International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 6, Issue 11 [November. 2017] PP: 25-34

Using Copulas To Estimate The Effect Of A Binary Endogenous VariableIn A Binary Response Model

Sulafah M.Binhimd[1],Mervat k. Abd Elaal [1, 2], AbeerH. Alsoruji[1]
[1] *Statistics Department, Faculty of Sciences King Abdulaziz University*
Jeddah, Kingdom of Saudi Arabia
&[2] Statistics Department, Faculty of Commerce Al-Azhar University, Girls Branch
Cairo, Egypt

***Abstract**: This paper presents a copula-based approach of a bivariate binary response model for estimating the effect of binary endogenous (treatment) variable in a binary response model.The coefficient of treatment in copula bivariate probit model may not be enterable. In order, to interpret the effect of treatment on the binary response, it is necessary to compute the average treatment effect from the estimated coefficients. The almost exclusively used approach to estimate the coefficient parameters is maximum likelihood estimation. The main point of this paper can be made using Gaussian, FGM, and Frank copulas in our simulations and show that the Gaussian copula bivariate probit model gives good estimation result of average treatment effect in FGM and Frank copulas data generating process.*

***Key Words:** Copula; Endogeneity; recursive bivariate probit model; sample average treatment effect.*

Date of Submission: 08-11-2017 Date of acceptance: 05-12-2017

I. Introduction

Endogeneity means that the two variables (binary treatment and binary response) may depend on correlated unobserved variables (the $\varepsilon^{'}$). Failing to take endogeneity into account may result in substantially biased results.Endogeneity can be controlled by using recursive bivariate probit model. This model controls for endogeneity by using two equation structural latent variable framework, where the first equation $y_1 = 1(x_1^T\alpha_1 + \varepsilon_1 > 0)$ determines whether or not the treatment is received, whereas the second equation $y_2 = 1(\gamma y_1 + x_2^T\alpha_2 + \varepsilon_2 > 0)$describes the binary response variable (y_2) as a function of a binary tretment (y_1). The model is completed by assuming that the latent errors ε_1 and ε_2 of the two equations follow a standard bivariate Gaussian distribution with correlation θ; $\theta \neq 0$ suggests that endogeneity is present, hence joint estimation of the two equations is required. If $\theta = 0$, then separate estimation of the second structural equation by a simple probit model identifies the structural treatment effectγ. Some applications are provided by Goldman et al.,[3], Gitto et al.,[2], Latif [7], and Li and Jensen [8].

The limitations of this model are the inability to deal with non-linear covariate effects and non-Gaussian dependencies between the treatment and response equations. To model flexibly covariate-response relationships, Chiband Greenberg [1] and MarraandRadice[10] introduced Bayesian and likelihood estimation approaches based on penalized regression splines, thereby allowing for a number of different flexible covariate-response structures. To deal with the problem of non-Gaussian dependence between treatment and response, Winkelmann [16] discussed a modification of the recursive bivariate probit that preserve the Gaussian assumption for the marginal distributions of the two equations while introducing non-Gaussian dependence between them using the Frank and Clayton copulas and he used maximum likelihood for estimated the parameters of the resulting copula bivariate probit model (CBP).

Radice, et al., [13] took a more general approach and extended the procedures discussed in Marra and Radice[10] and Winkelmann [16] to make it possible to deal simultaneously with endogeneity, flexible covariate effects and non-linear dependencies between two binary responses. In particular, they generalized the approach based on the assumption of bivariate normality presented in Marra and Radice[10] by allowing for non-normal dependencies between the two equations through Clayton, Frank, Gumbel, Joe, Ali-Mikhail-Haq and Farlie-Gumbel-Morgenstern copulas and the rotated versions of Clayton, Gumbel and Joe. They calculated the percentage bias for sample average treatment effect (SATE) for each model to select the best model in their simulations.

In empirical applications, researchers likely do not know the true data generating copula, because of its association to the familiar normal distribution, and the statistics packages often include functions that calculate univariate and bivariate normal probabilities, a researcher might be tempted to use the Gaussian copula. Zimmer [17] show in his paper the Gaussian copula might often be appropriate for estimating parameters of Poisson

models (count variable models) with binary endogenous variable when he use FGM and Frank copulas data generating process (DGPs).

