Smooth Tests of Copula Specification under General Censorship

Juan Lin∗ Ximing Wu†

October 24, 2013

Abstract

We propose a family of smooth tests of copula specification of multivariate copula models under general censorship. We consider both parametric and semiparametric testing procedures. We establish the large sample properties of both approaches. Our Monte Carlo simulations show that the parametric tests, when correctly specified, perform well but are sensitive to mis-specification of the marginal distributions. The semiparametric tests provide satisfactory overall performance, which is comparable to and sometimes superior to that of correctly-specified parametric tests. Our tests are distribution free and provide guidance on alternative copula specifications when the null hypothesis is rejected. We illustrate the usefulness of the proposed tests in two empirical applications.

Keywords: copula specification; smooth test; censorship; empirical copula distribution

JEL Classification: C12; C14; C34

---

∗Risk Management Institute, National University of Singapore, 21 Heng Mui Keng Terrace, I3 Building, Level 4, Singapore, 119613. email: rmilj@nus.edu.sg
†Corresponding author. Department of Agricultural Economics, Texas A&M University, College Station, TX, 77843. Tel: (979) 458-1355, fax: (979) 862-1563, email: xwu@tamu.edu
1 Introduction


It has been widely appreciated that the specification of copula models is of great importance since different copulas lead to multivariate models that may have very different dependence properties. For instance the Gaussian copula only allows symmetric dependence and does not accommodate tail dependence. A number of studies have attempted to address the issue of copula specification. For general overviews of copula specification tests, see Chen and Fan (2005, 2006), Genest et al. (2009), Berg (2009), Chen et al. (2010) and Fernández (2013). Recently, Chen (2007) developed moment based copula tests for multivariate dynamic models whose marginals and copulas are both parametrically specified. Lin and Wu (2012) proposed a family of smooth tests for semiparametric multivariate copula-based models. One particularly appealing feature of their methods is that the tests are not only diagnostic but also conductive to improved copula specifications. More specifically, when a null hypothesis is rejected under a selected set of moment conditions, alternative copula functions can be subsequently constructed by augmenting the null distribution with these extra moment conditions in the spirit of Efron and Tibshirani (1996). However, their methods are only applicable to complete data. In practice, censored data are frequently encountered in many areas, such as risk management, insurance and labor economics. For example, in the field of credit risk, a firm’s lifetime is censored if either default does not occur by the end of the observation period, or if the firm drops out of the sample for reasons other than default such as merger, acquisition or liquidation. Another example of censored data is the insurance losses, which are usually bounded from above because the amount of claim cannot exceed the insured’s policy limit. In labor economics, high earnings are typically topcoded (right censored) in order to protect the confidentiality of respondents. Duration data are often censored because an unemployment spell might still be continuing at the end of the data collection period.

In this paper, we extend the moment-based tests of Chen (2007) and Lin and Wu (2012) to allow for general data censorship in the test of copula specification. Similar to Chen et al. (2010), our tests allow various censoring mechanisms, including the simple random censoring, fixed censoring and no censoring. We adopt the two-stage estimation strategy of
copula models, estimating first the marginal distributions and then the copula coefficients. Regarding the estimation of the marginal distributions, we consider both parametric and nonparametric approaches. We establish the large sample properties of our tests. One advantage of our tests is that they are distribution free and thus do not require re-sampling to obtain critical values. They can be viewed as a generalization of the smooth test of Neyman (1937). In contrast to omnibus tests, moment-based tests face the nontrivial task of moment selection. We base our tests on a series of orthogonal basis functions such that the selection of moments can be proceeded hierarchically. The tests are flexible such that they can be tailored to focus on certain types of hypotheses, such as symmetry, tail dependence and so on.

Our Monte Carlo simulations demonstrate good finite sample size and power performance of our tests. We find that the performance of the semiparametric tests is comparable to that of the tests with correctly specified parametric marginal distributions. On the other hand when the marginal distributions are misspecified, the parametric tests are shown to be biased and dominated by the semiparametric tests. We applied the proposed methods to two empirical examples of censored data. In one example, the null hypothesis is rejected and we subsequently construct alternative copula densities suggested by our testing procedure.

The remainder of this paper is organized as follows. In section 2, we first briefly review Neyman’s smooth test of distributions and its extension to censored data. Section 3 proposes copula smooth tests for censored data under simple hypotheses without unknown parameters. In Section 4, we present general copula smooth tests with estimated marginal distributions and estimated copula parameters and derive their large sample properties. Finite sample performance of the proposed tests are investigated in section 5 through Monte Carlo simulations. Section 6 provides two empirical applications and the last section briefly concludes. All proofs are relegated to the appendix.

Throughout the paper, we use upper case letters to denote the cumulative distribution functions and corresponding lower case letters to denote the density functions. We use subscript $t$ to index observations and subscript $j$ to denote the coordinate of multivariate random vectors. We use a prime symbol to indicate the transpose. For simplicity, we focus on bivariate copulas in this study. Extensions to higher dimensional cases are straightforward.

2 Smooth test of distributions

In this section, we briefly review Neyman’s smooth test of distributions and its extension to censored data of univariate distributions.
2.1 Neyman’s Smooth Test

Let $Y_1, \ldots, Y_n$ be an iid random sample from an unknown continuous distribution. The hypothesis of interest is whether the sample is generated from a specific distribution $F_y$. Neyman’s test of distribution is based on the fact that under the null hypothesis $F_y(Y_1), \ldots, F_y(Y_n)$ are distributed according to the standard uniform distribution. Therefore, the test on a generic distribution $F_y$ is reduced to a test on uniformity upon the probability integral transformation. Neyman (1937) tackled the uniformity test by considering a flexible family of smooth alternatives. In particular, he embedded the uniform distribution in a family of exponential distributions with the following density

$$f_K(v) = \exp(\sum_{k=1}^{K} \lambda_k \phi_k(v) - \lambda_0), \quad v \in [0, 1],$$

where $\lambda_0 = \log \int \exp(\sum_{k=1}^{K} \lambda_k \phi_k(v)) dv$ ensures that $f$ integrates to unity, and $\phi_k$’s are normalized Legendre polynomials on $[0, 1]$ given by

$$\phi_k(v) = \frac{\sqrt{2k+1}}{k!} \frac{d^k}{dv^k}\{(v^2 - v)^k\}, \quad k = 1, \ldots, K,$$

which are orthonormal with respect to the standard uniform distribution.

Under the null hypothesis, $\lambda = (\lambda_1, \ldots, \lambda_K)' = 0$ and the test of uniformity is equivalent to a test on $\lambda = 0$. One can construct a likelihood ratio test on this hypothesis. Alternatively one can use a score test. Although asymptotically equivalent to the likelihood ratio test, the score test is locally optimal and particularly convenient since it does not require the estimation of $\lambda$. Let $\phi = (\phi_1, \ldots, \phi_K)'$ and

$$\hat{\phi} = \frac{1}{n} \sum_{t=1}^{n} \phi(F_y(Y_t)).$$

Neyman’s smooth test is constructed as

$$S = n\hat{\phi}' \hat{\phi}.$$

Under the null hypothesis, $S$ converges in distribution to the $\chi^2$ distribution with $K$ degrees of freedom as $n \to \infty$.

Compared to the omnibus tests, the smooth tests have attractive finite sample properties (see Rayner and Best (1990) for a comprehensive review of the smooth tests). In addition,
the smooth tests have one particular appealing feature. When the null hypothesis \( \lambda = 0 \) is rejected, it is natural to consider \( f_K \) as a plausible alternative density. In this sense, this test is not only diagnostic but also constructive for it provides useful pointers for subsequent analysis.

2.2 Neyman’s smooth test for censored data

Gray and Pierce (1985) extended Neyman’s (1937) smooth test of distribution to accommodate censored data. As in Neyman (1937), they specified their smooth alternatives through the density functions. In the context of failure time models, Peña (1998) proposed a smooth goodness-of-fit test, embedding the baseline hazard function in a larger family of hazard functions. In this study, we formulate the smooth alternatives through the density functions as Neyman (1937) and Gray and Pierce (1985) did. Let \( T_1, \ldots, T_n \) be an i.i.d. sample from an unknown continuous distribution. Let \( D \) denote the censoring variable. We assume that \( D \) is independent of the survival variable \( T \). Suppose \( n \) independently (but not necessarily identically) distributed observations \( (X_t, \delta_t)_{t=1}^n \) are available, where \( X_t = T_t \wedge D_t \) and \( \delta_t = I(T_t \leq D_t) \), in which \( a \wedge b = \min(a, b) \) for real numbers \( a \) and \( b \) and \( I(\cdot) \) is the indicator function.

The hypothesis of interest is whether a random sample \( T_1, \ldots, T_n \) are generated from a specific distribution \( F(t) \). Denote by \( f(t) \) the density function of \( T \) and by \( S(t) \) the survival function of \( T \). Under the null hypothesis, \( S(T_1), \ldots, S(T_n) \) are independently and identically distributed according to the standard uniform distribution. We consider a family of alternative densities for \( T \) of the form

\[
f_\phi(t; \lambda) = f(t) \exp(\sum_{k=1}^{K} \lambda_k \phi_k(S(t)) - \lambda_0),
\]

where \( \phi = (\phi_1, \ldots, \phi_K)' \) are a series of bounded real valued functions defined on \([0, 1]\) and \( \lambda_0 = \log(\int_0^1 \exp(\sum_{k=1}^{K} \lambda_k \phi_k(v))dv) \) is a normalization constant.

Define \( U_t = S(X_t) \). The average log-likelihood of the observations \( \{(X_t, \delta_t)\}_{t=1}^n \) under the alternative (1) is given by

\[
\frac{1}{n} \sum_{t=1}^{n} \left\{ \delta_t[\log f(X_t) + \lambda' \phi(U_t)] + (1 - \delta_t) \log \int_0^{X_t} f(t) \exp(\sum_{k=1}^{K} \lambda_k \phi_k(S(t)))dt - \lambda_0 \right\}
= \frac{1}{n} \sum_{t=1}^{n} \left\{ \delta_t[\log f(X_t) + \lambda' \phi(U_t)] + (1 - \delta_t) \log \int_0^{U_t} \exp(\lambda' \phi(v))dv - \lambda_0 \right\},
\]
where the equality follows via a simple change of variable \( U_t = S(X_t) \). Define for \( u \in (0, 1] \),

\[
\Phi(u, \delta) = \delta \phi(u) + (1 - \delta) \int_0^u \phi(v)dv/u - \int_1^1 \phi(v)dv
\]

\[
= \delta \phi(u) + (1 - \delta)E[\phi(v)|v < u] - E[\phi(v)].
\]

The first order condition of (2) with respect to \( \lambda \), evaluated at \( \lambda = 0 \) is given by

\[
\hat{\Phi}_n = \frac{1}{n} \sum_{t=1}^n \left\{ \delta_t \phi(U_t) + (1 - \delta_t) \int_0^{U_t} \phi(v)dv/U_t \right\} - \int_0^1 \phi(v)dv
\]

\[
= \frac{1}{n} \sum_{t=1}^n \Phi(U_t, \delta_t).
\]  

(3)

The uncensored observations contribute to the score function in the usual form of \( \phi(S(T_t)) - E[\phi(v)] \), while the contributions from the censored variables take the form \( E[\phi(S(T)|T > X_t)] - E[\phi(v)] \). Thus, (3) yields the conditional expectation of the score function for the complete data \( \{T_t\}_{t=1}^n \) given the observed data \( \{(X_t, \delta_t)\}_{t=1}^n \). This is similar to the idea of the EM algorithm for censored data proposed by Dempster et al. (1977).

It is straightforward to show that under the null hypothesis, \( n^{-1} \sum_{t=1}^n E[\Phi(U_t, \delta_t)] = 0 \). Let \( V_n = n^{-1} \sum_{t=1}^n \text{var}[\Phi(U_t, \delta_t)] \). It follows that under the null hypothesis, \( \hat{\Phi}_n \xrightarrow{p} 0 \) and \( \sqrt{n}\hat{\Phi}_n \xrightarrow{d} N(0, V_n) \) as \( n \to \infty \) under suitable regularity conditions. Next define \( \hat{V}_n = n^{-1} \sum_{t=1}^n \Phi(U_t, \delta_t)\Phi(U_t, \delta_t)' \). Neyman’s smooth test under the censorship is given by

\[
LM_n = n\hat{\Phi}_n' \hat{V}_n^{-1} \hat{\Phi}_n.
\]

Under the null hypothesis, \( LM_n \) converges in distribution to the \( \chi^2 \) distribution with \( K \) degrees of freedom as \( n \to \infty \).

3 Smooth tests of copula specification for censored data: simple hypothesis

A copula is a multivariate probability distribution for which the marginal probability of each variable is uniformly distributed. Copulas describe the dependence between random variables. Let \( (x_1, x_2) \) be a random vector from a bivariate distribution function \( F(x_1, x_2) \) with continuous marginal distributions \( F_1(x_1) \) and \( F_2(x_2) \). According to the theorem of Sklar
(1959), there exists a unique copula function $C$ such that

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

for all $x_1, x_2 \in \mathbb{R}$. Suppose $C$ is differentiable, there exists a corresponding copula density $c(v_1, v_2) \equiv \partial^2 / \partial v_1 \partial v_2 C(v_1, v_2)$.

