COPULA-BASED TIME SERIES WITH FILTERED NONSTATIONARITY

By

Xiaohong Chen, Zhijie Xiao, and Bo Wang

July 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2242

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Copula-Based Time Series With Filtered Nonstationarity*

Xiaohong Chen† Yale University

Zhijie Xiao‡ Boston College

Bo Wang§ Boston College

First version: March 2014; Revised: July 2020.

Abstract

Economic and financial time series data can exhibit nonstationary and nonlinear patterns simultaneously. This paper studies copula-based time series models that capture both patterns. We propose a procedure where nonstationarity is removed via a filtration, and then the nonlinear temporal dependence in the filtered data is captured via a flexible Markov copula. We study the asymptotic properties of two estimators of the parametric copula dependence parameters: the parametric (two-step) copula estimator where the marginal distribution of the filtered series is estimated parametrically; and the semiparametric (two-step) copula estimator where the marginal distribution is estimated via a rescaled empirical distribution of the filtered series. We show that the limiting distribution of the parametric copula estimator depends on the nonstationary filtration and the parametric marginal distribution estimation, and may be non-normal. Surprisingly, the limiting distribution of the semiparametric copula estimator using the filtered data is shown to be the same as that without nonstationary filtration, which is normal and free of marginal distribution specification. The simple and robust properties of the semiparametric copula estimators extend to models with misspecified copulas, and facilitate statistical inferences, such as hypothesis testing and model selection tests, on semiparametric copula-based dynamic models in the presence of nonstationarity. Monte Carlo studies and real data applications are presented.

JEL code: C14, C22.

Keywords: Residual copula, Cointegration, Unit Root, Nonstationarity, Nonlinearity, Tail Dependence, Semiparametric.

---

*Ron Gallant has positively influenced us for many years through his creative research and humorous yet insightful comments. We thank George Tauchen (Guest Editor), an Associate Editor, two anonymous referees, Yanqing Fan, Oliver Linton, Peter Robinson, James Stock and participants at the 2014 JSM and the 2018 FERM conference for helpful comments on early versions.

†Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520, USA. Tel: 203-432-5852. Email: xiaohong.chen@yale.edu

‡Department of Economics, Boston College, Chestnut Hill, MA 02467, USA. Tel: 617-552-1709. Email: xiaoz@bc.edu.

§Department of Economics, Boston College, Chestnut Hill, MA 02467, USA.

1. Introduction

Nonstationarity and nonlinearity are important empirical features in economic and financial time series. For many economic time series, nonstationary behavior is often the most dominant characteristic. Some series grow in a secular way over long periods of time, others appear to wander around as if they have no fixed population mean. Growth characteristics are especially evident in time series that represent aggregate economic behavior. Random wandering behavior is also evident in many financial time series. In addition, existing literature (e.g. Gallant, Rossi, Tauchen (1993), Granger (2002), Gallant (2009)) points out that the classical linear time series modelling based on the Gaussian distribution assumption clearly fails to explain the stylized facts observed in economic and financial data, and that it is highly undesirable to perform various economic policy evaluations, financial forecasts, and risk managements based on linear Gaussian models.

Econometric analysis that ignores either nonstationarity or nonlinearity may lead to erroneous inference for policy evaluations and financial applications. Arguably the most common nonstationarity in many economic time series are persistency and trending characteristics. Deterministic or stochastic trend components are usually used to capture these kinds of nonstationarity in time series. In the presence of a deterministic trend, detrending methods are commonly used to extract this trend and the residuals are then analyzed as a stationary time series. Unit root and cointegration models are widely used to model stochastic trends in economic time series. For stationary series, copula-based Markov models provide a rich source of potential nonlinear dynamics describing temporal dependence and tail dependence, without imposing any restrictions on marginal distributions. See, e.g., Joe (1997), Chen and Fan (2006a), Patton (2006, 2009, 2012), Ibragimov (2009), Cherubini, et al (2012) and the references therein. However, existing large sample theories for estimation and inference on the copula-based time series models rule out nonstationarity.

An important issue in practice is that nonstationarity and nonlinearity may occur simultaneously. In this paper, we study copula-based time series models that can capture nonstationarity and nonlinearity (and tail dependence). We propose a sequential procedure where nonstationarity is first removed via a filtration, and then the nonlinear temporal dependence (and the tail dependence) in the filtered data is captured by a copula-based first-order stationary Markov model. We are interested in simple estimation and inference on the copula dependence parameter for the deterministic or stochastic detrended data. We focus on the sequential approach due to its easy implementation in empirical applications.

An advantage of copula-based modeling approach is to leave the marginal distribution completely free of parametric assumptions. Nevertheless, many empirical researchers still like to assume marginal distribution belonging to a parametric family and estimate it parametrically before proceeding to estimate the copula dependence parameters. For the sake of comparison, we consider both the parametric (two-step) copula estimation where the marginal distribution of the filtered series belongs to a parametric family, and the semiparametric (two-step) copula estimation where the marginal distribution of the filtered series is nonparametric. Without nonstationary filtering and for observable stationary Markov data, both copula estimators are shown to be asymptotically normal, while the semiparametric copula estimator is obviously robust to misspecification of the marginal distribution. We show that the copula estimators using nonstationary filtered data have very different properties, however. In particular, the limiting distribution of the parametric (two-step) copula estimator is affected by the nonstationary filtration and the parametric marginal distribution estimation, and may be non-normal in the presence of stochastic trends (unit root or cointegration). While the parametric copula estimator using deterministic trend filtered data is shown to be asymptotically normal, its asymptotic variance still depends on the filtrating and the parametric marginal specification in a complicated way. Surprisingly, we show that the limiting distribution of the semiparametric (two-step) copula estimator using the filtered data is the same as that without nonstationary filtration, which is normal and free of marginal distribution specification. While this surprising result is first derived for models with correctly specified parametric copulas in Section 3, we show in Section 4 that the limiting distribution of the semiparametric copula estimator (for the pseudo-true parameters) is still not affected by the nonstationary filtration even in misspecified parametric copula models. The simple and robust properties of the semiparametric copula estimators greatly facilitate statistical inferences, such as hypothesis testing and model selection tests, on semiparametric copula-based dynamic models in the presence of nonstationarity.

Previously, Chen and Fan (2006b) uses parametric copula to generate contemporaneous dependence among multivariate standardized innovations of observed weakly-dependent multivariate time series, where the standardized innovations have no serial dependence. They also obtained a surprising result that the limiting distribution of their semiparametric two-step copula estimator does not depend on the stationary filtering in the first step. It is interesting that both papers establish the "no-filtering-effect" in semiparametric two-step copula parameter estimation. While Chen and Fan (2006b) consider the contemporaneous copula dependence among multivariate standardized innovations that are orthogonal to the dynamic filtering part, our paper studies the temporal copula dependence of univariate nonstationary filtered residuals, and there is dependence among the nonstationary (stochastic trending) and the stationary parts in our setting.

Monte Carlo studies reveal interesting finite sample behaviors of the parametric and the semiparametric copula estimators under various combinations of nonstationary filtration, correctly- and incorrectly- specified marginal distribution of the filtered series, and copula function specification (with or without tail dependence). Simulation evidences (in terms of biases and variances) indicate that the finite sample performance of parametric copula estimator is indeed very sensitive to different types of filtration and the parametric estimation of marginal distributions. The semiparametric copula estimator not only is robust to specification of marginal distributions, but also performs very similarly to the infeasible semiparametric estimator without nonstationary filtering. In comparison to the parametric copula estimator with correctly specified parametric marginal distributions, the semiparametric estimator has reasonably good sampling performance over a wide range of copula parameter values. Simulation patterns are consistent with the theoretical findings in our paper.

To illustrate the practical usefulness of our proposed models and method. We first apply our method to estimate the short term dynamics in the GNP time series after the cointegrating regression of GNP on consumption series. Our semiparametric copula estimation and testing using the filtered data enable us to detect both lower and upper tail dependence in the GNP series (of the USA). We next apply our method to the famous "CAY" time series that was first constructed in Lettau and Ludvigson (2001), which is the residual term from a cointegrating regression of consumption (c_t) on asset holding (a_t) and labor income (y_t). According to Lettau and Ludvigson (2001) and many subsequent work, the "CAY" time series contain important information of future returns at short horizons. Our semiparametric copula estimation and testing detects very significant lower tail dependence and relatively weak upper tail dependence in the "CAY" series.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents estimation of copula parameters for both the parametric and semiparametric models of the filtered data. It also obtains the large sample properties of the parametric and semiparametric copula estimators. Section 4 considers estimation under possibly misspecified copula models. It also discusses semiparametric copula model selection tests using nonstationary filtered data. Section 5 presents Monte Carlo studies and Section 6 provides empirical applications. Section 7 briefly concludes with future research. In the supplementary appendices, Appendix A displays tables summarizing the Monte Carlo results, and Appendix B contains the technical proofs. Notation: $BM(\omega^2)$ denotes a Brownian motion with variance ω^2. For a generic parameter, say, β, we denote the true parameter value by β^*, the pseudo-true value by $\bar{\beta}$ and the feasible estimator by $\widehat{\beta}$.

2. The Model

We assume that the observed scalar time series $\{Z_t\}_{t=1}^n$ can be modelled as

$$Z_t = X_t'\pi^* + Y_t, \tag{2.1}$$

where $X_t'\pi^*$ is the nonstationary component in which X_t is an observed d_x-dimensional vector of nonstationary regressors. For example, X_t may contain deterministic trends, unit root or near unit root nonstationary time series. Y_t is the latent stationary ergodic component that could exhibit nonlinear temporal dependence and/or tail dependence.

Estimation of the parameter π^* in model (2.1) is by now standard (usually an OLS regression of Z_t on X_t) and is not the focus of our paper. Instead we are interested in estimation of the copula parameter β that captures stationary nonlinear temporal dependence in $\{Y_t\}_{t=1}^n$. Unfortunately $\{Y_t\}_{t=1}^n$ is unobserved. We shall estimate the latent temporal dependence parameter β and study its asymptotic properties based on the filtered time series $\{\widehat{Y}_t\}_{t=1}^n$, where

$$\widehat{Y}_t = Z_t - X_t'\widehat{\pi},$$

and $\widehat{\pi}$ denotes some nonstationary filtering estimator for π^*. We state the basic regularity conditions on the nonstationary part and the stationary part as follows. The assumptions about the nonstationary part $\{X_t'\pi^*\}_{t=1}^n$ are the typical ones for trend, unit roots and cointegration, and the assumptions about the stationary part $\{Y_t\}_{t=1}^n$ are the same as those in Chen and Fan (2006a).