In the next section in this paper we introduce the recursive bivariate probit model. Section 3 introduces the concept of copula function in general and discuss the copula-based approach to allow non-normal dependence which is called CBP model. Section 4 discusses properties of three specific copulas (Gaussian, FGM, and Frank) and shows how they can be integrated into a bivariate binary response model. Section 5 introduced sample average treatment effect (SATE) for CBP model which is our basis comparison between the models and computed the average of the percentage absolute bias for SATE. Section 6 reports results from a series of Monte Carlo simulations where it is shown that The Gaussian copula might often be appropriate for estimating average treatment effect of binary response model with binary endogenous variable (treatment) when we use FGM and Frank copulas data generating process (DGPs) which all are symmetric copula, each are widely-used and famous among copula researchers, and each permit both positive and negative dependence. Section 7concludes.

II. The model

The model consists of two latent equations: for an observation i, $i = 1,2. \dots, n$,

$y_{1i}^* = x_{1i}^T\alpha_1 + \varepsilon_{1i}$ (1)

$y_{2i}^* = \gamma y_{1i} + x_{2i}^T\alpha_2 + \varepsilon_{2i}$ (2)

Where y_{1i}^* and y_{2i}^* are latent variables, y_{1i} and y_{2i} are binary variables observed according to the rule

$$\begin{cases} y_{ji} = 1 \, if y_{ji}^* > 0 \\ y_{ji} = 0 \, if y_{ji}^* \leq 0 \end{cases}; \qquad j = 1,2$$

Moreover, $x_{1i}^T = (1, x_{12i}, \dots, x_{1P_1i})$ is the ith row vector of the$n \times P_1$model matrix X_1 and α_1 is a parameter vector.

Similarly, x_{2i}^T is the ith row vector of the$n \times P_2$ model matrix X_2, α_2 is a coefficient vector, and γ is the parameter of endogenous binary variable y_{1i}. The covariates in equation (1) should contain at least one or more regressors (usually referred to as instruments) not included in the second equation. These regressors must induce variation in y_{1i}, have not to directly effect y_{2i}, and must be independent of $(\varepsilon_{1i}, \varepsilon_{2i})$ given covariates. Ieva et al., [6].

The variable y_1 in the second equation is the primary interest, but it is potentially endogenous. The endogeneity issue arises when the errors ε_1 and ε_2 are not independent. In order to estimate the coefficient parameters consistently, the dependence between ε_1 and ε_2 needs to be taken into consideration. The almost exclusively used approach is maximum likelihood estimation (MLE), where the log likelihood function has a general form $lnL = \sum_{i=1}^{n} lnPr(y_{1i}, y_{2i})$, where$Pr(y_{1i}, y_{2i})$is the joint probability of y_1 andy_2.

In order to implement MLE, we need to specify the joint distributionof the error terms, $F(\varepsilon_1, \varepsilon_2)$. A standard approach is to assume that these errors are jointly normally distributed. Under bivariate normality, $F(\varepsilon_1, \varepsilon_2) = \Phi_2(\varepsilon_1, \varepsilon_2, \rho)$, where $\Phi_2(.)$is thecumulative distribution function (cdf) of the bivariate normal distributionwith the coefficient of correlation ρ. The joint probabilityof y_1 and $y_2$is

$\Pr(y_1, y_2) = \Phi_2[s_1(x_{1i}^T\alpha_1), s_2(\gamma y_{1i} + x_{2i}^T\alpha_2), s_1 s_2 \rho]$ (3)

Where $s_j = 2y_j - 1, \; j = 1,2$. The model is called a (recursive)bivariate probit model Greene, [4] andMaddala, [9].

