3\left(y_{i3}|\boldsymbol{x}_{i3},\boldsymbol{\upsilon}_3\right)|\theta\right)+ \\ &+\sum_{i=1}^{n}\sum_{j=1}^{3} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right),\end{aligned} \tag{3.7}$$

where $\boldsymbol{\eta} = (\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2, \boldsymbol{\upsilon}_3, \theta)$ is the vector of model parameters, $\boldsymbol{\upsilon}_j$ is the margin j's parameter vector, $f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, $F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$ is the c.d.f. of y_{ij}, and $c\left(u_{i1}, u_{i2}, u_{i3}|\theta\right)$, with $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton survival copula, which is calculated from (3.6) as

$$\begin{aligned}c\left(u_{i1}, u_{i2}, u_{i3}|\theta\right) &= \frac{\partial^3 C\left(u_{i1}, u_{i2}, u_{i3}|\theta\right)}{\partial u_{i1}\partial u_{i2}\partial u_{i3}} = \\ &= (\theta+1)(2\theta+1)\left[(1-u_{i1})(1-u_{i2})(1-u_{i3})\right]^{-\theta-1}\times \\ &\quad\times\left[(1-u_{i1})^{-\theta}+(1-u_{i2})^{-\theta}+(1-u_{i3})^{-\theta}-2\right]^{-\frac{1}{\theta}-3}.\end{aligned}$$

Using copula methods, as well as the log-likelihood function form given by (3.7), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}} = \arg\max_{\boldsymbol{\upsilon}_j} \sum_{i=1}^{n} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij},\boldsymbol{\upsilon}_j\right), \tag{3.8}$$

for $j = 1, 2, 3$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$ by

$$\widehat{\theta}_{\mathrm{IFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\big(F_1\left(y_{i1}|\boldsymbol{x}_{i1},\widehat{\boldsymbol{\upsilon}}_{1,\mathrm{IFM}}\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2},\widehat{\boldsymbol{\upsilon}}_{2,\mathrm{IFM}}\right), F_3\left(y_{i3}|\boldsymbol{x}_{i3},\widehat{\boldsymbol{\upsilon}}_{3,\mathrm{IFM}}\right)|\theta\big). \tag{3.9}$$

However, the IFM method provides a biased estimate for the parameter θ in the presence of censored observations in the margins (as will be seen in Section 3.2.2.2), which occurs because there is a violation of Sklar's theorem in this case (see discussion in Section 3.1.1.1). In order to obtain an unbiased estimate for the association parameter θ, we can augment the semi-continuous/censored marginal distributions to achieve continuity. More specifically, we replace y_{ij} by the augmented data y_{ij}^{a}, or equivalently and more simply (thus, preferred by us), we can replace u_{ij} by the augmented uniform data u_{ij}^{a} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the cases of continuous margins. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence is achieved. The (frequentist) data augmentation technique we use here is partially based on Algorithm A2 presented in Wichitaksorn *et al.* (2012).

In the remaining part of this subsubsection, we discuss the proposed estimation method (an extension of the MIFM method proposed in Section 2.2.1.1 for the trivariate case) when using the Clayton survival copula to describe the nonlinear dependence structure of the trivariate SUR Tobit right-censored model with arbitrary margins (e.g., normal, power-normal and logistic distribution assumption for the marginal error terms). Nevertheless, the proposed approach can be extended to other copula functions by applying different sampling algorithms. For the cases where just a single dependent variable/margin is censored (i.e. when $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$, or $y_{i1} < d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$, or $y_{i1} < d_1$ and $y_{i2} < d_2$ and $y_{i3} = d_3$), the uniform data augmentation is performed through the (univariate) truncated conditional distribution of the Clayton survival copula. For the cases where two of the dependent variables/margins are censored (i.e. when $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$, or $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} = d_3$, or $y_{i1} < d_1$ and $y_{i2} = d_2$ and $y_{i3} = d_3$), the uniform data augmentation is performed through the (bivariate) truncated conditional distribution of the Clayton survival copula, e.g., by iterative (i.e. successive) conditioning. If the inverse conditional distribution of the copula used has a closed-form expression, which is the case of the Clayton survival copula (see Appendix A), we can generate random numbers from its truncated version by applying the method by Devroye (1986, p. 38-39). Otherwise, numerical root-finding procedures are required. As the (tridimensional) Clayton survival copula, as well as the (tridimensional) Clayton copula has the truncation dependence invariance property, the conditional distribution of u_{i1}, u_{i2} and u_{i3} in a sub-region of a Clayton survival copula, with one corner at $(1, 1, 1)$, can be written by means of a Clayton survival copula. That formulation enables a simple simulation scheme in the cases where all dependent variables/margins are censored (i.e. when $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} = d_3$). For copulas that do not have the truncation-invariance property, an iterative simulation scheme can be used.

The implementation of the trivariate Clayton survival copula-based SUR Tobit right-censored model with arbitrary margins via the proposed (extended) MIFM method can be described as follows. In particular, if the marginal error distributions are normal, then set $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)$; if marginal error distributions are power-normal, so $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = \left[\Phi\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / \sigma_j\right)\right]^{\alpha_j}$; and if marginal error distributions are logistic, then $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, s_j\right)$ and $H_j\left(z|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right) = G\left(\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right) = \left[1 + \exp\left\{-\left(z - \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j\right) / s_j\right\}\right]^{-1}$, for $j = 1, 2, 3$ and $z \in \mathbb{R}$.

Stage 1. Estimate the marginal parameters using (3.8). Set $\hat{\boldsymbol{\upsilon}}_{j,\text{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\text{IFM}}$, for $j = 1, 2, 3$.

Stage 2. Estimate the copula parameter using, e.g., (3.9). Set $\hat{\theta}_{\text{MIFM}}^{(1)} = \hat{\theta}_{\text{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} = d_3$, then draw $(u_{i1}^{\text{a}}, u_{i2}^{\text{a}}, u_{i3}^{\text{a}})$ from $C\left(u_{i1}^{\text{a}}, u_{i2}^{\text{a}}, u_{i3}^{\text{a}}|\hat{\theta}_{\text{MIFM}}^{(\omega)}\right)$ truncated to the region $(a_{i1}, 1) \times (a_{i2}, 1) \times (a_{i3}, 1)$. This can be performed relatively easily using the following steps.

1. Draw (p, q, r) from $C\left(p, q, r|\hat{\theta}_{\text{MIFM}}^{(\omega)}\right) = p+q+r-2+\left[(1-p)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+(1-q)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-1\right]^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}+\left[(1-p)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+(1-r)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-1\right]^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}+\left[(1-q)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+(1-r)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-1\right]^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}-\left[(1-p)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+(1-q)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+(1-r)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-2\right]^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}$. See Appendix A for the multidimensional Clayton survival copula data generation (conditional sampling).

2. Compute $a_{ij} = H_j\left(d_j|\boldsymbol{x}_{ij}, \hat{\boldsymbol{\upsilon}}_{j,\text{MIFM}}\right)$, for $j = 1, 2, 3$.

3. Set $u_{i1}^{\text{a}} = 1-\left\{\left[(1-a_{i1})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} + (1-a_{i2})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} + (1-a_{i3})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} - 2\right](1-p)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+2-(1-a_{i2})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-(1-a_{i3})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}\right\}^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}$.

4. Set $u_{i2}^{\text{a}} = 1-\left\{\left[(1-a_{i1})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} + (1-a_{i2})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} + (1-a_{i3})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}} - 2\right](1-q)^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}+2-(1-a_{i1})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}-(1-a_{i3})^{-\hat{\theta}_{\text{MIFM}}^{(\omega)}}\right\}^{-1/\hat{\theta}_{\text{MIFM}}^{(\omega)}}$.

5. Set $u^{\mathrm{a}}_{i3} = 1 - \left\{ \left[(1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-a_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-a_{i3})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2 \right] (1-r)^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + 2 - (1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - (1-a_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} \right\}^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

If $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$, then draw u^{a}_{i1} from $C\left(u^{\mathrm{a}}_{i1} | u_{i2}, u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i1}, 1)$. This can be done according to the following steps.

1. Compute $u_{ij} = H_j\left(y_{ij} | \boldsymbol{x}_{ij}, \hat{\boldsymbol{v}}_{j,\mathrm{MIFM}}\right)$, for $j = 2, 3$.

2. Compute $a_{i1} = H_1\left(d_1 | \boldsymbol{x}_{i1}, \hat{\boldsymbol{v}}_{1,\mathrm{MIFM}}\right)$.

3. Draw t from $Uniform\,(0,1)$.

4. Compute $v_{i1} = t + (1-t)\left\{ 1 - \left[\frac{(1-a_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-u_{i3})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 2}{(1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-u_{i3})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1} \right]^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} - 2} \right\}$.

5. Set $u^{\mathrm{a}}_{i1} = 1 - \left\{ (1-v_{i1})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} / \left(2\hat{\theta}^{(\omega)}_{\mathrm{MIFM}} + 1\right)} \left[(1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} + (1-u_{i3})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - 1 \right] + 2 - (1-u_{i2})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} - (1-u_{i3})^{-\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}} \right\}^{-1/\hat{\theta}^{(\omega)}_{\mathrm{MIFM}}}$.

If $y_{i1} < d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$, then draw u^{a}_{i2} from $C\left(u^{\mathrm{a}}_{i2} | u_{i1}, u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i2}, 1)$. This can be done by following the five steps of the previous case (i.e. $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$) by switching subscripts 1 and 2.

If $y_{i1} < d_1$ and $y_{i2} < d_2$ and $y_{i3} = d_3$, then draw u^{a}_{i3} from $C\left(u^{\mathrm{a}}_{i3} | u_{i1}, u_{i2}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i3}, 1)$. This can be done through the five steps of the penultimate case (i.e. $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$) by switching subscripts 1 and 3.

If $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2} | u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(a_{i1}, 1) \times (a_{i2}, 1)$. This can be performed relatively easily using the following steps (iterative conditioning).

1. Draw u^{a}_{i2} from $C\left(u^{\mathrm{a}}_{i2} | u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i2}, 1)$. This can be done in the same manner as in the case of just a single censored dependent variable/-margin in Section 2.2.1.1 (note that here C is the bidimensional Clayton survival copula given by (2.15)).

2. Draw u^{a}_{i1} from $C\left(u^{\mathrm{a}}_{i1} | u^{\mathrm{a}}_{i2}, u_{i3}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the interval $(a_{i1}, 1)$. This can be done according to the five steps of the second case (i.e. $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$).

If $y_{i1} = d_1$ and $y_{i2} < d_2$ and $y_{i3} = d_3$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i3})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i3} | u_{i2}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(a_{i1}, 1) \times (a_{i3}, 1)$. This can be done by following the steps of the previous case (i.e. $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$) by switching subscripts 2 and 3.

If $y_{i1} < d_1$ and $y_{i2} = d_2$ and $y_{i3} = d_3$, then draw $(u^{\mathrm{a}}_{i2}, u^{\mathrm{a}}_{i3})$ from $C\left(u^{\mathrm{a}}_{i2}, u^{\mathrm{a}}_{i3} | u_{i1}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(a_{i2}, 1) \times (a_{i3}, 1)$. This can be done by following the steps of the penultimate case (i.e. $y_{i1} = d_1$ and $y_{i2} = d_2$ and $y_{i3} < d_3$) by switching subscripts 1 and 3.

If $y_{i1} < d_1$ and $y_{i2} < d_2$ and $y_{i3} < d_3$, then set $u^{\mathrm{a}}_{ij} = u_{ij} = H_j\left(y_{ij} | \boldsymbol{x}_{ij}, \hat{\boldsymbol{v}}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, 3$.

Given the generated/augmented marginal uniform data u^{a}_{ij}, we estimate the association parameter θ by [6]

$$\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, u^{\mathrm{a}}_{i3} | \theta\right).$$

The algorithm terminates when it satisfies the stopping/convergence criterion: $|\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} - \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}| < \xi$, where ξ is the tolerance parameter (e.g., $\xi = 10^{-3}$).

3.2.1.2 Interval estimation

We propose the use of bootstrap methods (standard normal and percentile by Efron & Tibshirani (1993), and basic by Davison & Hinkley (1997)) to build confidence intervals for the parameters of the trivariate Clayton survival copula-based SUR Tobit right-censored model. It makes the analytic derivatives no longer required to compute the asymptotic covariance matrix associated with the vector of parameter estimates. For further details on our bootstrap approach, we refer to Section 3.1.1.2.

3.2.2 Simulation study

A simulation study was performed to investigate the behavior of the MIFM estimates, focusing on the copula association parameter estimate; and check the coverage probabilities of different confidence intervals (constructed using the three bootstrap methods mentioned in Section 3.2.1.2 and described in Section 3.1.1.2) for the trivariate Clayton survival copula-based SUR Tobit right-censored model parameters. Here, we considered some circumstances that might arise in the development of trivariate copula-based SUR

[6] The generated/augmented marginal uniform data u^{a}_{ij} should carry (ω) as a superscript $\left(\text{i.e. } u^{\mathrm{a}(\omega)}_{ij}\right)$, but we omit it so as not to clutter the notation.

Tobit right-censored models, involving the sample size, the censoring percentage (i.e. the percentage of d_1, d_2 and d_3 observations in margins 1, 2 and 3, respectively) in the dependent variables/margins and their interdependence degree. We also considered/assumed different distributions for the marginal error terms.

3.2.2.1 General specifications

In the simulation study, we applied the Clayton survival copula to model the nonlinear dependence structure of the trivariate SUR Tobit right-censored model. We set the true value for the association parameter θ at 0.67, 2 and 6, corresponding to a Kendall's tau association measure [7] of 0.25, 0.50 and 0.75, respectively. See Appendix A for the multidimensional Clayton survival copula data generation.

For $i = 1, ..., n$, the covariates for margin 1, $\boldsymbol{x}_{i1} = \left(x_{i1,0}, x_{i1,1}\right)^{'}$, were $x_{i1,0} = 1$ and $x_{i1,1}$ was randomly simulated from $N\left(2, 1^2\right)$. The covariates for margin 2, $\boldsymbol{x}_{i2} = \left(x_{i2,0}, x_{i2,1}\right)^{'}$, were generated as $x_{i2,0} = 1$ and $x_{i2,1}$ was randomly simulated from $N\left(1, 2^2\right)$. Finally, the covariates for margin 3, $\boldsymbol{x}_{i3} = \left(x_{i3,0}, x_{i3,1}\right)^{'}$, were generated as $x_{i3,0} = 1$ and $x_{i3,1}$ was randomly simulated from $Uniform\left(1, 3\right)$. The model errors ϵ_{i1}, ϵ_{i2} and ϵ_{i3} were assumed to follow the following distributions:

- **Normal**: i.e. $\epsilon_{i1} \sim N\left(0, \sigma_1^2\right)$, $\epsilon_{i2} \sim N\left(0, \sigma_2^2\right)$ and $\epsilon_{i3} \sim N\left(0, \sigma_3^2\right)$, where $\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 1$ are the standard deviations (scale parameters) for margins 1, 2 and 3, respectively. To ensure a percentage of censoring for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = d_3 = 5$ and assumed the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$, $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$ and $\boldsymbol{\beta}_3 = \left(\beta_{3,0}, \beta_{3,1}\right)^{'}$:

  - ■ $\boldsymbol{\beta}_1 = (0.7, 1)$, $\boldsymbol{\beta}_2 = (-0.6, 1)$ and $\boldsymbol{\beta}_3 = (-1.5, 2)$;
  - ■ $\boldsymbol{\beta}_1 = (1.5, 1)$, $\boldsymbol{\beta}_2 = (1.1, 1)$ and $\boldsymbol{\beta}_3 = (-0.7, 2)$;
  - ■ $\boldsymbol{\beta}_1 = (2, 1)$, $\boldsymbol{\beta}_2 = (2.1, 1)$ and $\boldsymbol{\beta}_3 = (-0.1, 2)$;
  - ■ $\boldsymbol{\beta}_1 = (2.5, 1)$, $\boldsymbol{\beta}_2 = (3, 1)$ and $\boldsymbol{\beta}_3 = (0.4, 2)$;
  - ■ $\boldsymbol{\beta}_1 = (3, 1)$, $\boldsymbol{\beta}_2 = (4, 1)$ and $\boldsymbol{\beta}_3 = (1, 2)$;

  respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y_{ij}^{*}, was randomly simulated from $N\left(\boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_j, \sigma_j^2\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y_{ij}^{*}, d_j\right\}$.

[7] The Kendall's tau for the tridimensional Clayton survival copula is $\tau_3 = \theta \;/\; (\theta + 2)$, which is the same for the tridimensional Clayton copula.

- **Power-normal**: i.e. $\epsilon_{i1} \sim PN\left(0, \sigma_1, \alpha_1\right)$, $\epsilon_{i2} \sim PN\left(0, \sigma_2, \alpha_2\right)$ and $\epsilon_{i3} \sim PN\left(0, \sigma_3, \alpha_3\right)$, where $\sigma_1 = 1$, $\sigma_2 = 2$ and $\sigma_3 = 1$ are the scale parameters for margins 1, 2 and 3, respectively; and $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$ are the shape parameters for margins 1, 2 and 3. To ensure a percentage of censoring for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = d_3 = 5$ and assumed the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$, $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$ and $\boldsymbol{\beta}_3 = \left(\beta_{3,0}, \beta_{3,1}\right)^{'}$:

    - $\boldsymbol{\beta}_1 = (1.1, 1)$, $\boldsymbol{\beta}_2 = (0.3, 1)$ and $\boldsymbol{\beta}_3 = (-1, 2)$;
    - $\boldsymbol{\beta}_1 = (2.1, 1)$, $\boldsymbol{\beta}_2 = (2.1, 1)$ and $\boldsymbol{\beta}_3 = (-0.1, 2)$;
    - $\boldsymbol{\beta}_1 = (2.6, 1)$, $\boldsymbol{\beta}_2 = (3.2, 1)$ and $\boldsymbol{\beta}_3 = (0.5, 2)$;
    - $\boldsymbol{\beta}_1 = (3.1, 1)$, $\boldsymbol{\beta}_2 = (4.2, 1)$ and $\boldsymbol{\beta}_3 = (1, 2)$;
    - $\boldsymbol{\beta}_1 = (3.7, 1)$, $\boldsymbol{\beta}_2 = (5.4, 1)$ and $\boldsymbol{\beta}_3 = (1.7, 2)$;

    respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y^*_{ij}, was randomly simulated from $PN\left(\boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j, \sigma_j, \alpha_j\right)$; therefore, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y^*_{ij}, d_j\right\}$.

- **Logistic**: i.e. $\epsilon_{i1} \sim L\left(0, s_1\right)$, $\epsilon_{i2} \sim L\left(0, s_2\right)$ and $\epsilon_{i3} \sim L\left(0, s_3\right)$, where $s_1 = 1$, $s_2 = 2$ and $s_3 = 1$ are the scale parameters for margins 1, 2 and 3, respectively. To ensure a percentage of censoring for all three margins of approximately 5%, 15%, 25%, 35% and 50%, we set $d_1 = d_2 = d_3 = 5$ and assumed the following true values for $\boldsymbol{\beta}_1 = \left(\beta_{1,0}, \beta_{1,1}\right)^{'}$, $\boldsymbol{\beta}_2 = \left(\beta_{2,0}, \beta_{2,1}\right)^{'}$ and $\boldsymbol{\beta}_3 = \left(\beta_{3,0}, \beta_{3,1}\right)^{'}$:

    - $\boldsymbol{\beta}_1 = (-0.3, 1)$, $\boldsymbol{\beta}_2 = (-2.5, 1)$ and $\boldsymbol{\beta}_3 = (-2.5, 2)$;
    - $\boldsymbol{\beta}_1 = (0.9, 1)$, $\boldsymbol{\beta}_2 = (-0.2, 1)$ and $\boldsymbol{\beta}_3 = (-1.2, 2)$;
    - $\boldsymbol{\beta}_1 = (1.7, 1)$, $\boldsymbol{\beta}_2 = (1.5, 1)$ and $\boldsymbol{\beta}_3 = (-0.4, 2)$;
    - $\boldsymbol{\beta}_1 = (2.3, 1)$, $\boldsymbol{\beta}_2 = (2.5, 1)$ and $\boldsymbol{\beta}_3 = (0.2, 2)$;
    - $\boldsymbol{\beta}_1 = (3, 1)$, $\boldsymbol{\beta}_2 = (4, 1)$ and $\boldsymbol{\beta}_3 = (1, 2)$;

    respectively. For $j = 1, 2, 3$, the latent dependent variable of margin j, y^*_{ij}, was randomly simulated from $L\left(\boldsymbol{x}^{'}_{ij}\boldsymbol{\beta}_j, s_j\right)$; thus, the observed dependent variable of margin j, y_{ij}, was obtained from $\min\left\{y^*_{ij}, d_j\right\}$.