In survival analysis, it is convenient to link the marginal survival functions to their joint survival function via the copula. Several authors have proposed copula specification tests for the censored data, see Frees and Valdez (1998), Klugman and Parsa (1999), Wang and Wells (2000), Denuit et al. (2006), Chen (2007) and Chen et al. (2010). In this study, we extend the smooth tests for copula specification proposed by Lin and Wu (2012) to accommodate censored data.

To fix idea, we consider first the simplest case where the marginal distributions are known and the hypothesized copula distribution is completely specified. Let $(T_1, T_2)$ denote a paired survival times with joint survival function $S(t_1, t_2) = \Pr(T_1 > t_1, T_2 > t_2)$ and continuous marginal survival functions $S_j(\cdot), j = 1, 2$. A straightforward application of Sklar’s (1959) theorem shows that there exists a unique copula function $C(\cdot)$ on $[0, 1]^2$ such that

$$S(t_1, t_2) = C(S_1(t_1), S_2(t_2)).$$

Let $c_0(v_1, v_2; \alpha)$, where $(v_1, v_2) \in [0, 1]^2$, be a class of parametric copula density functions characterized by a finite $p$-dimensional parameter $\alpha \in \mathcal{A}$. We are interested in the following simple hypothesis

$$H_0 : \Pr(c(v_1, v_2) = c_0(v_1, v_2; \alpha_0)) = 1,$$  

against the alternative hypothesis

$$H_1 : \Pr(c(v_1, v_2) = c_0(v_1, v_2; \alpha_0)) < 1,$$

for some $\alpha_0 \in \mathcal{A} \subset \mathbb{R}^p$.

Let $g = (g_1, \ldots, g_K)'$ be a series of linearly independent bounded real valued functions defined on $[0, 1]^2$. Similar to Neyman’s smooth test of uniformity, our smooth test of copula specification can be motivated by a smooth alternative. Consider a family of densities for $(T_1, T_2)$ of the form,

$$c_g(v_1, v_2; \alpha_0, \lambda) = c_0(v_1, v_2; \alpha_0) \exp(\lambda g(v_1, v_2) - \lambda_0)$$
with a normalization constant \( \lambda_0 = \log \int c_0(v_1, v_2; \alpha_0) \exp(\lambda' g(v_1, v_2)) dv_1 dv_2 \). This construction has an appealing information theoretic interpretation: it can be derived as the density that minimizes the conditional Kullback-Leibler information criterion between the target density and the null density subject to side moment conditions associated with \( g \) (cf. Efron and Tibshirani (1996)).

Under the assumption that \((v_1, v_2)\) are distributed according to a distribution in the family of (5), hypothesis (4) is equivalent to that \( \lambda = 0 \). Define

\[
\mu_g = E[g(U_1, U_2)] = \int g(v_1, v_2)c_0(v_1, v_2; \alpha_0) dv_1 dv_2,
\]

where the expectation is taken with respect to the unknown true copula distribution. One can then construct a score test on \( \lambda = 0 \) based on the discrepancy between \( \mu_g \) and its sample counterpart under censorship.

In the presence of right censorship, we use \((D_1, D_2)\) to denote the censoring variables. We assume that the censoring variables \((D_1, D_2)\) are independent of the survival variables \((T_1, T_2)\). Under right censorship, one observes \((X_1, X_2) = (T_1 \land D_1, T_2 \land D_2)\) and a pair of indicators, \((\delta_1, \delta_2) = (I\{T_1 \leq D_1\}, I\{T_2 \leq D_2\})\). Suppose \( n \) independent (but not necessarily identically distributed) observations \( \{(X_{1t}, X_{2t}, \delta_{1t}, \delta_{2t})\}_{t=1}^n \) are available, where \((X_{1t}, X_{2t}) = (T_{1t} \land D_{1t}, T_{2t} \land D_{2t})\) and \((\delta_{1t}, \delta_{2t}) = (I\{T_{1t} \leq D_{1t}\}, I\{T_{2t} \leq D_{2t}\})\). Define \( U_{jt} = S_j(X_{jt}), j = 1, 2, \) and denote by \( C_g(u_1, u_2; \alpha, \lambda) \) the distribution function of the density \( c_g(u_1, u_2; \alpha, \lambda) \). The average log-likelihood of the observations under alternative (5) is then given by \( L_g(\alpha_0, \lambda) = \frac{1}{n-1} \sum_{t=1}^n l_g(U_{1t}, U_{2t}, \delta_{1t}, \delta_{2t}; \alpha_0, \lambda) \), where

\[
l_g(u_{1t}, u_{2t}, \delta_{1t}, \delta_{2t}; \alpha, \lambda)
= \delta_1 \delta_2 \log c_g(u_{1t}, u_{2t}; \alpha, \lambda) + \delta_1(1 - \delta_2) \log \frac{\partial C_g(u_{1t}, u_{2t}; \alpha, \lambda)}{\partial u_1}
+ \delta_2(1 - \delta_1) \log \frac{\partial C_g(u_{1t}, u_{2t}; \alpha, \lambda)}{\partial u_2}
+ (1 - \delta_1)(1 - \delta_2) \log C_g(u_{1t}, u_{2t}; \alpha, \lambda)
= \delta_1 \delta_2 [\lambda' g(u_{1t}, u_{2t}) + \log c_0(u_{1t}, u_{2t}; \alpha)]
+ \delta_1(1 - \delta_2) \log \int_0^{u_{2t}} \exp(\lambda' g(u_{1t}, u_2))c_0(u_{1t}, u_2; \alpha) du_2
+ \delta_2(1 - \delta_1) \log \int_0^{u_{1t}} \exp(\lambda' g(u_{1}, u_{2t}))c_0(u_{1}, u_{2t}; \alpha) du_2
+ (1 - \delta_1)(1 - \delta_2) \log \int_0^{u_{2t}} \int_0^{u_{1t}} \exp(\lambda' g(u_1, u_2))c_0(u_{1}, u_{2}; \alpha) du_1 du_2 - \lambda_0. \tag{6}
\]
The first order condition of $L_g(\alpha_0, \lambda)$ with respect to $\lambda$, evaluated at $\lambda = 0$, is given by

$$
\hat{g}(\alpha_0) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{g}(U_{1t}, U_{2t}, \delta_{1t}, \delta_{2t}; \alpha_0),
$$

(7)

where

$$
\mathbf{g}(u_{1t}, u_{2t}, \delta_{1t}, \delta_{2t}; \alpha) = \delta_{1t}\delta_{2t}g(u_{1t}, u_{2t}) + \delta_{1t}(1 - \delta_{2t}) \frac{\int_{0}^{u_{2t}} g(u_{1t}, u_{2t})c_{0}(u_{1t}, u_{2t}; \alpha)du_{2}}{\int_{0}^{u_{2t}} c_{0}(u_{1t}, u_{2t}; \alpha)du_{2}}
$$

$$
+ \delta_{2t}(1 - \delta_{1t}) \frac{\int_{0}^{u_{1t}} g(u_{1t}, u_{2t})c_{0}(u_{1t}, u_{2t}; \alpha)du_{1}}{\int_{0}^{u_{1t}} c_{0}(u_{1t}, u_{2t}; \alpha)du_{1}}
$$

$$
+ (1 - \delta_{1t})(1 - \delta_{2t}) \frac{\int_{0}^{u_{2t}} \int_{0}^{u_{1t}} g(u_{1t}, u_{2t})c_{0}(u_{1t}, u_{2t}; \alpha)du_{1}du_{2}}{C_{0}(u_{1t}, u_{2t}; \alpha)}
$$

$$
- \int g(u_{1t}, u_{2t})c_{0}(u_{1t}, u_{2t}; \alpha)du_{1}du_{2}.
$$

(8)

To simplify notation, we henceforth use $\mathbf{g}(u_{1t}, u_{2t}; \alpha)$ to denote (8). Different types of observations contribute to the score function (7) differently. Under the null hypothesis,

• the score of the uncensored observations takes the usual form

$$
g(U_{1t}, U_{2t}) - \mu_{g} = g(S_{1}(T_{1t}), S_{2}(T_{2t})) - \mu_{g};
$$

• if an observation is censored in both margins, its contribution to the score takes the form

$$
\frac{\int_{0}^{U_{2t}} \int_{0}^{U_{1t}} g(u_{1t}, u_{2t})c_{0}(u_{1t}, u_{2t}; \alpha_{0})du_{1}du_{2}}{C_{0}(U_{1t}, U_{2t}; \alpha_{0})} - \mu_{g} = E[g(S_{1}(T_{1}), S_{2}(T_{2})) | T_{1} > X_{1t}, T_{2} > X_{2t}] - \mu_{g},
$$

which is the difference between the conditional mean of $g$ given the censoring points and the corresponding unconditional mean;

• if an observation is censored, say, in the first margin but not in the second, its contribution to the score function is given by

$$
\frac{\int_{0}^{U_{2t}} g(u_{1t}, U_{2t})c_{0}(u_{1t}, U_{2t}; \alpha_{0})du_{1}}{\int_{0}^{U_{1t}} c_{0}(u_{1t}, U_{2t}; \alpha_{0})du_{1}} - \mu_{g} = E[g(S_{1}(T_{1}), S_{2}(T_{2})) | T_{1} > X_{1t}] - \mu_{g},
$$

which is difference between the conditional mean of $g$ given that $T_{1} > X_{1t}$ and $T_{2} = T_{2t}$ and the unconditional mean.
Thus similar to (3), (8) is the conditional expectation of the score function for the complete data \( \{(T_{1t}, T_{2t})\}_{t=1}^{n} \) given the observed data \( \{(X_{1t}, X_{2t}, \delta_{1t}, \delta_{2t})\}_{t=1}^{n} \). This is again similar to the idea of the EM algorithm proposed by Dempster et al. (1977). Below we will use the EM algorithm to implement the maximum likelihood estimation of the alternative copula function when some of the data are censored.

Let

\[
I_n(\alpha_0) = \frac{1}{n} \sum_{t=1}^{n} \text{var}[g(U_{1t}, U_{2t}; \alpha_0)].
\] 

(9)

It follows that under the null hypothesis, \( \hat{g}(\alpha_0) \xrightarrow{p} 0 \) and \( \sqrt{n} \hat{g}(\alpha_0) \xrightarrow{d} N(0, I_n(\alpha_0)) \) as \( n \to \infty \) under suitable regularity conditions.

Next define \( \tilde{I}_n(\alpha_0) = n^{-1} \sum_{t=1}^{n} g(U_{1t}, U_{2t}; \alpha_0)g(U_{1t}, U_{2t}; \alpha_0)' \). We construct a smooth test for copula specification as follows:

\[
Q_n = n \hat{g}(\alpha_0)'[\tilde{I}_n(\alpha_0)]^{-1} \hat{g}(\alpha_0).
\] 

(10)

Under null hypothesis (4), the test statistic (10) can be shown to converge in distribution to the \( \chi^2 \) distribution with \( K \) degrees of freedom under suitable regularity conditions given in the next section.

### 4 Smooth tests of copula specification for censored data: composite hypothesis

In practice, the marginal distributions and copula coefficients are usually unknown and have to be estimated. Consequently, the hypothesis of interest becomes a composite one as follows:

\[
H_0 : \Pr(c(u_1, u_2) = c_0(u_1, u_2; \alpha_0)) = 1 \text{ for some } \alpha_0 \in A,
\] 

(11)

against the alternative hypothesis

\[
H_1 : \Pr(c(u_1, u_2) = c_0(u_1, u_2; \alpha)) < 1 \text{ for all } \alpha \in A.
\]

One can estimate the marginal distributions using parametric or nonparametric methods. In general, the tests using parametrically estimated marginal distributions are more efficient if the parametric margins are correctly specified. However, under incorrect distributional
assumptions, the tests can suffer persistent size distortions even if the sample size goes to infinity. In this study, we investigate both testing strategies theoretically and numerically.