It is important to understand that the bivariate probit model introduces two sources of dependent between $y_1$and y_2, related to the parameters γ and ρ, respectively. Although the joint model simplifies to two univariate probit equations under independence of the structural errors ($\rho = 0$), this does not mean that y_1 and y_2 are independent in this case. The reason is that the second probit equation of the base model gives the probability of y_2 conditional on y_1. Therefore, full independence of y_1 and y_2 requires $\rho = 0$and $\gamma = 0$. Winkelmann, [16].

III. The copula approach

Winkelmann [16] discusses the copula-based approach to allow non-normal dependence. In short, a copula $C(.)$is a function that binds univariate marginal distributions together to represent a joint distribution. Using a copula, the joint distribution of ε_1 and $\varepsilon_2$is $F(\varepsilon_1, \varepsilon_2) = C(F_1(\varepsilon_1), F_2(\varepsilon_2); \theta)$, where $F_j(.)$ is the cdfof a univariate distribution of ε_j and θ is a dependence parameter that governs the degree of dependence. A substantial advantage of the copula approach is that the marginal distributions may come from different families. So, copulas can be used to generate joint distribution functions for the two structural errors in the

bivariate probit model,$\varepsilon_1$and ε_2, keeping the normal marginals but do not require joint normality. Then $F(\varepsilon_1, \varepsilon_2) = C(F_1(\varepsilon_1), F_2(\varepsilon_2); \theta)$ is a joint distribution function for ε_1 and ε_2 with marginal distributions that are standard normal.Define $u = F_1(\varepsilon_1)$ and $v = F_2(\varepsilon_2)$with $u, v \in$ [0, 1]; then, copula function$C$ satisfies the following properties: (i) $C(1, v) = v$ and$C(u, 1) = u$, (ii) $C(0, v) = C(u, 0) = 0$.

In applications, researchers need to choose a copula from severalavailable copulas (the list of copulas is available in Nelsen, [12].Researchers also have freedom to choose themarginal distributions $F_1(.)$and $F_2(.)$. When the marginal distributionsare standard normal distributions and the copula is Gaussian,the copula-based model is the same as the bivariate probit model. See Hasebe, [5].

Under a copula representation with probit marginals, the data identify the four possible events ($y_{1i} = 1, y_{2i} = 1$), ($y_{1i} = 1, y_{2i} = 0$), ($y_{1i} = 0, y_{2i} = 1$), and ($y_{1i} = 0, y_{2i} = 0$) with probabilities given asMarra, et. al., (2014):

$P(y_{1i} = 1, y_{2i} = 1) = C[F_1(x_{1i}^T \alpha_1), F_2(\gamma + x_{2i}^T \alpha_2)]$ (4)
$P(y_{1i} = 0, y_{2i} = 1) = F_2(x_{2i}^T \alpha_2) - C[F_1(x_{1i}^T \alpha_1), F_2(x_{2i}^T \alpha_2)]$ (5)
$P(y_{1i} = 1, y_{2i} = 0) = F_1(x_{1i}^T \alpha_1) - C[F_1(x_{1i}^T \alpha_1), F_2(\gamma + x_{2i}^T \alpha_2)]$ (6)
$P(y_{1i} = 0, y_{2i} = 0) = 1 - F_1(x_{1i}^T \alpha_1) - F_2(x_{2i}^T \alpha_2) + C[F_1(x_{1i}^T \alpha_1), F_2(x_{2i}^T \alpha_2)]$(7)

The joint probability of y_1 and y_2 can be written as
$\Pr(y_1, y_2) =$
$(1 - y_1)(1 - y_2) + (1 - y_2)s_1 F_1(x_1^T \alpha_1) + (1 - y_1)s_2 F_2(\gamma y_1 + x_2^T \alpha_2) + s_1 s_2 C[F_1(x_1^T \alpha_1), F_2(\gamma y_1 + x_2^T \alpha_2)]$(8)
The likelihood function for n observations can be derived from (4)-(7) as follows:

$\prod_{i=1}^{n} P(y_{1i} = 1, y_{2i} = 1)^{y_{1i} y_{2i}} \times P(y_{1i} = 1, y_{2i} = 0)^{y_{1i}(1 - y_{2i})} \times P(y_{1i} = 0, y_{2i} = 1)^{(1 - y_{1i}) y_{2i}} \times P(y_{1i} = 0, y2i = 0(1 - y1i)(1 - y2i)$(9)