For each error distribution assumption (normal, power-normal and logistic), censoring percentage in the margins (5%, 15%, 25%, 35% and 50%) and degree of dependence among them (low: $\theta = 0.67$, moderate: $\theta = 2$ and high: $\theta = 6$), we generated 100 datasets of sizes $n = 200$, 800 and 2000. Then, for each dataset (original sample), we obtained 500 bootstrap samples through a parametric resampling plan (parametric bootstrap approach), i.e. we fitted a trivariate Clayton survival copula-based SUR Tobit right-censored model with the corresponding error distributions to each dataset using the (extended) MIFM approach, and then generated a set of 500 new datasets (the same size as the original dataset/sample) from the estimated parametric model. The computing language was written in R statistical programming environment (R Core Team, 2014) and ran on a virtual machine of the Cloud-USP at ICMC, with Intel Xeon processor E5500 series, 8 core (virtual CPUs), 32 GB RAM.

We assessed the performance of the proposed models and methods through the coverage probabilities of the nominally 90% standard normal, percentile and basic bootstrap confidence intervals, the Bias and the Mean Squared Error (MSE), in which the Bias and the MSE of each parameter η_h, $h = 1, ..., k$, are given by Bias $= M^{-1} \sum_{r=1}^{M} (\hat{\eta}_h^r - \eta_h)$ and MSE $= M^{-1} \sum_{r=1}^{M} (\hat{\eta}_h^r - \eta_h)^2$, respectively, where $M = 100$ is the number of replications (original datasets/samples) and $\hat{\eta}_h^r$ is the estimated value of η_h at the rth replication.

3.2.2.2 Simulation results

In this subsubsection, we present the main results obtained from the simulation study performed with samples (datasets) of different sizes, percentages of censoring in the margins and degrees of dependence among them, regarding the trivariate Clayton survival copula-based SUR Tobit right-censored model parameters estimated using the (extended) MIFM approach. Since both the (extended) MIFM and IFM methods provide the same marginal parameter estimates (the first stage of the proposed method is similar to the first stage of the usual one, as seen in Section 3.2.1.1), we focus here on the Clayton survival copula parameter estimate. For some asymptotic results (such as asymptotic normality) associated with the IFM method, see, e.g., Joe & Xu (1996). We also show the results related to the estimated coverage probabilities of the 90% confidence intervals for θ, obtained through bootstrap methods (standard normal, percentile and basic intervals).

Figures 3.10, 3.11 and 3.12 show the Bias and MSE of the observed MIFM estimates of θ for normal, power-normal and logistic marginal errors, respectively. From these figures, we observe that, regardless of the error distribution assumption, the percentage of censoring in the margins and their interdependence degree, the Bias and MSE of the MIFM estimator of θ are relatively low and tend to zero for large n, i.e. the MIFM estimator is asymptotically unbiased and consistent for the Clayton survival copula parameter.

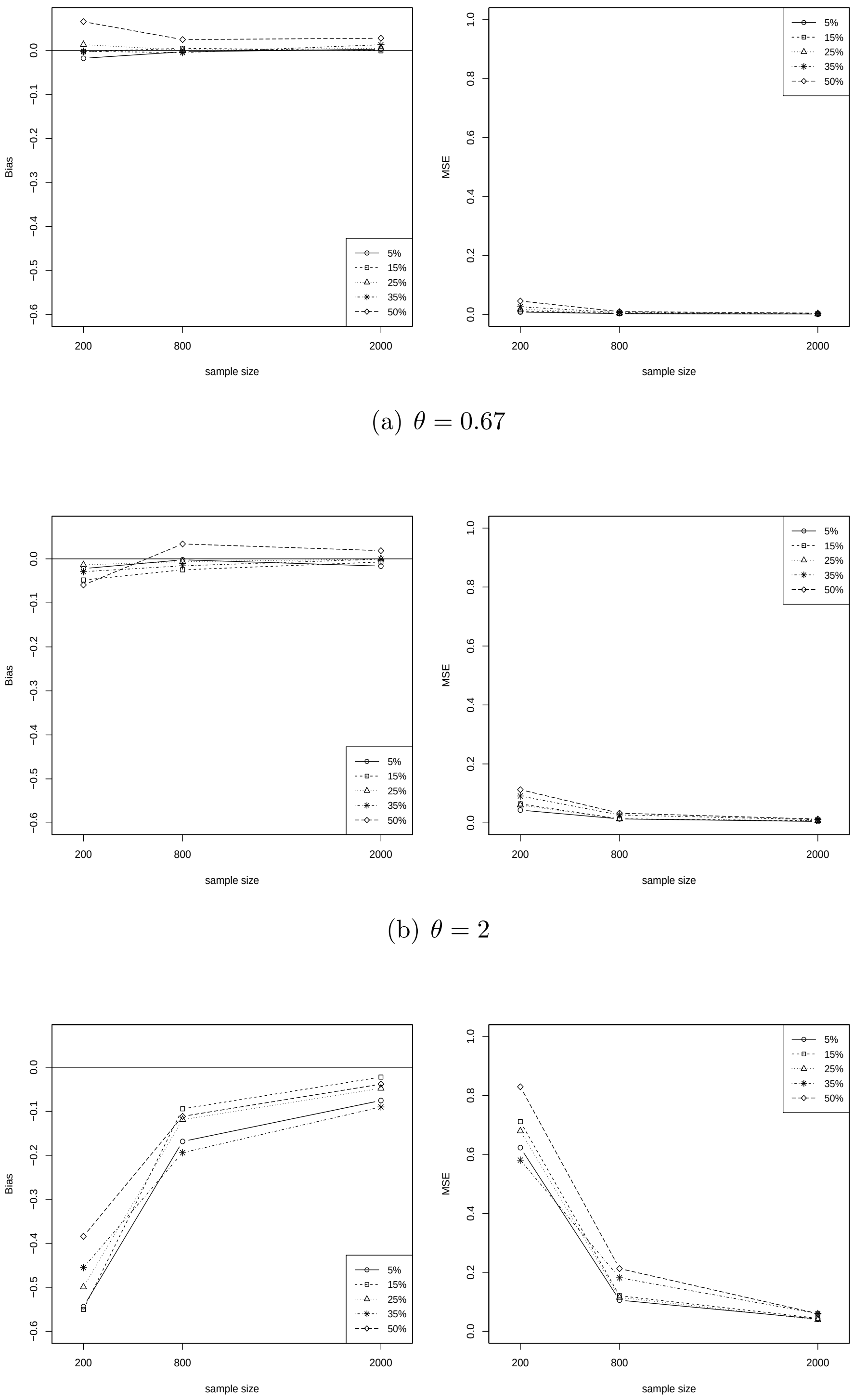
Figures 3.13, 3.14 and 3.15 show the estimated coverage probabilities of the bootstrap confidence intervals for θ for normal, power-normal and logistic marginal errors, respectively. Note that the estimated coverage probabilities are sufficiently high and close to the nominal value of 0.90, except for the percentile intervals in general, and for a few cases in which n is mainly small to moderate ($n = 200$ and 800) and the degree of dependence among the margins is high ($\theta = 6$) (see Figures 3.13(c), 3.14(c) and 3.15(c)).

Finally, Figures 3.16, 3.17 and 3.18 compare, via boxplots, the observed MIFM estimates of θ with its estimates obtained through the IFM method for normal, power-normal and logistic marginal errors, respectively, and for $n = 2000$. It can be seen from Figure 3.16 that there is a certain equivalence between the two estimation methods (with a slight advantage for the (extended) MIFM method over the IFM method, in terms of bias) when the degree of dependence among the margins is low, which is $\theta = 0.67$ (Figure 3.16(a)); however, the IFM method underestimates θ for dependence at a higher level, which is $\theta = 2$ and $\theta = 6$ (Figures 3.16(b) and 3.16(c), respectively). From Figure 3.17, we observe that the IFM method overestimates θ for dependence at a lower level, that is $\theta = 0.67$ (Figure 3.17(a)), and underestimates θ for dependence at a higher level, that is $\theta = 2$ and $\theta = 6$ (Figures 3.17(b) and 3.17(c), respectively). Similar behavior is observed for the plots in Figure 3.18. Note also from Figures 3.16, 3.17 and 3.18 that the difference (distance) between the distributions of the IFM and MIFM estimates often increases with the percentage of censoring in the margins.

3.2.3 Application

In this subsection, we illustrate the applicability of our proposed trivariate models and methods for the customer churn data described in Section 1.1.2.

In this application, the relationship among the reported log(time) to churn Product A, log(time) to churn Product B and log(time) to churn Product C (right-censored at $d_1 = d_2 = d_3 = 2.3$, or approximately 10 years) of 927 customers of a Brazilian commercial

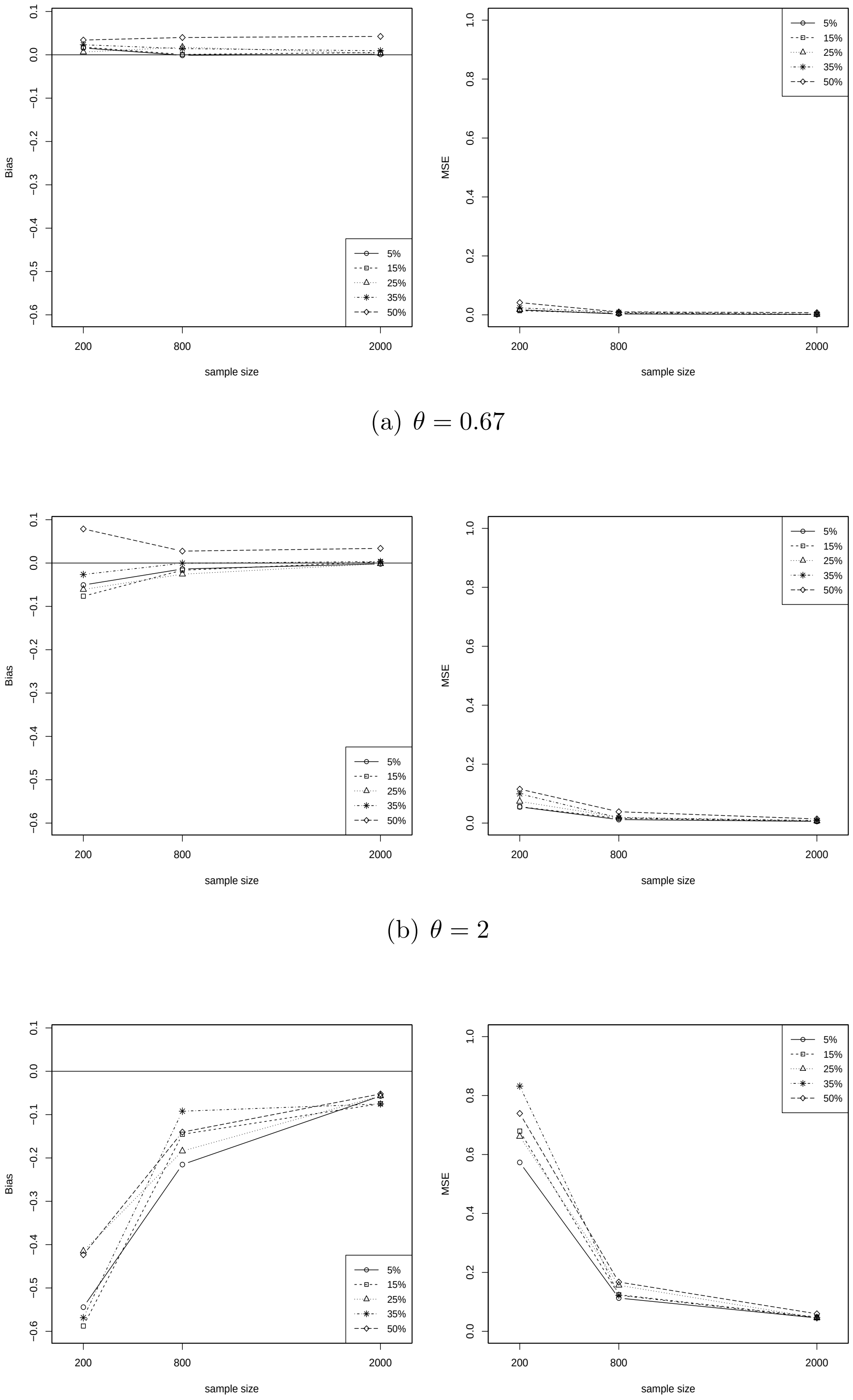


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.10: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (normal marginal errors).

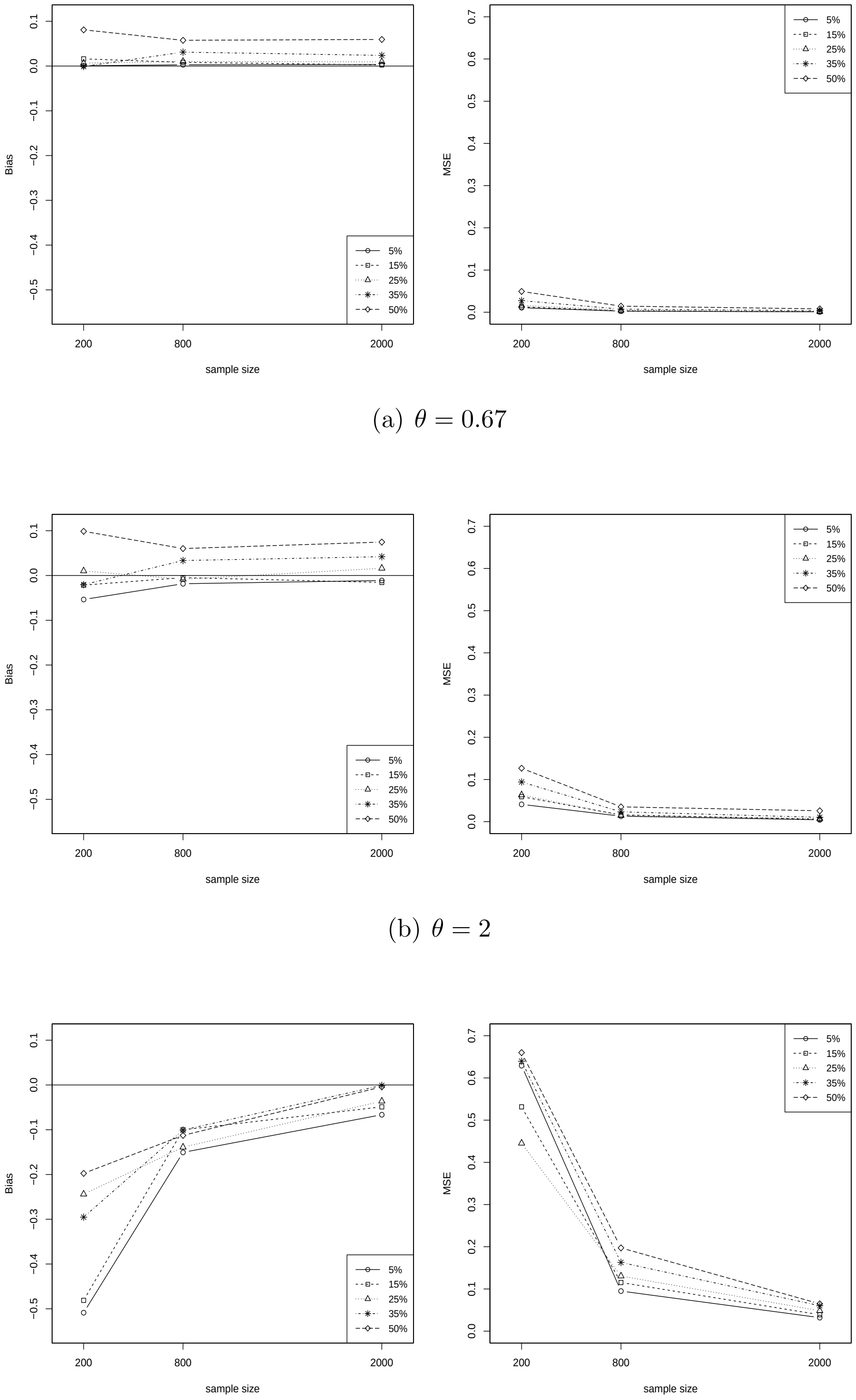


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.11: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (power-normal marginal errors).

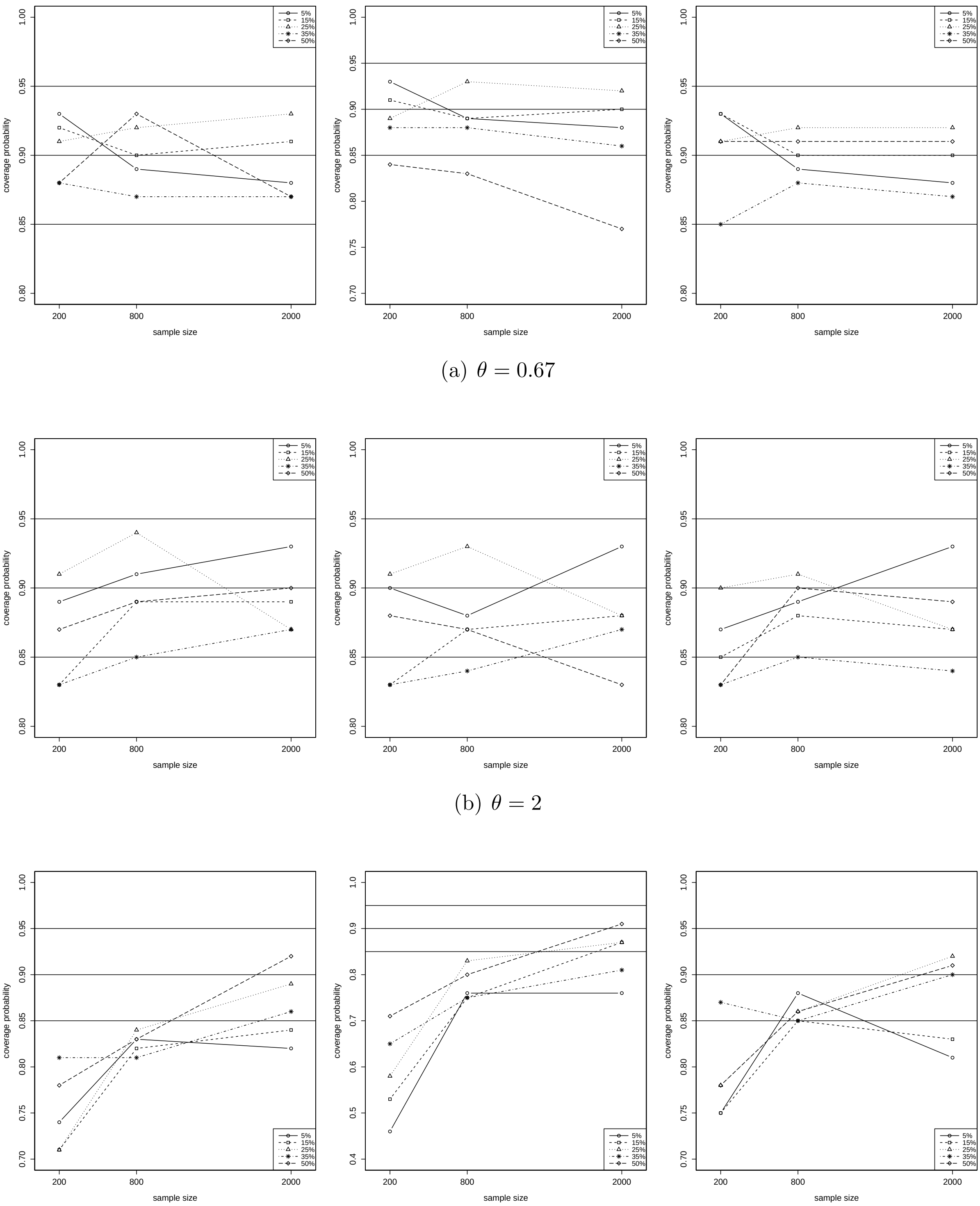


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.12: Bias and MSE of the MIFM estimate of the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (logistic marginal errors).