4.1 Tests with parametric marginal distributions

We first consider the case where the marginal distributions are assumed to be known up to a finite number of unknown parameters. Suppose that the densities of the marginal distributions are given by \( f_j(\cdot; \beta_j) \), where \( \beta_j = \beta_{j0} \in \mathcal{B} \), a compact subset of \( \mathbb{R}^q \), \( j = 1, 2 \). We use a two-stage parametric estimation procedure to estimate the copula parameter \( \alpha \). The average log-likelihood of the marginal distributions is, for \( j = 1, 2 \),

\[
\mathcal{L}(\beta_j) = n^{-1} \sum_{t=1}^n \ell(X_{jt}, \delta_j; \beta_j),
\]

where

\[
\ell(X_{jt}, \delta_j; \beta_j) = \delta_j \log f_j(X_{jt}; \beta_j) + (1 - \delta_j) \log S_j(X_{jt}; \beta_j).
\]

In the first stage, we estimate \((\hat{\beta}_1, \hat{\beta}_2)\) by the maximum likelihood estimates \((\hat{\beta}_1, \hat{\beta}_2)\) assuming independence, where

\[
\hat{\beta}_j = \arg \max_{\beta_j} \mathcal{L}(\beta_j), \ j = 1, 2.
\]

In the second stage, we take the estimated \((\hat{\beta}_1, \hat{\beta}_2)\) as given. Write \( U_{jt} = S_j(X_{jt}; \beta_{j0}) \) and \( \hat{U}_{jt} = S_j(X_{jt}; \hat{\beta}_j) \), \( t = 1, ..., n \). Under the null hypothesis, the average log-likelihood of the copula function is given by

\[
L(\hat{U}_{1t}, \hat{U}_{2t}; \alpha) = n^{-1} \sum_{t=1}^n l(\hat{U}_{1t}, \hat{U}_{2t}, \delta_{1t}, \delta_{2t}; \alpha),
\]

where

\[
l(u_{1t}, u_{2t}, \delta_{1t}, \delta_{2t}; \alpha) = \delta_{1t} \delta_{2t} \log c_0(u_{1t}, u_{2t}; \alpha) + \delta_{1t}(1 - \delta_{2t}) \log \frac{\partial C_0(u_{1t}, u_{2t}; \alpha)}{\partial u_1} + \delta_{2t}(1 - \delta_{1t}) \log \frac{\partial C_0(u_{1t}, u_{2t}; \alpha)}{\partial u_2} + (1 - \delta_{1t})(1 - \delta_{2t}) \log C_0(u_{1t}, u_{2t}; \alpha).
\]
We then estimate the copula parameters by the maximum likelihood estimation

\[ \hat{\alpha}_n = \arg\max_{\alpha} L(\hat{U}_{1t}, \hat{U}_{2t}; \alpha). \]

One can envision that a smooth test on hypothesis (11) can be constructed based on the estimated score

\[ \hat{g}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n) = \frac{1}{n} \sum_{t=1}^{n} g(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n), \quad (16) \]

which is analogous to (7) with \( U_{jt} \) replaced by \( \hat{U}_{jt} \), \( j = 1, 2 \) and \( \alpha_0 \) by \( \hat{\alpha}_n \). However, test (10) for simple hypotheses is not directly applicable here because of the presence of nuisance parameters in (16). There are two sets of nuisance parameters: the finite dimensional copula parameter \( \alpha \) and the parameters in the marginal distributions \( \beta_j, j = 1, 2 \).

In order to construct a test based on (16), we need to account for the influences of nuisance parameters. We start with the asymptotic distribution of \( \hat{\alpha}_n \), which is studied by Shih and Louis (1995) and Chen et al. (2010). To simplify notation, we now let \( l(u_1, u_2; \alpha) = l(u_1, u_2, \delta_1, \delta_2; \alpha) \). We also denote \( \ell_{\beta_j}(X_{jt}, \delta_j; \beta_j) = \partial \ell(X_{jt}, \delta_j; \beta_j)/\partial \beta_j, \ell_{\beta_j\beta_j}(X_{jt}, \delta_j; \beta_j) = \partial^2 \ell(X_{jt}, \delta_j; \beta_j)/\partial \beta_j \partial \beta_j' \), \( l_\alpha(u_1, u_2; \alpha) = \partial l(u_1, u_2; \alpha)/\partial \alpha \), \( l_{\alpha\alpha}(u_1, u_2; \alpha) = \partial^2 l(u_1, u_2; \alpha)/\partial \alpha \partial \alpha' \), \( l_{\alpha j}(u_1, u_2; \alpha) = \partial^2 l(u_1, u_2; \alpha)/\partial \alpha \partial u_j \), \( l_{\alpha j}(u_1, u_2; \alpha) = l_{\alpha j}(u_1, u_2; \alpha)(\partial u_j/\partial \beta_j) \), \( g_{\alpha}(u_1, u_2; \alpha) = \partial g(u_1, u_2; \alpha)/\partial \alpha \), \( g_{\beta}(u_1, u_2; \alpha) = \partial g(u_1, u_2; \alpha)(\partial u_j/\partial \beta_j) \), for \( j = 1, 2 \).

Next define, for \( j = 1, 2 \),

\[ B_{n,\beta_j} = -\frac{1}{n} \sum_{t=1}^{n} E[l_{\alpha\beta_j}(U_{1t}, U_{2t}; \alpha_0)], \quad (17) \]

\[ D_{n,\beta_j} = -\frac{1}{n} \sum_{t=1}^{n} E[\ell_{\beta_j\beta_j}(X_{jt}, \delta_{jt}; \beta_{j0})], \quad (18) \]

and

\[ W_{jt}^* \equiv W_{jt}^*(X_{jt}, \delta_{jt}; \alpha_0, \beta_{j0}) = -B_{n,\beta_j} D_{n,\beta_j}^{-1} \ell_{\beta_j}(X_{jt}, \delta_{jt}; \beta_{j0}). \quad (19) \]

We can then show the following.

**Theorem 1.** Under regularity conditions 1-8 given in the Appendix, the parametric estimator
\( \hat{\alpha}_n \overset{p}{\to} \alpha_0 \) and \( \sqrt{n}(\hat{\alpha}_n - \alpha_0) \overset{d}{\to} N(0, B_n^{-1}\Sigma_n^* B_n^{-1}) \), where

\[
B_n = -\frac{1}{n} \sum_{t=1}^{n} E[l_{aa}(U_{1t}, U_{2t}; \alpha_0)],
\]

\[
\Sigma^*_n = \frac{1}{n} \sum_{t=1}^{n} \text{var} \left\{ l_{aa}(U_{1t}, U_{2t}; \alpha_0) + \sum_{j=1}^{2} W^*_j \right\}.
\]

Under the null hypothesis, we have \( B_n = n^{-1} \sum_{t=1}^{n} E[l_{aa}(U_{1t}, U_{2t}; \alpha_0)l_{aa}(U_{1t}, U_{2t}; \alpha_0)'] \) by the information matrix equality. In addition, since the conditional expectation of \( l_{aa}(U_{1t}, U_{2t}; \alpha_0) \) with respect to either \( U_{jt} \) is zero, \( l_{aa}(U_{1t}, U_{2t}; \alpha_0) \) is uncorrelated with each \( W^*_j \). Therefore, the asymptotic variance of \( \hat{\alpha}_n \) can be simplified to

\[
B_n^{-1} + \frac{1}{n} B_n^{-1} \sum_{t=1}^{n} \text{var}\{W^*_1 + W^*_2\} B_n^{-1}.
\]

Obviously the asymptotic variance of \( \hat{\alpha}_n \) is reduced to \( B_n^{-1} \) when the marginal distributions are known.

Next let

\[
\hat{B}_n = \frac{1}{n} \sum_{t=1}^{n} l_{aa}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)l_{aa}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)',
\]

\[
\hat{W}^*_j = -\hat{B}_{n,\beta_j} \hat{D}_{n,\beta_j}^{-1} \ell_{\beta_j}(X_{jt}, \delta_{jt}; \hat{\beta}_j),
\]

where

\[
\hat{B}_{n,\beta_j} = \frac{1}{n} \sum_{t=1}^{n} l_{aa}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)l_{\beta_j}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)',
\]

\[
\hat{D}_{n,\beta_j} = \frac{1}{n} \sum_{t=1}^{n} \ell_{\beta_j}(X_{jt}, \delta_{jt}; \hat{\beta}_j)\ell_{\beta_j}(X_{jt}, \delta_{jt}; \hat{\beta}_j)'.
\]

The asymptotic variance of \( \hat{\alpha}_n \) can be consistently estimated by

\[
\hat{B}_n^{-} + \hat{B}_n^{-} \left\{ \frac{1}{n} \sum_{t=1}^{n} \left( \sum_{j=1}^{2} \hat{W}^*_j \right) \left( \sum_{j=1}^{2} \hat{W}^*_j \right)' \right\} \hat{B}_n^{-},
\]

where \( \hat{B}_n^{-} \) is the generalized inverse of \( \hat{B}_n \).
Suppose all conditions of Theorem 1 are satisfied and the regularity conditions and copula parameter are known, the asymptotic variance \( \Omega^* \) can be dropped, resulting in

\[
\Omega^*_n = n^{-1} \sum_{t=1}^{n} \text{var} \left\{ \hat{g}(U_{1t}, U_{2t}; \alpha_0) + \sum_{j=1}^{2} Z^*_j + \hat{g}' \beta^{-1}(l_\alpha(U_{1t}, U_{2t}; \alpha_0) + \sum_{j=1}^{2} W^*_j) \right\}.
\]  

**Remark 1.** Gray and Pierce (1985) indicate that in the smooth tests of censored data, the observed information matrix cannot be used in place of the expected information because the observed information matrix often fails to be positive definite at \((\hat{\alpha}, 0)\), where \(\hat{\alpha}\) is the maximum likelihood estimator of \(\alpha\). This presumably is due to that the likelihood defined in (6) as a function of \((\alpha, \lambda)\) may not be concave everywhere. To circumvent this problem, one can use the outer product of the score function to estimate the information matrix. By doing so, there is no need to assume the functional form of the censoring distribution in spite of the fact that the covariance matrix \( \Omega^*_n \) depends on the censoring distributions (Gray and Pierce, 1985). We adopt this estimation strategy in this study.

**Remark 2.** The asymptotic variance \( \Omega^*_n \) reflects the influence of nuisance parameters in \( \hat{g}(U_{1t}, U_{2t}; \hat{\alpha}_n) \). If the marginal distributions are known, the terms involving \( W \)'s and \( Z \)'s can be dropped, resulting in \( \Omega^*_n = n^{-1} \sum_{t=1}^{n} \text{var} \{ \hat{g}(U_{1t}, U_{2t}; \alpha_0) + \hat{g}' \beta^{-1}(l_\alpha(U_{1t}, U_{2t}; \alpha_0) \right\} \). When the copula parameter \( \alpha \) is known, the third term on the right hand side of (22) can be dropped, resulting in \( \Omega^*_n = n^{-1} \sum_{t=1}^{n} \text{var} \{ \hat{g}(U_{1t}, U_{2t}; \alpha_0) + \sum_{j=1}^{2} Z^*_j \right\} \). When both the marginal distributions and copula parameter are known, \( \Omega^*_n \) is further reduced to \( n^{-1} \sum_{t=1}^{n} \text{var} \{ \hat{g}(U_{1t}, U_{2t}; \alpha_0) \right\} \), which is exactly the variance (9) derived in the previous section for simple hypotheses.

Next we present a consistent estimator for \( \Omega^*_n \). Let

\[
\hat{\phi}_t = \hat{g}(U_{1t}, U_{2t}; \hat{\alpha}_n) + \sum_{j=1}^{2} \hat{Z}^*_j + \hat{g}' \beta^{-1}(l_\alpha(U_{1t}, U_{2t}; \hat{\alpha}_n) + \sum_{j=1}^{2} \hat{W}^*_j),
\]  

We obtain the following asymptotic distribution for \( \hat{g}(U_{1t}, U_{2t}; \hat{\alpha}_n) \).
where
\[
\hat{G}_\alpha = n^{-1} \sum_{t=1}^n g_\alpha (\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n), \quad \hat{G}_{\beta_j} = n^{-1} \sum_{t=1}^n g_{\beta_j} (\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n),
\]
\[
\hat{Z}^*_jt = \hat{G}'_{\beta_j} \hat{D}_{nj}^{-1} \ell_{\beta_j} (X_{jt}, \delta_{jt}; \hat{\beta}_j), \quad j = 1, 2.
\]

It follows that \( \Omega^*_n \) can be estimated consistently by
\[
\hat{\Omega}_n = \frac{1}{n} \sum_{t=1}^n \hat{\phi}_t \hat{\phi}'_t.
\]

We can now construct a smooth test of copula specification as follows.

**Theorem 3.** Suppose all conditions of Theorem 1 and 2 are satisfied. The parametric smooth test of copula specification for censored data is given by
\[
\hat{Q}_n = n \hat{g}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n) \hat{Q}^{-1} \hat{g}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n).
\]
Under null hypothesis (11), \( \hat{Q}_n \xrightarrow{d} \chi^2_K \) as \( n \to \infty \).

### 4.2 Tests with nonparametric marginal distributions

In this section, we relax parametric assumptions on the marginal distributions and instead estimate \( S_j(\cdot) \) by the Kaplan-Meier estimator,
\[
\tilde{S}_j(x) = \prod_{X_j(t) \leq x} \left( 1 - \frac{1}{n - t + 1} \right)^{\delta_j(t)},
\]
where \( X_{j(1)} \leq X_{j(2)} \leq \ldots \leq X_{j(n)} \) are the order statistics of \( \{X_{jt}\}_{t=1}^n \) for \( j = 1, 2 \), and \( \{\delta_{j(t)}\}_{t=1}^n \) are similarly defined. Lai and Ying (1991) establish the consistency of \( \tilde{S}_j(\cdot) \) under the assumption of independent censoring.