Due to the nonstationarity in X_t, we introduce appropriate re-standardization via a scaling matrix D_n to facilitate asymptotic analysis. Denoting $X_n(r) = n^{1/2}D_n^{-1}X_{[nr]}$ and $Y_n(r) = n^{-1/2}\sum_{t=1}^{[nr]} Y_t$ for $r \in [0,1]$, we make the following assumption concerning the nonstationary component and the related filtration.

Assumption X. In model (2.1), the elements in X_t can be either a deterministic trend function, or an unit root or local to unit root process such that

$$\begin{bmatrix} Y_n(r) \\ X_n(r) \end{bmatrix} \Rightarrow \begin{bmatrix} B_Y(r) \\ X(r) \end{bmatrix}, \, r \in [0,1] \quad \text{as} \quad n \to \infty,$$

where $B_Y(r)$ is a Brownian motion, $X(r)$ is a vector of stochastic or deterministic functions. And

$$D_n\left(\widehat{\pi} - \pi^*\right) \Rightarrow \xi \quad \text{as} \quad n \to \infty.$$

The limit of the standardized nonstationary component $n^{1/2}D_n^{-1}X_{[nr]}$, may be stochastic processes such as Brownian motions, or deterministic functions, or a mixture of both type. $B_Y(r)$ is a Brownian motion. In the case when $X(r)$ contains stochastic functions, $B_Y(r)$ and $X(r)$ may be correlated. The limiting distribution of the filtration parameter, ξ, is a function of $X(\cdot)$ and may not be a normal variate. We give below a few examples that are widely used in time series applications. In all these examples, we use the OLS filtration.

Example 1. Trending Time Series. X_t is a vector of deterministic trend function and $n^{1/2}D_n^{-1}X_{[nr]} \to X(r)$, where $X(r)$ is a piecewise continuous limiting trending function. Let $\widehat{\pi}$ be the OLS estimator of π^*,

$$D_n\left(\widehat{\pi} - \pi^*\right) \Rightarrow \xi_1,$$

where in general ξ_1 is a normal variate. In particular, let $B_Y(r) = BM(\omega_Y^2)$ denote the weak limit of $Y_n(r) = n^{-1/2}\sum_{t=1}^{[nr]} Y_t$, then

$$\xi_1 = \left[\int X(r)X(r)'dr\right]^{-1}\left[\int X(r)dB_Y(r)\right],$$

which is a mean zero normal random variable with variance-covariance matrix $\omega_Y^2 \left[\int X(r)X(r)'dr\right]^{-1}$. For example, if the observed time series $\{Z_t\}_{t=1}^n$ contains a linear trend:

$$Z_t = \pi_0^* + \pi_1^* t + Y_t,$$

then $X_t = (1,t)'$ and $X(r) = (1,r)'$, and the standardization matrix is $D_n = diag(n^{1/2}, n^{3/2})$.

Example 2. Time Series with a Root Close to Unity. $X_t = Z_{t-1}$ and $\pi = 1 + c/n$. Thus $X_t = Z_{t-1}$ can be a unit root ($c = 0$) or local to unit root process ($c < 0$). $D_n = n$, and $n^{-1/2}X_{[nr]} \Rightarrow X(r) = J_c(r) = \int_0^r e^{(r-s)c}dB_Y(s)$, where $J_c(r)$ is a Ornstein–Uhlenbeck process. If $c = 0$, $J_0(r) = B_Y(r)$ is simply a Brownian motion. The OLS filtration estimators $\widehat{\pi}$ converges at rate-n to a non-normal limit: $n(\widehat{\pi} - \pi) \Rightarrow \xi_2$, where

$$\xi_2 = \left[\int_0^1 J_c(r)^2 dr\right]^{-1} \left[\int_0^1 J_c(r)dB_Y(r) + \lambda\right],$$

with $\lambda = \sum_{h=1}^{\infty} E(Y_1 Y_{1+h})$.

Example 3 Cointegrated Time Series. $X_t = (X_{1t}', X_{2t}')'$, where X_{1t} is a vector of deterministic trend, and X_{2t} is a vector of stochastic nonstationary process, then

$$n^{1/2}D_{1n}^{-1}X_{1,[nr]} \to X_1(r),\ n^{-1/2}X_{2,[nr]} \Rightarrow B_2(r) = BM(\omega_2^2),$$

$X_1(r)$ is the limiting trending function, and $B_2(r)$ is a stochastic process. Let $D_n = diag\{D_{1n}, n, \cdots, n\}$,

$$n^{1/2}D_n^{-1}X_{[nr]} \to X(r) = \begin{bmatrix} X_1(r) \\ B_2(r) \end{bmatrix}.$$

The OLS filtration estimators $\widehat{\pi}$ has the following limit:

$$D_n(\widehat{\pi} - \pi) \Rightarrow \left[\int X(r)X(r)'dr\right]^{-1} \left[\int X(r)'dB_Y(r) + \Lambda_{XY}\right],$$

where $\Lambda_{XY}' = [0, \Lambda_{2Y}']$. In typical cointegration models, $\Lambda_{2Y} \neq 0$, $B_2(r)$ is correlated with $B_Y(r)$, and $\left[\int B_2(r)B_2(r)'dr\right]^{-1} \int B_2(r)dB_Y(r)$ is asymmetrically distributed.

The latent component, Y_t, is a stationary ergodic process that may display nonlinear dynamics captured by a copula function. For simplicity, we assume that $\{Y_t\}_{t=1}^n$ is a strictly stationary first-order Markov process (see, e.g., Chen and Fan 2006a). Higher order Markov process of $\{Y_t\}_{t=1}^n$ can be handled similarly (see, e.g., Ibragimov, 2009).

Under the assumption that $\{Y_t\}_{t=1}^n$ is a first-order stationary Markov process, its probabilistic properties are determined by the true joint distribution of Y_{t-1} and Y_t, say, $G^*(y_{t-1}, y_t)$. Suppose that

Y_t has continuous marginal distribution function $F^*(\cdot)$, then by Sklar's (1959) Theorem, there exists an unique copula function $C(\cdot,\cdot)$ such that

$$G^*(y_{t-1}, y_t) \equiv C(F^*(y_{t-1}), F^*(y_t)),$$

where the copula function $C(\cdot,\cdot)$ is a bivariate probability distribution function with uniform marginals. Denote the corresponding copula density of $C(u,v)$ by $c(u,v)$, and the density of the marginal distribution $F(\cdot)$ by $f(\cdot)$, the true conditional density of Y_t given Y_{t-1} is

$$p(y_t|y_{t-1}) = f^*(y_t)c(F^*(y_{t-1}), F^*(y_t)).$$

We assume the following basic conditions on the dynamics of the latent process $\{Y_t\}$.

Assumption DGP: $\{Y_t\}_{t=1}^n$ in model (2.1) is a stationary first-order Markov process generated from $(F^*(\cdot), C(\cdot,\cdot;\beta^*))$, where $F^*(\cdot)$ is the true invariant distribution that is absolutely continuous with respect to Lebesgue measure on the real line; $C(\cdot,\cdot;\beta^*)$ is the copula for (Y_{t-1}, Y_t), is absolutely continuous with respect to Lebesgue measure on $[0,1]^2$.

Assumption MX: The process $\{Y_t\}$ is absolutely regular with mixing coefficient $\beta(\tau) = O(\tau^{-\delta})$, for a constant $\delta > 0$.

See Chen and Fan (2006a), Chen, Wu and Yi (2009), Beare (2010), Longla and Peligrad (2012) and others about sufficient conditions that most commonly used copula-based Markov processes are geometric ergodic and hence absolutely regular (or beta-mixing) with exponentially decaying mixing coefficients.

3. Estimation Under Correctly-Specified Copulas

We are interested in estimation and inference on the copula dependence parameter β^*.

3.1. Feasible estimation of copula parameter using filtered data $\widehat{Y}_t$

Let $\widehat{Y}_t$ be the filtered time series, and $\widehat{F}(\cdot)$ be a feasible estimator of the marginal distribution $F^*(\cdot)$ using $\widehat{Y}_t$. In this paper we propose and study the properties of the following feasible copula estimator

$$\widehat{\beta} = \arg\max_{\beta} \widehat{Q}_n(\widehat{F}, \beta), \text{ where } \widehat{Q}_n(\widehat{F}, \beta) = \frac{1}{n}\sum_{t=2}^{n} \log c(\widehat{F}(\widehat{Y}_{t-1}), \widehat{F}(\widehat{Y}_t), \beta). \tag{3.1}$$

3.1.1. Parametric marginal case

We first consider the parametric case where the marginal distribution of Y_t belongs to a parametric family. Denote the unknown true marginal density function and the distribution function of Y_t by

$f(\cdot, \alpha^*)$ and $F(\cdot, \alpha^*)$, where α is an k_1-dimensional vector of unknown parameters. We could then estimate the true marginal $F^*(\cdot)$ by $F(\cdot, \widehat{\alpha})$ where

$$\widehat{\alpha} = \arg\max_{\alpha} \sum_{t=1}^{n} \log f(\widehat{Y}_t, \alpha), \tag{3.2}$$

and estimate the copula parameter β^* by the following "parametric copula estimator":

$$\widehat{\beta}_P = \arg\max_{\beta} \widehat{Q}_n(\beta), \text{ where } \widehat{Q}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} \log c(F(\widehat{Y}_{t-1}, \widehat{\alpha}), F(\widehat{Y}_t, \widehat{\alpha}), \beta).$$

3.1.2. Nonparametric marginal case

In practice, the exact form of marginal distribution is usually beyond our knowledge and thus the parametric model of marginal distribution may be misspecified. We now consider a semiparametric model where the marginal distribution is estimated nonparametrically based on the filtered time series $\widehat{Y}_t$. We use the so-called rescaled empirical distribution function (EDF) to estimate $F^*(\cdot)$:

$$\widehat{F}_n(y) = \frac{1}{n+1} \sum_{t=1}^{n} 1\left(\widehat{Y}_t \le y\right),$$

and estimate the copula parameter β^* by the following "semiparametric copula estimator":

$$\widehat{\beta}_{SP} = \arg\max_{\beta} \widehat{\mathcal{L}}_n(\beta), \text{ where } \widehat{\mathcal{L}}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} \log c(\widehat{F}_n(\widehat{Y}_{t-1}),\ \widehat{F}_n(\widehat{Y}_t), \beta).$$