The unknown parameters $\xi = (\alpha_1 \alpha_2 \gamma \theta)$ of CBP can be easily eatimated with the maximum likelihood method. Numerical optimization methods can be used to maximize the log-likelihood function Winkelmann, [16]. These can employ analytical first derivatives that have a relatively dissolvable form. For example,

$\frac{\partial P(y_1 = 0, y_2 = 0 | \xi)}{\partial \alpha_2} = c_v[\Phi(x_{1i}^T \alpha_1), \Phi(\gamma + x_{2i}^T \alpha_2)]\Phi(\gamma + x_{2i}^T \alpha_2) x_{2i}^T$ (10)

Where $c_v = \partial C(u.v)/\partial v$.

IV. Gaussian, Farlie-Gumbel-Morgenstern, and Frank copulas

In this paper we will use the following three copulas.

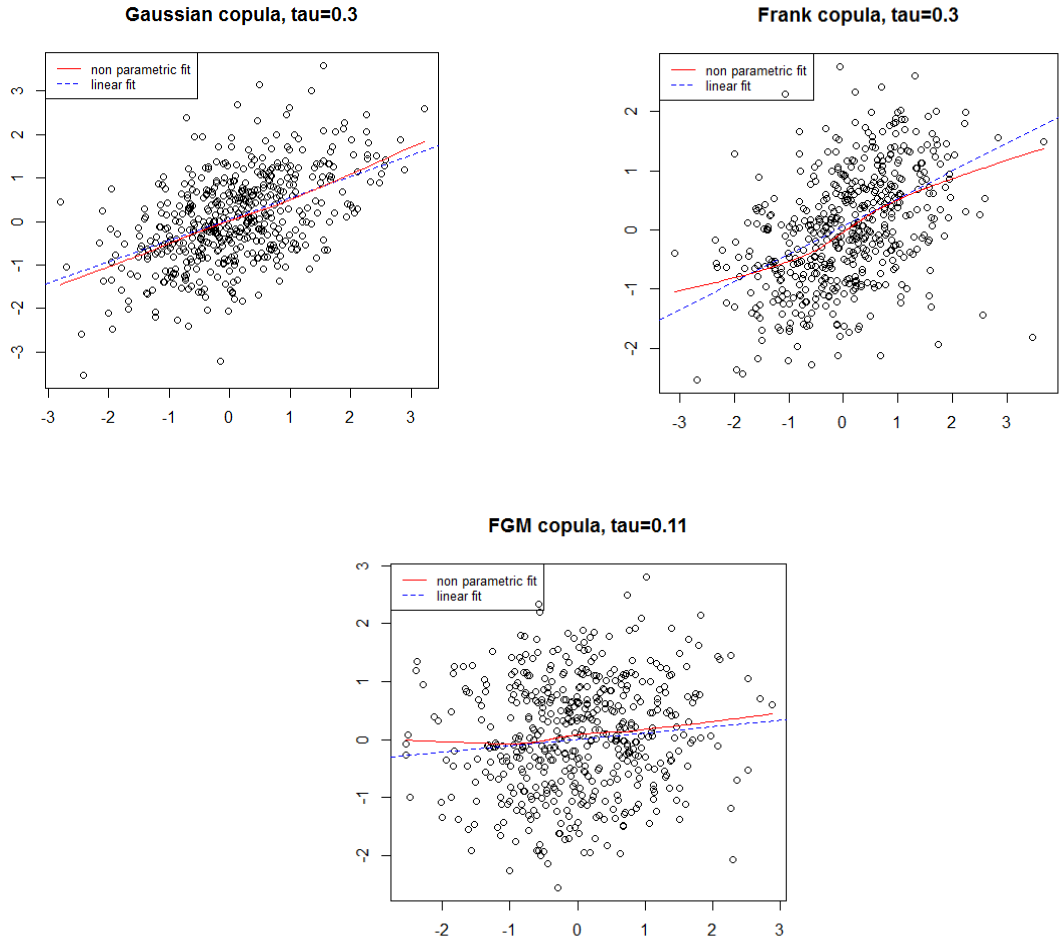
$C(F_1(\varepsilon_1), F_2(\varepsilon_2); \theta)$	Domain for θ
Gaussian $\Phi_2(\Phi^{-1}(F_1(\varepsilon_1)), \Phi^{-1}(F_2(\varepsilon_2)), \theta)$	$-1 < \theta < 1$
FGM $F_1(\varepsilon_1)F_2(\varepsilon_2)\left(1 + \theta(1 - F_1(\varepsilon_1))(1 - F_2(\varepsilon_2))\right)$	$-1 < \theta < 1$
Frank $-\theta^{-1} log\left(1 + \frac{(e^{-\theta F_1(\varepsilon_1)} - 1)(e^{-\theta F_2(\varepsilon_2)} - 1)}{(e^{-\theta} - 1)}\right)$	$-\infty < \theta < \infty; \theta \neq 0$