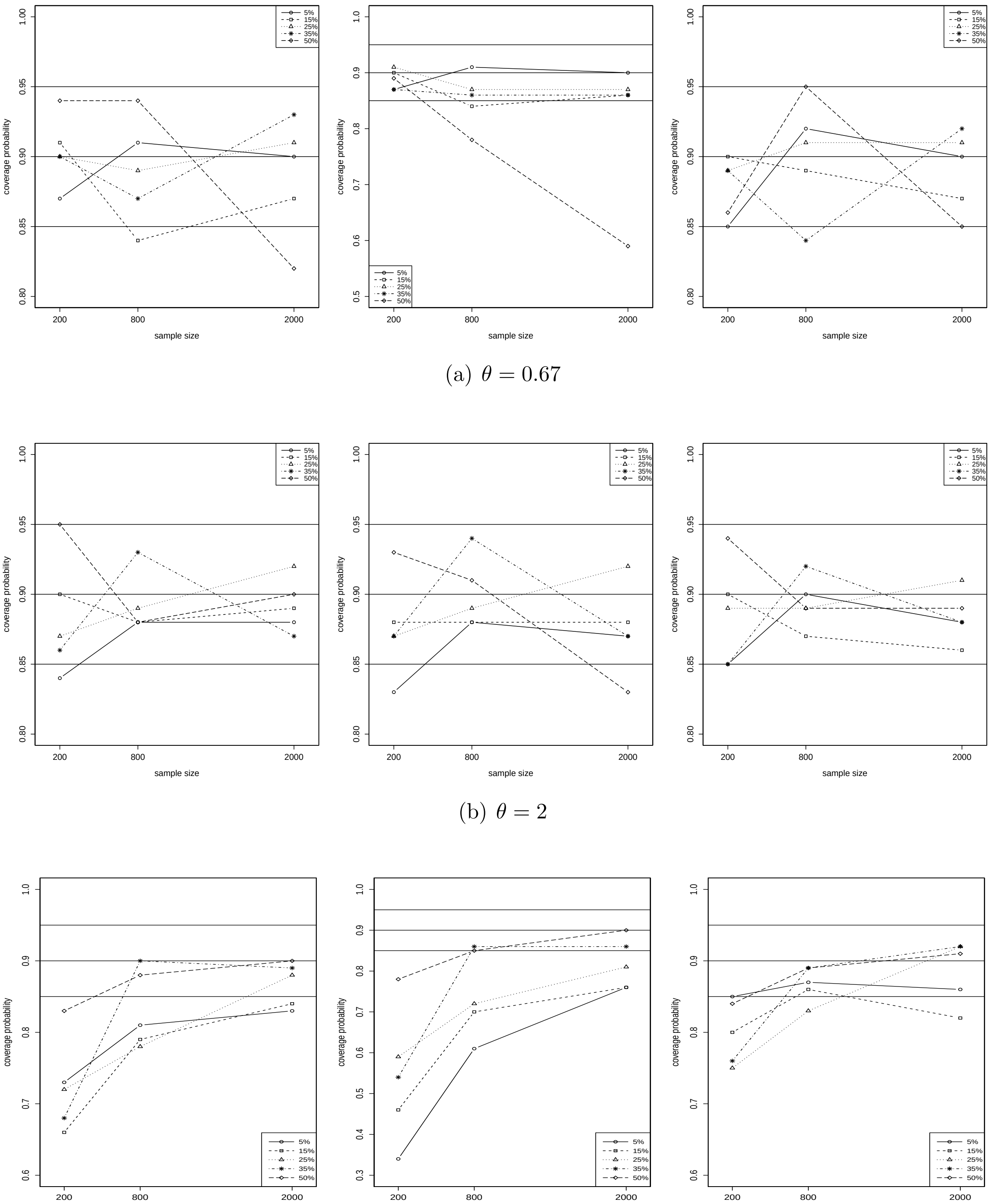


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.13: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

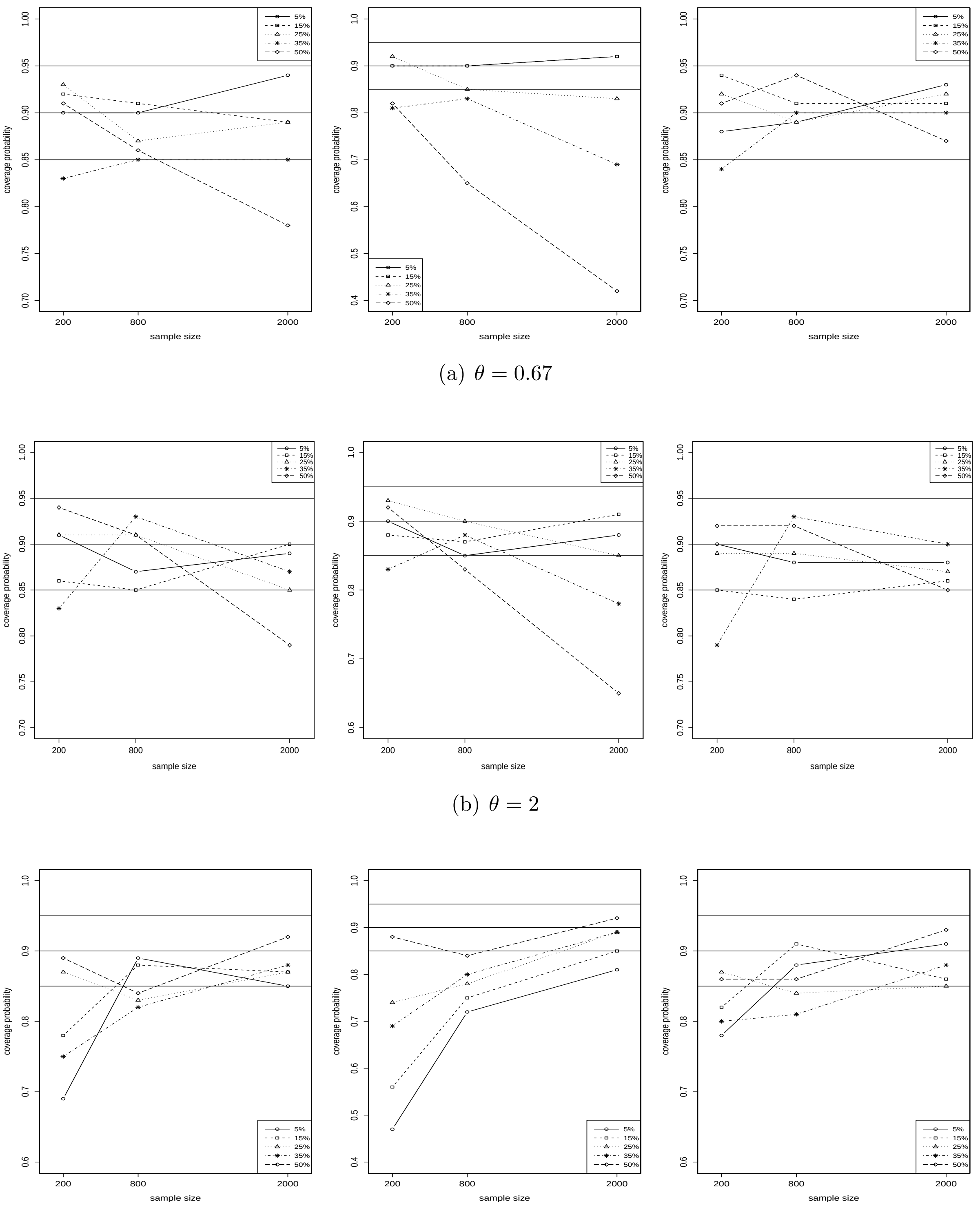


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.14: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (power-normal marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

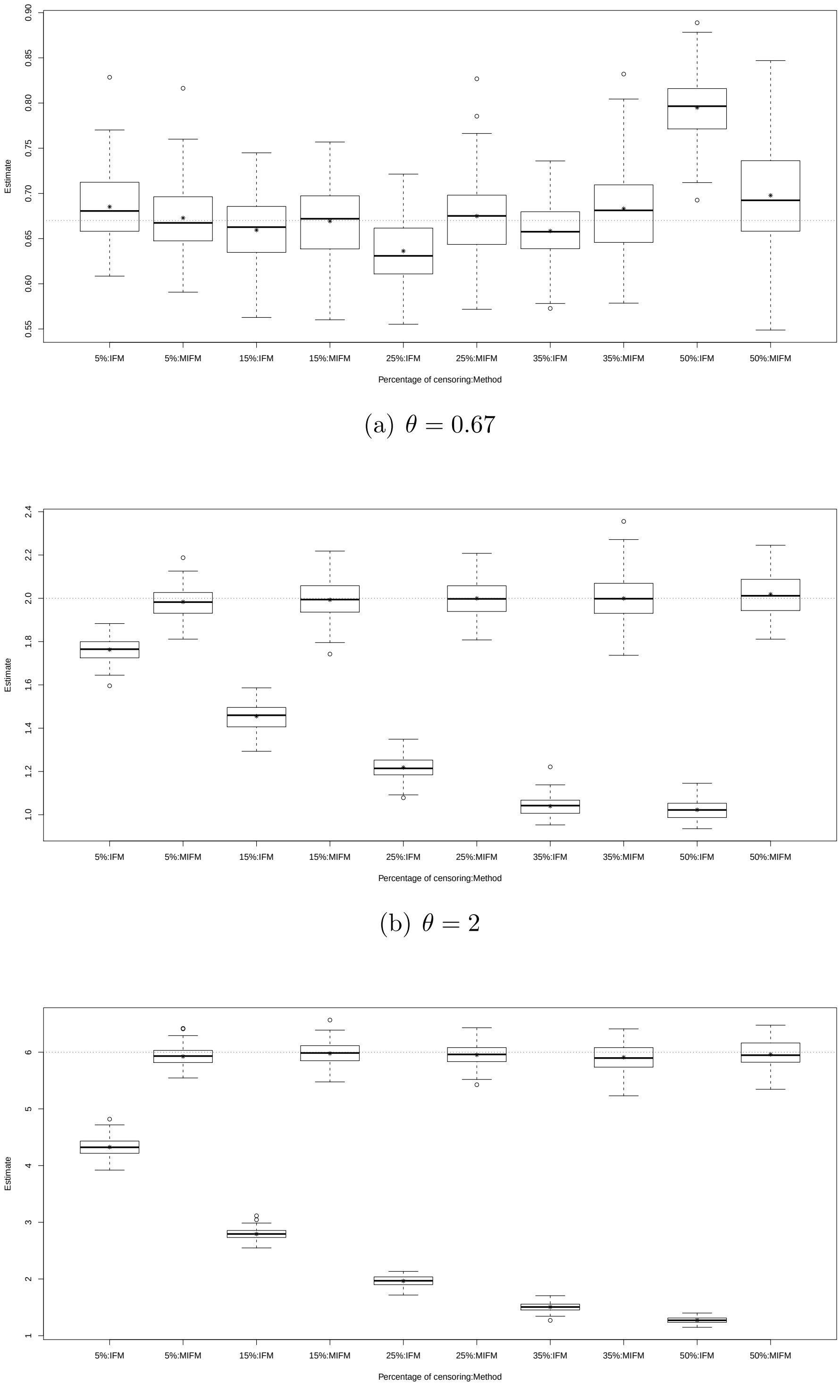


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.15: Coverage probabilities (CPs) of the 90% standard normal (panels on the left), percentile (middle panels) and basic (panels on the right) confidence intervals for the Clayton survival copula parameter versus sample size, percentage of censoring in the margins and degree of dependence among them (logistic marginal errors). The horizontal line at CP = 0.90 and the two horizontal lines at CP = 0.85 and 0.95 correspond, respectively, to the lower and upper bounds of the 90% confidence interval of the CP = 0.90. Thus, if a confidence interval has exact coverage of 0.90, roughly 90% of the observed coverages should be between these lines.

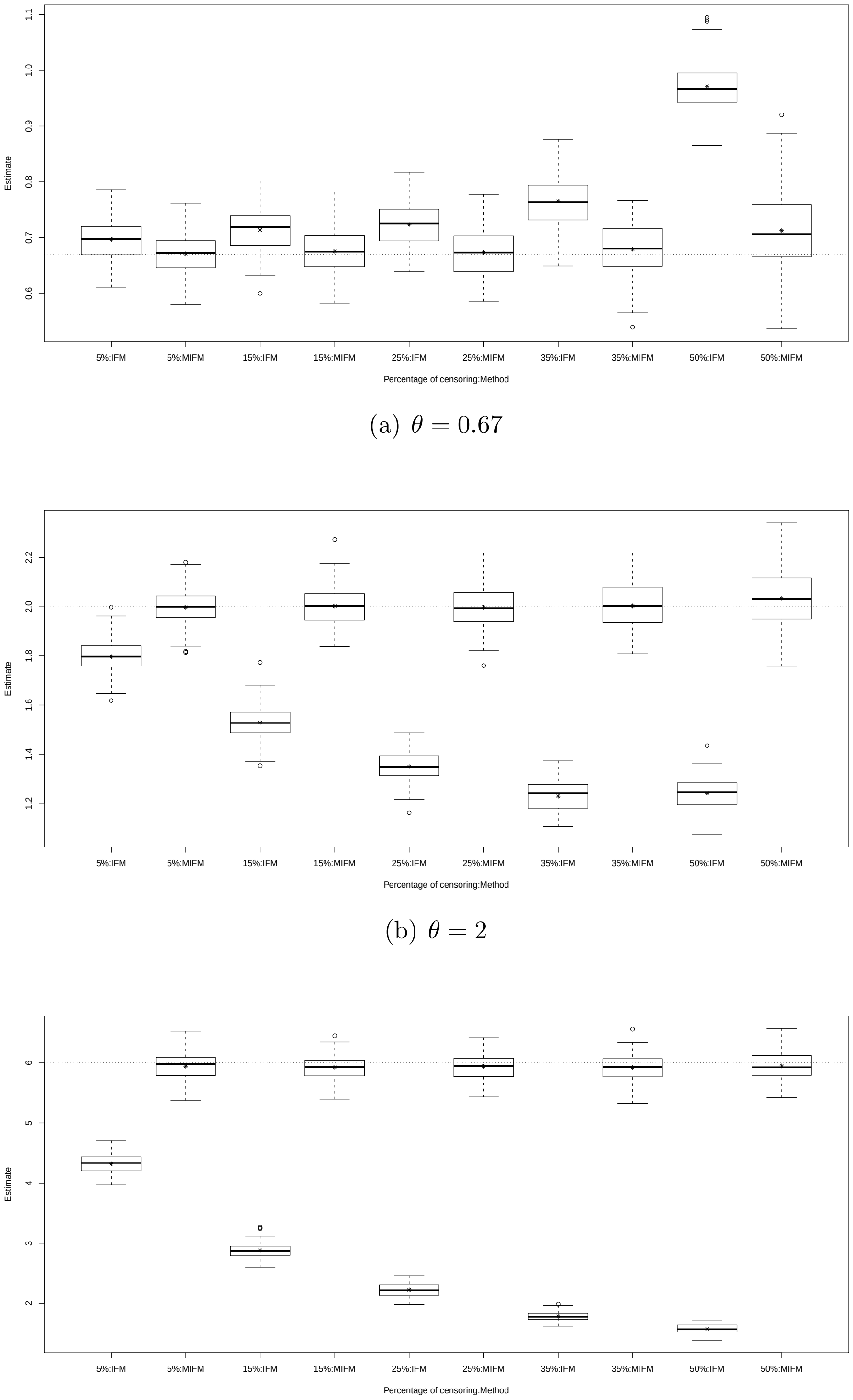


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.16: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.

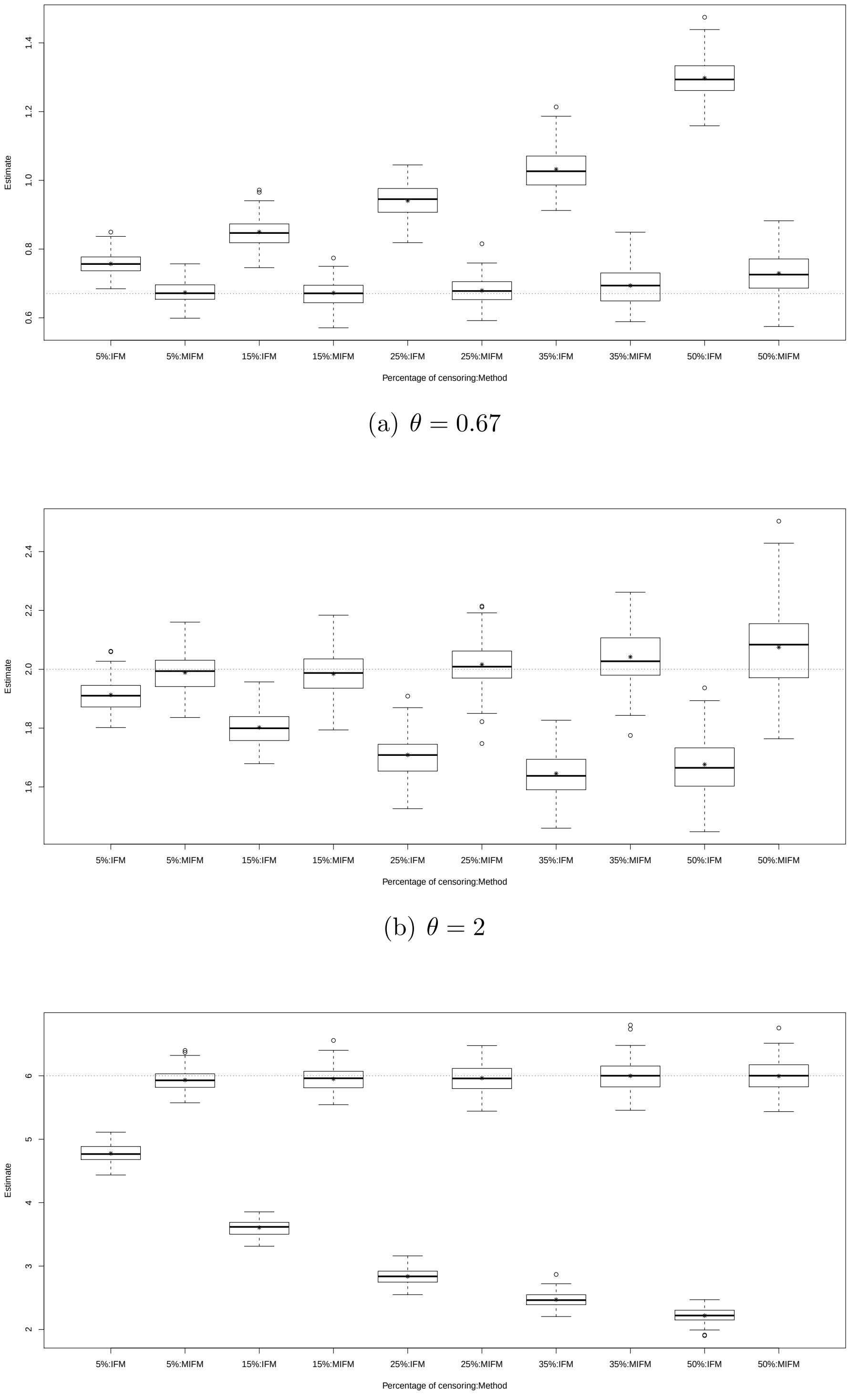


(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.17: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (power-normal marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.