3.2. Infeasible estimation of copula parameter using Y_t

For comparison purpose, we review an infeasible estimator, $\widetilde{\beta}$, of β^* assuming that Y_t is observed. Let $\widetilde{F}(\cdot)$ be an infeasible estimator of the true marginal distribution $F^*(\cdot)$ using Y_t. Then a pseudo maximum likelihood estimator of β^* using observed Y_t is given by

$$\widetilde{\beta} = \arg\max_{\beta} Q_n(\widetilde{F}, \beta), \text{ where } Q_n(\widetilde{F}, \beta) = \frac{1}{n} \sum_{t=2}^{n} \log c(\widetilde{F}(Y_{t-1}), \widetilde{F}(Y_t), \beta).$$

Again, $\widetilde{\beta}_P$ denotes the parametric copula estimator using the infeasible parametric marginal estimator $\widetilde{F} = F(\cdot, \widetilde{\alpha})$, where[1]

$$\widetilde{\alpha} = \arg\max_{\alpha} \sum_{t=1}^{n} \log f(Y_t, \alpha).$$

[1] Previously, Joe and Xu (1996) and Joe (2005) studied two-step parametric estimation of copula parameter β for iid data $\{(Y_{1,i}, ..., Y_{m,i})\}_{i=1}^{n}$ of a multivariate random vector $(Y_1, ..., Y_m)$ whose concurrent copula density $c(F_1(Y_1; \alpha_1), ..., F_m(Y_m; \alpha_m); \beta)$ links different parametric marginal distributions $F_j(Y_j; \alpha_j), j = 1, ..., m$.

And $\widetilde{\beta}_{SP}$ denotes the semiparametric copula estimator using the infeasible rescaled estimator for $F^*(\cdot)$:

$$\widetilde{F}(y) = F_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} 1\left(Y_t \leq y\right).$$

Chen and Fan (2006a) has proposed and studied the asymptotic properties of $\widetilde{\beta}_{SP}$ for first-order stationary Markov process Y_t.

Comparing $\widehat{\beta}$ and $\widetilde{\beta}$, the infeasible estimator $\widetilde{\beta}$ assumes that Y_t is observed so that it is not affected by filtration of nonstationarity. In addition to $\widetilde{\beta}$ and $\widetilde{\beta}_{SP}$, we also compare our estimators with the ideal infeasible estimator $\breve{\beta}$, which is the maximum likelihood estimator of β^* assuming Y_t is observed with a completely known marginal distribution $F^*(\cdot)$:

$$\breve{\beta} = \arg\max_{\beta} Q_n(F^*, \beta), \text{ where } Q_n(F^*, \beta) = \frac{1}{n}\sum_{t=2}^{n} \log c(F^*(Y_{t-1}), F^*(Y_t), \beta). \tag{3.3}$$

In the next two subsections, we show that although the parameter estimators $\widehat{\beta}_P$ and $\widetilde{\beta}_P$ could have different asymptotic properties, the semiparametric estimators $\widehat{\beta}_{SP}$ and $\widetilde{\beta}_{SP}$ have the same asymptotic distribution.

3.3. Asymptotic properties of parametric copula estimator

In this subsection we establish the consistency and limiting distribution for the feasible parametric copula estimators. We introduce some notation in the parametric case. Let $g\left(Y_{t-1}, Y_t, \alpha, \beta\right) = \log c(F(Y_{t-1}, \alpha), F(Y_t, \alpha), \beta)$ and $g_\beta\left(s_1, s_2, \alpha, \beta\right) = \partial g\left(s_1, s_2, \alpha, \beta\right)/\partial\beta$. For $i = 1, 2$, $j = 1, 2$, we define

$$\begin{aligned}
\frac{\partial g_\beta\left(s_1, s_2, \alpha, \beta\right)}{\partial\alpha} &= g_{\beta\alpha}\left(s_1, s_2, \alpha, \beta\right), \ \frac{\partial g_\beta\left(s_1, s_2, \alpha, \beta\right)}{\partial\beta} = g_{\beta\beta}\left(s_1, s_2, \alpha, \beta\right), \\
\frac{\partial g_\beta\left(s_1, s_2, \alpha, \beta\right)}{\partial s_j} &= g_{\beta j}\left(s_1, s_2, \alpha, \beta\right), \ \frac{\partial g_{\beta\beta}\left(s_1, s_2, \alpha, \beta\right)}{\partial s_j} = g_{\beta\beta j}\left(s_1, s_2, \alpha, \beta\right), \\
\frac{\partial g_{\beta\beta}\left(s_1, s_2, \alpha, \beta\right)}{\partial\alpha} &= g_{\beta\beta\alpha}\left(s_1, s_2, \alpha, \beta\right), \ \frac{\partial g_{\beta\alpha}\left(s_1, s_2, \alpha, \beta\right)}{\partial s_j} = g_{\beta\alpha j}\left(s_1, s_2, \alpha, \beta\right), \\
\frac{\partial g_{\beta i}\left(s_1, s_2, \alpha, \beta\right)}{\partial s_j} &= g_{\beta ij}\left(s_1, s_2, \alpha, \beta\right), \ \frac{\partial g_{\beta i}\left(s_1, s_2, \alpha, \beta\right)}{\partial\alpha} = g_{\beta i\alpha}\left(s_1, s_2, \alpha, \beta\right).
\end{aligned}$$

For convenience, we also denote $\ell(u, v, \beta) = \log c(u, v, \beta)$, and

$$\frac{\partial\ell(u, v, \beta)}{\partial\beta} = \ell_\beta\left(u, v, \beta\right), \frac{\partial\ell(u, v, \beta)}{\partial u} = \ell_1\left(u, v, \beta\right), \frac{\partial\ell(u, v, \beta)}{\partial v} = \ell_2\left(u, v, \beta\right),$$

$$\frac{\partial\ell_\beta\left(u, v, \beta\right)}{\partial u} = \ell_{\beta 1}\left(u, v, \beta\right), \frac{\partial\ell_\beta\left(u, v, \beta\right)}{\partial v} = \ell_{\beta 2}\left(u, v, \beta\right), \frac{\partial\ell_\beta\left(u, v, \beta\right)}{\partial\beta} = \ell_{\beta\beta}\left(u, v, \beta\right).$$

For consistency in the parametric case, we make the following assumptions.

Assumption ID1: (1) $\mathcal{A}$ and $\mathfrak{B}$ are compact subsets of $\mathcal{R}^{k_1}$ and $\mathcal{R}^k$. (2). $q(\alpha)$ =E[$\log f(Y_t, \alpha)$] has a unique maximizer $\alpha^* \in \mathcal{A}$; and $Q(\beta)$ =E[$\ell(F(Y_{t-1}, \alpha^*), F(Y_t, \alpha^*), \beta)$] has a unique maximizer $\beta^* \in \mathfrak{B}$. (3) $f(y, \alpha)$ is continuous in $\alpha \in \mathcal{A}$, and $g(\alpha, \beta)$ =E[$g(Y_{t-1}, Y_t, \alpha, \beta)$] is Lipschitz continuous in $\alpha \in \mathcal{A}$ and $\beta \in \mathfrak{B}$.

Assumption M1 (1) $\mathrm{E}[\sup_\alpha |\log f(Y_t, \alpha)|] < \infty$, and $\mathrm{E}\big[\sup_{\beta \in \mathfrak{B}, \alpha \in \mathcal{A}_\delta} |g(Y_{t-1}, Y_t, \alpha, \beta)|\big] < \infty$. (2) $f(y, \alpha)$ is uniformly continuous in y, uniformly over $\alpha \in \mathcal{A}$, in the sense that for any $\epsilon > 0$, there exists $\delta > 0$, such that if $|y_1 - y_2| < \delta$, then

$$\sup_{\alpha \in \mathcal{A}} |\log f(y_1, \alpha) - \log f(y_2, \alpha)| < \epsilon.$$

Similarly, $g(s_1, s_2, \alpha, \beta)$ is uniformly continuous in (s_1, s_2, α), uniformly over $\beta \in \mathcal{B}$, in the sense that for any $\epsilon > 0$, there exists $\delta > 0$, such that if $|s_1' - s_1''| + |s_2' - s_2''| + |\alpha' - \alpha''| < \delta$, then

$$\sup_{\beta \in \mathcal{B}} \left| g(s_1', s_2', \alpha', \beta) - g(s_1'', s_2'', \alpha'', \beta) \right| < \epsilon.$$

Theorem 1: Under Assumptions DGP, MX, ID1, M1, and X, $\widehat{\beta}_P = \beta^* + o_p(1)$.

We introduce additional notation and assumptions for convenience of developing the limiting distribution of $\widehat{\beta}_P$. Denote

$$\Omega_\beta = E\left[\ell_\beta\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) \ell_\beta\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)'\right]$$

and

$$\Omega_\alpha = \mathrm{E}\left[\frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha} \frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha'}\right], \quad H_\alpha = -\mathrm{E}\left[\frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial \alpha \partial \alpha'}\right].$$

Assumption ID2: (1). $\widehat{\beta}_P = \beta^* + o_p(1)$ and $\beta^* \in int(\mathfrak{B})$ (2) $\partial \widehat{Q}_n(\widehat{\beta}_P)/\partial \beta = o_p(n^{-1/2})$. (3) $\ell_\beta(s_1, s_2, \beta)$ is Lipschitz continuous in β, $\ell_{\beta j}(s_1, s_2, \beta)$ are continuous in (s_1, s_2, β). (3). $H_\beta = -\mathrm{E}\ell_{\beta\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) = \Omega_\beta$ is positive definite. (4). $f(\cdot, \alpha^*)$ and $F(\cdot, \alpha^*)$, are differentiable in α^*. (5) $H_\alpha = \Omega_\alpha$ is positive definite, $\sqrt{n}\left(\widetilde{\alpha} - \alpha^*\right) \Rightarrow N\left(0, \Omega_\alpha\right)$.