(For the Gaussian copula, Φ^{-1} denotes the quantile function of the standard normal distribution and Φ_2 denotes the cdf of bivariate standard normal distribution).

We consider these copulas for three reasons.First, each are widely-used and famous among copula researchers. Second, dependence in each is symmetric in both tails. Third, unlike some other copulas, each permits negative dependence between the marginals.

In order, to compare the differences in the dependence structure implied by the FGM and Frank copulas with that of the bivariate normal distribution, correlation is not a good indicator. First, it detects only linear dependence, whereas dependence in copulas is nonlinear in general. Second, it is not invariant to transformation of the marginal distributions. Therefore, other measures of dependence have been suggested. A common one is Kendall's τ, a measure of the degree of concordance. Hence τ for the two random pairs (U_1, V_1) and (U_2, V_2) from the joint distribution of U and V is defined as$\tau = P[(U_1 - U_2)(V_1 - V_2) > 0] - P[(U_1 - U2V1 - V2 < 0]$. Kendall's τ can vary between $-1$and 1. It is zero if U and V are independent. Not all copulas cover the full spectrum of possible τs. If they do so, they are called comprehensive. Gaussian and Frank copulas are comprehensive.

Dependence in Gaussian copula is symmetric in both tails of the distribution. Similarly, dependence in Frank copula is also symmetric in both tails. However, compared to the Gaussian copula, dependence in the Frank copula is weaker in both tails and stronger in the center of the distribution. This suggests that the Frank copula is best suited for applications in which tail dependence is relatively weak. Finally, the FGM copula exhibits symmetry in both tails, but it cannot accommodate variables with large dependence. The FGM copula is popular in applied settings due to its simplicity and because it allows negative dependence, but it is only appropriate for applications with weak dependence. Understanding dependence structures of different copulas is imperative in empirical applications Trivedi and Zimer[15].The following figures show a sample of 500 draws from the Gaussian, FGM, and Frank copulas, with standard normal marginals,τ=0.3 for Gaussian and Frank copulas and τ = 0.11 for FGM copula (FGM copula cannot handle large dependence).



V. Sample average treatment effect (SATE) for CBP model

Since latent variables do not typically have well-defined units of measurements, parameter γ in equation (2) may not be interpretable. For this reason, the effect of y_{1i} on the probability that $y_{2i} = 1$ is of primary interest. In other words, the aim is to investigate how the binary endogenous variable (treatment)

changes the expected response. Thus, treatment effect is given by the difference between the expected response with treatment andthe expected response without treatment. This can be done by using the average treatment effect. Given estimates for the model components, the ATE can be estimated as follows:

$SATE = \frac{1}{n}\sum_{i=1}^{n} P(y_{2i} = 1|y_{1i} = 1) - P(y_{2i} = 1|y_{1i} = 0)$ (11)