(a) $\theta = 0.67$

(b) $\theta = 2$

(c) $\theta = 6$

Figure 3.18: Comparison between the IFM and MIFM estimates of the Clayton survival copula parameter, for $n = 2000$ (logistic marginal errors). The averages of the parameter estimates are shown with a star symbol. The dotted horizontal line represents the true value of the Clayton survival copula parameter.

bank is modeled by the trivariate SUR Tobit right-censored model with normal, power-normal and logistic marginal errors through the Clayton survival copula (see Sections 1.1.2 and 1.3 for the reasons for this model choice). We include age and income as the covariates and use them for all margins in all three candidate models.

Tables 3.5, 3.6 and 3.7 show the MIFM estimates for the parameters of the trivariate Clayton survival copula-based SUR Tobit right-censored model with normal, power-normal and logistic marginal errors, respectively, as well as the 90% confidence intervals obtained through the standard normal, percentile and basic bootstrap methods. Tables 3.5, 3.6 and 3.7 also present the log-likelihood, AIC and BIC criterion values for the three fitted models. Note that the trivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors has the smallest AIC and BIC criterion values and therefore provides the best fit to the customer churn data. From the Lilliefors (Kolmogorov-Smirnov) normality tests of augmented marginal residuals [8], we obtain p-values equal to 0.5991, 0.1831 and 0.9974 for Product A, Product B and Product C models, respectively. Hence, the normality assumption for the marginal errors is valid. The results reported in Table 3.5 reveal significant positive effects of age and income on log of time to churn Products A, B and C. The MIFM estimate of the Clayton survival copula parameter $\left(\widehat{\theta}_{\text{MIFM}} = 2.5514\text{, obtained after 5 iterations}\right)$ and its 90% bootstrap-based confidence intervals reveal that the relationship among the log(time) to churn Product A, log(time) to churn Product B and log(time) to churn Product C is positive (the estimated Kendall's tau is $\widehat{\tau}_3 = \widehat{\theta}_{\text{MIFM}}/\big(\widehat{\theta}_{\text{MIFM}} + 2\big) = 0.5606$) and significant at the 10% level (the lower limits of the 90% bootstrap-based confidence intervals for θ are greater than and far above zero), justifying joint estimation of the censored equations through the Clayton survival copula to improve statistical efficiency. Moreover, the estimated trivariate tail dependence coefficients for Clayton survival copula, $\hat{\lambda}_{\text{U}}^{1|23} = 0.8531$ and $\hat{\lambda}_{\text{U}}^{12|3} = 0.6501$, obtained from $(3\,/\,2)^{-1/\hat{\theta}_{\text{MIFM}}}$ and $3^{-1/\hat{\theta}_{\text{MIFM}}}$, respectively (the trivariate upper tail dependence coefficients for Clayton survival copula are equal to the trivariate lower tail dependence coefficients for Clayton copula), show the positive dependence at the upper tail of the joint distribution, i.e. for high times or log of times to churn Products A, B and C.

[8] The augmented residuals are the differences between the augmented observed and predicted responses, i.e. $e_{ij}^{\text{a}} = y_{ij}^{\text{a}} - \boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}$, for $i = 1, ..., n$ and $j = 1, 2, 3$, where $y_{ij}^{\text{a}} = \boldsymbol{x}_{ij}^{'}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}} + \hat{\sigma}_{j,\text{MIFM}}\Phi^{-1}\left(u_{ij}^{\text{a}}\right)$, with $\Phi^{-1}(.)$ being the inverse function of the $N(0,1)$ c.d.f.; or simply, $e_{ij}^{\text{a}} = \hat{\sigma}_{j,\text{MIFM}}\Phi^{-1}\left(u_{ij}^{\text{a}}\right)$.

Table 3.5: Estimation results of trivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors for the customer churn data.

		90% Confidence Intervals		
Product A	Estimate	Standard Normal	Percentile	Basic
Intercept	0.1775	[0.0130; 0.3420]	[0.0129; 0.3253]	[0.0296; 0.3420]
Age	0.0226	[0.0188; 0.0263]	[0.0189; 0.0264]	[0.0187; 0.0263]
Income	4×10^{-5}	[1×10^{-5}; 6×10^{-5}]	[1×10^{-5}; 7×10^{-5}]	[1×10^{-5}; 6×10^{-5}]
σ_1	0.9928	[0.9507; 1.0349]	[0.9466; 1.0322]	[0.9534; 1.0390]
		90% Confidence Intervals		
Product B	Estimate	Standard Normal	Percentile	Basic
Intercept	0.2233	[0.0862; 0.3604]	[0.0820; 0.3531]	[0.0934; 0.3645]
Age	0.0238	[0.0206; 0.0270]	[0.0206; 0.0272]	[0.0204; 0.0270]
Income	8×10^{-5}	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1×10^{-4}]
σ_2	0.9098	[0.8715; 0.9480]	[0.8701; 0.9469]	[0.8726; 0.9494]
		90% Confidence Intervals		
Product C	Estimate	Standard Normal	Percentile	Basic
Intercept	0.0707	[-0.0874; 0.2288]	[-0.0937; 0.2164]	[-0.0751; 0.2351]
Age	0.0248	[0.0212; 0.0283]	[0.0214; 0.0281]	[0.0214; 0.0281]
Income	7×10^{-5}	[5×10^{-5}; 1×10^{-4}]	[5×10^{-5}; 1×10^{-4}]	[4×10^{-5}; 1×10^{-4}]
σ_3	0.9666	[0.9238; 1.0094]	[0.9228; 1.0091]	[0.9241; 1.0104]
θ	2.5514	[2.3320; 2.7708]	[2.3611; 2.8119]	[2.2908; 2.7416]
Log-likelihood	-2627.9800			
AIC	5281.9600			
BIC	5344.7760			

Table 3.6: Estimation results of trivariate Clayton survival copula-based SUR Tobit right-censored model with power-normal marginal errors for the customer churn data.

		90% Confidence Intervals		
Product A	Estimate	Standard Normal	Percentile	Basic
Intercept	0.5195	[-0.2412; 1.2803]	[-0.2642; 1.1544]	[-0.1153; 1.3033]
Age	0.0229	[0.0190; 0.0267]	[0.0190; 0.0266]	[0.0191; 0.0268]
Income	4×10^{-5}	[1×10^{-5}; 6×10^{-5}]	[1×10^{-5}; 7×10^{-5}]	[1×10^{-5}; 6×10^{-5}]
σ_1	0.8594	[0.5830; 1.1358]	[0.6175; 1.1165]	[0.6024; 1.1013]
α_1	0.6481	[-0.1319; 1.4282]	[0.2575; 1.4866]	[-0.1904; 1.0387]
		90% Confidence Intervals		
Product B	Estimate	Standard Normal	Percentile	Basic
Intercept	0.2230	[-0.5447; 0.9907]	[-0.5798; 0.8902]	[-0.4442; 1.0258]
Age	0.0237	[0.0203; 0.0270]	[0.0204; 0.0270]	[0.0203; 0.0269]
Income	8×10^{-5}	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1×10^{-4}]
σ_2	0.9113	[0.6534; 1.1692]	[0.6772; 1.1846]	[0.6380; 1.1454]
α_2	1.0060	[-0.8961; 2.9080]	[0.4037; 2.3885]	[-0.3766; 1.6082]
		90% Confidence Intervals		
Product C	Estimate	Standard Normal	Percentile	Basic
Intercept	-0.3828	[-1.3423; 0.5767]	[-1.3539; 0.4218]	[-1.1875; 0.5883]
Age	0.0245	[0.0210; 0.0280]	[0.0208; 0.0277]	[0.0213; 0.0282]
Income	7×10^{-5}	[4×10^{-5}; 1×10^{-4}]	[4×10^{-5}; 1×10^{-4}]	[4×10^{-5}; 1×10^{-4}]
σ_3	1.1259	[0.8260; 1.4258]	[0.8648; 1.4313]	[0.8204; 1.3870]
α_3	1.6500	[-0.4990; 3.7991]	[0.7104; 3.9180]	[-0.6180; 2.5896]
θ	2.4520	[2.2312; 2.6729]	[2.2560; 2.6937]	[2.2104; 2.6481]
Log-likelihood	-2636.6990			
AIC	5305.3980			
BIC	5382.7090			

Table 3.7: Estimation results of trivariate Clayton survival copula-based SUR Tobit right-censored model with logistic marginal errors for the customer churn data.

		90% Confidence Intervals		
Product A	Estimate	Standard Normal	Percentile	Basic
Intercept	0.1566	[-0.0066; 0.3198]	[0.0112; 0.3320]	[-0.0188; 0.3020]
Age	0.0231	[0.0195; 0.0268]	[0.0193; 0.0267]	[0.0196; 0.0270]
Income	4×10^{-5}	[1×10^{-5}; 6×10^{-5}]	[1×10^{-5}; 6×10^{-5}]	[1×10^{-5}; 6×10^{-5}]
s_1	0.5750	[0.5463; 0.6037]	[0.5465; 0.6042]	[0.5458; 0.6035]
		90% Confidence Intervals		
Product B	Estimate	Standard Normal	Percentile	Basic
Intercept	0.1592	[0.0125; 0.3058]	[0.0164; 0.3018]	[0.0165; 0.3019]
Age	0.0252	[0.0219; 0.0285]	[0.0221; 0.0284]	[0.0220; 0.0283]
Income	8×10^{-5}	[6×10^{-5}; 1.1×10^{-4}]	[6×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]
s_2	0.5363	[0.5103; 0.5622]	[0.5088; 0.5631]	[0.5094; 0.5637]
		90% Confidence Intervals		
Product C	Estimate	Standard Normal	Percentile	Basic
Intercept	0.0830	[-0.0697; 0.2358]	[-0.0632; 0.2337]	[-0.0677; 0.2292]
Age	0.0242	[0.0208; 0.0276]	[0.0205; 0.0275]	[0.0209; 0.0279]
Income	8×10^{-5}	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]	[5×10^{-5}; 1.1×10^{-4}]
s_3	0.5668	[0.5398; 0.5939]	[0.5400; 0.5954]	[0.5382; 0.5937]
θ	2.3808	[2.1812; 2.5804]	[2.2204; 2.6140]	[2.1476; 2.5411]
Log-likelihood	-2666.6660			
AIC	5359.3320			
BIC	5422.1470			

For comparison purposes, we also fit the basic trivariate SUR Tobit right-censored model (which is the trivariate SUR Tobit right-censored model whose dependence among the marginal error terms ϵ_{i1}, ϵ_{i2} and ϵ_{i3}, $i = 1, ..., n$, is modeled through the trivariate normal distribution) using the MCECM algorithm of Huang (1999) adapted for right-censored trivariate normal data. The estimation results (obtained after 4 iterations) are presented in Table 3.8. The standard errors were derived from the bootstrap estimate of the covariance matrix (bootstrap standard errors). Note that, with the exception of the intercept term in the Product C model, all of the parameter estimates are significant at the 10% level. Moreover, the marginal parameter estimates obtained through the (adapted) MCECM and (extended) MIFM methods are similar (see Tables 3.5 and 3.8). However, the trivariate Clayton survival copula-based SUR Tobit right-censored model with normal marginal errors overcomes the basic trivariate SUR Tobit right-censored model in both AIC and BIC criterion. This indicates that the gain for introducing the Clayton survival copula to model the nonlinear dependence structure of the trivariate SUR Tobit right-censored model was substantial for this dataset.

Table 3.8: Estimation results of basic trivariate SUR Tobit right-censored model for the customer churn data.

Product A	Estimate	Standard Error
Intercept	0.2232 *	0.0875
Age	0.0210 *	0.0019
Income	4×10^{-5} *	1×10^{-5}
σ_1	0.9524 *	0.0216
Product B	Estimate	Standard Error
Intercept	0.2804 *	0.0804
Age	0.0224 *	0.0018
Income	6×10^{-5} *	1×10^{-5}
σ_2	0.8730 *	0.0192
Product C	Estimate	Standard Error
Intercept	0.0909	0.0925
Age	0.0247 *	0.0021
Income	6×10^{-5} *	1×10^{-5}
σ_3	0.9694 *	0.0237
σ_{12} †	0.6125 *	0.0302
σ_{13} ‡	0.5936 *	0.0326
σ_{23} §	0.6126 *	0.0318
Log-likelihood	-2916.1670	
AIC	5862.3330	
BIC	5934.8120	

* Denotes significant at the 10% level.

† Denotes the covariance between Products A and B.

‡ Denotes the covariance between Products A and C.

§ Denotes the covariance between Products B and C.

3.3 Final remarks

In this chapter, we extended the bivariate models and methods proposed in the previous chapter to the trivariate case in a straightforward way. Again, our decision for two parametric families of copula (Clayton copula for the trivariate SUR Tobit model, and Clayton survival copula for the trivariate SUR Tobit right-censored model), as well as non-normal (power-normal and logistic) distribution assumption for the marginal error terms, were mainly motivated by the real data at hand (U.S. salad dressing, tomato and lettuce consumption data, and Brazilian commercial bank customer churn data). Furthermore, some advantages arose from these copula choices, regarding the development of the (extended) MIFM method for obtaining the estimates of the trivariate models' parameters. Indeed, the tridimensional generalizations of the Clayton and Clayton survival copulas that we used here are the simplest ones and present the whole trivariate dependence structure with only one single copula parameter θ. Moreover, these tridimensional copulas implicitly assume that the order of margins within the copula function is exchangeable. This means that, e.g., $C\left(u_{i1}, u_{i2}, u_{i3} | \theta\right) = C\left(u_{i3}, u_{i1}, u_{i2} | \theta\right)$, which is not plausible for many applications (cf. McNeil *et al.*, 2005, p. 224; Savu & Trede, 2010). A more flexible method is provided by hierarchical Archimedean copula (HAC), discussed by Joe (1997), Embrechts, Lindskog & McNeil (2003), Whelan (2004), Savu & Trede (2010) and Okhrin, Okhrin & Schmid (2013). In contrast to the usual Archimedean copula, the HAC defines the whole dependence structure in a recursive way, i.e. by aggregating one dimension step by step starting from a low-dimensional copula.

In the simulation studies, we assessed the performance of our proposed trivariate models and methods, obtaining satisfactory results (unbiased estimates of the copula parameter, high and near the nominal value coverage probabilities of the standard normal and basic bootstrap confidence intervals) regardless of the error distribution assumption, the censoring percentage in the margins and their degree of interdependence.

Besides the basic bootstrap method, another alternative to the percentile method, which in general yielded confidence intervals for the copula parameter with low coverage probabilities, could be the Bias-Corrected and Accelerated (BCa) method by Efron (1987), which adjusts for both bias and skewness in the bootstrap distribution. However, this bootstrap method is more computationally expensive (it requires much more computer memory and time) than the ones considered in this chapter.

Finally, we pointed out the applicability of our proposed trivariate models and methods for real datasets, where we found that the gain for introducing the copulas to model the nonlinear dependence structure of the trivariate SUR Tobit models was substantial for these datasets.

In the next chapter, we will briefly present a generalization of the models and methods proposed in this thesis for the multivariate case.

Chapter 4

Multivariate Copula-based SUR Tobit Models

In this chapter, we present a straightforward generalization of the models and methods proposed in the previous chapters for the multivariate case. We first present the multivariate Clayton copula-based SUR Tobit model, which is the SUR Tobit model with $m \geq 2$ left-censored (at zero point) dependent variables whose dependence among them is modeled through the multidimensional Clayton copula. Then, we present the multivariate Clayton survival copula-based SUR Tobit right-censored model, i.e. the SUR Tobit model with $m \geq 2$ right-censored (at point $d_j > 0$, $j = 1, 2, \ldots, m$) dependent variables whose dependence structure among them is modeled by the multidimensional Clayton survival copula. Brief discussions concerning the model implementation through the proposed (generalized) MIFM method, as well as the confidence intervals construction from the bootstrap distribution of model parameters, are made for each proposed multivariate model.

4.1 Multivariate Clayton copula-based SUR Tobit model formulation

The SUR Tobit model with $m \geq 2$ left-censored (at zero point) dependent variables, or simply multivariate SUR Tobit model, can be expressed as

$$y_{ij}^{*} = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_{j} + \sigma_{j}\epsilon_{ij},$$

$$y_{ij} = \begin{cases} y_{ij}^{*} & \text{if } y_{ij}^{*} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, 2, ..., n$ and $j = 1, 2, \ldots, m$, where n is the number of observations, y_{ij}^* is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^* whenever y_{ij}^* is above zero and zero otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients, σ_j is the scale parameter of margin j, and ϵ_{ij} is the margin j's error that follows some standard distribution.

Generally, the dependence among the error terms ϵ_{i1}, ϵ_{i2}, $\ldots$, ϵ_{im} is modeled through a multivariate distribution, especially the multivariate normal distribution (basic multivariate SUR Tobit model). However, applying a multivariate distribution to the multivariate SUR Tobit model is limited to the linear relationship among marginal distributions through the correlation coefficients. Moreover, estimation methods for high-dimensional SUR Tobit models are often computationally demanding and difficult to implement. To overcome these restrictions, we can use a copula function to model the nonlinear dependence structure in the multivariate SUR Tobit model.

Thus, for the censored outcomes $y_{i1}, y_{i2}, \ldots, y_{im}$, the multivariate copula-based SUR Tobit distribution is given by

$$F\left(y_{i1}, y_{i2}, \ldots, y_{im}\right) = C\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right),$$

where u_{ij} is the c.d.f. of y_{ij}, i.e. $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$, with $\boldsymbol{\upsilon}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ being the margin j's parameter vector, for $j = 1, 2 \ldots, m$; and θ is the copula parameter (or copula parameter vector), which is assumed to be scalar.

Suppose that C is the multidimensional Clayton copula, which takes the form

$$C\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right) = \left(\sum_{j=1}^{m} u_{ij}^{-\theta} - m + 1\right)^{-\frac{1}{\theta}} \tag{4.1}$$

(Cherubini *et al.*, 2004, p. 150), with $\theta \in (0, \infty)$. The dependence among the margins increases as the value of θ increases, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect positive dependence. This multidimensional Archimedean copula shows lower tail dependence and is characterized by zero upper tail dependence (De Luca & Rivieccio, 2012; Di Bernardino & Rullière, 2014).

4.1.1 Inference

In this subsection, we briefly discuss inference (point and interval estimation) for the parameters of the multivariate Clayton copula-based SUR Tobit model.

4.1.1.1 Estimation through the (generalized) MIFM method

Following Trivedi & Zimmer (2005) and Anastasopoulos *et al.* (2012), we can write the log-likelihood function for the multivariate Clayton copula-based SUR Tobit model in the following form

$$\begin{aligned}\ell(\boldsymbol{\eta}) = &\sum_{i=1}^{n} \log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \boldsymbol{\upsilon}_1\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \boldsymbol{\upsilon}_2\right), \dots, F_m\left(y_{im}|\boldsymbol{x}_{im}, \boldsymbol{\upsilon}_m\right)|\theta\right) + \\ &+ \sum_{i=1}^{n}\sum_{j=1}^{m} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right),\end{aligned} \tag{4.2}$$

where $\boldsymbol{\eta} = (\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2, \dots, \boldsymbol{\upsilon}_m, \theta)$ is the vector of model parameters, $f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, and $c\left(u_{i1}, u_{i2}, \dots, u_{im}|\theta\right)$, with $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton copula, which is calculated from (4.1) as

$$\begin{aligned} c\left(u_{i1}, u_{i2}, \dots, u_{im}|\theta\right) &= \frac{\partial^m C\left(u_{i1}, u_{i2}, \dots, u_{im}|\theta\right)}{\partial u_{i1}\partial u_{i2}\dots\partial u_{im}} = \\ &= \theta^m \frac{\Gamma\left(\frac{1}{\theta}+m\right)}{\Gamma\left(\frac{1}{\theta}\right)} \left(\prod_{j=1}^{m} u_{ij}^{-\theta-1}\right)\left(\sum_{j=1}^{m} u_{ij}^{-\theta} - m + 1\right)^{-\frac{1}{\theta}-m}\end{aligned} \tag{4.3}$$

(Cherubini *et al.*, 2004, p. 225), where $\Gamma(.)$ is the gamma function.