Denote \( (\tilde{U}_{1t}, \tilde{U}_{2t}) = (\tilde{S}_1(X_{1t}), \tilde{S}_2(X_{2t})) \). Given \( (\tilde{U}_{1t}, \tilde{U}_{2t}) \), we proceed to estimate copula parameters via the maximum likelihood estimation
\[
\tilde{\alpha}_n = \arg \max_{\alpha} \frac{1}{n} \sum_{t=1}^n l(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha),
\]
where \( l(u_{1t}, u_{2t}; \alpha) \) is defined in (15).
Similar to the test with parametric margins, the test with nonparametric margins is based on the score function

\[ \tilde{g}(\tilde{\alpha}_n) = \frac{1}{n} \sum_{t=1}^{n} g(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n), \]  

where \( g(u_{1t}, u_{2t}; \alpha) \) is defined in (8). We need to account for the influences of the nuisance parameters, including the finite dimensional copula parameter \( \alpha \) and the infinite-dimensional marginal survival function \( S_j, j = 1, 2 \). We start with the asymptotic distribution of \( \tilde{\alpha}_n \), which is studied by Shih and Louis (1995) and Chen et al. (2010). Define

\[ W_{jt} \equiv W_j(X_{jt}, \delta_{jt}; \alpha_0) = E\{l_\alpha(U_{1s}, U_{2s}; \alpha_0)I_j(X_{jt}, \delta_{jt})(X_{js})|X_{jt}, \delta_{jt}\}, \]

\[ I_j(X_{jt}, \delta_{jt})(X_{js}) = -S_j(X_{js}) \left[ \int_{-\infty}^{X_{js}} \frac{dN_{jt}(u)}{P_{n,j}(u)} - \int_{-\infty}^{X_{js}} \frac{I\{X_{jt} \geq u\}dH_j(u)}{P_{n,j}(u)} \right], \]

where \( H_j(x) \equiv -\log(S_j(x)) \) is the cumulative hazard function of \( X_j, N_{jt}(x) = \delta_{jt}I(X_{jt} \leq x) \), \( dN_{jt}(x) = N_{jt}(x) - N_{jt}(x-) \), and \( P_{nj}(x) = n^{-1} \sum_{k=1}^{n} P(X_{jk} \geq x) \).

The asymptotic distribution of \( \tilde{\alpha}_n \) is given by the following theorem.

**Theorem 4.** Suppose \( E[l(U_{1t}, U_{2t}; \alpha)] \) has a unique maximum at \( \alpha_0 \). Under Conditions 1-4, 6-8, and 13-14 given in the Appendix, the semiparametric estimator \( \tilde{\alpha}_n \xrightarrow{p} \alpha_0 \) and \( \sqrt{n}(\tilde{\alpha}_n - \alpha_0) \xrightarrow{d} N(0, B_n^{-1}\Sigma_n B_n^{-1}) \), where \( B_n \) is defined in Theorem 1 and

\[ \Sigma_n = n^{-1} \sum_{t=1}^{n} \text{var}\{l_\alpha(U_{1t}, U_{2t}; \alpha_0)\} + 2 \sum_{j=1}^{2} W_{jt}. \]

Since the conditional expectation of \( l_\alpha(U_{1t}, U_{2t}; \alpha_0) \) with respect to either \( U_{jt} \) is zero, \( l_\alpha(U_{1t}, U_{2t}; \alpha_0) \) is uncorrelated with each \( W_{jt} \). Therefore, the asymptotic variance of \( \sqrt{n}(\tilde{\alpha}_n - \alpha_0) \) can be simplified to

\[ B_n^{-1} + \frac{1}{n} B_n^{-1} \sum_{t=1}^{n} \text{var}\{W_{1t} + W_{2t}\}B_n^{-1}. \]

If the marginal distributions are known, the asymptotic variance of \( \tilde{\alpha}_n \) is reduced to \( B_n^{-1} \).
Next define

$$
\tilde{B}_n = \frac{1}{n} \sum_{t=1}^{n} l_{\alpha}(\tilde{U}_1, \tilde{U}_2; \tilde{\alpha}_n) l_{\alpha}(\tilde{U}_1, \tilde{U}_2; \tilde{\alpha}_n)';
$$

$$
\tilde{W}_{jt} = \frac{1}{n} \sum_{s=1, s \neq t}^{n} l_{\alpha j}(\tilde{U}_1s, \tilde{U}_2s; \tilde{\alpha}_n) \tilde{I}_j(X_{jt}, \delta_{jt})(X_{js}),
$$

where for $j = 1, 2$,

$$
\tilde{I}_j(X_{jt}, \delta_{jt})(X_{js}) = -\tilde{S}_j(X_{js}) \left[ \frac{I(X_{jt} \leq X_{js}) \delta_{jt}}{n^{-1} \sum_{k=1}^{n} I(X_{jk} \geq X_{jt})} - \frac{1}{n} \sum_{l=1}^{n} \frac{I(X_{js} \geq X_{jt}) I(X_{jt} \geq X_{jt}) \delta_{jt}}{[n^{-1} \sum_{k=1}^{n} I(X_{jk} \geq X_{jt})]^2} \right].
$$

As indicated in Chen et al. (2010), an alternative expression for $\tilde{I}_j(X_{jt}, \delta_{jt})(X_{js})$ is

$$
\tilde{I}_j(X_{jt}, \delta_{jt})(X_{js}) = -\tilde{S}_j(X_{js}) \left[ \frac{I\{X_{jt} \leq X_{js}, \delta_{jt} = 1\}}{P_{nj}(X_{jt})} - \sum_{X_{jt} \leq X_{js}} \frac{I(X_{jt} \geq X_{jt}) \Delta \tilde{H}_j(X_{jt})}{P_{nj}(X_{jt})} \right],
$$

where $P_{nj}(x) = n^{-1} \sum_{k=1}^{n} I(X_{jk} \geq x)$ and $\Delta \tilde{H}_j(x) = \frac{I\{\tilde{Y}_j(x) > 0\}}{\tilde{Y}_j(x)} d\tilde{N}_j(x)$, in which $\tilde{Y}_j(x) = \sum_{k=1}^{n} I\{X_{jk} \geq x\}$ and $\tilde{N}_j(x) = \sum_{k=1}^{n} N_{jk}(x)$. $\Delta \tilde{H}_j(x)$ is so-called Nelson’s estimator.

The asymptotic variance of $\tilde{\alpha}_n$ can be consistently estimated by

$$
\tilde{B}^{-1}_n + \tilde{B}_n \left\{ \frac{1}{n} \sum_{t=1}^{n} (\sum_{j=1}^{2} \tilde{W}_{jt}) (\sum_{j=1}^{2} \tilde{W}_{jt})' \right\} \tilde{B}_n^{-1},
$$

where $\tilde{B}_n^{-1}$ is the generalized inverse of $\tilde{B}_n$.

Further define for $j = 1, 2$

$$
Z_{jt} \equiv Z_j(X_{jt}, \delta_{jt}; \alpha_0) = E[g_j(U_{1s}, U_{2s}; \alpha_0) I_j(X_{jt}, \delta_{jt})(X_{js}) \mid X_{jt}, \delta_{jt}],
$$

where $g_j$’s are defined in (A.5) and (A.6). We obtain the following asymptotic distribution of $\tilde{g}(\tilde{\alpha}_n)$.

**Theorem 5.** Suppose all conditions of Theorem 4 are satisfied and the regularity condition 15 given in the Appendix holds. Under null hypothesis (11), $\tilde{g}(\tilde{\alpha}_n) \buildrel p \over \rightarrow 0$ and $\sqrt{n} \tilde{g}(\tilde{\alpha}_n) \buildrel d \over \rightarrow \mathcal{N}(0, \tilde{B}_n^{-1}).$
Remark 3. The various simplifications of variance for the parametric test discussed in Remark 2 apply to its semiparametric counterpart (29) as well.

Next we present a consistent estimator for $\Omega_n$. Define $\tilde{G}_\alpha = n^{-1} \sum_{t=1}^{n} g_{\alpha}(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n)$ and

$$
\tilde{Z}_{jt} = \frac{1}{n} \sum_{s=1, s \neq t}^{n} g_j(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n) I_j(X_{jt}, \delta_{jt})(X_{js}).
$$

Let

$$
\tilde{\varphi}_t = g(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n) + \sum_{j=1}^{2} \tilde{Z}_{jt} + \tilde{G}'_\alpha B^{-1}_n(l_\alpha(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n) + \sum_{j=1}^{2} \tilde{W}_{jt}).
$$

It follows that $\Omega_n$ can be estimated consistently by

$$
\hat{\Omega}_n = \frac{1}{n} \sum_{t=1}^{n} \tilde{\varphi}_t \tilde{\varphi}'_t.
$$

We can now construct a smooth test of copula specification as follows.

Theorem 6. Suppose all conditions of Theorem 4 and 5 are satisfied. The semiparametric smooth test of copula specification for censored data is given by

$$
\tilde{Q}_n = n\tilde{g}(\tilde{\alpha}_n)'\hat{\Omega}_n^{-1}\tilde{g}(\tilde{\alpha}_n).
$$

Under null hypothesis (11), $\tilde{Q}_n \overset{d}{\rightarrow} \chi^2_K$ as $n \rightarrow \infty$.

5 Simulations

We conduct a series of Monte Carlo simulations to assess the finite-sample performance of the proposed tests. We choose the Clayton copula, which is commonly used in bivariate survival analyses, as the null copula distribution. We generate the marginal distributions from the Weibull distribution with shape parameter 2 and scale parameter 1. We consider the following three types of alternative copula distributions:
(1) Radially symmetric copulas: the Gaussian copula \((C_N)\), the Student’s \(t\) copula with four degrees of freedom \((C_{t_4})\), and the Plackett copula \((C_P)\).

(2) Symmetric but not radially symmetric copula: the Gumbel-Hougaard copula \((C_G)\).

(3) Asymmetric copula: the asymmetric Gumbel-Hougaard copula \((C_{AG})\), which is constructed based on the symmetric Gumbel-Hougaard copula using Khoudraji’s device (Khoudraji, 1995) as follows:

\[
C_{AG}(v_1, v_2; \lambda, k, \alpha) = v_1^{1-\lambda}v_2^{1-k}C_G(v_1^{\lambda}, v_2^{k}; \alpha)
\]

with \(\alpha = 4\) and \(k = 0.95\).

For each copula, we consider two scenarios with the Kendall’s tau \(\tau = 0.3\) and 0.6. The parameters for copulas are calculated by inversion of Kendall’s tau. We consider 30% censoring. To achieve 30% censoring, we let the two censoring variables be independently and identically distributed uniformly over \((0, 3)\). We run simulations with sample size \(n = 300, 500\) and each experiment is repeated 1,000 times.

One appealing feature of the proposed smooth test is its flexibility such that one can tailor the moment functions, \(g(v_1, v_2)\), according to his research needs. Several general considerations apply here. First, the functions shall be linearly independent. Second, as is well known about Neyman’s smooth test, its finite sample power suffers when the degree of freedom, \(K\), is large. Usually, a small number of terms is used. For instance, Neyman (1937) recommended \(K \leq 4\); Thomas and Pierce (1979) suggested \(K = 2\) for composite tests. Third, ideally one shall choose the moment functions that best capture deviations from the null distributions.

Taking all these factors into consideration, when there is no a priori reason to focus on a particular direction of deviation, we select the moment functions from the normalized Legendre polynomials. We focus on the first three Legendre polynomials given by, for \(v \in [0, 1]\),

\[
\psi_1(v) = \sqrt{3}(2v - 1), \quad \psi_2(v) = \sqrt{5}(6v^2 - 6v + 1), \\
\psi_3(v) = \sqrt{7}(20v^3 - 30v^2 + 12v - 1).
\]

Our \(g\) functions then consist of elements of the tensor products of these basis functions under the restriction that the maximum number of terms is 4. Table 1 summarizes the various configurations considered in this study.
In Table 1, consider the tensor product of \( \psi \). Other terms are defined similarly. A few explanations are in order for the configurations listed in Table 1. Consider the tensor product of \( \langle \psi_1, \psi_2, \psi_3 \rangle \) as a 3 by 3 matrix with the \((i,j)\)th entry being \( \psi_i \psi_j \). The first group ‘Singleton’ contains single entry of the matrix. The group ‘Diagonal’ contains multiple entries on the diagonal or anti-diagonal. The group ‘Others’ contains other configurations. \( g_o_1 \) and \( g_o_2 \) contain entries in the upper right and lower right triangular matrix with \( i + j \leq 4 \); \( g_o_3 \) contains all entries with \( i, j \leq 2 \); \( g_o_4 \) contains entries with \( i + j \) being an even number and \( 1 \leq i, j \leq 3 \); \( g_o_5 \) contains all entries with \( 2 \leq i, j \leq 3 \).