Where

$$P(y_{2i} = 1|y_{1i} = 1) = \frac{C[F_1(x_{1i}^T\alpha_1), F_2(\gamma + x_{2i}^T\alpha_2); \theta]}{F_1(x_{1i}^T\alpha_1)}$$

$$P(y_{2i} = 1|y_{1i} = 0) = \frac{F_2(x_{2i}^T\alpha_2) - C[F_1(x_{1i}^T\alpha_1), F_2(x_{2i}^T\alpha_2)]}{1 - F_1(x_{1i}^T\alpha_1)}$$

(since$P(y_{1i}{}^* \leq 0) = P(x_{1i}^T\alpha_1 + \varepsilon_{1i} \leq 0) = P(\varepsilon_{1i} \leq -x_{1i}^T\alpha_1) = F_1(-x_{1i}^T\alpha_1) = 1 - F_1(x_{1i}^T\alpha_1)$)
Radice, et. al., (2015)

We computed the average of the percentage absolute bias for SATE as:

$\frac{abs\,(SATE - ATE)}{abs\,(ATE)} \times 100$percent (12)

Where SATE is sample average treatment effect that calculated from CBP model and ATE is the true value of average treatment effect. Small average of the percentage absolute bias for SATE refers to the best model see Terza, et. al., [14].

VI. Simulation Study

In simulating, the DGPs are all based on the following recursive bivariate probit model:

$y_1 = 1(\alpha_1 x_1 + \varepsilon_1 > 0)$

$y_2 = 1(\alpha_0 + \gamma y_1 + \alpha_2 x_2 + \varepsilon_2 > 0)$

With true value of parameters listed as $\alpha_1 = 1$, $\alpha_0 = 0.5$, $\gamma = 0.5$,$\alpha_2 =$ -0.5 and x_1, x_2 are both independent identically distributed Gaussian with mean zero and standard deviation of 1. Note that x_1 is instrument variable that affects y_1 but not directly effect on y_2. The binary variable y_1 is endogenous with respect to y_2, because the two are drawn from a copula with nonzero dependence parameter θ. Dependence parameters are not directly comparable among different copulas, so a standard practice is to convert dependence parameters to measures of concordance, such as Kendall's Tau.

In the simulation, the dependence parameters for the FGM, and Frank copulas have been chosen to yield the same value of Kendall's Tau. Sample sizes were set to 1000, 5000 and 10000 and the number of replicates to 250.

To closely approximate the true average treatment effect (ATE), we averaged
$\Phi(\alpha_0 + \gamma + \alpha_2 x_2) - \Phi(\alpha_0 + \alpha_2 x_2)$over a simulated super sample of size five million at the true parameter values, where Φ denotes the comulativedensity function of standard normal distribution.

Table 1 and 2 in appendix show results for FGM and Frank DGPs with sample sizes 1000, 5000, and 10000 and the dependence parameters θ for the FGM and Frank copulas are set, respectively, to 0.675 and 1.40, which assures that Kendall's Tau approximately 0.15. The Gaussian copula bivariate probit model, which do attempt to address endogeneity bias, appears to accurately estimate of ATE for FGM and Frank copulas DGPs. The models were fitted using SemiParBIVProbit (list (eq1, eq2), BivD = D), where eq1 and eq2 were specified according to the simulated y_1 and y_2 above, and D was equal to "N" for normal (Gaussian) copula, "FGM" for FGM copula and "Frank" for Frank copula. The sample average treatment effect for each replicate and fitted model was extracted using AT( ) from the package SemiParBIVProbit. For each model and case considered, we calculated the percentage bias for SATE as defined in eq(12).

Table 3 and 4 in appendix show DGPs with stronger dependence: a Frank with $\theta = 5$ (Kendall's Tau = 0.45), and a Frank with $\theta = 7$ (Kendall's Tau = 0.56). (The FGM copula cannot handle large dependence). That increase in dependence does not appear to affect the accuracy of the Gaussian copula bivariate probit model estimates.

VII. Conclusion

This paper presents a copula-based approach of a bivariate binary response model for estimating the effect of binary endogenous (treatment) variable in a binary response model. In our study, we show that the Gaussian CBP model suffices when simulated the model errors using FGM and Frank copula with low and high dependence parameter θ in our DGPs. We used the ATE as our basis comparison and we found that the mean absolute percentage bias for ATE when using Gaussian CBP model in FGM and Frank DGPs can be small and still small as the sample size increases. This paper do not prove that the Gaussian CBP model suffices in all setting. Perhaps DGPs with asymmetric dependence would pose problems for the Gaussian copula.