Assumption M2 (1) the derivatives of $g_\beta(s_1, s_2, \alpha, \beta)$ are uniformly continuous in $(s_1, s_2, \alpha, \beta)$. (2) the following limits hold in probability:

$$P_{nj} = \frac{1}{n} \sum_{t=2}^{n} g_{\beta j}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) X_{t-2+j}' D_n^{-1} n^{1/2} = P_j + o_p(1), \quad j = 1, 2,$$

$$P_{n3} = n^{-1} \sum_{t=2}^{n} g_{\beta\alpha}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) = P_3 + o_p(1).$$

$$H_{n\alpha Y} = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial \alpha \partial Y} \left(X_t' D_n^{-1} n^{1/2}\right) = H_{\alpha Y} + o_p(1).$$

Theorem 2: Under Assumptions DGP, MX, ID2, M2, and X, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1}\Omega_\beta^{\#}H_\beta^{-1}\right) - H_\beta^{-1}\left(P_1 + P_2 + P_3\Omega_\alpha^{-1}H_{\alpha Y}\right)\xi$$

where

$$\begin{aligned}\Omega_\beta^{\#} &= \lim_n \operatorname{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\left(\ell_\beta\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) + P_3\Omega_\alpha^{-1}\frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha}\right)\right) \\ &= \Omega_\beta + P_3\Omega_\alpha^{-1}P_3'.\end{aligned}$$

An immediate result from Theorem 2 is: in the presence of nonstationarity, the limiting distribution of the parametric copula estimator may not be normal even asymptotically.

From the proof of Theorem 2, we can decompose the limiting distribution of the parametric copula estimator $\widehat{\beta}$ into three components: The first part is $N\left(0, H_\beta^{-1}\Omega_\beta H_\beta^{-1}\right) = N\left(0, \Omega_\beta\right)$, the normal limit of the ideal infeasible estimator when Y_t is observed with a completely known marginal $F^*(Y_t) = F(Y_t, \alpha^*)$ (or a known α^*); The second part is $N(0, H_\beta^{-1}P_3\Omega_\alpha^{-1}P_3'H_\beta^{-1})$, the normal limit from the parametric estimation of marginal parameter α^* using Y_t; The third part is $H_\beta^{-1}\left(P_1 + P_2 + P_3\Omega_\alpha^{-1}H_{\alpha Y}\right)\xi$, the effect of nonstationary filtration $\widehat{Y}_t$. The first two parts are normal random variates but the third part may not be normal. Unless $P_1 + P_2 + P_3\Omega_\alpha^{-1}H_{\alpha Y} = o_p(1)$, the nonstationary filtration will affect the limiting distribution of the parametric copula estimator $\widehat{\beta}_P$. In particular, the filtration affects the limiting distribution of $\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right)$ directly through $\widehat{Y}_t$ and indirectly through $\widehat{\alpha}$. Unless X_t is purely deterministic, the limiting distribution of $\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right)$ is not normal and is generally affected by nuisance parameters in a complicated way.

Remark 1. *Recall the simple asymptotic normality result for the ideal infeasible estimator $\breve{\beta}$, assuming Y_t is observed with a completely known marginal distribution $F^*(\cdot)$, is given by*

$$\sqrt{n}\left(\breve{\beta} - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1}\Omega_\beta H_\beta^{-1}\right) = N\left(0, H_\beta^{-1}\right) = N\left(0, \Omega_\beta\right).$$

From the proof of Theorem 2, we have

$$\sqrt{n}\left(\widetilde{\beta}_P - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1}\Omega_\beta^{\#}H_\beta^{-1}\right).$$

Since $\Omega_\beta^{\#} - \Omega_\beta$ is positive definite, even assuming observable Y_t, there is still efficiency loss of the infeasible parametric copula estimator $\widetilde{\beta}_P$ using a consistent parametric estimator of marginal distribution $F^*()$. Nevertheless, according to Theorem 2, it is unclear which one, $\widetilde{\beta}_P$ vs $\widehat{\beta}_P$, is more efficient.

Example 1 (Continued). Trending Time Series. X_t is a vector of deterministic trend with a limiting trending function $X(r)$. Let

$$\eta = \sum_{j=1}^{2} \mathrm{E}g_{\beta j}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) \ + P_3\Omega_\alpha^{-1}\mathrm{E}\left[\frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial \alpha \partial Y}\right], \tag{3.4}$$

and

$$\eta_X = \eta \int_0^1 X(r)'dr \left(\int_0^1 X(r)X(r)'dr \right)^{-1},$$

notice that,

$$P_{nj} \to P_j = \mathrm{E}g_{\beta j}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) \int_0^1 X(r)'dr, \; j = 1, 2,$$

$$H_{n\alpha Y} \to H_{\alpha Y} = \mathrm{E}\left[\frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial \alpha \partial Y} \right] \int_0^1 X(r)'dr,$$

we have

$$P_1 + P_2 + P_3 \Omega_\alpha^{-1} H_{\alpha Y} = \eta \int_0^1 X(r)'dr,$$

and

$$\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1} \overline{\Omega}_\beta^{\#} H_\beta^{-1}\right),$$

where

$$\overline{\Omega}_\beta^{\#} = \lim_{n\to\infty} \mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=2}^{n} g_\beta\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) + P_{n3}\Omega_\alpha^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha} - \eta_X \sum_t D_n^{-1} X_t Y_t \right).$$

In this example, since the nonstationary component is deterministic and thus is uncorrelated with Y_t, the limiting distribution of $D_n\left(\widehat{\pi} - \pi\right)$ coming from nonstationary filtration is normal, and thus the limiting distribution of the parametric copula estimator in this case $\widehat{\beta}_P$ is normal although it is affected by the filtration asymptotically which is reflected in the formula of the limiting variance matrix $\overline{\Omega}_\beta^{\#}$.

Example 2 (Continued). Unit Root. Suppose that the time series Z_t is a process with unit root. Then $X_t = Z_{t-1}$, $\pi^* = 1$, and the filtration process is an autoregression

$$Z_t = \widehat{\pi} Z_{t-1} + \widehat{Y}_t,$$

$$n\left(\widehat{\pi} - \pi^*\right) \Rightarrow \xi_2 = \left[\int_0^1 B_Y(r)^2 dr \right]^{-1} \left[\int_0^1 B_Y(r) dB_Y(r) + \lambda \right]$$

with $\lambda = \sum_{h=1}^{\infty} E(Y_1 Y_{1+h})$. Then,

$$\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1} \Omega_\beta^{\#} H_\beta^{-1}\right) - \eta H_\beta^{-1} h(B_Y(r))$$

where η is defined as (3.4), and

$$h(B_Y(r)) = \int_0^1 B_Y(r) dr \left[\int_0^1 B_Y(r)^2 dr \right]^{-1} \left[\int_0^1 B_Y(r) dB_Y(r) + \lambda \right].$$

In this example, the limiting distribution ξ_2 coming from nonstationary filtration is non-normal, and thus the limiting distribution of the parametric copula estimator $\widehat{\beta}_P$ is not normal because it is affected by the filtration asymptotically.

Example 3 (Continued). Cointegrated Time Series. $X_t = (X_{1t}', X_{2t}')'$, where X_{1t} is a vector of deterministic trend, and X_{2t} is a vector of unit root process, then

$$P_{nj} \to P_j = \mathrm{E} g_{\beta j}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\left[\int_0^1 X_1(r)' dr, \int_0^1 B_2(r)' dr\right], \; j = 1, 2,$$

and

$$H_{n\alpha Y} \to H_{\alpha Y} = \mathrm{E}\left[\frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial\alpha\partial Y}\right]\int_0^1 X(r)' dr.$$

Then,

$$\sqrt{n}\left(\widehat{\beta}_P - \beta^*\right) \Rightarrow N\left(0, H_\beta^{-1}\Omega_\beta^{\#} H_\beta^{-1}\right) - \eta H_\beta^{-1} h_3(X_1, B_2, B_Y)$$

where

$$\begin{aligned} h_3(X_1, B_2, B_Y) &= \left[\int_0^1 X_1(r)' dr, \int_0^1 B_2(r)' dr\right]\left[\left[\begin{array}{cc} \int X_1(r) X_1(r)' dr & \int X_1(r) B_2(r)' dr \\ \int B_2(r) X_1(r)' dr & \int B_2(r) B_2(r)' dr \end{array}\right]\right]^{-1} \\ &\quad \times \left[\begin{array}{c} \int_0^1 X_1(r) dB_Y(r) \\ \int_0^1 B_2(r) dB_Y(r) + \Lambda_{2Y} \end{array}\right] \end{aligned}$$

In this example, since the nonstationary component contains a vector of stochastic nonstationary process X_{2t} which is usually correlated with Y_t, and a bias term Λ_{2Y}, the limiting distribution coming from nonstationary filtration is not normal. Thus the limiting distribution of the parametric copula estimator in this case $\widehat{\beta}_P$ is not normal.

3.4. Asymptotic properties of semiparametric copula estimator

We denote the space of continuous probability distributions over the support of Y_t as $\mathcal{F}$, then $F \in \mathcal{F}$. For an appropriate positive weighting function $w(\cdot)$ (whose property is specified below in Assumption SP), we define a weighted metric $\|\cdot\|_w$ as

$$\|F - F^*\|_w = \sup_y \left|\{F(y) - F^*(y)\}/w(F^*(y))\right|.$$

For a small $\delta > 0$, let $\mathcal{F}_\delta = \{F \in \mathcal{F} : \|F - F^*\|_w \leq \delta\}$. Then, $F^* \in \mathcal{F}_\delta$, and $F_n \in \mathcal{F}_\delta$ with probability approaching 1 as $n \to \infty$.

Assumption SP: (1) There exists $\overline{Y}$, for $|y| > \overline{Y}$, and any sequence $\delta_n = o(1)$, $|F(y+\delta_n) - F(y)| \leq F(y)(1+o(1))$. (2) $w(\cdot)$ is a continuous function on $[0,1]$ which is strictly positive on $(0,1)$, symmetric at $u = 0.5$, and increasing on $(0, 1/2]$, satisfying $w(u) \geq \zeta\,[u(1-u)]^\mu \log(1/(u(1-u)))^{\mu_1}$ with $\zeta > 0$, $\mu_1 > 0$, $\mu < 1/2q$, $q > 1$.