In what follows, we refer to the smooth tests with moment functions \( g_s_1 \) as \( Q_s_1 \) and all other tests are defined similarly. We conduct two sets of experiments: (1) the margins are estimated parametrically; (2) the margins are estimated nonparametrically. The empirical sizes and powers of the tests with parametric margins are reported in Table 2. The bold columns under the header \( C_C \) correspond to the results under the Clayton copula and reflect the empirical size. To illustrate the consequence of misspecified marginal distributions on the test, we also calculate the empirical size when the marginal distributions are mistakenly estimated under the assumption of the exponential distribution. The results are reported in \textit{italic} under the column header \( C^m_C \). The rest of the table reflects powers of the test against various alternatives. Overall, when the margins are correctly specified, the tests exhibit correct sizes, centering about the nominal value of 5%. In contrast, when the margins are misspecified, the empirical size can significantly deviate from the nominal size depending on the choice of moments.

We report the simulation results for the copula models with nonparametric margins in Table 3. Comparing with the results reported in Table 2, the semiparametric tests seem to have higher powers. We note some common patterns across the two types of tests and
various alternative distributions. Firstly, both sizes and powers improve with sample sizes. Secondly, tests with low order instruments (first and second moments) seem to provide ample powers against the alternative distributions. Thirdly, tests with two or three instruments have satisfactory powers and generally are more powerful than tests with a single instrument. Lastly, almost all tests show higher power under $\tau = 0.6$ than under $\tau = 0.3$.

A closer examination of the results reveals the following additional insights.

- When the alternative distributions are radially symmetric, including the Gaussian, t and Plackett copulas, the terms $\psi_1\psi_2$ and $\psi_2\psi_1$ contribute significant discriminant powers, while some other terms like $\psi_1\psi_1$, $\psi_2\psi_3$ and $\psi_3\psi_2$ also show good discriminant powers. For the t copula, which is known to have fat tails, terms like $\psi_2\psi_2$ and $\psi_3\psi_3$ show higher discriminant powers than for the other two copulas in this group.

- For the Gumbel copula: all singleton tests that contain the $\psi_2$ term, including $\psi_1\psi_2$, $\psi_2\psi_1$, $\psi_2\psi_2$, $\psi_2\psi_3$, and $\psi_2\psi_3$, seem to have good powers. Since every test with multiple instruments contains at least one $\psi_2$ term, they all have significant powers.

- As for the asymmetric Gumbel-Hougaard copula, the terms $\psi_1\psi_2$ and $\psi_2\psi_3$ are consistently powerful across different degrees of dependence, even with the small sample size at $n = 300$. In addition, $\psi_1\psi_1$ and $\psi_2\psi_1$ also show good discriminant powers. Consequently, tests including these terms show significant powers.

- The instruments $\psi_1\psi_3$ and $\psi_3\psi_1$ seem to offer little power against all alternative copulas.

Overall, the simulation results suggest that tests with multiple instruments outperform those with a single instrument in most cases. This is important because in many practical cases, researchers have no a priori guidance on how to select a particular direction for specification tests. It is reassuring that a combination of a small number of terms appears to be a safe testing strategy that delivers satisfactory performance. Regarding the choice between parametric and nonparametric tests, since the parametric forms of the marginal distributions are usually not known, the nonparametric tests seem to be a more robust strategy, given the reported severe bias of the parametric tests under misspecification. As a matter of fact, our simulations show that even under correct parametric specifications, the parametric tests are often outperformed by the nonparametric tests, sometimes by significant margins. (See Charpentier et al. (2007) on reasons why the empirical distributions rather than the parametric distributions may be preferred in copula estimations and inferences, even when the parametric distributions are correctly specified.)
Table 2: Empirical sizes and powers of parametric tests

<table>
<thead>
<tr>
<th>τ</th>
<th>C_C</th>
<th>C_C^m</th>
<th>C_N</th>
<th>C_{t4}</th>
<th>C_P</th>
<th>C_G</th>
<th>C_AG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 300</td>
<td>n = 500</td>
<td>n = 300</td>
<td>n = 500</td>
<td>n = 300</td>
<td>n = 500</td>
<td>n = 300</td>
</tr>
<tr>
<td>0.3</td>
<td>Q_{S_1}</td>
<td>4.2</td>
<td>4.5</td>
<td>100</td>
<td>100</td>
<td>33.2</td>
<td>52.1</td>
</tr>
<tr>
<td></td>
<td>Q_{S_2}</td>
<td>5.3</td>
<td>3.6</td>
<td>5.1</td>
<td>7.1</td>
<td>13.4</td>
<td>20.3</td>
</tr>
<tr>
<td></td>
<td>Q_{S_3}</td>
<td>4.9</td>
<td>4.6</td>
<td>90.2</td>
<td>97.1</td>
<td>5.4</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Q_{S_4}</td>
<td>4.8</td>
<td>4.6</td>
<td>55.9</td>
<td>88.8</td>
<td>55.7</td>
<td>78.8</td>
</tr>
<tr>
<td></td>
<td>Q_{S_5}</td>
<td>5.7</td>
<td>4.6</td>
<td>55.2</td>
<td>87.5</td>
<td>57.1</td>
<td>78.4</td>
</tr>
<tr>
<td></td>
<td>Q_{S_6}</td>
<td>4.3</td>
<td>5.1</td>
<td>100</td>
<td>100</td>
<td>4.9</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td>Q_{S_7}</td>
<td>3.6</td>
<td>5.7</td>
<td>100</td>
<td>100</td>
<td>5.5</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>Q_{S_8}</td>
<td>5.3</td>
<td>4.9</td>
<td>54.4</td>
<td>84.8</td>
<td>26.1</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Q_{S_9}</td>
<td>4.3</td>
<td>5.7</td>
<td>57.4</td>
<td>85.7</td>
<td>26.7</td>
<td>40.8</td>
</tr>
<tr>
<td></td>
<td>Q_{D_1}</td>
<td>5.1</td>
<td>4.7</td>
<td>100</td>
<td>100</td>
<td>26</td>
<td>45.6</td>
</tr>
<tr>
<td></td>
<td>Q_{D_2}</td>
<td>6.1</td>
<td>4.3</td>
<td>100</td>
<td>100</td>
<td>24.2</td>
<td>40.8</td>
</tr>
<tr>
<td></td>
<td>Q_{D_3}</td>
<td>5.7</td>
<td>5</td>
<td>68.7</td>
<td>93.5</td>
<td>59.7</td>
<td>83.2</td>
</tr>
<tr>
<td></td>
<td>Q_{D_4}</td>
<td>4.1</td>
<td>4.7</td>
<td>100</td>
<td>100</td>
<td>9</td>
<td>14.1</td>
</tr>
<tr>
<td></td>
<td>Q_{O_1}</td>
<td>5.2</td>
<td>5</td>
<td>100</td>
<td>100</td>
<td>47</td>
<td>74.2</td>
</tr>
<tr>
<td></td>
<td>Q_{O_2}</td>
<td>5.6</td>
<td>4.7</td>
<td>64.9</td>
<td>90.2</td>
<td>56.2</td>
<td>71.6</td>
</tr>
<tr>
<td></td>
<td>Q_{O_3}</td>
<td>4.8</td>
<td>4.9</td>
<td>100</td>
<td>100</td>
<td>48.1</td>
<td>78.3</td>
</tr>
<tr>
<td></td>
<td>Q_{O_4}</td>
<td>5.4</td>
<td>4.7</td>
<td>100</td>
<td>100</td>
<td>21</td>
<td>39.2</td>
</tr>
<tr>
<td></td>
<td>Q_{O_5}</td>
<td>6</td>
<td>5.9</td>
<td>92.5</td>
<td>99.9</td>
<td>26</td>
<td>35.8</td>
</tr>
</tbody>
</table>

Note: The null hypothesis is that the data are generated according to the Clayton copula. The bold and italic entries under the headers C_C and C_C^m represent the sizes in percentage under correct and incorrect specifications of the marginal distributions respectively. The remaining columns reflect the empirical powers in percentage. The DGP of the marginals is the Weibull distribution with the shape parameter 2 and scale parameter 1. The nominal size in all cases is 5%. The number of replications is 1,000.
Table 3: Empirical sizes and powers of semiparametric tests

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$C_C$</th>
<th>$C_N$</th>
<th>$C_{t_4}$</th>
<th>$C_P$</th>
<th>$C_G$</th>
<th>$C_{AG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n=300$</td>
<td>$n=500$</td>
<td>$n=300$</td>
<td>$n=500$</td>
<td>$n=300$</td>
<td>$n=500$</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q_{S_1}$</td>
<td>5.4</td>
<td>4.4</td>
<td>62.4</td>
<td>87.5</td>
<td>31.8</td>
<td>59.5</td>
</tr>
<tr>
<td>$Q_{S_2}$</td>
<td>4.7</td>
<td>5.9</td>
<td>11.7</td>
<td>19</td>
<td>67.7</td>
<td>91.7</td>
</tr>
<tr>
<td>$Q_{S_3}$</td>
<td>4.8</td>
<td>5.2</td>
<td>5.2</td>
<td>4.9</td>
<td>19.1</td>
<td>33.9</td>
</tr>
<tr>
<td>$Q_{S_4}$</td>
<td>4.4</td>
<td>5.5</td>
<td>70.6</td>
<td>87.3</td>
<td>69</td>
<td>89.3</td>
</tr>
<tr>
<td>$Q_{S_5}$</td>
<td>5.7</td>
<td>5.2</td>
<td>67.9</td>
<td>86.9</td>
<td>72.3</td>
<td>90.4</td>
</tr>
<tr>
<td>$Q_{S_6}$</td>
<td>5.9</td>
<td>5.8</td>
<td>6.8</td>
<td>7.9</td>
<td>5.9</td>
<td>7.2</td>
</tr>
<tr>
<td>$Q_{S_7}$</td>
<td>4.9</td>
<td>6.1</td>
<td>3.9</td>
<td>7.3</td>
<td>7.3</td>
<td>7.2</td>
</tr>
<tr>
<td>$Q_{S_8}$</td>
<td>5.5</td>
<td>4.5</td>
<td>29.7</td>
<td>43.2</td>
<td>53.5</td>
<td>77.1</td>
</tr>
<tr>
<td>$Q_{S_9}$</td>
<td>5.5</td>
<td>5.1</td>
<td>31.1</td>
<td>42.5</td>
<td>53.4</td>
<td>75.9</td>
</tr>
<tr>
<td>$Q_{D_1}$</td>
<td>6.1</td>
<td>5.1</td>
<td>51.9</td>
<td>84.5</td>
<td>74.2</td>
<td>97.1</td>
</tr>
<tr>
<td>$Q_{D_2}$</td>
<td>5.2</td>
<td>5.1</td>
<td>46.4</td>
<td>79.4</td>
<td>68.9</td>
<td>93.1</td>
</tr>
<tr>
<td>$Q_{D_3}$</td>
<td>4.1</td>
<td>6.1</td>
<td>71.7</td>
<td>91.8</td>
<td>80.9</td>
<td>95.2</td>
</tr>
<tr>
<td>$Q_{D_4}$</td>
<td>6.2</td>
<td>5.2</td>
<td>9.1</td>
<td>12.6</td>
<td>58.1</td>
<td>85.1</td>
</tr>
<tr>
<td>$Q_{D_5}$</td>
<td>6.6</td>
<td>4.9</td>
<td>60.2</td>
<td>86.9</td>
<td>76.6</td>
<td>95.4</td>
</tr>
<tr>
<td>$Q_{O_1}$</td>
<td>5.7</td>
<td>4.7</td>
<td>66</td>
<td>88.7</td>
<td>83.2</td>
<td>97.4</td>
</tr>
<tr>
<td>$Q_{O_2}$</td>
<td>8.8</td>
<td>4.6</td>
<td>62.5</td>
<td>87.6</td>
<td>80</td>
<td>96.4</td>
</tr>
<tr>
<td>$Q_{O_3}$</td>
<td>5</td>
<td>5.2</td>
<td>43.8</td>
<td>76.6</td>
<td>63.3</td>
<td>93.3</td>
</tr>
<tr>
<td>$Q_{O_4}$</td>
<td>6.6</td>
<td>6</td>
<td>30.3</td>
<td>44.5</td>
<td>67.3</td>
<td>88.9</td>
</tr>
<tr>
<td>$Q_{O_5}$</td>
<td>4.8</td>
<td>5.9</td>
<td>72.9</td>
<td>94.7</td>
<td>49.8</td>
<td>79.4</td>
</tr>
<tr>
<td>$Q_{S_1}$</td>
<td>4.8</td>
<td>5.9</td>
<td>98.7</td>
<td>100</td>
<td>99.4</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{S_2}$</td>
<td>6.1</td>
<td>5.7</td>
<td>60.3</td>
<td>87.3</td>
<td>90.6</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{S_3}$</td>
<td>6.3</td>
<td>6.3</td>
<td>99.1</td>
<td>100</td>
<td>98.4</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{S_4}$</td>
<td>4.6</td>
<td>5.5</td>
<td>99.1</td>
<td>100</td>
<td>98.5</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{S_5}$</td>
<td>8.7</td>
<td>6.9</td>
<td>18.3</td>
<td>36.8</td>
<td>21.2</td>
<td>39</td>
</tr>
<tr>
<td>$Q_{S_6}$</td>
<td>8.5</td>
<td>5.5</td>
<td>15.9</td>
<td>31.8</td>
<td>20.7</td>
<td>42</td>
</tr>
<tr>
<td>$Q_{S_7}$</td>
<td>6.2</td>
<td>5.6</td>
<td>96.5</td>
<td>99.9</td>
<td>98.2</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{S_8}$</td>
<td>4.5</td>
<td>5.9</td>
<td>97.2</td>
<td>100</td>
<td>98.4</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{D_1}$</td>
<td>4.9</td>
<td>4.5</td>
<td>95.6</td>
<td>100</td>
<td>98</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{D_2}$</td>
<td>4.8</td>
<td>4.6</td>
<td>95.1</td>
<td>100</td>
<td>98.5</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{D_3}$</td>
<td>4.9</td>
<td>5.9</td>
<td>99.2</td>
<td>100</td>
<td>98.9</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{D_4}$</td>
<td>7.2</td>
<td>5.5</td>
<td>98.5</td>
<td>100</td>
<td>99.6</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{O_1}$</td>
<td>7</td>
<td>5.4</td>
<td>99.2</td>
<td>100</td>
<td>99.8</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{O_2}$</td>
<td>7.2</td>
<td>5</td>
<td>99.5</td>
<td>100</td>
<td>99.7</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{O_3}$</td>
<td>6.7</td>
<td>6.2</td>
<td>98.8</td>
<td>100</td>
<td>99.5</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{O_4}$</td>
<td>7.7</td>
<td>6.3</td>
<td>94.3</td>
<td>100</td>
<td>98.6</td>
<td>100</td>
</tr>
<tr>
<td>$Q_{O_5}$</td>
<td>8.3</td>
<td>7.2</td>
<td>98.3</td>
<td>100</td>
<td>99.5</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: The null hypothesis is that the data are generated according to the Clayton copula. The bold entries under the header $C_C$ represent the sizes in percentage and the remaining columns reflect the empirical powers in percentage. The DGP of the marginals is the Weibull distribution with the shape parameter 2 and scale parameter 1. The nominal size in all cases is 5%. The number of replications is 1,000.
6 Empirical examples