For model estimation, the use of copula methods, as well as the log-likelihood function form given by (4.2), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\text{IFM}} = \arg\max_{\boldsymbol{\upsilon}_j} \sum_{i=1}^{n} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right), \tag{4.4}$$

for $j = 1, 2, \dots, m$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\text{IFM}}$ by

$$\widehat{\theta}_{\text{IFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\big(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \widehat{\boldsymbol{\upsilon}}_{1,\text{IFM}}\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \widehat{\boldsymbol{\upsilon}}_{2,\text{IFM}}\right), \dots, F_m\left(y_{im}|\boldsymbol{x}_{im}, \widehat{\boldsymbol{\upsilon}}_{m,\text{IFM}}\right)|\theta\big). \tag{4.5}$$

However, as seen in Sections 2.1.2.2 and 3.1.2.2, the above-described IFM method provides a biased estimate for the parameter θ, since there is a violation of Sklar's theorem (Sklar, 1959) in the cases with the presence of censored observations in the margins (semi-continuous/censored margins).

Thus, in order to facilitate the implementation of copula models with semi-continuous margins, the semi-continuous marginal distributions could be augmented to achieve continuity (and thus satisfy the Sklar's theorem!). More specifically, we can use a (frequentist) data augmentation technique to simulate the latent (i.e. unobserved) dependent variables in the censored margins (Wichitaksorn *et al.*, 2012). Then, we replace y_{ij} by the augmented data y^{a}_{ij}, or equivalently and more simply, we replace u_{ij} by the augmented uniform data u^{a}_{ij} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the continuous margin cases. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence occurs.

In the remaining part of this subsubsection, we discuss the proposed estimation method (a generalization of the MIFM method proposed in Sections 2.1.1.1 and 3.1.1.1) when using the Clayton copula to model the nonlinear dependence structure of the multivariate SUR Tobit model. However, the proposed approach can be extended to other copula functions by applying different sampling algorithms.

Let margin j's error ϵ_{ij} have a standard distribution $H_j\left(.\right)$ and consider the upper bounds given by $b_{ij} = H_j\left(-\boldsymbol{x}'_{ij}\hat{\boldsymbol{\beta}}_{j,\mathrm{MIFM}}\,/\,\hat{\sigma}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, \ldots, m$. The implementation of the multivariate Clayton copula-based SUR Tobit model through the proposed (generalized) MIFM method can be briefly described as follows.

Stage 1. Estimate the marginal parameters using (4.4). Set $\hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$, i.e. $\left(\hat{\boldsymbol{\beta}}_{j,\mathrm{MIFM}}, \hat{\sigma}_{j,\mathrm{MIFM}}\right) = \left(\hat{\boldsymbol{\beta}}_{j,\mathrm{IFM}}, \hat{\sigma}_{j,\mathrm{IFM}}\right)$, for $j = 1, 2, \ldots, m$.

Stage 2. Estimate the copula parameter using, e.g., (4.5). Set $\hat{\theta}^{(1)}_{\mathrm{MIFM}} = \hat{\theta}_{\mathrm{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = y_{i2} = \cdots = y_{im} = 0$, then draw $(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, \ldots, u^{\mathrm{a}}_{im})$ from $C\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, \ldots, u^{\mathrm{a}}_{im} | \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(0, b_{i1}) \times (0, b_{i2}) \times \cdots \times (0, b_{im})$. This can be performed relatively easily using the truncation dependence invariance property of the (multidimensional) Clayton copula (Sungur, 2002).

If $y_{i1} = \cdots = y_{i,s-1} = y_{i,s+1} = \cdots = y_{im} = 0$ and $y_{is} > 0$, then draw $\left(u^{\mathrm{a}}_{i1}, \ldots, u^{\mathrm{a}}_{i,s-1}, u^{\mathrm{a}}_{i,s+1}, \ldots, u^{\mathrm{a}}_{im}\right)$ from $C\left(u^{\mathrm{a}}_{i1}, \ldots, u^{\mathrm{a}}_{i,s-1}, u^{\mathrm{a}}_{i,s+1}, \ldots, u^{\mathrm{a}}_{im} | u_{is}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\right)$ truncated to the region $(0, b_{i1}) \times \cdots \times (0, b_{i,s-1}) \times (0, b_{i,s+1}) \times \cdots \times (0, b_{im})$. This can be performed through iterative conditioning (conditional sampling) by successive application of the method by Devroye (1986, p. 38-39).

$\vdots$

If $y_{is} = 0$ and $y_{i1} > 0, \ldots, y_{i,s-1} > 0, y_{i,s+1} > 0, \ldots, y_{im} > 0$, then draw u^{a}_{is} from $C\Big(u^{\mathrm{a}}_{is} | u_{i1}, \ldots, u_{i,s-1}, u_{i,s+1}, \ldots, u_{im}, \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}\Big)$ truncated to the interval $(0, b_{is})$. This can be done by applying the method by Devroye (1986, p. 38-39).

If $y_{i1} > 0, y_{i2} > 0, \ldots, y_{im} > 0$, then set $u^{\mathrm{a}}_{ij} = u_{ij} = H_j\left(\left(y_{ij} - \boldsymbol{x}'_{ij}\hat{\boldsymbol{\beta}}_{j,\mathrm{MIFM}}\right) / \hat{\sigma}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, \ldots, m$.

Given the generated/augmented marginal uniform data u^{a}_{ij}, we estimate the association parameter θ by

$$\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u^{\mathrm{a}}_{i1}, u^{\mathrm{a}}_{i2}, \ldots, u^{\mathrm{a}}_{im} | \theta\right).$$

The algorithm stops if a termination criterion is fulfilled, e.g. if $|\hat{\theta}^{(\omega+1)}_{\mathrm{MIFM}} - \hat{\theta}^{(\omega)}_{\mathrm{MIFM}}| < \xi$, where ξ is the tolerance parameter.

4.1.1.2 Interval estimation

We propose the use of bootstrap methods for computing confidence intervals for the parameters of the multivariate Clayton copula-based SUR Tobit model. It makes the analytic derivatives no longer required to compute the asymptotic covariance matrix associated with the vector of parameter estimates.

Our proposed bootstrap approach is described as follows. Let η_h, $h = 1, ..., k$, be any component of the parameter vector $\boldsymbol{\eta}$ of the multivariate Clayton copula-based SUR Tobit model (see Section 4.1.1.1). By using a parametric resampling plan, we obtain the bootstrap estimates $\hat{\eta}^*_{h1}, \hat{\eta}^*_{h2}, ..., \hat{\eta}^*_{hB}$ of η_h through the (generalized) MIFM method. Hinkley (1988) suggests that the minimum value of the number of bootstrap samples, B, will depend on the parameter being estimated, but that it will often be 100 or more. Then, we can derive confidence intervals from the bootstrap distribution through the following two methods, for instance.

- **Basic bootstrap** (Davison & Hinkley, 1997, p. 194). The $100\,(1 - 2\alpha)\,\%$ basic confidence interval is defined by

$$\left[2\hat{\eta}_h - \hat{\eta}^{*(1-\alpha)}_h, 2\hat{\eta}_h - \hat{\eta}^{*(\alpha)}_h\right],$$

where $\hat{\eta}_h^{*(\alpha)}$ and $\hat{\eta}_h^{*(1-\alpha)}$ are, respectively, the $100\,(\alpha)$th and $100\,(1-\alpha)$th percentiles of the bootstrap distribution of $\hat{\eta}_h^*$, and $\hat{\eta}_h$ is the original estimate (i.e. from the original data) of η_h, obtained through the proposed (generalized) MIFM method. If there is a parameter constraint (such as $\eta_h > 0$), then the $100\,(1-2\alpha)\,\%$ basic confidence interval may include invalid parameter values.

- **Standard normal interval** (Efron & Tibshirani, 1993, p. 154). Since most statistics are asymptotically normally distributed, in large samples we can use the standard error estimate, $\widehat{se}_h$, as well as the normal distribution, to yield a $100\,(1-2\alpha)\,\%$ confidence interval for η_h based on the original estimate $\hat{\eta}_h$:

$$\left[\hat{\eta}_h - z^{(1-\alpha)}\widehat{se}_h, \hat{\eta}_h - z^{(\alpha)}\widehat{se}_h\right],$$

where $z^{(\alpha)}$ represents the $100\,(\alpha)$th percentile point of a standard normal distribution, and $\widehat{se}_h$ is the hth entry on the diagonal of the bootstrap-based covariance matrix estimate of the parameter vector estimate $\hat{\boldsymbol{\eta}}$, which is given by

$$\widehat{\boldsymbol{\Sigma}}_{\text{boot}} = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)\left(\hat{\boldsymbol{\eta}}_b^* - \overline{\hat{\boldsymbol{\eta}}}^*\right)',$$

where $\hat{\boldsymbol{\eta}}_b^*$, $b = 1, ..., B$, is the bootstrap estimate of $\boldsymbol{\eta}$ and

$$\overline{\hat{\boldsymbol{\eta}}}^* = \left(\frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{1b}^*, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{2b}^*, \ldots, \frac{1}{B}\sum_{b=1}^{B}\hat{\eta}_{kb}^*\right).$$

Among other bootstrap methods that could be applied to build confidence intervals for the multivariate Clayton copula-based SUR Tobit model parameters, we can cite the Bias-Corrected and Accelerated (BCa) method by Efron (1987) and the percentile method by Efron & Tibshirani (1993, p. 171). However, we do not encourage the use of the percentile method in the high-dimensional setting since it usually yields confidence intervals for the copula association parameter whose coverage probabilities are lower than the nominal level (as seen in Section 3.1.2.2). The use of the BCa method should also be avoided due to its computational cost (it requires much more computer memory and time).

4.2 Multivariate Clayton survival copula-based SUR Tobit right-censored model formulation

The SUR Tobit model with $m \geq 2$ right-censored dependent variables, or simply multivariate SUR Tobit right-censored model, can be expressed as

$$y_{ij}^{*} = \boldsymbol{x}_{ij}^{'}\boldsymbol{\beta}_{j} + \sigma_{j}\epsilon_{ij},$$

$$y_{ij} = \begin{cases} y_{ij}^{*} & \text{if } y_{ij}^{*} < d_{j}, \\ d_{j} & \text{otherwise,} \end{cases}$$

for $i = 1, ..., n$ and $j = 1, 2, \ldots, m$, where n is the number of observations, d_j is the censoring point/threshold of margin j (which we assume to be known and constant), y_{ij}^{*} is the latent (i.e. unobserved) dependent variable of margin j, y_{ij} is the observed dependent variable of margin j (which is defined to be equal to the latent dependent variable y_{ij}^{*} whenever y_{ij}^{*} is below d_j and d_j otherwise), $\boldsymbol{x}_{ij}$ is the $k \times 1$ vector of covariates, $\boldsymbol{\beta}_j$ is the $k \times 1$ vector of regression coefficients, σ_j is the scale parameter of margin j and ϵ_{ij} is the margin j's error that follows some standard distribution.

Generally, the dependence among the error terms ϵ_{i1}, ϵ_{i2}, $\ldots$, ϵ_{im} is modeled through a multivariate distribution, especially the multivariate normal distribution (basic multivariate SUR Tobit right-censored model). Nevertheless, applying a multivariate distribution to the multivariate SUR Tobit right-censored model is limited to the linear relationship among marginal distributions through the correlation coefficients. Furthermore, estimation methods for high-dimensional SUR Tobit right-censored models are often computationally demanding and difficult to implement. To overcome these restrictions, we can apply a copula function to model the nonlinear dependence structure in the multivariate SUR Tobit right-censored model.

Therefore, for the censored outcomes y_{i1}, y_{i2}, $\ldots$, y_{im}, the multivariate copula-based SUR Tobit right-censored distribution is given by

$$F\left(y_{i1}, y_{i2}, \ldots, y_{im}\right) = C\left(u_{i1}, u_{i2}, \ldots, u_{im} | \theta\right),$$

where u_{ij} is the c.d.f. of y_{ij}, i.e. $u_{ij} = F_j\left(y_{ij} | \boldsymbol{x}_{ij}, \boldsymbol{v}_j\right)$, with $\boldsymbol{v}_j = \left(\boldsymbol{\beta}_j, \sigma_j\right)$ being the margin j's parameter vector, for $j = 1, 2 \ldots, m$, and θ is the copula parameter (or copula parameter vector), which is assumed to be scalar.

Suppose that C is the multidimensional Clayton survival copula with a single parameter $\theta > 0$. It takes the form of

$$C\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right) = 1 - \sum_{j=1}^{m} \left(1 - u_{ij}\right) + \sum_{S \subset \{1,\ldots,m\}, |S| \geq 2} \left(-1\right)^{|S|} C_{|S|}\left(1 - u_{il}, l \in S|\theta\right) \tag{4.6}$$

(Joe, 2014, p. 28), where $|S|$ is the cardinality of S and $C_{|S|}$ denotes the $|S|$-dimensional Clayton copula which is given by (4.1). The dependence among the margins increases as the value of θ increases, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect positive dependence. This multidimensional copula shows upper tail dependence and is characterized by zero lower tail dependence.

4.2.1 Inference

In this subsection, we briefly discuss inference (point and interval estimation) for the parameters of the multivariate Clayton survival copula-based SUR Tobit right-censored model.

4.2.1.1 Estimation through the (generalized) MIFM method

Following Trivedi & Zimmer (2005) and Anastasopoulos *et al.* (2012), we can write the log-likelihood function for the multivariate Clayton survival copula-based SUR Tobit right-censored model in the form

$$\begin{aligned}\ell\left(\boldsymbol{\eta}\right) = &\sum_{i=1}^{n} \log c\left(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \boldsymbol{\upsilon}_1\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \boldsymbol{\upsilon}_2\right), \ldots, F_m\left(y_{im}|\boldsymbol{x}_{im}, \boldsymbol{\upsilon}_m\right)|\theta\right) + \\ &+ \sum_{i=1}^{n}\sum_{j=1}^{m} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right),\end{aligned} \tag{4.7}$$

where $\boldsymbol{\eta} = \left(\boldsymbol{\upsilon}_1, \boldsymbol{\upsilon}_2, \ldots, \boldsymbol{\upsilon}_m, \theta\right)$ is the vector of model parameters, $f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$ is the p.d.f. of y_{ij}, and $c\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right)$, with $u_{ij} = F_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right)$, is the p.d.f. of the Clayton survival copula calculated from (4.6) as

$$\begin{aligned}c\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right) = &\frac{\partial^m C\left(u_{i1}, u_{i2}, \ldots, u_{im}|\theta\right)}{\partial u_{i1} \partial u_{i2} \ldots \partial u_{im}} = \\ &= \theta^m \frac{\Gamma\left(\frac{1}{\theta} + m\right)}{\Gamma\left(\frac{1}{\theta}\right)} \left[\prod_{j=1}^{m} \left(1 - u_{ij}\right)^{-\theta-1}\right] \left[\sum_{j=1}^{m} \left(1 - u_{ij}\right)^{-\theta} - m + 1\right]^{-\frac{1}{\theta}-m},\end{aligned}$$

which is similar to the p.d.f. of the Clayton copula (given by (4.3)).

Using copula methods, as well as the log-likelihood function form given by (4.7), enables the use of the (classical) two-stage ML/IFM method by Joe & Xu (1996), which estimates the marginal parameters $\boldsymbol{\upsilon}_j$ at a first step through

$$\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}} = \arg\max_{\boldsymbol{\upsilon}_j} \sum_{i=1}^{n} \log f_j\left(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\upsilon}_j\right), \tag{4.8}$$

for $j = 1, 2, \ldots, m$, and then estimates the association parameter θ given $\widehat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$ by

$$\widehat{\theta}_{\mathrm{IFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\big(F_1\left(y_{i1}|\boldsymbol{x}_{i1}, \widehat{\boldsymbol{\upsilon}}_{1,\mathrm{IFM}}\right), F_2\left(y_{i2}|\boldsymbol{x}_{i2}, \widehat{\boldsymbol{\upsilon}}_{2,\mathrm{IFM}}\right), \ldots, F_m\left(y_{im}|\boldsymbol{x}_{im}, \widehat{\boldsymbol{\upsilon}}_{m,\mathrm{IFM}}\right)|\theta\big). \tag{4.9}$$

Nevertheless, as seen in Sections 2.2.2.2 and 3.2.2.2, the IFM method provides a biased estimate for the parameter θ in the presence of censored observations in the margins. This occurs because there is a violation of Sklar's theorem in this case.

In order to obtain an unbiased estimate for the association parameter θ, we can augment the semi-continuous/censored marginal distributions to achieve continuity (and thus satisfy the Sklar's theorem!). More specifically, we replace y_{ij} by the augmented data y_{ij}^{a}, or equivalently and more simply, we replace u_{ij} by the augmented uniform data u_{ij}^{a} at the second stage of the IFM method and proceed with the copula parameter estimation as usual for the continuous margin cases. This process (uniform data augmentation and copula parameter estimation) is then repeated until convergence is achieved.

In the remaining part of this subsubsection, we discuss the proposed estimation method (a generalization of the MIFM method proposed in Sections 2.2.1.1 and 3.2.1.1) when using the Clayton survival copula to model the nonlinear dependence structure of the multivariate SUR Tobit right-censored model. Nevertheless, the proposed approach can be extended to other copula functions by applying different sampling algorithms.

Let margin j's error ϵ_{ij} have a standard distribution $H_j\left(.\right)$ and consider the lower bounds given by $a_{ij} = H_j\left(\left(d_j - \boldsymbol{x}_{ij}^{\prime}\hat{\boldsymbol{\beta}}_{j,\mathrm{MIFM}}\right) / \hat{\sigma}_{j,\mathrm{MIFM}}\right)$, for $j = 1, 2, \ldots, m$. The implementation of the multivariate Clayton survival copula-based SUR Tobit right-censored model through the proposed (generalized) MIFM method can be briefly described as follows.

Stage 1. Estimate the marginal parameters using (4.8). Set $\hat{\boldsymbol{\upsilon}}_{j,\mathrm{MIFM}} = \hat{\boldsymbol{\upsilon}}_{j,\mathrm{IFM}}$, i.e. $\left(\hat{\boldsymbol{\beta}}_{j,\mathrm{MIFM}}, \hat{\sigma}_{j,\mathrm{MIFM}}\right) = \left(\hat{\boldsymbol{\beta}}_{j,\mathrm{IFM}}, \hat{\sigma}_{j,\mathrm{IFM}}\right)$, for $j = 1, 2, \ldots, m$.

Stage 2. Estimate the copula parameter using, e.g., (4.9). Set $\hat{\theta}^{(1)}_{\text{MIFM}} = \hat{\theta}_{\text{IFM}}$ and then consider the algorithm below.

For $\omega = 1, 2, ...,$

For $i = 1, 2, ..., n,$

If $y_{i1} = d_1, y_{i2} = d_2, \ldots, y_{im} = d_m$, then draw $(u^{\text{a}}_{i1}, u^{\text{a}}_{i2}, \ldots, u^{\text{a}}_{im})$ from $C\Big(u^{\text{a}}_{i1}, u^{\text{a}}_{i2}, \ldots, u^{\text{a}}_{im} | \hat{\theta}^{(\omega)}_{\text{MIFM}}\Big)$ truncated to the region $(a_{i1}, 1) \times (a_{i2}, 1) \times \cdots \times (a_{im}, 1)$. This can be performed relatively easily using the truncation dependence invariance property of the (multidimensional) Clayton survival copula.