Reference

[1]. Chib, S. and E. Greenberg. (2007). Semiparametric modeling and estimation of instrumental variable models. Journal of Computational and Graphical Statistics 16:86–114.

[2]. Gitto, L., Santoro, D., &Sobbrio, G. (2006). Choice of dialysis treatment and type of medical unit (private vs public), application of a recursive bivariate probit. Health Economics, 15, 1251–1256.

[3]. Goldman, D. P., Bhattacharya, J., McCaffrey, D. F., Duan, N., Leibowitz, A. A., Joyce, G. F., & Morton, S. C. (2001). Effect of insurance on mortality in an HIV-positive population in care. Journal of the American Statistical Association, 96, 883–894.

[4]. Greene, W.H., 2008. Econometric Analysis, sixth ed. Pearson/Prentice Hall, UpperSaddle River, NJ.

[5]. Hasebe, T. (2013). Binary effects of a bivariate binary choice model. Economics Letters 121 ,298-301.

[6]. Ieva, F. and Marra, G. and Paganoni, A. M. and Radice, R. (2014). A semiparametric Bivariate Probit Model for Joint Modeling of Outcomes in STEMI Patients. Computational and Mathematical Methods in Medicine, volume 2014, Article ID 240435.

[7]. Latif, E. (2009). The impact of diabetes on employment in Canada. Health Economics, 18, 577–589.

[8]. Li, Y. & Jensen, G. A. (2011). The impact of private long-term care insurance on the use of long-term care. Inquiry, 48(1), 34–50.

[9]. Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics. CAMBRIDGE UNIVARSITY PRESS.

[10]. Marra, G. and R. Radice. (2011). Estimation of a semiparametric recursive bivariate probit model in the presence of endogeneity. Canadian Journal ofStatistics 39:259–279.

[11]. Marra, G., Papageorgiou, G., and R. Radice. (2014). Estimation of a semiparametric recursive bivariate probit model with Nonparametric Mixing. Australian & New Zealand Journal of Statistics. 55(3), 321-342. Australian & New Zealand Journal of StatisAustrali&

[12]. Nelsen, R. B. (2006), An introduction to copulas, second edition. Springer Verlag, New York.

[13]. Radice, Rosalba, GiampieroMarra and MalgorzataWojtys.(2015). "Copula Regression Spline Models for Binary Outcomes." Statistics and Computing, DOI 10.1007/s11222-015-9581-6.

[14]. Terza, J. V., Bradford, W. D. and Dismuke, C. E. (2007). The use of linear Instrumental Variables Methods in Health Services Research and Health Economics: A Cautionary Note. Health Research and Educational Trust, DOI: 10.1111/j.1475-6773.2007.00807.x

[15]. Trivedi, P. K. and Zimer, D. M. (2005). Copula Modeling: An Introduction for Practitioners. Foundation and Trends in econometrics.

[16]. Winkelmann, R. (2012). Copula Bivariate Probit Models: With Application to Medical Expenditures. Health Econ, vol. 21.

[17]. Zimmer, D. M. (2016). Using copulas to estimate the coefficient of a binary endogenous regressor in account regression model.

Appendix

Table 1. Simulation for FGM copula

N3 = 10000							
FGM CBP model	1.001149	0.4982284	-0.4995238	0.5035147	0.1427619	0.6743269	