In this section, we apply the proposed tests to two real world examples. To save space, we report only results of the nonparametric tests with more than one testing instrument, which are recommended based on our simulation results. In addition to testing for copula specification, when a null hypothesis is rejected, we also estimate alternative copula densities suggested by our tests and compare them to the hypothesized null distribution.

6.1 Couples’ wage and salary earnings

Our first example concerns the dependence structure of married couples’ total wage and salary earnings. We randomly selected the earnings of 1,500 couples from the March 2001 Supplement to the Current Population Survey (CPS). The data were restricted to dual-earner couples in which both husband and wife were aged 25 to 55 at the time of the survey, eligible to labor force, and reporting non-negligible individual annual earnings ($1,000 or more). In this dataset, there are 142 observations where either husband’s earnings or wife’s earnings is right censored at some top-coded values, and 8 observations where both spouses’ earnings are censored. For the details of topcoding rules for earnings, please refer to March 2001 Current Population Survey (Bureau of the Census) \(^1\).

We first estimate the marginal distributions by the Kaplan-Meier estimator. The scatterplot of the couples’ earnings on the logarithm scale and the rank transformed data are displayed in Figure 1a and 1b. Only uncensored observations are plotted. Figure 1a shows that more data points are clustered below the diagonal line, suggesting that a husband is more likely to earn more than his wife. Figure 1b shows apparent positive right-tail dependence between the couples’ wage and salary earnings: men with high earnings tend to pair with women with high earnings, which is consistent with the theory of “positive assortative mating”. At the same time, the left-tail dependence is also not negligible.

We tested the null hypothesis of the Gumbel copula on the couples’ earnings data. The two-step estimator of the parameter of the Gumbel copula is \( \hat{\alpha}_n = 1.2247 \). The P-values of constructed tests are reported in Table 6. The null hypothesis of Gumbel copula is rejected at the 5% level by all tests using more than one moment function.

For each test that rejects the null hypothesis, we can subsequently construct an alternative copula density by incorporating the set of moment functions involved in the test. For instance, we denote by \( c_{D_3} \) the alternative copula density corresponding to test \( Q_{D_3} \). It

\(^1\)http://www.census.gov/prod/techdoc/cps/cpsmar01.pdf
follows that
\[ c_{D3}(v_1, v_2; \tilde{\alpha}_n, \tilde{\lambda}) = c_G(v_1, v_2; \tilde{\alpha}_n) \exp(\tilde{\lambda}_1 \psi_1(v_1) \psi_2(v_2) + \tilde{\lambda}_2 \psi_2(v_1) \psi_1(v_2) - \tilde{\lambda}_0), \]
where
\[ \tilde{\lambda}_0 = \log \int c_G(v_1, v_2; \tilde{\alpha}_n) \exp(\tilde{\lambda}_1 \psi_1(v_1) \psi_2(v_2) + \tilde{\lambda}_2 \psi_2(v_1) \psi_1(v_2)) dv_1 dv_2. \]
All other densities are constructed in a similar manner. It is not a trivial task to obtain the maximum likelihood estimates \( \tilde{\alpha}_n, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_K \) due to the complicate functional form of the log-likelihood function (6). In this study we adopt a two-step estimation approach. In the first step, we estimate the parameter \( \alpha \) by restricting \( \lambda_k = 0, k = 1, \ldots, K \). In the second step, we fix the estimated \( \tilde{\alpha}_n \) and use the EM algorithm to evaluate the log-likelihood function. At the maximization step, we use a Newton-type algorithm (see Wu, 2003, 2010) to update the parameters, \( \tilde{\lambda}_1, \ldots, \tilde{\lambda}_K \). We report the log likelihood, the AIC and BIC values of each estimated copula density in Table 5. For comparison, we also report results on the null distribution. Both the AIC and BIC favor \( c_{D3} \).

Figure 2 reports the contour plots of the Gumbel copula density \( c_G(1.2247) \) and the alternative copula density function \( c_{D3} \). The Gumbel copula is symmetric about the diagonal. In contrast, the alternative copula density captures some salient features of the data, which are not consistent with the Gumbel copula. The alternative copula density is apparently asymmetric about the diagonal. Its upper right tail density is lower than that of the Gumbel copula while its lower left tail density is higher than that of the Gumbel copula.

Table 4: P-values of copula tests on couples’ earnings

<table>
<thead>
<tr>
<th>Test</th>
<th>( Q_{D_1} )</th>
<th>( Q_{D_2} )</th>
<th>( Q_{D_3} )</th>
<th>( Q_{D_4} )</th>
<th>( Q_{O_1} )</th>
<th>( Q_{O_2} )</th>
<th>( Q_{O_3} )</th>
<th>( Q_{O_4} )</th>
<th>( Q_{O_5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-value</td>
<td>2.67e-04</td>
<td>4.712e-06</td>
<td>1.576e-04</td>
<td>0.01</td>
<td>7.667e-05</td>
<td>1.88e-04</td>
<td>4.554e-07</td>
<td>3.381e-07</td>
<td>5.312e-04</td>
</tr>
</tbody>
</table>

Table 5: Estimation results of alternative copula density functions for couples’ earnings

<table>
<thead>
<tr>
<th>Density</th>
<th>( c_G )</th>
<th>( c_{D_1} )</th>
<th>( c_{D_2} )</th>
<th>( c_{D_3} )</th>
<th>( c_{D_4} )</th>
<th>( c_{O_1} )</th>
<th>( c_{O_2} )</th>
<th>( c_{O_3} )</th>
<th>( c_{O_4} )</th>
<th>( c_{O_5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>log ( L )</td>
<td>54.265</td>
<td>54.376</td>
<td>58.083</td>
<td>73.756</td>
<td>65.633</td>
<td>54.598</td>
<td>73.759</td>
<td>73.76</td>
<td>63.207</td>
<td>72.608</td>
</tr>
</tbody>
</table>
(a) original data
(b) rank transformed data

Figure 1: Scatterplot of couples’ wage and salary earnings

Figure 2: Contour plots of the Gumbel copula density function $c_G(1.2247)$ (left panel) and of alternative copula density function $c_{D_3}$ (right panel)
6.2 LOSS and ALAE data

Our second example concerns the dependence structure of the indemnity payment (LOSS) and the allocated loss adjustment expense (ALAE) of insurance companies. The data set consists of 1,500 observations, of which 34 observations of the losses are right censored because the amount of claim cannot exceed the policy limit. The ALAE data are not censored. The data were collected by the US Insurance Services Office and have been analyzed by Frees and Valdez (1998), Genest et al. (1995), Klugman and Parsa (1999), Denuit et al. (2004) and Chen and Fan (2005). Lin and Wu (2012) have also studied this data set, but since their specification test applies to uncensored data only, they restrict their analysis to the subset of 1,466 complete data. We now apply our proposed test to the full data of 1,500 observations. We first estimate the marginal distribution by the Kaplan-Meier estimator. The scatterplots of the original data and transformed data are depicted in Figure 3a and 3b. Both figures show strong right tail dependence but weak left tail dependence between losses and ALAE’s. The transformed data suggest a dependence structure symmetric about the diagonal of unit square.

We conduct the specification test under the Gumbel copula null hypothesis. The two-step estimator of the Gumbel copula yields $\tilde{\alpha}_n = 1.4440$. The P-values of the tests are reported in Table 6. It transpires that the previous studies have done a rather admirable job in selecting the Gumbel-Houggard copula: all smooth tests considered in this analysis fail to reject the Gumbel copula null hypothesis at the 5% confidence level.

![Figure 3: Scatter plots of LOSS and ALAE data](image-url)
Table 6: P-values of copula tests on the losses and ALAE’s

<table>
<thead>
<tr>
<th>Test</th>
<th>$Q_{D_1}$</th>
<th>$Q_{D_2}$</th>
<th>$Q_{D_3}$</th>
<th>$Q_{D_4}$</th>
<th>$Q_{O_1}$</th>
<th>$Q_{O_2}$</th>
<th>$Q_{O_3}$</th>
<th>$Q_{O_4}$</th>
<th>$Q_{O_5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-value</td>
<td>0.441</td>
<td>0.115</td>
<td>0.168</td>
<td>0.659</td>
<td>0.274</td>
<td>0.17</td>
<td>0.283</td>
<td>0.185</td>
<td>0.124</td>
</tr>
</tbody>
</table>

7 Conclusion

In this paper, we extend the moment-based smooth tests of copula specification of multivariate models of Lin and Wu (2012) to accommodate censored data. Our tests can be characterized as score tests on moment conditions of empirical copula distributions under the null hypothesis. We investigate two possible strategies of estimating the marginal distributions – parametric and nonparametric – and compare their finite sample performances. We find that the tests with nonparametric margins are more robust given the reported bias of the parametric tests under misspecification. Our Monte Carlo simulations and empirical examples demonstrate the efficacy and usefulness of our methods.

References


England.

Cherubini, U., Mulinacci, S., Gobbi, F., Romagnoli, S., 2011. Dynamic Copula methods in

, 1–38.

for loss-alae data in non-life insurance. IS Discussion Papers 423.


Efron, B., Tibshirani, R., 1996. Using specially designed exponential families for density

applications to risk management. Handbook of heavy tailed distributions in finance 8,
329–384.

Fermanian, J.D., 2013. An overview of the goodness-of-fit test problem for copulas, in:
Copulae in Mathematical and Quantitative Finance. Springer, pp. 61–89.

actuarial journal 2, 1–25.

Genest, C., Ghoudi, K., Rivest, L., 1995. A semiparametric estimation procedure of depend-
ence parameters in multivariate families of distributions. Biometrika 82(3), 543–552.


Statistics 13, 552–563.


Appendix

To simplify notation, we let \( \sum_t = \sum_{t=1}^{n} \), \( \sum_j = \sum_{j=1}^{2} \).

We first list some conditions needed to establish the large sample properties of the proposed tests.

**Condition 1.**

(i) The sequence of survival variables, \( \{(T_{1t}, T_{2t})\}_{t=1}^{n} \), is an i.i.d. sample from an unknown survival function \( S(t_1, t_2) \) with continuous marginal survival functions \( S_j(\cdot), j = 1, 2 \).

(ii) The sequence of censoring variables \( \{(D_{1t}, D_{2t})\}_{t=1}^{n} \) is an independent sample with joint survival functions \( \{G_t(x_1, x_2)\}_{t=1}^{n} = \{P(D_{1t} > x_1, D_{2t} > x_2)\}_{t=1}^{n} \) and marginal survival functions \( G_{jt}(\cdot) \), \( j = 1, 2 \).

(iii) The censoring variables \( (D_{1t}, D_{2t}) \) are independent of survival variables \( (T_{1t}, T_{2t}) \) and there is no mass concentration at 0 in the sense that \( \limsup_{n \to \infty} n^{-1} \sum_t (1 - G_{jt}(\eta)) \to 0 \) as \( \eta \to 0 \).