If $y_{i1} = d_1, \ldots, y_{i,s-1} = d_{s-1}, y_{i,s+1} = d_{s+1}, \ldots, y_{im} = d_m$, and $y_{is} < d_s$, then draw $\left(u^{\text{a}}_{i1}, \ldots, u^{\text{a}}_{i,s-1}, u^{\text{a}}_{i,s+1}, \ldots, u^{\text{a}}_{im}\right)$ from $C\left(u^{\text{a}}_{i1}, \ldots, u^{\text{a}}_{i,s-1}, u^{\text{a}}_{i,s+1}, \ldots, u^{\text{a}}_{im} | u_{is}, \hat{\theta}^{(\omega)}_{\text{MIFM}}\right)$ truncated to the region $(a_{i1}, 1) \times \cdots \times (a_{i,s-1}, 1) \times (a_{i,s+1}, 1) \times \cdots \times (a_{im}, 1)$. This can be performed through iterative conditioning (conditional sampling) by successive application of the method by Devroye (1986, p. 38-39).

$\vdots$

If $y_{is} = d_s$ and $y_{i1} < d_1, \ldots, y_{i,s-1} < d_{s-1}, y_{i,s+1} < d_{s+1}, \ldots, y_{im} < d_m$, then draw u^{a}_{is} from $C\left(u^{\text{a}}_{is} | u_{i1}, \ldots, u_{i,s-1}, u_{i,s+1}, \ldots, u_{im}, \hat{\theta}^{(\omega)}_{\text{MIFM}}\right)$ truncated to the interval $(a_{is}, 1)$. This can be done by applying the method by Devroye (1986, p. 38-39).

If $y_{i1} < d_1, y_{i2} < d_2, \ldots, y_{im} < d_m$, then set $u^{\text{a}}_{ij} = u_{ij} = H_j\left(\left(y_{ij} - \boldsymbol{x}'_{ij}\hat{\boldsymbol{\beta}}_{j,\text{MIFM}}\right) / \hat{\sigma}_{j,\text{MIFM}}\right)$, for $j = 1, 2, \ldots, m$.

Given the generated/augmented marginal uniform data u^{a}_{ij}, we estimate the association parameter θ by

$$\hat{\theta}^{(\omega+1)}_{\text{MIFM}} = \arg\max_{\theta} \sum_{i=1}^{n} \log c\left(u^{\text{a}}_{i1}, u^{\text{a}}_{i2}, \ldots, u^{\text{a}}_{im} | \theta\right).$$

The algorithm terminates when it satisfies the stopping/convergence criterion: $|\hat{\theta}^{(\omega+1)}_{\text{MIFM}} - \hat{\theta}^{(\omega)}_{\text{MIFM}}| < \xi$, where ξ is the tolerance parameter.

4.2.1.2 Interval estimation

We can build confidence intervals for the parameters of the multivariate Clayton survival copula-based SUR Tobit right-censored model using the same bootstrap approach (a parametric resampling plan, standard normal and basic bootstrap methods) as described in Section 4.1.1.2.

4.3 Final remarks

In this chapter, we presented a straightforward generalization of the models and methods proposed in the previous chapters for the multivariate setting.

Regarding model estimation, we only gave some general guidelines to implement the multivariate copula-based SUR Tobit models (multivariate Clayton copula-based SUR Tobit model and multivariate Clayton survival copula-based SUR Tobit right-censored model) through the proposed (generalized) MIFM method.

The multidimensional generalizations of the Clayton and Clayton survival copulas that we considered here are the simplest ones and they present the whole dependence structure with only one single copula parameter θ, independent of the dimension of the model. Consequently, the substructure of the dependence is hidden/invisible. Moreover, they implicitly assume the exchangeability of the order of the marginal distributions within the copula functions, which is very restrictive for many applications (cf. McNeil *et al.*, 2005, p. 224; Savu & Trede, 2010). In view of these limitations, we could employ more flexible methods, like the hierarchical Archimedean copula (HAC), discussed by Joe (1997), Embrechts *et al.* (2003), Whelan (2004), Savu & Trede (2010) and Okhrin *et al.* (2013). In contrast to the usual Archimedean copula, the HAC defines the whole dependence structure in a recursive way, i.e. by aggregating one dimension step by step starting from a low-dimensional copula.

Chapter 5

Conclusions

In Section 5.1 of this last chapter, we summarize our main results. Moreover, since during the course of our work we identified open problems and possible extensions of our results, in Section 5.2 we suggest potential topics for further researches.

5.1 Concluding remarks

The starting point of this thesis was the bivariate SUR Tobit model. We extended the analysis of the SUR Tobit model with two left-censored or right-censored dependent variables by modeling its nonlinear dependence structure through copulas and assuming non-normal marginal error distributions. Our decision for two parametric families of copula (Clayton copula for the bivariate SUR Tobit model, and Clayton survival copula for the bivariate SUR Tobit right-censored model), as well as non-normal (power-normal and logistic) distribution assumption for the marginal error terms, were mainly motivated by the real data at hand (U.S. consumption data and Brazilian commercial bank customer churn data). The ability to capture/model the tail dependence, especially the lower (case of the Clayton copula) or upper (case of the Clayton survival copula) tail where some data are censored, is one of the attractive features of copulas.

Since some most commonly used classical procedures for bivariate copula-based model implementation (the IFM method, proposed by Joe & Xu (1996)) and interval estimation using resampling techniques (delete-one jackknife method - normal approach), are troublesome in the cases where both margins are censored/semi-continuous (the IFM method results in a biased estimate of the copula association parameter, and the jackknife method overestimates the standard error of the copula association parameter estimate), our study used a (frequentist) data augmentation technique to generate the unobserved/censored values (thus obtaining continuous margins) and proceeded with the bivariate copula-based SUR Tobit model implementation through the proposed MIFM estimation method (which is a modified version of the IFM method). The MIFM method, as well as the IFM method, is more computationally attractive (feasible) than the full maximum likelihood approach, since each maximization step has a small number of parameters, which reduces the computational difficulty. Moreover, the two-stage procedure is considerably less time consuming than its one-stage counterpart. Here, some advantages arose from our copula choices, regarding the development of the MIFM method for obtaining the estimates of the bivariate models' parameters. First, the Clayton copula and its survival copula are known to be preserved under truncation (truncation dependence invariance property), which enabled simple simulation schemes in the cases where both dependent variables/margins were censored (for copulas that do not have the truncation-invariance property, an iterative simulation scheme could be used). Second, the existence of closed-form expressions for the inverse of the conditional Clayton (see, e.g., Armstrong, 2003) and Clayton survival (see Appendix A) copulas' distributions enabled simple simulation schemes when just a single dependent variable/margin was censored, by applying the method by Devroye (1986, p. 38-39) (if the inverse conditional distribution of the copula used does not have a closed-form expression, then numerical root-finding procedures are required). We also proposed the use of bootstrap methods (standard normal and percentile) for obtaining confidence intervals for the model parameters.

In the simulation studies, we assessed the performance of our proposed bivariate models and methods, obtaining satisfactory results (unbiased estimates of the copula parameter, high and near the nominal value coverage probabilities of the bootstrap-based confidence intervals) regardless of the error distribution assumption, the censoring percentage in the margins and their degree of interdependence.

We also pointed out the applicability of our proposed bivariate models and methods for real data sets, where we found that the gain for introducing the copulas was substantial for these datasets.

Although it is relatively rare to analyze the SUR Tobit with over two dimensions, unless it is modeled in the longitudinal setting (see, e.g., Baranchuk & Chib (2008) for an example of the longitudinal Tobit model), our proposed models and methods were successfully extended/applied to high-dimensional SUR Tobit models.

5.2 Further researches

The topics addressed in this work open some potential subjects for further researches. Firstly, we considered only the same normally-, power-normally- or logistically-distributed marginal errors. However, the flexibility in coupling different marginal distributions is an important feature of copulas in general. It would allow us to apply not necessarily the same, as well as many other distributions for the multivariate SUR Tobit models' marginal errors, e.g., scale mixtures of normal (SMN) distributions, as proposed in Garay (2014). The Student-t, Pearson type VII, slash, contaminated normal, among others distributions, are contained in this class of symmetric distributions. We could also use other copula families exhibiting left tail dependence, like the Gumbel survival copula, the copula of equation (4.2.12) of Nelsen's book (see Nelsen, 2006, p. 116) and the Student-t copula, in addition to the Clayton copula; as well as other copula families exhibiting upper tail dependence, like the Gumbel and Student-t copulas, in addition to the Clayton survival copula. Since these copulas do not have neither the truncation-invariance property nor closed-form expression for the inverse conditional distribution, iterative simulation schemes and numerical root-finding procedures are required when using the MIFM approach. These consist in the subjects to our further study.

The copulas used in this work were found to be acceptable by visual inspection of the data. However, a formal way to evaluate the appropriateness or adequacy of a model is using goodness-of-fit tests. Thus, the derivation of goodness-of-fit tests for copula models in the framework of SUR models with limited dependent variables will be the subject of our future research. Furthermore, the multidimensional generalizations of the copulas that we considered in this work are the simplest ones, presenting the whole complex multivariate dependence structure with only one single copula parameter θ, independent of the dimension of the model. This is certainly not an acceptable assumption in many practical applications. In order to consider/assume more flexible dependence structures, we could use the hierarchical Archimedean copulas (HACs), discussed by Joe (1997), Embrechts *et al.* (2003), Whelan (2004), Savu & Trede (2010) and Okhrin *et al.* (2013). In contrast to the usual Archimedean copulas, the HACs define the whole dependence structure in a recursive way, i.e. by aggregating one dimension step by step starting from a low-dimensional copula. Therefore, we leave to further research the issue of extending the MIFM approach to the multivariate setting for multiparameter copulas.

We also propose in future work to perform misspecification studies in order to verify if we can distinguish among the copula-based SUR Tobit models with arbitrary margins in the light of the data based on some model selection criteria such as AIC and BIC.

Finally, we leave to future studies the derivation of other asymptotic properties (such as asymptotic normality) for the copula parameter estimate obtained through the MIFM method in our framework of SUR models with limited (partially observed or left- and/or right-censored) dependent variables.

Appendix A

The multidimensional or m-dimensional ($m \geq 2$) Clayton survival copula with parameter $\theta > 0$ takes the form

$$C_m\left(u_1, u_2, \ldots, u_m | \theta\right) = 1 - \sum_{j=1}^{m} \left(1 - u_j\right) + \sum_{S \subset \{1,\ldots,m\}, |S| \geq 2} (-1)^{|S|}\, C_{|S|}^{Clayton}\left(1 - u_l, l \in S | \theta\right)$$

(Joe, 2014, p. 28), where $|S|$ is the cardinality of S and $C_{|S|}^{Clayton}$ denotes the $|S|$-dimensional Clayton copula.

The following algorithm generates a random variate $(u_1, u_2, \ldots, u_m)$ from the Clayton survival copula.

- Simulate m independent random variables $(v_1, v_2, \ldots, v_m)$ from $Uniform\,(0, 1)$.

- Set $u_1 = v_1$.

- Set $v_2 = C_2\left(u_2 | u_1, \theta\right)$, hence

  $$v_2 = \frac{\partial C_2\left(u_1, u_2 | \theta\right)}{\partial u_1} = 1 - \left[\frac{\left(1 - u_1\right)^{-\theta} + \left(1 - u_2\right)^{-\theta} - 1}{\left(1 - u_1\right)^{-\theta}}\right]^{-\frac{1}{\theta} - 1}.$$

  Finally,

  $$u_2 = C_2^{-1}\left(v_2 | u_1, \theta\right) = 1 - \left\{1 + \left(1 - u_1\right)^{-\theta}\left[\left(1 - v_2\right)^{-\frac{\theta}{\theta+1}} - 1\right]\right\}^{-\frac{1}{\theta}}.$$

- Set

  $$\begin{aligned} v_3 = C_3\left(u_3 | u_1, u_2, \theta\right) &= \frac{\partial^2 C_3\left(u_1, u_2, u_3 | \theta\right) / \partial u_1 \partial u_2}{\partial^2 C_2\left(u_1, u_2 | \theta\right) / \partial u_1 \partial u_2} = \\ &= 1 - \left[\frac{\left(1 - u_1\right)^{-\theta} + \left(1 - u_2\right)^{-\theta} + \left(1 - u_3\right)^{-\theta} - 2}{\left(1 - u_1\right)^{-\theta} + \left(1 - u_2\right)^{-\theta} - 1}\right]^{-\frac{1}{\theta} - 2} \end{aligned}$$

  and solve it in u_3.

- $\ldots$

- Solve in u_m the equation

$$v_m = 1 - \left[\frac{(1-u_1)^{-\theta} + (1-u_2)^{-\theta} + \cdots + (1-u_m)^{-\theta} - m + 1}{(1-u_1)^{-\theta} + (1-u_2)^{-\theta} + \cdots + (1-u_{m-1})^{-\theta} - m + 2} + \right]^{-\frac{1}{\theta} - m + 1},$$

so we have

$$\begin{aligned} u_m = 1 - & \Big\{ 1 + \Big[(1-u_1)^{-\theta} + (1-u_2)^{-\theta} + \cdots + (1-u_{m-1})^{-\theta} - m + 2 \Big] \times \\ & \times \Big[(1-v_m)^{\frac{\theta}{\theta(1-m)-1}} - 1 \Big] \Big\}^{-\frac{1}{\theta}}. \end{aligned}$$

Appendix B

In this appendix, we include the R codes that were used in the bivariate examples throughout the thesis. To avoid repetition, only the R codes for the best fittings (according to the AIC and BIC criterion) are presented.