Table 2.
Simulation for
Frank copula

N3 = 10000	N1 = 1000				N2 = 5000						
Frank CBP model		True	Gaussian CBP model	FGM CBP model		True	Gaussian CBP model	FGM CBP model		True	Gaussian CBP model
0.9989947	α_1	1	1.005187	1.006079	α_1	1	1.001004	1.001296	α_1	1	1.000907
0.4971741	α_0	0.5	0.4985238	0.4984622	α_0	0.5	0.4994790	0.5021456	α_0	0.5	0.4948743
-0.4992915	α_2	-0.5	-0.5013292	-0.5032975	α_2	-0.5	-0.4985607	-0.5002859	α_2	-0.5	-0.4977632
0.5011487	γ	0.5	0.4986569	0.5084590	γ	0.5	0.4940744	0.4976153	γ	0.5	0.5012719
0.1423380	AT	0.142	0.1409108	0.1432472	AT	0.142	0.1402814	0.1408629	AT	0.142	0.1425581
0.366753	Bias %		0.6310619	1.0165108	Bias %		1.0748947	0.6648772	Bias %		0.5305609

Table 3.
Simulation for
Frank copula

N3 = 10000	N1 = 1000				N2 = 5000						
Frank CBP model		True	Gaussian CBP model	Frank CBP model		True	Gaussian CBP model	Frank CBP model		True	Gaussian CBP model
0.9989414	α_1	1	1.004674	1.004919	α_1	1	1.000531	1.000802	α_1	1	0.9987166
0.5003007	α_0	0.5	0.4942799	0.5001402	α_0	0.5	0.4940852	0.4996772	α_0	0.5	0.4916808
-0.5010164	α_2	-0.5	-0.5040388	-0.5055079	α_2	-0.5	-0.5003193	-0.5018835	α_2	-0.5	-0.4977256
0.5007102	γ	0.5	0.5162705	0.5126300	γ	0.5	0.5060013	0.5034354	γ	0.5	0.5034607
0.1417918	AT	0.142	0.1452648	0.1438769	AT	0.142	0.1436160	0.1424872	AT	0.142	0.1433988
0.01841709	Bias %		2.430508	1.451819	Bias %		1.2678683	0.4719571	Bias %		1.114701

Table 4.
Simulation for
Frank copula

N3 = 10000	N1 = 1000				N2 = 5000						
Frank CBP model		True	Gaussian CBP model	Frank CBP model		True	Gaussian CBP model	Frank CBP model		True	Gaussian CBP model
1.000247	α_1	1	1.004855	1.005482	α_1	1	0.9962119	0.9973054	α_1	1	0.9977290
0.5003403	α_0	0.5	0.4904211	0.5010822	α_0	0.5	0.4879841	0.5003209	α_0	0.5	0.4875702
-0.5000577	α_2	-0.5	-0.4947408	-0.4992289	α_2	-0.5	-0.4961265	-0.5006472	α_2	-0.5	-0.4964391
0.5005560	γ	0.5	0.4994923	0.5016702	γ	0.5	0.4997089	0.4991133	γ	0.5	0.5020173
0.1418082	AT	0.142	0.1413659	0.1407484	AT	0.142	0.1427083	0.1413707	AT	0.142	0.1433287
0.006862183	Bias %		0.3187422	0.7541368	Bias %		0.6278360	0.3153538	Bias %		1.06530686

N1 = 1000				N2 = 5000					
	True	Gaussian CBP model	Frank CBP model		True	Gaussian CBP model	Frank CBP model		Gaussian CBP model
α_1	1	1.000907	1.000953	α_1	1	0.9998941	1.0009036	α_1	0.999240
α_0	0.5	0.4842278	0.4960109	α_0	0.5	0.4858691	0.4980750	α_0	0.4876629
α_2	-0.5	-0.5018673	-0.5066532	α_2	-0.5	-0.4945183	-0.4991811	α_2	-0.4953325
γ	0.5	0.5042183	0.5070663	γ	0.5	0.5043205	0.5064246	γ	0.4996052
AT	0.142	0.1428195	0.1421949	AT	0.142	0.1439929	0.1432966	AT	0.1427984
Bias %		0.7062491	0.2658192	Bias %		1.533679	1.042656	Bias %	0.691374113

Sulafah M.Binhimd "Using Copulas To Estimate The Effect Of A Binary Endogenous VariableIn A Binary Response Model." International Journal of Engineering Inventions (IJEI), vol. 6, no. 10, 2017, pp. 25-34.