We first establish an important Lemma for a weighted empirical process that is of independent interest to handle filtration for time series. Consider $b = (b_1, \cdots, b_n)'$, let

$$Z_n(y, b) = \frac{1}{\sqrt{n+1}}\sum_{t=1}^n\left[1\left(Y_t \leq y + n^{-1/2} b_t\right) - F^*(y + n^{-1/2} b_t)\right]$$

and denote $|b| = \max_t |b_t|$.

LEMMA 1. Under Assumptions DGP, MX, SP, and X, for any given $B > 0$,

$$\sup_{|b|\leq B} \sup_y \left| \frac{Z_n(y,b) - Z_n(y,0)}{w(F^*(y))} \right| = o_p(1).$$

We modify the assumptions ID1 and M1 to facilitate asymptotic analysis in the semiparametric case.

Assumption ID3: (1). $\mathfrak{B}$ is a compact subset of $\mathcal{R}^k$. (2) $\mathrm{E}[\ell_\beta(F^*(Y_{t-1}), F^*(Y_t), \beta)] = 0$ if and only if $\beta = \beta^* \in \mathfrak{B}$. (3) $\ell_\beta(s_1, s_2, \beta)$ is Lipschitz continuous in β, $\ell_{\beta j}(s_1, s_2, \beta)$ are continuous in (s_1, s_2, β).

Assumption M3 (1). $\mathrm{E}\big[\sup_{\beta\in\mathfrak{B}} \|\ell_\beta(F^*(Y_{t-1}), F^*(Y_t), \beta)\| \log(1 + \|\ell_\beta(F^*(Y_{t-1}), F^*(Y_t), \beta)\|)\big] < \infty$. (2). $\mathrm{E}\big[\sup_{\beta\in\mathfrak{B}, F\in\mathcal{F}_\delta} \|\ell_{\beta j}(F(Y_{t-1}), F(Y_t), \beta)\| \, w(F^*(Y_{t-2+j}))\big] < \infty$, $j = 1, 2$. (3). $\sup_y |f(y)/w(F^*(y))| < \infty$.

Theorem 3 below gives the consistency of the semiparametric estimator.

Theorem 3: Under Assumptions DGP, SP MX, ID3, M3, and X, $\widehat{\beta}_{SP} = \beta^* + o_p(1)$.

The following additional assumptions are added for asymptotic normality of $\widehat{\beta}_{SP}$.

Assumption ID4: (1). Assumption ID3 is satisfied with $\beta^* \in int(\mathfrak{B})$, (2) $H_\beta = -\mathrm{E}\ell_{\beta\beta}(F^*(Y_{t-1}), F^*(Y_t), \beta^*)$ is positive definite. (3). $\sup_y |(F_n(y) - F^*(y))/w(F^*(y))| = O_p(n^{-1/2})$.

Assumption M4 (1). Let $F_\eta = F^* + \eta[F - F^*]$ for $\eta \in [0,1]$ and $F \in \mathcal{F}_\delta$, the interchange of differentiation and integration of $\ell_\beta\left(F_\eta(Y_{t-1}), F_\eta(Y_t), \beta_\eta\right)$ w.r.t $\eta \in (0,1)$ is valid.

(2) $\mathrm{E}\Big[\sup_{\|\beta-\beta^*\|\leq\delta, F\in\mathcal{F}_\delta} \|\ell_\beta(F(Y_{t-1}), F(Y_t), \beta)\|^2 \log(1 + \|\ell_\beta(F(Y_{t-1}), F(Y_t), \beta)\|)\Big] < \infty$,
$\mathrm{E}\Big[\sup_{\|\beta-\beta^*\|\leq\delta, F\in\mathcal{F}_\delta} \|\ell_{\beta\beta}(F(Y_{t-1}), F(Y_t), \beta)\|^2\Big] < \infty$.

(3). $\mathrm{E}\Big[\sup_{\|\beta-\beta^*\|\leq\delta, F\in\mathcal{F}_\delta} \|\ell_{\beta j}(F(Y_{t-1}), F(Y_t), \beta)\| \, w(F^*(Y_{t-2+j}))\Big]^2 < \infty$, $j = 1, 2$.
$\mathrm{E}\Big[\sup_{\|\beta-\beta^*\|\leq\delta, F\in\mathcal{F}_\delta} |\ell_{\beta ij}(F(Y_{t-1}), F(Y_t), \beta) \, w(F^*(Y_{t+i-2})) w(F^*(Y_{t+j-2}))|\Big] < \infty$, $i, j = 1, 2$.
$\mathrm{E}\Big[\sup_{\|\beta-\beta^*\|\leq\delta, F\in\mathcal{F}_\delta} \|\ell_{\beta\beta j}(F(Y_{t-1}), F(Y_t), \beta) \, w(F^*(Y_{t+j-2}))\|\Big] < \infty$, $i, j = 1, 2$.

Denote

$$\mathcal{G}_n = \frac{1}{\sqrt{n}} \sum_{t=2}^n \left\{\ell_\beta(F^*(Y_{t-1}), F^*(Y_t), \beta^*) + G_0(Y_t) + G_1(Y_{t-1})\right\},$$

where

$$G_j(Y_{t-j}) = \int_0^1 \int_0^1 \left[1(F^*(Y_{t-j}) \leq v_{2-j}) - v_{2-j}\right] \ell_{\beta,2-j}(v_1, v_2; \beta^*) \, c(v_1, v_2; \beta^*) \, dv_1 dv_2, j = 0, 1.$$

Let

$$\Omega_\beta^+ = \lim_{n\to\infty} Var(\mathcal{G}_n) = \Omega_\beta + Var(G_0(Y_t) + G_1(Y_{t-1})).$$

Theorem 4: Under Assumptions DGP, SP, MX, ID4, M4, and X, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\beta}_{SP} - \beta^*\right) = \sqrt{n}\left(\widetilde{\beta}_{SP} - \beta^*\right) + o_p(1) \Rightarrow N\left(0, H_\beta^{-1}\Omega_\beta^+ H_\beta^{-1}\right).$$

In contrast to Theorem 2, which shows that the nonstationary filtration affects the limiting distribution of the parametric copula estimator $\widehat{\beta}_P$, Theorem 4 shows that the nonstationary filtration does not affect the limiting distribution of the semiparametric copula estimator $\widehat{\beta}_{SP}$, which is the same as that of the infeasible semiparametric copula estimator $\widetilde{\beta}_{SP}$ using Y_t.

From the proof of Theorem 4 in the Appendix, we can again decompose the distribution of the semiparametric copula estimator $\widehat{\beta}_{SP}$ into three components: The first part is $N\left(0, H_\beta^{-1}\Omega_\beta H_\beta^{-1}\right) = N(0, \Omega_\beta)$, the normal limit of the ideal infeasible estimator $\breve{\beta}$ when Y_t is observed with a completely known marginal distribution $F^*(\cdot)$; The second part, denoted as $A_{n2}+A_{n4}$ in the Appendix, is from the nonparametric estimation of the unknown marginal distribution using Y_t, and is also asymptotically normal; The third part, denoted as $A_{n1} + A_{n3}$ in the Appendix, is the effect of nonstationary filtration $\widehat{Y}_t$. We show in the Appendix that $A_{n1} + A_{n3} = o_p(1)$, thanks to the fact that the nonparametric marginal distribution estimator enters the copula score function in a symmetric manner that absorbed and cancelled the filtration effects. Therefore, the distribution of $\sqrt{n}\left(\widehat{\beta}_{SP} - \beta^*\right)$ is only asymptotically affected by the first two parts. Consequently, the limiting distribution of $\sqrt{n}\left(\widehat{\beta}_{SP} - \beta^*\right)$ is the same as that of $\sqrt{n}\left(\widetilde{\beta}_{SP} - \beta^*\right)$, which is always normal.

Remark 2. *Chen and Fan (2006b) studied the following class of semiparametric copula-based multivariate dynamic models*

$$\begin{aligned}
Z_t &= (Z_{1,t}, ..., Z_{d,t}), \quad Z_{j,t} = \mu_{j,t}(\theta^*) + \sigma_{j,t}(\theta^*) Y_{j,t}, \\
\mu_{j,t}(\theta^*) &= E[Z_{j,t}|\mathcal{I}_{t-1}], \quad \sigma_{j,t}^2(\theta^*) = VarE[Z_{j,t}|\mathcal{I}_{t-1}], \\
Y_t &= (Y_{1,t}, ..., Y_{d,t}) \quad \textit{is independent of } \mathcal{I}_{t-1}, \textit{ and } \{Y_t\}_{t=1}^n \textit{ is i.i.d. over } t
\end{aligned}$$

where the joint distribution of the multivariate standardized innovation $Y_t = (Y_{1,t}, ..., Y_{d,t})$ has the concurrent copula density $c(F_1(Y_{1,t}), ..., F_d(Y_{d,t}); \beta)$ that links marginal distributions $F_j(Y_{j,t}), j = 1, ..., d$ of individual standardized innovation at the same time period t. Chen and Fan (2006b) established that the asymptotic distribution of the semiparametric (two-step) copula parameter estimator using the filtered standardized innovation $\widehat{Y}_t$ is the same as that based on true multivariate standardized innovation Y_t, and hence is not affected by the estimation of the dynamic conditional mean and volatility parameters θ. Although results look similar, we should stress that the result behind Chen and Fan (2006b) crucially depends on the independence between $Y_t = (Y_{1,t}, ..., Y_{d,t})$ and the dynamic part $\mathcal{I}_{t-1}$ of the observed time series Z_t. However, in the presence of nonstationarity (say, unit-root or cointegration) as in our paper, X_t can be correlated with the residual term Y_t, and hence our Theorem 4 could not be explained by that in Chen and Fan (2006b).

3.5. Semiparametric inference on copula parameters

The simple and robust asymptotic properties of the semiparametric (two-step) copula estimator greatly simplify all kinds of statistical inferences on copula models for latent $\{Y_t\}$. In this section, we briefly mention the Wald test for restrictions on the copula dependence parameters β using the asymptotic results of Theorem 4.