B.1. U.S. salad dressing and lettuce consumption data

```
######### Functions to fit the bivariate Clayton copula-based SUR Tobit model with logistic marginal errors
######### to the salad dressing and lettuce consumption data using the MIFM method, as well as to build
######### confidence intervals through the standard normal and percentile bootstrap methods

##### Load required R packages

library(”AER”)
library(”stats”)
library(compiler)
enableJIT(3)

##### Create/define the following functions in R

#### Step 1: defining the components of the loglikelihood - tobit margins and copula

dtobito=function(theta,y,x){
  l=length(theta)
  n=length(y)
  I=rep(1,n)
  for(i in 1:n){
    if(y[i]==0) I[i]=0
  }
  f=dlogis(y, location=x%*%theta[-l], scale=theta[l], log=FALSE)
  F=plogis(0, location=x%*%theta[-l], scale=theta[l], lower.tail=TRUE, log.p=FALSE)
  (f^I)*(F^(1-I))
}

loglik.tobito=function(theta,y,x){
  sum(log(dtobito(theta,y,x)))
}

loglik.cop=function(a,u){
  somalog=0
  for(i in 1:nrow(u)){
    somalog=somalog+(log(a+1)-(a+1)*(log(u[i,1])+log(u[i,2]))-((2*a+1)/a)*log(u[i,1]^(-a)+u[i,2]^(-a)-1))
  }
  somalog
}

#### Step 2: calculating the probability integral transformed margins

ptobito=function(theta,y,x){
  l=length(theta)
  n=length(y)
  acum=numeric(n)
  for(i in 1:n){
    if(y[i]==0) acum[i]=plogis(0, location=x[i,]%*%theta[-l], scale=theta[l], lower.tail=TRUE, log.p=FALSE)
    else acum[i]=plogis(y[i], location=x[i,]%*%theta[-l], scale=theta[l], lower.tail=TRUE, log.p=FALSE)
  }
  acum
}

probtrans=function(theta,y,x){
  ptobito(theta,y,x)
}

#### Step 3: composing the loglikelihood function

myloglik=function(thetas,y,xmat){
  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  theta1=thetas[1:l1]
  theta2=thetas[(l1+1):(l1+l2)]
  a=thetas[-(1:(l1+l2))]
  u=cbind(probtrans(theta1,y[,1],xmat[[1]]), probtrans(theta2,y[,2],xmat[[2]]))
  loglik=loglik.tobito(theta1,y[,1],xmat[[1]])+loglik.tobito(theta2,y[,2],xmat[[2]])+loglik.cop(a,u)
  loglik
}

#### Step 4: defining a function to generate response variables from given parameter vector, design matrices and
#### copula structure

qtobito=function(theta,p,x){
  l=length(theta)
  n=length(p)
  acum0=numeric(n)
  quan=numeric(n)
  for(i in 1:n){
    acum0[i]=plogis(0, location=x[i,]%*%theta[-l], scale=theta[l], lower.tail=TRUE, log.p=FALSE)
    if(p[i]<=acum0[i]) quan[i]=0
    else quan[i]=qlogis(p[i], location=x[i,]%*%theta[-l], scale=theta[l], lower.tail=TRUE, log.p=FALSE)
  }
  quan
}

rCCopula=function(n,a){
  u=runif(n)
  t=runif(n)
  v=((t^(-a/(a+1))-1)*(u^(-a))+1)^(-1/a)
  cbind(u,v)
}

genY=function(thetas,xmat){
  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  theta1=thetas[1:l1]
  theta2=thetas[(l1+1):(l1+l2)]
  a=thetas[-(1:(l1+l2))]
  n=nrow(xmat[[1]])
  u=rCCopula(n,a)
  y1=qtobito(theta1, u[,1], xmat[[1]])
  y2=qtobito(theta2, u[,2], xmat[[2]])
  cbind(y1,y2)
}

mifm=function(theta1,theta2,a,y,ua,xmat) {

  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  n=nrow(xmat[[1]])
  ll1=length(theta1)
  ll2=length(theta2)

  y4=y
  erro=TRUE

  while(erro) {

    erro=FALSE

    for (i in 1:n){

      if(y[i,1]==0 && y[i,2]==0){
        data1=rCCopula(1,a)
        p=data1[,1]
        q=data1[,2]
        b1=plogis(0, location=xmat[[1]][i,]%∗%theta1[−ll1], scale=theta1[ll1], lower.tail=TRUE, log.p=FALSE)
        b2=plogis(0, location=xmat[[2]][i,]%∗%theta2[−ll2], scale=theta2[ll2], lower.tail=TRUE, log.p=FALSE)
        ua[i,1]=((b1ˆ(−a)+b2ˆ(−a)−1)∗(pˆ(−a))+1−(b2ˆ(−a)))ˆ(−1/a)
        ua[i,2]=((b1ˆ(−a)+b2ˆ(−a)−1)∗(qˆ(−a))+1−(b1ˆ(−a)))ˆ(−1/a)
        y4[i,1]=xmat[[1]][i,]%∗%theta1[−ll1]+theta1[ll1]∗qlogis(ua[i,1], location=0, scale=1, lower.tail=TRUE, log.p=FALSE)
        y4[i,2]=xmat[[2]][i,]%∗%theta2[−ll2]+theta2[ll2]∗qlogis(ua[i,2], location=0, scale=1, lower.tail=TRUE, log.p=FALSE)
      }

      if(y[i,1]==0 && y[i,2]>0){
        u2=plogis(y[i,2], location=xmat[[2]][i,]%∗%theta2[−ll2], scale=theta2[ll2], lower.tail=TRUE, log.p=FALSE)
        b1=plogis(0, location=xmat[[1]][i,]%∗%theta1[−ll1], scale=theta1[ll1], lower.tail=TRUE, log.p=FALSE)
        v1=runif(1)∗((b1ˆ(−a)+u2ˆ(−a)−1)ˆ(−(a+1)/a))∗(u2ˆ(−a−1))
        ua[i,1]=((v1ˆ(−a/(a+1))−1)∗(u2ˆ(−a))+1)ˆ(−1/a)
        y4[i,1]=xmat[[1]][i,]%∗%theta1[−ll1]+theta1[ll1]∗qlogis(ua[i,1], location=0, scale=1, lower.tail=TRUE, log.p=FALSE)
      }

      if(y[i,1]>0 && y[i,2]==0){
        u1=plogis(y[i,1], location=xmat[[1]][i,]%∗%theta1[−ll1], scale=theta1[ll1], lower.tail=TRUE, log.p=FALSE)
        b2=plogis(0, location=xmat[[2]][i,]%∗%theta2[−ll2], scale=theta2[ll2], lower.tail=TRUE, log.p=FALSE)
        v2=runif(1)∗((b2ˆ(−a)+u1ˆ(−a)−1)ˆ(−(a+1)/a))∗(u1ˆ(−a−1))
        ua[i,2]=((v2ˆ(−a/(a+1))−1)∗(u1ˆ(−a))+1)ˆ(−1/a)
        y4[i,2]=xmat[[2]][i,]%∗%theta2[−ll2]+theta2[ll2]∗qlogis(ua[i,2], location=0, scale=1, lower.tail=TRUE, log.p=FALSE)
      }

    }

    udat2=ua
    y5=y4

    fit.ifm2=try(optim(a0, fn=loglik.cop, u=udat2, method=”L−BFGS−B”, lower=0.0001, upper=Inf, control=list(fnscale=−1,
                maxit=100000)), TRUE)

    if(inherits(fit.ifm2,”try−error”)){erro=TRUE}

  } #end while(erro)

  aa=fit.ifm2$par
  saida=list()
  saida[[1]]=aa; saida[[2]]=udat2; saida[[3]]=y5
  saida

}

##### Import a local txt file named consumo.txt (salad dressing and lettuce consumption dataset)

dados=read.table(”C:\\Users\\Aluno\\Desktop\\Dados\\consumo.txt”, header = TRUE)

n = nrow(dados)

attach(dados)

sex=factor(SEX, levels=c(2,1))
race=factor(RACEN, levels=c(1,2,3,4))
pctpov=PCTPOVN
fat2=FAT2N
veg5=VEG5N
region=factor(REGION, levels=c(3,1,2,4))
age=factor(AGEN, levels=c(5,1,2,3,4))

summary(pctpov); sd(pctpov)
summary(fat2); sd(fat2)
summary(veg5); sd(veg5)
summary(fat2[which(fat2>0)]); sd(fat2[which(fat2>0)])
summary(veg5[which(veg5>0)]); sd(veg5[which(veg5>0)])

length(which(sex==1))/n
length(which(age==1))/n
length(which(age==2))/n
length(which(age==3))/n
length(which(age==4))/n
length(which(age==5))/n
length(which(region==1))/n
length(which(region==2))/n
length(which(region==3))/n
length(which(region==4))/n

par(mfrow=c(1,2))
hist(fat2, main=””, xlab=”Quantity (100 g)”, freq=FALSE)
hist(veg5, main=””, xlab=”Quantity (200 g)”, freq=FALSE)

tau=cor(cbind(fat2,veg5),method=”kendall”)[1,2]
a0=2∗tau/(1−tau)

k=21

B=500

### Fit the bivariate Clayton copula−based SUR Tobit model with logistic marginal errors to the salad dressing and
### lettuce consumption data

xmat=list(model.matrix(˜age+region+pctpov), model.matrix(˜age+region+pctpov))

y=cbind(fat2, veg5)

# censoring percentage in the margins
cont1=0
cont2=0
for(i in 1:n){
  if(y[i,1]==0) cont1=cont1+1
  if(y[i,2]==0) cont2=cont2+1
}
cens1=cont1/n
cens2=cont2/n

# two−stage parametric ML method − IFM method − by Joe and Xu (1996)

# stage 1
tobito1=tobit(y[,1]˜xmat[[1]][,−1],left=0,right=Inf,dist=”logistic”)
est1=summary(tobito1)$coefficients
theta1hat=c(est1[1,1], est1[2,1], est1[3,1], est1[4,1], est1[5,1], est1[6,1], est1[7,1], est1[8,1], est1[9,1], exp(est1[10,1]))

tobito2=tobit(y[,2]˜xmat[[2]][,−1],left=0,right=Inf,dist=”logistic”)
est2=summary(tobito2)$coefficients
theta2hat=c(est2[1,1], est2[2,1], est2[3,1], est2[4,1], est2[5,1], est2[6,1], est2[7,1], est2[8,1], est2[9,1], exp(est2[10,1]))

par(mfrow=c(1,2))

# scatter plot of y1 versus y2
plot(y[,1],y[,2])

# stage 2
udat=cbind(probtrans(theta1hat,y[,1],xmat[[1]]), probtrans(theta2hat,y[,2],xmat[[2]]))

# scatter plot of udat[,1] versus udat[,2]
plot(udat[,1],udat[,2], xlab=expression(u[1]), ylab=expression(u[2]))

fit.ifm=optim(a0, fn=loglik.cop, u=udat, method=”L−BFGS−B”, lower=0.0001, upper=Inf, control=list(fnscale=−1,
  maxit=100000))
thetas.ifm=c(theta1hat, theta2hat, fit.ifm$par)

thetas.est=thetas.ifm

# two−stage parametric ML method − IFM method − by Joe and Xu (1996) with augmented data (MIFM method)

# stage 2
ua1=udat

l1=ncol(xmat[[1]])+1
l2=ncol(xmat[[2]])+1
```

```

theta1=thetas.ifm[1:l1]
theta2=thetas.ifm[(l1+1):(l1+l2)]

phi=numeric()
loglike=numeric()

phi[1]=fit.ifm$par

parm.margins=c(theta1,theta2)

loglike[1]=myloglik(c(parm.margins,phi[1]),y,xmat)

out=mifm(theta1,theta2,phi[1],y,ua1,xmat)
phi[2]=out[[1]]
ua11=out[[2]]
ya=out[[3]]

loglike[2]=myloglik(c(parm.margins,phi[2]),y,xmat)

w=1

eps=0.001

while (abs(phi[w+1]−phi[w]) >= eps){
  out2=mifm(theta1,theta2,phi[w+1],y,ua1,xmat)
  phi[w+2]=out2[[1]]
  ua11=out2[[2]]
  ya=out2[[3]]
  loglike[w+2]=myloglik(c(parm.margins,phi[w+2]),y,xmat)
  w=w+1
}

niter=length(phi)
phi.est.mifm = phi[niter]

plot(phi, xlab=”Iteration”, ylab=expression(hat(theta)[MIFM]), type=”b”)
plot(loglike, xlab=”Iteration”, ylab=”Log−likelihood”, type=”b”)

thetas.ifm2=c(theta1, theta2, phi.est.mifm)

thetas.est2=thetas.ifm2

# histograms of y1 and y2 (augmented data)
hist(ya[,1], main=” ”, xlab=”Quantity (100 g)”, freq=FALSE); abline(v=0,lty=2)
hist(ya[,2], main=” ”, xlab=”Quantity (200 g)”, freq=FALSE); abline(v=0,lty=2)

# scatter plot of y1 versus y2 (augmented data)
plot(ya[,1],ya[,2], xlab=”Salad dressings (100 g)”, ylab=”Lettuce (200 g)”); abline(h=0, v=0,lty=2)

# scatter plot of u1 versus u2 (augmented data)
plot(ua11[,1],ua11[,2], xlab=expression(u[1]), ylab=expression(u[2]))

# kolmogorov−smirnov tests of augmented marginal residuals
res1=ya[,1]−xmat[[1]]%∗%theta1[−l1]
res2=ya[,2]−xmat[[2]]%∗%theta2[−l2]

hist(res1, main=””, xlab=”Residuals”, freq=FALSE); hist(res2, main=””, xlab=”Residuals”, freq=FALSE)

ks.test(res1, ”plogis”, mean(res1), theta1hat[10]); ks.test(res2, ”plogis”, mean(res2), theta2hat[10])

# AIC and BIC criterion values
AIC=−2∗loglike[niter]+2∗k
BIC=−2∗loglike[niter]+k∗log(n)

# Parametric bootstrap approach: generate y1 and y2 values using thetas.ifm2 in genY() function

thetas.boot = matrix(numeric(k),B,k)
niter.boot=numeric(B)
phi.est.mifm.boot=numeric(B)

for(b in 1:B){

  y.boot=genY(thetas.ifm2,xmat)

  # two-stage parametric ML method - IFM method - by Joe and Xu (1996)

  # stage 1
  tobito1.boot=tobit(y.boot[,1]~xmat[[1]][,-1],left=0,right=Inf,dist="logistic")
  est1.boot=summary(tobito1.boot)$coefficients
  theta1hat.boot=c(est1.boot[1,1], est1.boot[2,1], est1.boot[3,1], est1.boot[4,1], est1.boot[5,1], est1.boot[6,1], est1.boot[7,1],
        est1.boot[8,1], est1.boot[9,1], exp(est1.boot[10,1]))

  tobito2.boot=tobit(y.boot[,2]~xmat[[2]][,-1],left=0,right=Inf,dist="logistic")
  est2.boot=summary(tobito2.boot)$coefficients
  theta2hat.boot=c(est2.boot[1,1], est2.boot[2,1], est2.boot[3,1], est2.boot[4,1], est2.boot[5,1], est2.boot[6,1], est2.boot[7,1],
        est2.boot[8,1], est2.boot[9,1], exp(est2.boot[10,1]))

  # stage 2
  udat.boot=cbind(probtrans(theta1hat.boot,y.boot[,1],xmat[[1]]), probtrans(theta2hat.boot,y.boot[,2],xmat[[2]]))

  fit.ifm.boot=optim(a0, fn=loglik.cop, u=udat.boot, method="L-BFGS-B", lower=0.0001, upper=Inf, control=list(fnscale=-1,
        maxit=100000))
  thetas.ifm.boot=c(theta1hat.boot, theta2hat.boot, fit.ifm.boot$par)

  # two-stage parametric ML method - IFM method - by Joe and Xu (1996) with augmented data (MIFM method)

  # stage 2

  ua.boot=udat.boot

  theta1.boot=thetas.ifm.boot[1:l1]
  theta2.boot=thetas.ifm.boot[(l1+1):(l1+l2)]

  phi.boot=numeric()
  loglike.boot=numeric()

  phi.boot[1]=fit.ifm.boot$par

  parm.margins.boot=c(theta1.boot,theta2.boot)

  loglike.boot[1]=myloglik(c(parm.margins.boot,phi.boot[1]),y.boot,xmat)

  out.boot=mifm(theta1.boot,theta2.boot,phi.boot[1],y.boot,ua.boot,xmat)
  phi.boot[2]=out.boot[[1]]
  ua11.boot=out.boot[[2]]

  loglike.boot[2]=myloglik(c(parm.margins.boot,phi.boot[2]),y.boot,xmat)

  w=1

  while (abs(phi.boot[w+1]-phi.boot[w]) >= eps){
    out2.boot=mifm(theta1.boot,theta2.boot,phi.boot[w+1],y.boot,ua.boot,xmat)
    phi.boot[w+2]=out2.boot[[1]]
    ua11.boot=out2.boot[[2]]
    loglike.boot[w+2]=myloglik(c(parm.margins.boot,phi.boot[w+2]),y.boot,xmat)
    w=w+1
  }

  niter.boot[b]=length(phi.boot)
  phi.est.mifm.boot[b]=phi.boot[niter.boot[b]]
  thetas.ifm2.boot=c(theta1.boot, theta2.boot, phi.est.mifm.boot[b])
  thetas.boot[b,]=thetas.ifm2.boot

  print(b)

}

# Bootstrap confidence intervals

# Standard normal interval

cov.boot=matrix(numeric(k),k,k)
mean.boot=as.matrix(apply(thetas.boot, 2, mean), k, 1, byrow=TRUE)

for(b in 1:B){
  thetas.boot2=as.matrix(thetas.boot[b,], k, 1, byrow=TRUE)
  cov.boot=cov.boot+((thetas.boot2-mean.boot)%*%t(thetas.boot2-mean.boot))
}

cov.boot=(1/(B-1))*cov.boot
var.boot=diag(cov.boot); se.boot=sqrt(var.boot)
inf4=thetas.ifm2-1.645*se.boot; sup4=thetas.ifm2+1.645*se.boot

# Percentile interval

inf5=numeric(k); sup5=numeric(k)

for(j in 1:k){
  percentis=quantile(thetas.boot[,j], probs=c(0.05,0.95))
  inf5[j]=percentis[[1]]; sup5[j]=percentis[[2]]
}
```

B.2. Brazilian commercial bank customer churn data (Products A and B)

```
######### Functions to fit the bivariate Clayton survival copula-based SUR Tobit right-censored model with
######### normal marginal errors to the customer churn data (Products A and B) using the MIFM method, as well
######### as to build confidence intervals through the standard normal and percentile bootstrap methods

##### Load required R packages

library("AER")
library("nortest")
library(compiler)
enableJIT(3)

##### Create/define the following functions in R

#### Step 1: defining the components of the loglikelihood - tobit margins and copula

dtobito=function(theta,y,x,d){
  l=length(theta)
  n=length(y)
  I=rep(1,n)
  for(i in 1:n){
    if(y[i]>=d) I[i]=0
  }
  f=1/theta[l]*dnorm((y-(x%*%theta[-l]))/theta[l], mean=0, sd=1, log=FALSE)
  S=1-pnorm((d-x%*%theta[-l])/theta[l], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
  (f^I)*(S^(1-I))
}

loglik.tobito=function(theta,y,x,d){
  sum(log(dtobito(theta,y,x,d)))
}

loglik.cop=function(a,u){
  somalog=0
  for(i in 1:nrow(u)){
    somalog=somalog+(log(a+1)-(a+1)*log(1-u[i,1])-(a+1)*log(1-u[i,2])-((2*a+1)/a)*log((1-u[i,1])^(-a)+(1-u[i,2])^(-a)-1))
  }
  somalog
}

#### Step 2: calculating the probability integral transformed margins

ptobito=function(theta,y,x,d){
  l=length(theta)
  n=length(y)
  acum=numeric(n)
  for(i in 1:n){
    if(y[i]>=d) acum[i]=pnorm((d-(x[i,]%*%theta[-l]))/theta[l], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
    else acum[i]=pnorm((y[i]-(x[i,]%*%theta[-l]))/theta[l], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
  }
  acum
```

```
}

probtrans=function(theta,y,x,d){
  ptobito(theta,y,x,d)
}

#### Step 3: composing the loglikelihood function

myloglik=function(thetas,y,xmat,d1,d2){
  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  theta1=thetas[1:l1]
  theta2=thetas[(l1+1):(l1+l2)]
  a=thetas[-(1:(l1+l2))]
  u=cbind(probtrans(theta1,y[,1],xmat[[1]],d1), probtrans(theta2,y[,2],xmat[[2]],d2))
  loglik=loglik.tobito(theta1,y[,1],xmat[[1]],d1)+loglik.tobito(theta2,y[,2],xmat[[2]],d2)+loglik.cop(a,u)
  loglik
}

#### Step 4: defining a function to generate response variables from given parameter vector, design matrices and
#### copula structure

qtobito=function(theta,p,x,d){
  l=length(theta)
  n=length(p)
  acum0=numeric(n)
  quan=numeric(n)
  for(i in 1:n){
    acum0[i]=pnorm((d-(x[i,]%*%theta[-l]))/theta[l], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
    if(p[i]>=acum0[i]) quan[i]=d
    else quan[i]=qnorm(p[i], mean=x[i,]%*%theta[-l], sd=theta[l], lower.tail=TRUE, log.p=FALSE)
  }
  quan
}

rCrCopula=function(n,a){
  u=runif(n)
  t=runif(n)
  v=1-((1+(((1-u)^(-a))*((1-t)^(-a/(a+1))-1)))^(-1/a))
  cbind(u,v)
}

genY=function(thetas,xmat,d1,d2){
  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  theta1=thetas[1:l1]
  theta2=thetas[(l1+1):(l1+l2)]
  a=thetas[-(1:(l1+l2))]
  n=nrow(xmat[[1]])
  u=rCrCopula(n,a)
  y1=qtobito(theta1, u[,1], xmat[[1]], d1)
  y2=qtobito(theta2, u[,2], xmat[[2]], d2)
  cbind(y1,y2)
}

mifm=function(theta1,theta2,a,y,ua,xmat,d1,d2) {

  l1=ncol(xmat[[1]])+1
  l2=ncol(xmat[[2]])+1
  n=nrow(xmat[[1]])
  ll1=length(theta1)
  ll2=length(theta2)

  y4=y
  erro=TRUE

  while(erro) {

    erro=FALSE

    for (i in 1:n){

      if(y[i,1]==d1 && y[i,2]==d2){
        data1=rCrCopula(1,a)
        p=data1[,1]
        q=data1[,2]
        a1=pnorm((d1−(xmat[[1]][i,]%∗%theta1[−ll1]))/theta1[ll1], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        a2=pnorm((d2−(xmat[[2]][i,]%∗%theta2[−ll2]))/theta2[ll2], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        ua[i,1]=1−(((1−a1)ˆ(−a)+(1−a2)ˆ(−a)−1)∗((1−p)ˆ(−a))+1−((1−a2)ˆ(−a)))ˆ(−1/a)
        ua[i,2]=1−(((1−a1)ˆ(−a)+(1−a2)ˆ(−a)−1)∗((1−q)ˆ(−a))+1−((1−a1)ˆ(−a)))ˆ(−1/a)
        y4[i,1]=xmat[[1]][i,]%∗%theta1[−ll1]+theta1[ll1]∗qnorm(ua[i,1], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        y4[i,2]=xmat[[2]][i,]%∗%theta2[−ll2]+theta2[ll2]∗qnorm(ua[i,2], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
      }

      if(y[i,1]==d1 && y[i,2]<d2){
        u2=pnorm((y[i,2]−(xmat[[2]][i,]%∗%theta2[−ll2]))/theta2[ll2], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        a1=pnorm((d1−(xmat[[1]][i,]%∗%theta1[−ll1]))/theta1[ll1], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        U=runif(1)
        v1=U+(1−U)∗(1−((1−u2)ˆ(−(a+1)))∗(((1−u2)ˆ(−a)+(1−a1)ˆ(−a)−1)ˆ(−(a+1)/a)))
        ua[i,1]=1−((1+(((1−u2)ˆ(−a))∗((1−v1)ˆ(−a/(a+1))−1)))ˆ(−1/a))
        y4[i,1]=xmat[[1]][i,]%∗%theta1[−ll1]+theta1[ll1]∗qnorm(ua[i,1], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
      }

      if(y[i,1]<d1 && y[i,2]==d2){
        u1=pnorm((y[i,1]−(xmat[[1]][i,]%∗%theta1[−ll1]))/theta1[ll1], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        a2=pnorm((d2−(xmat[[2]][i,]%∗%theta2[−ll2]))/theta2[ll2], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
        U=runif(1)
        v2=U+(1−U)∗(1−((1−u1)ˆ(−(a+1)))∗(((1−u1)ˆ(−a)+(1−a2)ˆ(−a)−1)ˆ(−(a+1)/a)))
        ua[i,2]=1−((1+(((1−u1)ˆ(−a))∗((1−v2)ˆ(−a/(a+1))−1)))ˆ(−1/a))
        y4[i,2]=xmat[[2]][i,]%∗%theta2[−ll2]+theta2[ll2]∗qnorm(ua[i,2], mean=0, sd=1, lower.tail=TRUE, log.p=FALSE)
      }

    }

    udat2=ua
    y5=y4

    fit.ifm2=try(optim(a0, fn=loglik.cop, u=udat2, method=”L−BFGS−B”, lower=0.0001, upper=Inf, control=list(fnscale=−1,
                  maxit=100000)), TRUE)

    if(inherits(fit.ifm2,”try−error”)){erro=TRUE}

  } #end while(erro)

  aa=fit.ifm2$par
  saida=list()
  saida[[1]]=aa; saida[[2]]=udat2; saida[[3]]=y5
  saida

}

##### Import a local txt file named churning.txt (Products A and B)

dados=read.table(”C:\\Users\\Aluno\\Desktop\\Dados\\churning.txt”, header = TRUE)

n=nrow(dados)

attach(dados)

d1=2.3
d2=2.3

summary(l1a); sd(l1a)
summary(l1a[l1a<d1]); sd(l1a[l1a<d1])
summary(l2a); sd(l2a)
summary(l2a[l2a<d2]); sd(l2a[l2a<d2])
summary(idade); sd(idade)
summary(renda); sd(renda)

par(mfrow=c(1,2))
hist(l1a, main=””, xlab=”log(time) to churn Product A”, freq=FALSE)
hist(l2a, main=””, xlab=”log(time) to churn Product B”, freq=FALSE)

tau=cor(cbind(l1a,l2a),method=”kendall”)[1,2]
a0=2∗tau/(1−tau)