**Condition 2.**

Let \( A \) be a compact subset of \( \mathcal{R}^p \). For every \( \epsilon > 0 \),

\[
\liminf_{\alpha \in A : \|\alpha - \alpha_0\| \geq \epsilon} \frac{1}{n} \sum_{t=1}^{n} [E(l(U_{1t}, U_{2t}; \alpha_0)) - E(l(U_{1t}, U_{2t}; \alpha))] > 0
\]

**Condition 3.** The true (unknown) copula function has continuous partial derivatives.

**Condition 4.**

(i) For any \( (u_1, u_2) \in (0, 1)^2 \), \( l(u_1, u_2; \alpha) \) is a continuous function of \( \alpha \in A \).

(ii) Let \( L_t = \sup_{\alpha \in A} \|l(U_{1t}, U_{2t}; \alpha)\| \) and \( L_{t\alpha} = \sup_{\alpha \in A} |l_\alpha(U_{1t}, U_{2t}; \alpha)| \). Then,

\[
\lim_{M \to \infty} \lim_{n \to \infty} \sup \frac{1}{n} \sum_{t=1}^{n} E\{L_t I(L_t \geq M) + L_{t\alpha} I(L_{t\alpha} \geq M)\} = 0.
\]

(iii) For any \( \eta > 0, \epsilon > 0 \), there is \( M > 0 \) such that \( |l(u_1, u_2; \alpha)| \leq M|l(\bar{u}_1, \bar{u}_2; \alpha)| \) for all \( \alpha \in A \) and all \( u_j \in [\eta, 1) \) such that \( 1 - u_j \geq \epsilon (1 - \bar{u}_j), j = 1, 2 \).
Condition 5. (i) $B_n = -n^{-1} \sum_i E\{l_{aa}(U_{1t}, U_{2t}; \alpha_0)\}$ has all its eigenvalues bounded below and above by some finite positive constants;

(ii) $\Sigma_n^* = n^{-1} \sum_i \text{var}\{l_a(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}^*\}$ has all its eigenvalues bounded below and above by some finite positive constants, where $W_{jt}^*$ is defined in (19).

(iii) $\{l_a(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}\}_{n=1}^n$ satisfies Lindeberg condition.

Condition 6. Functions $l_{aa}(u_1, u_2; \alpha)$ and $l_{a\beta_j}(u_1, u_2; \alpha)$ for $j = 1, 2$ are well-defined and continuous in $(u_1, u_2; \alpha) \in (0, 1)^2 \times A$. $\ell_{\beta_j\beta_j}(u_1, u_2; \beta_j)$ is well-defined and continuous in $(u_1, u_2; \beta_j) \in (0, 1)^2 \times \mathcal{B}$.

Condition 7. (i) $|l_a(u_1, u_2; \alpha_0)| \leq q(u_1(1-u_1))^{-a_1}u_2(1-u_2)^{-a_2}$ for some $q > 0$ and $a_j \geq 0$ such that $\limsup n^{-1} \sum_t E[(U_{1t}(1-U_{1t}))^{-2a_1}(U_{2t}(1-U_{2t}))^{-2a_2}] < \infty$;

(ii) $|l_{aj}(u_1, u_2; \alpha_0)| \leq \text{constant}\{u_j(1-u_j)\}^{-b_j}\{u_k(1-u_k)\}^{-a_k}$ for some $b_j, a_k$ and $j \neq k$ such that $\limsup n^{-1} \sum_t E[(U_{jt}(1-U_{jt}))^{\xi_j}u_j(1-u_j)^{\xi_j}] < \infty$ for some $\xi_j \in (0, 1/2)$.

Condition 8. (i) Let $L_{taj} = \sup_{\alpha \in A} |l_{aj}(U_{1t}, U_{2t}; \alpha)|$ and $L_{taa} = \sup_{\alpha \in A} |l_{aa}(U_{1t}, U_{2t}; \alpha)|$. Then, $\lim_{K \to \infty} \limsup_{n \to \infty} n^{-1} \sum_t E\{L_{taj}I(L_{taj} \geq K) + L_{taa}I(L_{taa} \geq K)\} = 0$.

(ii) For any $\eta > 0$ and any $\epsilon > 0$, there is $K > 0$, such that $|l_a(u_1, u_2; \alpha)| + |l_{aa}(u_1, u_2; \alpha)| \leq K\{|l_a(\bar{u}_1, \bar{u}_2; \alpha)| + |l_{aa}(\bar{u}_1, \bar{u}_2; \alpha)|\}$ for all $\alpha \in A$ and all $u_j \in [\eta, 1)$ such that $1 - u_j \geq \epsilon(1 - \bar{u}_j)$, $j = 1, 2$.

In contrast to the censoring mechanism in Shih and Louis (1995), Condition 1(ii) allows the censoring variables $\{(D_{1t}, D_{2t})\}_{t=1}^\infty$ to be non-identically distributed. In addition, no assumptions are made on the joint survival function $G_t(x_1, x_2)$ of the censoring variables $(D_{1t}, D_{2t})$. Therefore, general censoring mechanism is allowed in this framework, for example, both variables are censored, one censored and the other uncensored, or one random censoring and the other fixed censoring. Condition 7 and 8 allows the score function and its partial derivatives with respect to $u_j$, $j = 1, 2$ to blow up at the boundaries. The condition characterizes many popular copula functions such as the Gaussian, $t$ and Clayton copulas.

Proof of Theorem 1. Under Conditions 1-4, by using the uniform consistency of the maximum likelihood estimators $\hat{\beta}_j$, $j = 1, 2$ and following the logic of the proof of Proposition 3.1 in Chen et al. (2010), we can obtain the uniform consistency of $\hat{\alpha}_n$, $j = 1, 2$. 

31
Recall that $L(\hat{U}_{1t}, \hat{U}_{2t}; \alpha)$ defined in (14) is the average log-likelihood function of $\alpha$. Expanding the score function $L_\alpha(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n) = \partial L(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)/\partial \alpha$ in a Taylor series around $\alpha_0$ and rearranging the terms, we get

$$\hat{\alpha}_n - \alpha_0 = (\hat{B}_n)^{-1} \frac{1}{n} \sum_t \{l_\alpha(U_{1t}, U_{2t}; \alpha_0) + \sum_j l_{\alpha\beta}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)(\hat{\beta}_j - \beta_{j0})\},$$

where $\hat{B}_n = -L_{\alpha\alpha}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n) = -\partial^2 L(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n)/\partial \alpha \partial \alpha'$ and $\hat{U}_{jt} = S_j(X_{jt}, \hat{\beta}_j)$, in which $\hat{\beta}_j$ is on the line segment between $\beta_{j0}$ and $\hat{\beta}_j$ and $\hat{\alpha}_n$ is on the line segment between $\alpha_0$ and $\hat{\alpha}_n$. Under Condition 6 and 8, following the same logic of the proof of Proposition 3.2 in Chen et al. (2010), we can obtain $\hat{B}_n \to B_n$ and $-n^{-1}\sum l_{\alpha\beta}(\hat{U}_{1t}, \hat{U}_{2t}; \hat{\alpha}_n) \to B_{n, \beta_j}$, where $B_n$ is defined in Theorem 1 and $B_{n, \beta_j}$ is defined in (17).

Therefore,

$$\hat{\alpha}_n - \alpha_0 = B_n^{-1}[L_\alpha(U_{1t}, U_{2t}; \alpha_0) - B_{n, \beta_1}(\hat{\beta}_1 - \beta_{10}) - B_{n, \beta_2}(\hat{\beta}_2 - \beta_{20})] + o_p(n^{1/2}). \quad (A.1)$$

Now we consider $\hat{\beta}_j - \beta_{j0}$. Recall that $\mathcal{L}(\hat{\beta}_j)$ defined in (12) is the average log-likelihood function of $\beta_j$. Expanding the score function $\mathcal{L}_{\beta_j}(\hat{\beta}_j) = \partial \mathcal{L}(\hat{\beta}_j)/\partial \beta_j$ in a Taylor series around $\beta_{j0}$ and rearranging the terms, we get

$$\hat{\beta}_j - \beta_{j0} = (\hat{D}_{n,jj})^{-1} \frac{1}{n} \sum t \ell_{\beta_j}(X_{jt}, \delta_{jt}; \beta_{j0})$$

where $\hat{D}_{n,jj} = -\mathcal{L}_{\beta_j\beta_j}(\hat{\beta}_j) = -\partial^2 \mathcal{L}(\hat{\beta}_j)/\partial \beta_j \partial \beta_j'$, in which $\hat{\beta}_j$ is on the line segment between $\beta_{j0}$ and $\hat{\beta}_j$. Under Condition 6, by the consistency of $\hat{\beta}_j$, this implies that $\hat{D}_{n,jj} \to D_{n,jj}$ in probability as $n \to \infty$, where $D_{n,jj}$ is defined in (18).

Therefore,

$$\hat{\beta}_j - \beta_{j0} = (D_{n,jj})^{-1} \frac{1}{n} \sum t \ell_{\beta_j}(X_{jt}, \delta_{jt}; \beta_{j0}) + o_p(n^{1/2}). \quad (A.2)$$

Substituting (A.2) into (A.1), we have,

$$\hat{\alpha}_n - \alpha_0 = B_n^{-1} \frac{1}{n} \sum t \{l_\alpha(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}^*\} + o_p(n^{1/2}). \quad (A.3)$$

where for $j = 1, 2$, $W_{jt}^*$ is defined in (19).

Therefore, $\text{var}(\sqrt{n}(\hat{\alpha}_n - \alpha_0)) = B_n^{-1} \Sigma_n^* B_n^{-1}$, where $\Sigma_n^* = n^{-1} \sum_t \text{var} \{l_\alpha(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}^*\}$. 

32
Proof of Theorem 2. For any \((u_1, u_2) \in (0,1)^2\), \(g(u_1, u_2)\) has continuous partial derivatives with respect to \(u_1\) and \(u_2\); \(g(u_1, u_2; \alpha)\) is a continuous function of \(\alpha\); \(E[\sup_{\alpha \in A} |g(U_1, U_2; \alpha)|] < \infty\).

**Condition 9.** \(\Omega_n^* = n^{-1} \sum \var{g(U_1t, U_2t; \alpha_0) + \sum_j Z_j^* + \mathcal{G}_n^{-1}B_n^{-1}[\alpha(U_1t, U_2t; \alpha_0) + \sum_j W_j^*]}\) is finite and positive definite, where \(W_j^*\) and \(Z_j^*\) are defined in (19) and (21), respectively, \(B_n\) is defined in Theorem 1 and \(\mathcal{G}_n\) is defined in (20).

**Condition 10.** \(\Omega_n^* = n^{-1} \sum \var{g(U_1t, U_2t; \alpha_0) + \sum_j Z_j^* + \mathcal{G}_n^{-1}B_n^{-1}[\alpha(U_1t, U_2t; \alpha_0) + \sum_j W_j^*]}\)

**Condition 11.** For any \((u_1, u_2) \in (0,1)^2\), \(\mathcal{G}_\beta_j(u_{1t}, u_{2t}; \alpha)\) is a continuous function of \(\beta_j\) in a neighborhood of \(\beta_{j0}\); \(\sup_{\beta_j \in B: \|\beta_j - \beta_{j0}\| = o(1)} \|\mathcal{G}_\beta_j(u_{1t}, u_{2t}; \alpha)\| < \infty\).

**Condition 12.** For any \((u_1, u_2) \in (0,1)^2\), \(\mathcal{G}_\alpha(u_{1t}, u_{2t}; \alpha)\) is continuous function of \(\alpha\) in a neighborhood of \(\alpha_0\); \(\sup_{\alpha \in A: \|\alpha - \alpha_0\| = o(1)} \|\mathcal{G}_\alpha(u_{1t}, u_{2t}; \alpha)\| < \infty\).

**Proof of Theorem 2.** For \(j = 1, 2\), define \(g_j(u_1, u_2) = \partial g(u_1, u_2)/\partial u_j\), \(C_j(u_1, u_2; \alpha) = \partial C_0(u_1, u_2; \alpha)/\partial u_j\), \(C_{jj}(u_1, u_2; \alpha) = \partial C_0(u_1, u_2; \alpha)/\partial u_j^2\), \(C_{j\alpha}(u_1, u_2; \alpha) = \partial C_0(u_1, u_2; \alpha)/\partial u_j \partial \alpha\), \(c_0(u_1, u_2; \alpha)\), \(c_j(u_1, u_2; \alpha)\), \(c_{jj}(u_1, u_2; \alpha)\) and \(c_{j\alpha}(u_1, u_2; \alpha)\) are defined similarly. Also define \(C_{12}(u_1, u_2; \alpha) = \partial C_0(u_1, u_2; \alpha)/\partial u_1 \partial u_2\).