Consider the general linear restriction $\mathrm{H}_{01} : R\beta^* = r$. A leading example is the significance test for β: H_{02}: $\beta_j^* = \beta_{0j}$. Notice that under the null H_{01} and regularity assumptions,

$$\sqrt{n}\left(R\widehat{\beta}_{SP} - r\right) \Rightarrow N\left(0, RH_\beta^{-1}\Omega_\beta^+ H_\beta^{-1}R'\right),$$

where H_β and Ω_β^+ are defined in Theorem 4. Thus, under H_{01}, as $n \to \infty$,

$$n\left(R\widehat{\beta}_{SP} - r\right)' \left[RH_\beta^{-1}\Omega_\beta^+ H_\beta^{-1}R'\right]^{-1} \left(R\widehat{\beta}_{SP} - r\right) \Rightarrow \chi^2_{d_r},$$

where d_r is the number of restrictions.

In order to construct the Wald test, we need to estimate $\Omega_\beta^+ = \lim_{n\to\infty}\mathrm{Var}(\mathcal{G}_n)$, and $H_\beta = -\mathrm{E}\ell_{\beta\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)$. We may estimate H_β by the sample analog:

$$\widehat{H}_\beta = -\frac{1}{n}\sum_{t=2}^{n} \ell_{\beta\beta}\left(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \widehat{\beta}_{SP}\right),$$

and estimate Ω_β^+ by a nonparametric kernel estimator:

$$\widehat{\Omega}_\beta^+ = \sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\widehat{\gamma}_n(h).$$

with

$$\widehat{\gamma}_n(h) = \frac{1}{\sqrt{n}} \sum_{\substack{t=2 \\ 2\le t, t+h\le n}}^{n} S_t\left(\widehat{F}_n, \widehat{\beta}\right) S_{t+i}\left(\widehat{F}_n, \widehat{\beta}\right),$$

where

$$S_{t+i}\left(\widehat{F}_n, \widehat{\beta}\right) = \ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t+i-1}), \widehat{F}_n(\widehat{Y}_{t+i}), \widehat{\beta}\right) + \widehat{G}_0(\widehat{Y}_{t+i}) + \widehat{G}_1(\widehat{Y}_{t+i-1})$$

$$\widehat{G}_j(\widehat{Y}_{t-j}) = \int_0^1\int_0^1 \left[1\left(\widehat{F}_n(\widehat{Y}_{t-j}) \le v_{2-j}\right) - v_{2-j}\right] \ell_{\beta,2-j}\left(v_1, v_2; \widehat{\beta}\right) c\left(v_1, v_2; \widehat{\beta}\right) dv_1 dv_2, j = 0, 1.$$

We define the Wald test statistic as

$$W_n = n\left(R\widehat{\beta}_{SP} - r\right)' \left[R\widehat{H}_\beta^{-1}\widehat{\Omega}_\beta^+ \widehat{H}_\beta^{-1}R'\right]^{-1} \left(R\widehat{\beta}_{SP} - r\right)$$

We assume the following bandwidth condition for the consistency of covariance estimator for Ω_β^+.

Assumption BW: As $n \to \infty$, $M \to \infty$ and $M = o(n^{1/3})$.

Theorem 5. Under Assumptions DGP, SP, MX, ID3, M3, X, and BW, we have: (1) $\widehat{\Omega}_\beta^+ = \Omega_\beta^+ + o_p(1)$**.** (2) Under H_{01}, $W_n \Rightarrow \chi^2_{d_r}$ where d_r is the number of linearly independent restrictions.

4. Semiparametric Estimation Under Copula-Misspecification

4.1. Semiparametric two-step estimation of pseudo-true copula parameters

Our previous analysis considers the case where the copula function is correctly specified. In some applications, economic or finance theory may shed little light on the specification of a parametric copula model. Although in practice one may select a copula to capture the main source of nonlinear correlation by eye spotting a simple plot of $\widehat{F}_n(\widehat{Y}_t)$ against $\widehat{F}_n(\widehat{Y}_{t-1})$ to roughly exam the dependence in data, the copula model is in general an approximation and maybe potentially misspecified. In practice, there might be multiple parametric copula functions that can generate the similar observed tail dependence structure. For this reason, in this section we consider our model when the copula functions are potentially misspecified.

Theorem 4 shows that the nonstationary filtration does not affect the limiting distribution of the semiparametric copula estimator for correctly specified copula functions. Since Monte Carlo results reveal the good finite sample performance of semiparametric copula estimator, we shall focus on semiparametric copula estimator allowing for misspecified copula functions in this section.

Suppose that the *true* copula function that captures the dependence in Y_t is given by $C^*(u,v)$, but we consider a copula function $C(u,v,\beta)$ and estimate β by $\widehat{\beta}$ which maximizes

$$\widehat{L}_n(\beta) = \frac{1}{n}\sum_{t=2}^{n} \log c(\widehat{F}_n(\widehat{Y}_{t-1}),\ \widehat{F}_n(\widehat{Y}_t), \beta),$$

where $\widehat{F}_n(\widehat{Y}_t)$ is the EDF of Y_t estimated based on the filtered time series $\left\{\widehat{Y}_t\right\}$ as in Section 3.2.

The infeasible semiparametric estimator based on unobserved Y_t maximize

$$L_n(\beta) = \frac{1}{n}\sum_{t=2}^{n} \log c(F_n(Y_{t-1}), F_n(Y_t), \beta).$$

where

$$F_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} 1\left(Y_t \le y\right).$$

The maximizer of $L_n(\beta)$ will converge to the pseudo-true value $\overline{\beta}$ of the copula dependence parameter defined as the minimizer of the Kullback-Leibler Information Criterion (KLIC) between the candidate parametric copula density and the true unknown copula density,

$$\overline{\beta} = \arg\min_{\beta} \mathrm{KLIC}\left(c^*, c\left(\cdot, \beta\right)\right)$$

where following White (1982),

$$\mathrm{KLIC}\left(c^*, c\left(\cdot, \beta\right)\right) = \mathrm{E}\log c^*(F^*(Y_{t-1}), F^*(Y_t)) - \mathrm{E}\log c(F^*(Y_{t-1}), F^*(Y_t), \beta).$$

In the special case when the class of copula functions $C(u,v,\beta)$ is *correctly* specified, $C^*(u,v) = C(u,v,\overline{\beta})$. In this section, we show that, even in the misspecified case, the nonstationary filtration does not affect the limiting distribution of the semiparametric estimator when it is centered around the pseudo-true parameter $\overline{\beta}$. Similar to Theorem 2 for the correctly specified case, the limiting distribution of parametric copula estimators based on filtered time series under copula misspecification are again affected by the preliminary filtration, and may not be asymptotic normal in the presence of a nonstationary component.

We still denote $\ell(u,v,\beta) = \log c(u,v,\beta)$ and define its derivatives in the same way as in Section 3, but keep in mind that the copula function is misspecified.

We make the following regularity assumptions, which are parallel to the assumptions in Section 3.4, but modified to accommodate the misspecified copula.

Assumption ID5: (1). $\overline{\beta} \in \mathfrak{B}$, $\mathfrak{B}$ is a compact subset of $\mathcal{R}^k$. (2) $Q(\beta)=$E$[\ell(F^*(Y_{t-1}), F^*(Y_t), \beta)]$ has a unique maximizer $\overline{\beta}$ on $\mathfrak{B}$. (3) $Q(\beta)$ is Lipschitz continuous in $\beta \in \mathfrak{B}$.

Theorem 6. Under Assumptions DGP, MX, ID5, M3, and X, $\widehat{\beta} = \overline{\beta} + o_p(1)$.

Assumption ID6: (1). Assumption ID5 is satisfied with $\overline{\beta} \in int(\mathfrak{B})$, (2). $\overline{H}_\beta = -$E$[\ell_{\beta\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \overline{\beta}\right)]$ is positive definite. (3). $\sup_y |(F_n(y) - F^*(y))/w(F^*(y))| = O_p(n^{-1/2})$.

Assumption M6: Assumption M4 holds for the misspecified log density $\ell(u,v,\beta)$ around the pseudo-true value $\overline{\beta}$.

Let $\overline{\Omega}_\beta = \lim_{n\to\infty}\text{Var}\left(\overline{\mathcal{G}}_n\right)$ where $\overline{\mathcal{G}}_n = n^{-1/2}\sum_{j=2}^n \overline{\ell}_\beta\left(U_{j-1}, U_j, \overline{\beta}\right)$ and $U_t = F^*(Y_t)$,

$$\overline{\ell}_\beta\left(U_{j-1}, U_j, \overline{\beta}\right) = \ell_\beta\left(U_{j-1}, U_j, \overline{\beta}\right) + \sum_{i=0}^{1} \text{E}\left[\ell_{\beta,2-i}\left(U_{t-1}, U_t, \overline{\beta}\right)\left[1\left(U_j \le U_{t-i}\right) - U_{t-i}\right]\middle| U_j\right]$$

Theorem 7. Under Assumptions DGP, MX, ID6, M6, and X, as $n \to \infty$,

$$\sqrt{n}\left(\widehat{\beta}_{SP} - \overline{\beta}\right) = \sqrt{n}\left(\widetilde{\beta}_{SP} - \overline{\beta}\right) + o_p(1) \Rightarrow N\left(0, \overline{H}_\beta^{-1}\overline{\Omega}_\beta\overline{H}_\beta^{-1}\right).$$

Theorem 7 shows that, in the case of misspecified copula, the nonstationary filtration does not affect the limiting distribution of the semiparametric copula estimator $\widehat{\beta}_{SP}$ (centered at the pseudo-true parameter $\overline{\beta}$), which is again normal, the same as that of the infeasible semiparametric copula estimator $\widetilde{\beta}_{SP}$ using Y_t, under misspecification.