```

```
k=9

B=500

######## Fit the bivariate Clayton survival copula-based SUR Tobit right-censored model with normal
######## marginal errors to the customer churn data (Products A and B)

xmat=list(model.matrix(~idade+renda), model.matrix(~idade+renda))

y=cbind(l1a, l2a)

# censoring percentage in the margins
cont1=0
cont2=0
for(i in 1:n){
  if(y[i,1]==d1) cont1=cont1+1
  if(y[i,2]==d2) cont2=cont2+1
}
cens1=cont1/n
cens2=cont2/n

# two-stage parametric ML method - IFM method - by Joe and Xu (1996)

# stage 1
tobito1=tobit(y[,1]~xmat[[1]][,-1],left=-Inf,right=d1,dist="gaussian")
est1=summary(tobito1)$coefficients
theta1hat=c(est1[1,1], est1[2,1], est1[3,1], exp(est1[4,1]))

tobito2=tobit(y[,2]~xmat[[2]][,-1],left=-Inf,right=d2,dist="gaussian")
est2=summary(tobito2)$coefficients
theta2hat=c(est2[1,1], est2[2,1], est2[3,1], exp(est2[4,1]))

par(mfrow=c(1,2))

# scatter plot of y1 versus y2
plot(y[,1],y[,2])

# stage 2
udat=cbind(probtrans(theta1hat,y[,1],xmat[[1]],d1), probtrans(theta2hat,y[,2],xmat[[2]],d2))

# scatter plot of udat[,1] versus udat[,2]
plot(udat[,1],udat[,2], xlab=expression(u[1]), ylab=expression(u[2]))

fit.ifm=optim(a0, fn=loglik.cop, u=udat, method="L-BFGS-B", lower=0.0001, upper=Inf, control=list(fnscale=-1,
  maxit=100000))
thetas.ifm=c(theta1hat, theta2hat, fit.ifm$par)

thetas.est=thetas.ifm

# two-stage parametric ML method - IFM method - by Joe and Xu (1996) with augmented data (MIFM method)

# stage 2
ua1=udat

l1=ncol(xmat[[1]])+1
l2=ncol(xmat[[2]])+1

theta1=thetas.ifm[1:l1]
theta2=thetas.ifm[(l1+1):(l1+l2)]

phi=numeric()
loglike=numeric()

phi[1]=fit.ifm$par

parm.margins=c(theta1,theta2)

loglike[1]=myloglik(c(parm.margins,phi[1]),y,xmat,d1,d2)

out=mifm(theta1,theta2,phi[1],y,ua1,xmat,d1,d2)
phi[2]=out[[1]]
ua11=out[[2]]
```

```
loglike[2]=myloglik(c(parm.margins,phi[2]),y,xmat,d1,d2)

eps=0.001

w=1

while (abs(phi[w+1]−phi[w]) >= eps){
  out2=mifm(theta1,theta2,phi[w+1],y,ua1,xmat,d1,d2)
  phi[w+2]=out2[[1]]
  ua11=out2[[2]]
  ya=out2[[3]]
  loglike[w+2]=myloglik(c(parm.margins,phi[w+2]),y,xmat,d1,d2)
  w=w+1
}

niter=length(phi)
phi.est.mifm=phi[niter]

plot(phi, xlab=”Iteration”, ylab=expression(hat(theta)[MIFM]), type=”b”)
plot(loglike, xlab=”Iteration”, ylab=”Log−likelihood”, type=”b”)

thetas.ifm2=c(theta1, theta2, phi.est.mifm)

thetas.est2=thetas.ifm2

# histograms of y1 and y2 (augmented data)
hist(ya[,1], main=” ”, xlab=”log(time) to churn Product A”, freq=FALSE); abline(v=d1,lty=2)
hist(ya[,2], main=” ”, xlab=”log(time) to churn Product B”, freq=FALSE); abline(v=d2,lty=2)

# scatter plot of y1 versus y2 (augmented data)
plot(ya[,1],ya[,2], xlab=”log(time) to churn Product A”, ylab=”log(time) to churn Product B”); abline(h=d1, v=d2,lty=2)

# scatter plot of u1 versus u2 (augmented data)
plot(ua11[,1],ua11[,2], xlab=expression(u[1]), ylab=expression(u[2]))

# lilliefors tests of augmented marginal residuals
res1=ya[,1]−xmat[[1]]%∗%theta1[−l1]
res2=ya[,2]−xmat[[2]]%∗%theta2[−l2]

hist(res1, main=””, xlab=”Residuals”, freq=FALSE); hist(res2, main=””, xlab=”Residuals”, freq=FALSE)

lillie.test(res1); lillie.test(res2)

# AIC and BIC criterion values
AIC=−2∗loglike[niter]+2∗k
BIC=−2∗loglike[niter]+k∗log(n)

# Parametric bootstrap approach: generate y1 and y2 values using thetas.ifm2 in genY() function

thetas.boot=matrix(numeric(k),B,k)
niter.boot=numeric(B)
phi.est.mifm.boot=numeric(B)

for(b in 1:B){

  y.boot=genY(thetas.ifm2,xmat,d1,d2)

  # two−stage parametric ML method − IFM method − by Joe and Xu (1996)

  # stage 1
  tobito1.boot=tobit(y.boot[,1]˜xmat[[1]][,−1],left=−Inf,right=d1,dist=”gaussian”)
  est1.boot=summary(tobito1.boot)$coefficients
  theta1hat.boot=c(est1.boot[1,1], est1.boot[2,1], est1.boot[3,1], exp(est1.boot[4,1]))

  tobito2.boot=tobit(y.boot[,2]˜xmat[[2]][,−1],left=−Inf,right=d2,dist=”gaussian”)
  est2.boot=summary(tobito2.boot)$coefficients
  theta2hat.boot=c(est2.boot[1,1], est2.boot[2,1], est2.boot[3,1], exp(est2.boot[4,1]))

  # stage 2
  udat.boot=cbind(probtrans(theta1hat.boot,y.boot[,1],xmat[[1]],d1), probtrans(theta2hat.boot,y.boot[,2],xmat[[2]],d2))

  fit.ifm.boot=optim(a0, fn=loglik.cop, u=udat.boot, method=”L−BFGS−B”, lower=0.0001, upper=Inf, control=list(fnscale=−1,
         maxit=100000))
  thetas.ifm.boot=c(theta1hat.boot, theta2hat.boot, fit.ifm.boot$par)

  # two-stage parametric ML method - IFM method - by Joe and Xu (1996) with augmented data (MIFM method)

  # stage 2
  ua.boot=udat.boot

  theta1.boot=thetas.ifm.boot[1:l1]
  theta2.boot=thetas.ifm.boot[(l1+1):(l1+l2)]

  phi.boot=numeric()
  loglike.boot=numeric()

  phi.boot[1]=fit.ifm.boot$par

  parm.margins.boot=c(theta1.boot,theta2.boot)

  loglike.boot[1]=myloglik(c(parm.margins.boot,phi.boot[1]),y.boot,xmat,d1,d2)

  out.boot=mifm(theta1.boot,theta2.boot,phi.boot[1],y.boot,ua.boot,xmat,d1,d2)
  phi.boot[2]=out.boot[[1]]
  ua11.boot=out.boot[[2]]

  loglike.boot[2]=myloglik(c(parm.margins.boot,phi.boot[2]),y.boot,xmat,d1,d2)

  w=1

  while (abs(phi.boot[w+1]-phi.boot[w]) >= eps){
    out2.boot=mifm(theta1.boot,theta2.boot,phi.boot[w+1],y.boot,ua.boot,xmat,d1,d2)
    phi.boot[w+2]=out2.boot[[1]]
    ua11.boot=out2.boot[[2]]
    loglike.boot[w+2]=myloglik(c(parm.margins.boot,phi.boot[w+2]),y.boot,xmat,d1,d2)
    w=w+1
  }

  niter.boot[b]=length(phi.boot)
  phi.est.mifm.boot[b]=phi.boot[niter.boot[b]]
  thetas.ifm2.boot=c(theta1.boot, theta2.boot, phi.est.mifm.boot[b])
  thetas.boot[b,]=thetas.ifm2.boot

  print(b)

}

# Bootstrap confidence intervals

# Standard normal interval

cov.boot=matrix(numeric(k),k,k)
mean.boot=as.matrix(apply(thetas.boot, 2, mean), k, 1, byrow=TRUE)

for(b in 1:B){
  thetas.boot2=as.matrix(thetas.boot[b,], k, 1, byrow=TRUE)
  cov.boot=cov.boot+((thetas.boot2-mean.boot)%*%t(thetas.boot2-mean.boot))
}

cov.boot=(1/(B-1))*cov.boot
var.boot=diag(cov.boot); se.boot=sqrt(var.boot)
inf4=thetas.ifm2-1.645*se.boot; sup4=thetas.ifm2+1.645*se.boot

# Percentile interval

inf5=numeric(k); sup5=numeric(k)

for(j in 1:k){
  percentis=quantile(thetas.boot[,j], probs=c(0.05,0.95))
  inf5[j]=percentis[[1]]; sup5[j]=percentis[[2]]
}
```

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki, editors, *2nd International Symposium on Information Theory*, pages 267–281.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6), 716–723.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics*, **24**(1-2), 3–61.

Anastasopoulos, P. C., Shankar, V. N., Haddock, J. E. & Mannering, F. L. (2012). A multivariate tobit analysis of highway accident-injury-severity rates. *Accident Analysis & Prevention*, **45**, 110–119.

Armstrong, M. (2003). Copula catalogue. Part 1: Bivariate Archimedean copulas. Unpublished paper available at: http://www.cerna.ensmp.fr/Documents/MA-CopulaCatalogue.pdf.

Baranchuk, N. & Chib, S. (2008). Assessing the role of option grants to CEOs: How important is heterogeneity? *Journal of Empirical Finance*, **15**(2), 145–166.

Bolfarine, H., Santos, B., Correia, L., Martínez, G., Goméz, H. & Bazán, J. (2013). Modelos de regressão com respostas limitadas e censuradas. In: 13a Escola de Modelos de Regressão, Maresias, São Sebastião, SP.

Brown, E. & Lankford, H. (1992). Gifts of money and gifts of time: Estimating the effects of tax prices and available time. *Journal of Public Economics*, **47**(3), 321–341.

Cameron, A. C. & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Carpenter, J. & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, **19**(9), 1141–1164.

Caudill, S. B. & Mixon, F. G. (2009). More on testing the normality assumption in the tobit model. *Journal of Applied Statistics*, **36**(12), 1345–1352.

Chen, S. & Zhou, X. (2011). Semiparametric estimation of a bivariate Tobit model. *Journal of Econometrics*, **165**(2), 266–274.

Cherubini, U., Luciano, E. & Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons, Chichester.

Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, **51**(1-2), 79–99.

Chib, S. & Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika*, **85**(2), 347–361.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–152.

Conover, W. J. (1971). *Practical Nonparametric Statistics*. John Wiley & Sons, New York, first edition.

Cook, R. D. & Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society, Series B*, **43**(2), 210–218.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, **39**(5), 829–844.

Davidson, R. & MacKinnon, J. G. (2003). *Econometric Theory and Methods*. Oxford University Press, New York.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press, Cambridge.

De Luca, G. & Rivieccio, G. (2012). Multivariate tail dependence coefficients for Archimedean copulae. *Advanced Statistical Methods for the Analysis of Large Data-Sets*, pages 287–296. In: A. Di Ciaccio, M. Coli, and J. M. Angulo Ibañez, eds., Springer, Berlin.

Devroye, L. (1986). *Non-Uniform Random Variate Generation.* Springer-Verlag, New York.

Di Bernardino, E. & Rullière, D. (2014). On tail dependence coefficients of transformed multivariate Archimedean copulas. Unpublished paper available at: https://hal.archives-ouvertes.fr/hal-00992707v2/document.

Dixon, M. (1999). 39 experts predict the future. *America's Community Banker*, **8**(7), 20–31.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, **82**(397), 171–185.

Efron, B. & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

Embrechts, P., Lindskog, F. & McNeil, A. J. (2003). Modelling dependence with copulas and applications to risk management. *Handbook of Heavy Tailed Distributions in Finance*, pages 329 – 384. In: S. T. Rachev, ed., Elsevier, North-Holland.

Garay, A. W. M. (2014). *Modelos de regressão para dados censurados sob distribuições simétricas*. Ph.D. thesis, Universidade de São Paulo, São Paulo.

Genest, C., Nešlehová, J. & Ben Ghorbal, N. (2011). Estimators based on kendall's tau in multivariate copula models. *Australian & New Zealand Journal of Statistics*, **53**(2), 157–177.

Georges, P., Lamy, A. G., Nicolas, E., Quibel, G. & Roncalli, T. (2001). Multivariate survival modelling: a unified approach with copulas. Working paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais, France.

Geweke, J. (1991). Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In

Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, pages 571–578. Seattle, USA, April 22-24.

Goethals, K., Janssen, P. & Duchateau, L. (2008). Frailty models and copulas: similarities and differences. *Journal of Applied Statistics*, **35**(9), 1071–1079.

Goldberger, A. S. (1964). *Econometric Theory*. John Wiley & Sons, New York.

Greene, W. H. (2003). *Econometric Analysis*. Prentice Hall, New Jersey, fifth edition.

Gumbel, E. J. (1961). Bivariate logistic distributions. *Journal of the American Statistical Association*, **56**(294), 335–349.

Gupta, R. D. & Gupta, R. C. (2008). Analyzing skewed data by power normal model. *Test*, **17**(1), 197–210.

Hildebrandt, A. (2007). *Empirical Evidence on Shareholder Value Effects of Corporate Restructuring*. GRIN Verlag.

Hinkley, D. V. (1988). Bootstrap methods. *Journal of the Royal Statistical Society, Series B*, **50**(3), 321–337.

Holden, D. (2004). Testing the normality assumption in the tobit model. *Journal of Applied Statistics*, **31**(5), 521–532.

Huang, C. J., Sloan, F. A. & Adamache, K. W. (1987). Estimation of seemingly unrelated Tobit regressions via the EM algorithm. *Journal of Business and Economic Statistics*, **5**(3), 425–430.

Huang, H. C. (1999). Estimation of the SUR Tobit model via the MCECM algorithm. *Economics Letters*, **64**(1), 25–30.

Huang, H. C. (2001). Bayesian analysis of the SUR Tobit model. *Applied Economics Letters*, **8**(9), 617–622.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hal, London.

Joe, H. & Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.

Kamakura, W. A. & Wedel, M. (2001). Exploratory Tobit factor analysis for multivariate censored data. *Multivariate Behavioral Research*, **36**(1), 53–82.

Kimeldorf, G. & Sampson, A. R. (1975). Uniform representations of bivariate distributions. *Communications in Statistics - Theory and Methods*, **4**(7), 617–627.

Lee, L. F. (1993). Multivariate Tobit models in econometrics. *Handbook of Statistics*, **11**, 145–175. In: G.S. Maddala and C.R. Rao, eds., North-Holland.

Malik, H. J. & Abraham, B. (1973). Multivariate logistic distributions. *The Annals of Statistics*, **1**(3), 588–590.

Martínez-Floréz, G., Bolfarine, H. & Gómez, H. W. (2013). The alpha-power tobit model. *Communications in Statistics - Theory and Methods*, **42**(4), 633–643.

McNeil, A. J., Frey, R. & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, New Jersey.

Meng, X. & Rubin, D. B. (1996). Efficient methods for estimating and testing seemingly unrelated regressions in the presence of latent variables and missing observations. *Bayesian Analysis in Statistics and Econometrics*, pages 215–227. In: D. A. Berry, K. M. Chaloner, and J. K. Geweke, eds., John Wiley & Sons, New York.

Moulton, L. H. & Halsey, N. A. (1995). A mixture model with detection limits for regression analyses of antibody response to vaccine. *Biometrics*, **51**(4), 1570–1578.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, second edition.

Oakes, D. (2005). On the preservation of copula structure under truncation. *The Canadian Journal of Statistics*, **33**(3), 465–468.

Okhrin, O., Okhrin, Y. & Schmid, W. (2013). On the structure and estimation of hierarchical Archimedean copulas. *Journal of Econometrics*, **173**(2), 189 – 204.

Omori, Y. & Miyawaki, K. (2010). Tobit model with covariate dependent thresholds. *Computational Statistics & Data Analysis*, **54**(11), 2736 – 2752.

Panagiotelis, A., Czado, C. & Joe, M. (2012). Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association*, **107**(499), 1063–1072.

Pitt, M., Chan, D. & Kohn, R. (2006). Efficient Bayesian inference for Gaussian copula regression models. *Biometrika*, **93**(3), 537–554.

R Core Team (2014). *R: A Language and Environment for Statistical Computing. (Version 3.1.0)*. R Foundation for Statistical Computing, Vienna, Austria.

Reichheld, F. F. & Sasser, W. E. (1990). Zero defections: quality comes to service. *Harvard Business Review*, **68**(5), 105–111.

Romeo, J. S., Tanaka, N. I. & Pedroso-de Lima, A. C. (2006). Bivariate survival modeling: a Bayesian approach based on copulas. *Lifetime Data Analysis*, **12**(2), 205–222.

Savu, C. & Trede, M. (2010). Hierarchies of Archimedean copulas. *Quantitative Finance*, **10**(3), 295–304.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2), pp. 461–464.

Shih, J. H. & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, **51**(4), 1384–1399.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

Slater, S. F. & Narver, J. C. (2000). Intelligence generation and superior customer value. *Journal of the Academy of Marketing Science*, **28**(1), 120–127.

Smith, M. S. & Khaled, M. A. (2012). Estimation of copula models with discrete margins via Bayesian data augmentation. *Journal of the American Statistical Association*, **107**(497), 290–303.

Su, L. J. & Arab, L. (2006). Salad and raw vegetable consumption and nutritional status in the adult US population: results from the Third National Health and Nutrition

Examination Survey. *Journal of the American Dietetic Association*, **106**(9), 1394–1404.

Sungur, E. A. (1999). Truncation invariant dependence structures. *Communications in Statistics - Theory and Methods*, **28**(11), 2553–2568.

Sungur, E. A. (2002). Some results on truncation dependence invariant class of copulas. *Communications in Statistics - Theory and Methods*, **31**(8), 1399–1422.

Taylor, M. R. & Phaneuf, D. (2009). Bayesian estimation of the impacts of food safety information on household demand for meat and poultry. In *Proceedings of the 2009 AAEA & ACCI Joint Annual Meeting*. Milwaukee, July 26-28.

Thode, H. (2002). *Testing For Normality*. Marcel Dekker, New York.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, **26**(1), 24–36.

Trivedi, P. K. & Zimmer, D. M. (2005). Copula modeling: An introduction for practitioners. *Foundations and Trends in Econometrics*, **1**(1), 1–111.

USDA (2000). Continuing survey of food intakes by individuals 1994-1996. CD-ROM. Agricultural Research Service, Washington, DC.

Wales, T. J. & Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, **21**(3), 263–285.

Wang, G., Liu, L., Peng, Y., Nie, G., Kou, G. & Shi, Y. (2010). Predicting credit card holder churn in banks of china using data mining and mcdm. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, volume 3, pages 215–218.

Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance*, **4**(3), 339–352.

Wichitaksorn, N., Choy, S. T. B. & Gerlach, R. (2012). Estimation of bivariate copula-based seemingly unrelated Tobit models. Discipline of Business Analytics, University of Sydney Business School, NSW 2006, Australia. Eletronic copy available at: http://ssrn.com/abstract=2122388.

Yen, S. T. & Lin, B. H. (2008). Quasi-maximum likelihood estimation of a censored equation system with a copula approach: meat consumption by U.S. individuals. *Agricultural Economics*, **39**(2), 207–217.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, **57**(298), 348–368.

Zellner, A. & Ando, T. (2010). Bayesian and non-Bayesian analysis of the seemingly unrelated regression model with Student-t errors, and its application for forecasting. *International Journal of Forecasting*, **26**(2), 413–434.