Expanding \(\hat{g}(\hat{U}_{1t}, \hat{U}_{2t}, \hat{\alpha}_n)\) defined in (16) in a Taylor series around \(\alpha_0\) yields,

\[
\hat{g}(\hat{U}_{1t}, \hat{U}_{2t}, \hat{\alpha}_n) = \frac{1}{n} \sum_t \left\{ g(U_{1t}, U_{2t}; \alpha_0) + g'(U_{1t}, U_{2t}; \alpha_0)\hat{\alpha}_n - \alpha_0 \right\} \\
+ \sum_j g'_{\beta_j}(U_{1t}, U_{2t}; \hat{\alpha}_n)\hat{\beta}_j - \beta_{j0} \right\} \tag{A.4}
\]

Let \(\mathcal{G}_\alpha\) and \(\mathcal{G}_\beta_j\) be defined as in (20) and (21). Under Condition 11 and 12, \(n^{-1} \sum_{t} \mathcal{G}_\alpha(U_{1t}, U_{2t}; \hat{\alpha}_n) \to \)
\( g_{\alpha}, n^{-1} \sum_t g_{ij}(U_{1t}, U_{2t}; \alpha_n) \rightarrow G_{ij}, j = 1, 2, \) and

\[
\begin{align*}
\mathfrak{g}_{\alpha}(u_{1t}, u_{2t}; \alpha) &= \delta_{1t}(1 - \delta_{2t}) \int_0^{u_{2t}} \frac{g(u_{1t}, u_{2}) c_\alpha(u_{1t}, u_{2}; \alpha) du_2}{C_1(u_{1t}, u_{2t}; \alpha)} \\
&\quad - \delta_{1t}(1 - \delta_{2t}) \int_0^{u_{2t}} (u_{1t}, u_{2}) c_0(u_{1t}, u_{2}; \alpha) du_2 C_{1\alpha}(u_{1t}, u_{2t}; \alpha) \\
&\quad + \delta_{2t}(1 - \delta_{1t}) \int_0^{u_{1t}} g(u_{1}, u_{2t}) c_\alpha(u_{1}, u_{2t}; \alpha) du_1 \\
&\quad - \delta_{2t}(1 - \delta_{1t}) \int_0^{u_{1t}} g(u_{1}, u_{2t}) c_0(u_{1}, u_{2t}; \alpha) du_1 C_{2\alpha}(u_{1t}, u_{2t}; \alpha) \\
&\quad + (1 - \delta_{1t})(1 - \delta_{2t}) \int_0^{u_{2t}} \int_0^{u_{1t}} g(u_{1}, u_{2}) c_\alpha(u_{1}, u_{2}; \alpha) du_1 du_2 \\
&\quad - (1 - \delta_{1t})(1 - \delta_{2t}) \int_0^{u_{2t}} \int_0^{u_{1t}} g(u_{1}, u_{2}) c_0(u_{1}, u_{2}; \alpha) du_1 du_2 C_{2}(u_{1t}, u_{2t}; \alpha) \\
&\quad - \int g(u_{1}, u_{2}) c_\alpha(u_{1}, u_{2}; \alpha) du_1 du_2,
\end{align*}
\]

\[
\begin{align*}
\mathfrak{g}_{1}(u_{1t}, u_{2t}, \alpha) &= \delta_{1t} \delta_{2t} g_{1}(u_{1t}, u_{2t}) + \delta_{1t}(1 - \delta_{2t}) \int_0^{u_{2t}} (g_{1}(u_{1t}, u_{2}) c_0(u_{1t}, u_{2}; \alpha) + g(u_{1t}, u_{2}) c_1(u_{1t}, u_{2}; \alpha)) du_2 \\
&\quad - \delta_{1t}(1 - \delta_{2t}) \int_0^{u_{2t}} g(u_{1t}, u_{2}) c_0(u_{1t}, u_{2}; \alpha) du_2 C_{11}(u_{1t}, u_{2t}; \alpha) \\
&\quad + \delta_{2t}(1 - \delta_{1t}) \int_0^{u_{1t}} g(u_{1t}, u_{2}) c_\alpha(u_{1t}, u_{2}; \alpha) du_1 \\
&\quad - \delta_{2t}(1 - \delta_{1t}) \int_0^{u_{1t}} g(u_{1t}, u_{2}) c_0(u_{1t}, u_{2}; \alpha) du_1 C_{1\alpha}(u_{1t}, u_{2t}; \alpha) \\
&\quad + (1 - \delta_{1t})(1 - \delta_{2t}) \int_0^{u_{2t}} g(u_{1t}, u_{2}) c_0(u_{1t}, u_{2}; \alpha) du_2 \\
&\quad - (1 - \delta_{1t})(1 - \delta_{2t}) \int_0^{u_{2t}} \int_0^{u_{1t}} g(u_{1}, u_{2}) c_\alpha(u_{1}, u_{2}; \alpha) du_1 du_2 C_{11}(u_{1t}, u_{2t}; \alpha) \\
&\quad \quad (A.5)
\end{align*}
\]
Proof of Theorem 3.

Consider a generic function $H(u_1, u_2; \alpha) : (0, 1)^2 \to \mathbb{R}$. We have by the
Proof of Theorem 4.

Under Conditions 1-4 and Condition 13, the consistency of \( \tilde{H} \) of Chen et al. (2010).

Conditions 6-8 and Conditions 13-14, the asymptotic normality of \( \tilde{\Theta} \) obtained according to Proposition 3.1 of Chen et al. (2010). Next under Conditions 1-4, the convergence Theorem. Similarly, \( \hat{\Theta} \) follows readily from the asymptotic normality of \( \hat{\Theta} \).

If \( \hat{\Theta} \) is finite and positive definite, where \( \hat{\Theta} \) is given by (23). Using essentially the same argument, one can show that \( \hat{\Omega} \) estimates \( \Omega^\alpha \) consistently under Condition 10. The result of this theorem then follows readily from the asymptotic normality of \( \hat{\Theta} \) given in Theorem 2.

\( \boxdot \)

**Condition 13.** If \( \{T_{jt}\}_{t=1}^n \) are subject to non-trivial censoring (i.e., \( D_{jt} \neq \infty \)), then \( \tilde{S}_j \) is truncated at the tail in the sense that for some \( \tau_j \), \( \tilde{S}_j(x_j) = \tilde{S}_j(\tau_j) \) for all \( x_j \geq \tau_j \) and \( \lim \inf n^{-1} \sum_{t=1}^n G_{jt}(\tau_j) S_{jt}(\tau_j) > 0 \).

**Condition 14.** (i) \( B_n = -n^{-1} \sum_{t=1}^n E\{l_{\alpha}(U_{1t}, U_{2t}; \alpha_0)\} \) has all its eigenvalues bounded below and above by some finite positive constants;

(ii) \( \Sigma_n = n^{-1} \sum_{t=1}^n \text{var}\{l_{\alpha}(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}\} \) has all its eigenvalues bounded below and above by some finite positive constants, where \( W_{jt} \) is defined in (26);

(iii) \( \{l_{\alpha}(U_{1t}, U_{2t}; \alpha_0)\} + \sum_j W_{jt}\}_{t=1}^n \) satisfies Lindeberg condition.

**Proof of Theorem 4.** Under Conditions 1-4 and Condition 13, the consistence of \( \hat{\alpha}_n \) is readily obtained according to Proposition 3.1 of Chen et al. (2010). Next under Conditions 1-4, Conditions 6-8 and Conditions 13-14, the asymptotic normality of \( \hat{\alpha}_n \) is given by Proposition 3.2 of Chen et al. (2010).

\( \boxdot \)

**Condition 15.** \( \Omega_n = n^{-1} \sum_i \text{var}\{g(U_{1t}, U_{2t}; \alpha_0) + \sum_j Z_{jt} + \mathcal{G}_\alpha B^{-1}_n[l_{\alpha}(U_{1t}, U_{2t}; \alpha_0) + \sum_j W_{jt}]\} \) is finite and positive definite, where \( W_{jt} \) and \( Z_{jt} \) are defined in (26) and (28), respectively.
Proof of Theorem 5. In the proofs that follow we make use of Lemma 1 of Chen et al. (2010). We present the results below for the reader’s convenience.

Lemma 1 (Chen et al., 2010) Suppose that Conditions 1 to 5 are satisfied. Then: (i) the marginal Kaplan-Meier estimators are uniformly strongly consistent: \( \sup_{x \leq \tau_j} |\tilde{S}_j(x) - S_j(x)| \to 0 \) a.s. for \( j = 1, 2 \); (ii) they can be expressed as martingale integrals:

\[
\tilde{S}_j(x) - S_j(x) = -S_j(x) \int_{-\infty}^{x} \frac{\tilde{S}_j(u-)}{S_j(u)} \sum_t dM_{jt}(u) \sum_t I(X_{jt} \geq u)
\]

where \( M_{jt}(x) = N_{jt}(x) - \int_{-\infty}^{x} J_{jt}(u) dH_j(u) \), in which \( N_{jt}(x) = \delta_{jt} I(X_{jt} \leq x) \), \( J_{jt}(x) = I(X_{jt} \geq x) \), and \( H_j(x) = -\log S_j(x) \).

By Theorem 4, we have \( \|\tilde{\alpha}_n - \alpha_0\| = o_p(1) \). Applying Taylor’s expansion to \( \tilde{g}(\tilde{\alpha}_n) \) with respect to \( \tilde{\alpha}_n \) at \( \alpha_0 \) yields

\[
\tilde{g}(\tilde{\alpha}_n) = \frac{1}{n} \sum_t g(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha_0) + \frac{1}{n} \sum_t g'_\alpha(\tilde{U}_{1t}, \tilde{U}_{2t}; \tilde{\alpha}_n)(\tilde{\alpha}_n - \alpha_0) \tag{A.10}
\]

where \( \tilde{\alpha}_n \) is between \( \alpha_0 \) and \( \tilde{\alpha}_n \).

Under Condition 12, we have

\[
\sup_{\alpha \in A: \|\alpha - \alpha_0\| = o(1)} \|\frac{1}{n} \sum_t g_\alpha(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha) - \mathcal{G}_\alpha\| = o(1)
\]

It follows that (A.10) can be written as,

\[
\tilde{g}(\tilde{\alpha}_n) = \frac{1}{n} \sum_t g(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha_0) + \mathcal{G}'_\alpha(\tilde{\alpha}_n - \alpha_0) + o_p(1) \tag{A.11}
\]

Next Theorem 4 indicates that \( \tilde{\alpha}_n \) can be expressed as an asymptotically linear estimator such that

\[
\tilde{\alpha}_n - \alpha_0 = B_n^{-1} \frac{1}{n} \sum_{t=1}^{n} l_\alpha(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha_0) + o_p(1) \tag{A.12}
\]

Plugging (A.12) into (A.11) yields,

\[
\tilde{g}(\tilde{\alpha}_n) = \frac{1}{n} \sum_t \{g(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha_0) + \mathcal{G}'_\alpha B_n^{-1} l_\alpha(\tilde{U}_{1t}, \tilde{U}_{2t}; \alpha_0)\} + o_p(1) \tag{A.13}
\]
By the mean-value theorem, we expand (A.13) to obtain

\[
\tilde{g}(\tilde{\alpha}_n) = \frac{1}{n} \sum_t \{ g(U_{1t}, U_{2t}; \alpha_0) + G'_\alpha B_n^{-1} l_\alpha(U_{1t}, U_{2t}; \alpha_0) \} + J_n + o_p(1),
\]

where

\[
J_n = \frac{1}{n} \sum_t \sum_j [ g_j(\bar{U}_{1t}, \bar{U}_{2t}; \alpha_0) + G'_\alpha B_n^{-1} l_{\alpha j}(\bar{U}_{1t}, \bar{U}_{2t}; \alpha_0) ] [(\tilde{U}_{jt} - U_{jt})]
\]

in which \((\bar{U}_{1t}, \bar{U}_{2t})\) lie on the line segment between \((U_{1t}, U_{2t})\) and \((\tilde{U}_{1t}, \tilde{U}_{2t})\).

By Lemma 1, we have

\[
J_n = -\frac{1}{n} \sum_t \sum_j \{ g_j(\bar{U}_{1t}, \bar{U}_{2t}; \alpha_0) + G'_\alpha B_n^{-1} l_{\alpha j}(\bar{U}_{1t}, \bar{U}_{2t}; \alpha_0) \} S_j(X_{jt}) \int_{-\infty}^{X_{jt}} \frac{\tilde{S}_j(u-)}{S_j(u-)} \sum_s dM_{js}(u) \sum_s I(X_{js} \geq u).
\]

Then following the same logic of the proof in Proposition 3.2 of Chen et al. (2010), we can see that \(J_n\) is asymptotically a sum of independent zero-mean random vectors. Given Condition 14, Theorem 5 now follows from the standard multivariate central limit theorem for independent but non-identically distributed random variables.

\[\square\]

Proof of Theorem 6. This theorem can be derived by following the proofs of Theorem 3 in this paper and Proposition 3.1 in Chen et al. (2010). For brevity, the proof is not reproduced here.

\[\square\]