4.2. Semiparametric inference on copula model selection

We next consider copula model selection using the asymptotic result derived in this Section. In practice, there might be more than one copula functions that can generate the similar observed dependence structure, and we want to select a copula function among candidate copula functions. Suppose that there are two candidate classes of parametric copula models given by $C_j\left(u_1, u_2, \beta_j\right)$, $j = 1, 2$. We are interested in selecting a copula model from these two candidates. Corresponding to the j-th copula, the conditional log likelihood of Y_t given Y_{t-1} is given by

$$\log f^*(y_t) + \log c_j(F^*\left(y_{t-1}\right), F^*\left(y_t\right), \beta_j).$$

Notice that the first term $\log f^*(y_t)$ is not dependent on the copula, we may consider the following log-likelihood ratio:

$$LR = \log \frac{c_2(F^*\left(y_{t-1}\right), F^*\left(y_t\right), \beta_2)}{c_1(F^*\left(y_{t-1}\right), F^*\left(y_t\right), \beta_1)}.$$

If we consider the hypothesis H_0: Copula model $C_1\left(u_1, u_2, \beta_1\right)$ is not worse than copula model $C_2\left(u_1, u_2, \beta_2\right)$; vs. H_1: Copula model $C_1\left(u_1, u_2, \beta_1\right)$ is worse than copula model $C_2\left(u_1, u_2, \beta_2\right)$. Then, under H_0, LR is small (negative). Otherwise, it is large (positive). In practice, neither F nor Y_t are observed, and have to be replaced by appropriate estimates. Let $\widehat{\beta}_j$ $(j = 1, 2.)$ be the semiparametric estimator $\left(\widehat{\beta}_{SP}\right)$ using the filtered time series $\left\{\widehat{Y}_t\right\}_{t=1}^n$ and based on model j, we construct the following pseudo-likelihood-ratio (PLR) statistic

$$\widehat{LR}_n = \frac{1}{n}\sum_{t=2}^{n} \log \frac{c_2(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \widehat{\beta}_2)}{c_1(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \widehat{\beta}_1)},$$

based on

$$\widehat{F}_n(\widehat{Y}_t) = \frac{1}{n+1}\sum_{j=1}^{n} 1\left(\widehat{Y}_j \leq \widehat{Y}_t\right).$$

For convenience of asymptotic analysis, we introduce the following infeasible PLR statistic LR_n based on unobserved $\{Y_t\}_{t=1}^n$,

$$LR_n = \frac{1}{n}\sum_{t=2}^{n} \log \frac{c_2(F_n(Y_{t-1}), F_n(Y_t), \widetilde{\beta}_2)}{c_1(F_n(Y_{t-1}), F_n(Y_t), \widetilde{\beta}_1)},$$

where $\widetilde{\beta}_j$ $(j = 1, 2.)$ are the semiparametric copula estimators based on model j and $\{Y_t\}_{t=1}^n$ and

$$F_n(Y_t) = \frac{1}{n+1}\sum_{j=1}^{n} 1\left(Y_j \leq Y_t\right).$$

The following theorem shows that the PLR statistic $\widehat{LR}_n$ is asymptotically equivalent to the infeasible PLR test LR_n.

Theorem 8: Under Assumptions DGP, SP, MX, ID4, M4, and X, as $n \to \infty$, (i) If

$$\Pr\left\{(Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \overline{\beta}_1) \neq c_2(F^*(Y_1), F^*(Y_2), \overline{\beta}_2)\right\} > 0,$$

where $\overline{\beta}_j$ are the pseudo-true values of the copula dependence parameters,

$$\sqrt{n}\left(\widehat{LR}_n - LR_n\right) = o_p(1).$$

(ii) If $\Pr\left\{(Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \overline{\beta}_1) = c_2(F^*(Y_1), F^*(Y_2), \overline{\beta}_2)\right\} = 1$,

$$n\left(\widehat{LR}_n - LR_n\right) = o_p(1).$$

Theorem 8 shows that, under our assumptions, the limiting distribution of the pseudo-likelihood-ratio (PLR) test $\widehat{LR}_n$ is the same as the infeasible PLR statistic LR_n based on unobserved Markov series $\{Y_t\}_{t=1}^n$. Thus, Chen and Fan (2006b) can be slightly modified to conduct PLR copula model selection test for latent Markov series $\{Y_t\}$ using nonstationary filtered data. In particular, when the two copula models are generalized non-nested in the sense

$$\Pr\left\{(Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \overline{\beta}_1) \neq c_2(F^*(Y_1), F^*(Y_2), \overline{\beta}_2)\right\} > 0,$$

the null hypothesis H_0 is a composite hypothesis, and we may consider the least favorable configuration (LFC) which satisfies

$$\mathrm{E}\left[\log \frac{c_2(F^*(Y_{t-1}), F^*(Y_t), \overline{\beta}_2)}{c_1(F^*(Y_{t-1}), F^*(Y_t), \overline{\beta}_1)}\right] = 0.$$

Thus, under the LFC and other regularity Assumptions,

$$\sqrt{n}\widehat{LR}_n \Rightarrow N\left(0, \omega^2\right), \text{ as } n \to \infty,$$

with

$$\omega^2 = \lim \mathrm{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=2}^{n} s(U_{t-1}, U_t, \overline{\beta}_2, \overline{\beta}_1) + \sum_{j=1}^{2}\left[\frac{1}{\sqrt{n}}\sum_{l=2}^{n}\left\{g_{2j}\left(U_l, \overline{\beta}_2\right) - g_{1j}\left(U_l, \overline{\beta}_1\right)\right\}\right]\right),$$

where

$$s(U_{t-1}, U_t, \overline{\beta}_2, \overline{\beta}_1) = \log \frac{c_2(U_{t-1}, U_t, \overline{\beta}_2)}{c_1(U_{t-1}, U_t, \overline{\beta}_1)}, \quad U_t = F^*(Y_t),$$

and

$$g_{ij}\left(U_l, \overline{\beta}_i\right) = \mathrm{E}\left\{\left[\frac{\partial \log c_i(U_{t-1}, U_t, \overline{\beta}_i)}{\partial U_{t-2+j}}\right][(1(U_l \leq U_{t-2+j}) - U_{t-2+j})]\middle| U_l\right\}.$$

Let $\widehat{\omega}^2$ be a consistent long-run variance estimator of ω^2 based on

$$\widehat{s}_t(\widehat{\beta}_1, \widehat{\beta}_2) = \log \frac{c_2(\widehat{F}\left(\widehat{Y}_{t-1}\right), \widehat{F}\left(\widehat{Y}_t\right), \widehat{\beta}_2)}{c_1(\widehat{F}\left(\widehat{Y}_{t-1}\right), \widehat{F}\left(\widehat{Y}_t\right), \widehat{\beta}_1)}$$

and for $i = 1, 2,\ j = 1, 2,$

$$\widehat{g}_{t,ij}\left(\widehat{\beta}_i\right) = \frac{1}{n}\sum_{l=2}^{n}\left[\frac{\partial \log c_i(\widehat{F}\left(\widehat{Y}_{l-1}\right), \widehat{F}\left(\widehat{Y}_l\right), \widehat{\beta}_i)}{\partial U_{l-2+j}}\right]\left[\left(1(\widehat{F}\left(\widehat{Y}_t\right) \leq \widehat{F}\left(\widehat{Y}_{l-2+j}\right) - \widehat{F}\left(\widehat{Y}_{l-2+j}\right)\right)\right].$$

Then we may consider the following testing statistic

$$\mathcal{L}_n = \frac{\sqrt{n}\widehat{LR}_n}{\widehat{\omega}}.$$

Under the LFC and generalized non-nested case,

$$\mathcal{L}_n \to N(0,1), \text{ as } n \to \infty.$$

Many applications using non-nested copula models, the above model selection test is directly applicable. For theoretical completeness, we could also consider generalized nested case, in which $c_1(F^*(Y_1), F^*(Y_2), \overline{\beta}_1) = c_2(F^*(Y_1), F^*(Y_2), \overline{\beta}_2)$, a.s.. Denote

$$H_{jn} = -\frac{1}{n}\sum_{t=2}^{n}\frac{\partial^2 \log c_j(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \overline{\beta}_j)}{\partial\beta\partial\beta'} \to\ \overline{H}_{j,\beta} = -\mathrm{E}\left[\frac{\partial^2 \log c_j(F^*(Y_{t-1}), F^*(Y_t), \overline{\beta}_j)}{\partial\beta\partial\beta'}\right]$$

and let $U_t = F^*(Y_t)$, and $\overline{\mathcal{G}}_{j,n} = n^{-1/2}\sum_{j=2}^{n}\overline{\ell}_{j,\beta}\left(U_{j-1}), U_j), \overline{\beta}_j\right)$, $j = 1, 2$, where

$$\begin{aligned}\overline{\ell}_{j,\beta}\left(U_{j-1}), U_j), \overline{\beta}_j\right) &= \frac{\partial \log c_j\left(U_{j-1}), U_j), \overline{\beta}_j\right)}{\partial\beta_j} \\ &+ \sum_{i=0}^{1}\mathrm{E}\left[\frac{\partial^2 \log c_j\left(U_{t-1}, U_t, \overline{\beta}_j\right)}{\partial\beta_j\partial U_{t-i}}\left[1\left(U_j \leq U_{t-i}\right) - U_{t-i}\right]\middle| U_j\right]\end{aligned}$$

then

$$\begin{bmatrix}\overline{\mathcal{G}}_{2,n} \\ \overline{\mathcal{G}}_{1,n}\end{bmatrix} \Rightarrow N\left(0, \begin{bmatrix}\overline{\Omega}_{2,\beta} & \overline{\Omega}_{2,1} \\ \overline{\Omega}'_{2,1} & \overline{\Omega}_{1,\beta}\end{bmatrix}\right).$$

Under the null, $2n\widehat{LR}_n$ converges to a weighted sum of independent χ^2_1 random variables in which the weights $(\lambda_1, \cdots, \lambda_{k_1+k_2})$ is the vector of eigenvalues of the following matrix

$$\begin{bmatrix}\overline{\Omega}_{2,\beta} & \overline{\Omega}_{2,1} \\ \overline{\Omega}'_{2,1} & \overline{\Omega}_{1,\beta}\end{bmatrix}\begin{bmatrix}\overline{H}^{-1}_{2,\beta} & \\ & -\overline{H}^{-1}_{1,\beta}\end{bmatrix}.$$

5. Monte Carlo Studies

In this section, we exam the finite sample performance of the parametric and semiparametric copula estimators based on filtered time series $\left\{\widehat{Y}_t\right\}$. We compare the sampling performance of the semiparametric estimator $\widehat{\beta}_{SP}$ with the parametric estimator $\widehat{\beta}_P$ under correct and incorrect specifications of the marginal distribution F^* (of the latent Y_t); in particular, $\widehat{\beta}_{P*}$ signifies the $\widehat{\beta}_P$ under correct specification and $\widehat{\beta}_{P1}$ signifies the $\widehat{\beta}_P$ under incorrect specification of F^*. In addition, for comparison purpose, we also look at two infeasible copula estimators based on the true values of $\{Y_t\}$: the infeasible parametric estimator $\widetilde{\beta}_{P*}$ under correct specification of F^*, and the infeasible semiparametric estimator $\widetilde{\beta}_{SP}$ using $\{Y_t\}$ process (no filtration is needed).

DGP designs: The observed time series $\{Z_t\}_{t=1}^n$ is generated by $Z_t = X_t'\pi^* + Y_t$, where $\{Y_t\}_{t=1}^n$ is a latent first-order stationary Markov process generated from a copula function $C(\cdot,\cdot;\beta)$ and a marginal distribution F^* such that the joint distribution of (Y_{t-1}, Y_t) is given by

$$G^*(y_{t-1}, y_t) \equiv C(F^*(y_{t-1}), F^*(y_t); \beta^*).$$

In the Monte Carlo studies, we have examed various combinations of three kinds of filtering part $X_t'\pi^*$, four kinds of copula functions $C(\cdot,\cdot;\beta)$ with a range value of the copula parameter β, and two kinds of marginal distributions F^*.

Three types of $X_t'\pi^*$: (1) X_t is a deterministic trend process; in particular we use a linear trend, i.e. $X_t = (1,t)'$, and $\{Z_t\}$ are generated by

$$Z_t = \pi_0^* + \pi_1^* t + Y_t \quad \text{with} \quad \pi_0^* = 0.2, \pi_1^* = 0.3. \tag{5.1}$$

(2) Z_t (and thus $X_t = Z_{t-1}$) is an unit root process:

$$Z_t = \pi^* Z_{t-1} + Y_t \quad \text{with} \quad \pi^* = 1. \tag{5.2}$$

(3) X_t is an I(1) process and is cointegrated with Z_t,

$$X_t = X_{t-1} + \varepsilon_t, \quad \text{with} \quad Z_t = \pi^* X_t + Y_t, \quad \text{with} \quad \pi^* = 1. \tag{5.3}$$

Two types of true marginal distributions: (i) symmetric one: student-$t(3)$ distribution; (ii) asymmetric one: re-centered Chi-square with d.f. 3.

Four types of copula functions: (A) The Gaussian Copula**.** Let $\Phi_\beta(\cdot,\cdot)$ be the distribution function of bivariate normal distribution with mean zeros, variances 1, and correlation coefficient β, and Φ be the CDF of a univariate standard normal. The bivariate Gaussian copula is given by

$$\begin{aligned} C(u,v;\beta) &= \Phi_\beta(\Phi^{-1}(u), \Phi^{-1}(v)) \\ &= \frac{1}{2\pi\sqrt{1-\beta^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left\{-\frac{(s^2 - 2\beta st + t^2)}{2(1-\beta^2)}\right\} dsdt. \end{aligned}$$

If the marginal distribution of Y_t is $F^*(\cdot)$. denote $U_t = F^*(Y_t)$, then the joint distribution of U_t and U_{t-1} is

$$C(u_{t-1}, u_t; \beta) = \Phi_\beta(\Phi^{-1}(u_{t-1}), \Phi^{-1}(u_t)).$$

(B). The Frank copula:

$$C(u, v; \beta) = \log(\beta^{-1}) \frac{\beta^{u+v}}{1-\beta} \left[1 - \frac{(1-\beta^u)(1-\beta^v)}{1-\beta}\right]^{-2}, \text{ if } \beta > 0,\ \beta \neq 1.$$

(C). The Clayton copula:

$$C(u, v; \beta) = [u^{-\beta} + v^{-\beta} - 1]^{-1/\beta}, \quad \text{where} \quad \beta > 0.$$

(D) The Gumbel copula:

$$C(u, v; \beta) = \exp\left\{-((-\ln u)^\beta + (-\ln v)^\beta)^{1/\beta}\right\} \quad \text{for} \quad 1 \leq \beta < \infty.$$

Gaussian and Frank copulas have zero tail dependence. Clayton copula has zero upper tail dependence but positive lower tail dependence ($2^{-1/\beta}$) that increases with β. Gumbel copula has zero lower tail dependence but positive upper tail dependence ($2 - 2^{1/\beta}$) that increases with β. The overall temporal dependence in Y_t measured as Kendall's tau is all increasing with copula parameter β in all these copula models. Finally, the Y_t generated according to all these copula functions are automatically beta-mixing with exponential decay. See, e.g., Chen, Wu and Yi (2009).

For all the above models, we investigate the finite sample performance of the five copula estimators mentioned at the beginning of this section: the three feasible ones $\widehat{\beta}_{SP}$, $\widehat{\beta}_{P*}$ and $\widehat{\beta}_{P1}$ use the nonstationary filtered data; and the two infeasible ones $\widetilde{\beta}_{SP}$ and $\widetilde{\beta}_{P*}$ use the true Y_t process (without filtration). Recall that $\widehat{\beta}_{SP}$ and $\widetilde{\beta}_{SP}$ have the same asymptotic normal distribution, which does not depend on the filtration or the functional form of F^*. The infeasible $\widetilde{\beta}_{P*}$ is asymptotically normal, with the limiting distribution independent of the filtration but does depend on the parametric estimation of F^*. The two feasible parametric estimators $\widehat{\beta}_{P*}$ and $\widehat{\beta}_{P1}$ have complex limiting distributions that depend on both the filtration and the parametric estimation of F^*, while they are asymptotically normal under deterministic trend filtration, are generally non-normal under stochastic trend (the unit root and cointegration) filtration.

In Appendix A we present all the monte Carlo tables. For each table, the number of Monte Carlo repetition is 2000 and the simulated sample size is $n = 500$ (although we considered a larger sample size of $n = 2000$ in a few tables as well). The Monte Carlo bias, variance, and the Ratio of MSE of an estimator over the MSE of $\widehat{\beta}_{P*}$ denoted by "Rmse", are reported in each table.

All the simulations reveal the following patterns. First, the semiparametric copula estimator $\widehat{\beta}_{SP}$ performs well in terms of finite sample bias, variance, "Rmse" compared to the correctly specified parametric estimator $\widehat{\beta}_{P*}$ in most situations. Second, for all the cases when there is no strong lower tail dependence, both the semiparametric copula estimator $\widehat{\beta}_{SP}$ and the correctly specified parametric copula estimator $\widehat{\beta}_{P*}$ perform much better than the parametric copula estimator $\widehat{\beta}_{P1}$ using incorrectly specified parametric marginals. The parametric copula estimator for copula dependence parameter β^* is very sensitive to the specification of parametric marginals, while the semiparametric copula estimator is truly robust to functional form of marginals as well as the nonstationary filtering. Third, the feasible semiparametric estimator $\widehat{\beta}_{SP}$ and its infeasible version $\tilde{\beta}_{SP}$ are reasonably close, corroborating the asymptotic results - the efficiency loss from filtration in the semiparametric estimators are of second order magnitude. The feasible parametric estimator $\widehat{\beta}_{P*}$ and its infeasible version $\widetilde{\beta}_{P*}$ are less close to each other, signaling that the parametric estimator is sensitive to nonstationary filtration. Forth, an interesting exception is the case for Clayton copula with very strong lower tail dependence (or large parameter value β^*). In this case, the infeasible parametric copula estimator $\widetilde{\beta}_{P*}$ performs much better than the feasible parametric estimator $\widehat{\beta}_{P*}$ and the semiparametric estimators, $\widehat{\beta}_{SP}$ and $\tilde{\beta}_{SP}$. The performance of $\widehat{\beta}_{SP}$ is again similar to the infeasible $\tilde{\beta}_{SP}$ for Clayton copula with very strong lower tail dependence, which has been shown to perform poorly (due to big bias) in Chen, Wu and Yi (2009).[2] We plan to investigate this issue in future research.

6. Empirical Applications

In this section, we consider two empirical applications to highlight the potentials of our proposed models and methods.

6.1. An application to macro time series

An important literature in empirical macroeconomic analysis is the study of long-run properties and short term dynamics of GNP. Many studies (e.g. Blanchard 1981, Kydland and Prescott 1980, etc) argue that GNP reverts to a long term trend following a shock, and that fluctuations in output represent temporary deviations from the trend. Various macroeconomic theories are designed to produce and understand the dynamics of transitory fluctuations that deviates from the long run trend. Studies on the transitory shocks provide important information on the prediction of variation in GNP growth. (see, e.g. Cochrane (1994), King, Plosser, Stock and Watson (1991)).

A time series that provides a good estimate of the "trend" in GNP is "consumption". Cochrane (1994) provides empirical evidence on the role of consumption as an measurement of long run compo-

[2] Chen, Wu and Yi (2009) had shown that Clayton copula generated Markov process $\{Y_t\}$ is beta-mixing with exponential decay. Ibragimov and Lentzas (2017) provided simulation evidence that, in finite samples, the time series plot of the Clayton copula generated stationary Markov process $\{Y_t\}$ may exhibit a spurious long memory-like behavior when the lower tail dependence is strong. This might explain the poor finite sample performance in this case

nent in GNP. In this section, we apply our model to estimate the short term dynamics in GNP time series based on the cointegrating regression of GNP on consumption. In particular, we consider the following trending cointegrating regression

$$Z_t = a_0 + a_1 t + a_2 X_t + Y_t \tag{6.1}$$

where Z_t is the logarithm of real GNP and X_t is the logarithm of real consumption. The permanent component of the GNP series is characterized by a linear time trend combined with a stochastic trend X_t. We assume that the latent process $\{Y_t\}$ is a stationary first-order Markov process generated from a flexible copula $C(\cdot, \cdot; \beta)$.

All data are from FRED® Economic Data.[3] We consider quarterly time series from 1947 Q1 to 2019 Q2, with length 290. Consumption is defined as the sum of nondurables and services. We first exam the nonstationarity of these series. In particular, we apply the ADF test to these series based on the following regression

$$Z_t = b_0 + \delta t + \rho Z_{t-1} + \sum_{i=1}^{p} b_i \Delta Z_{t-i} + \varepsilon_t$$

The ADF testing statistics of the GDP and consumption time series are -1.530622 (lag length = 3), and 0.206161 (lag length = 3) respectively, both are smaller (in absolute value) than the 5% critical value (-3.43), thus the null hypothesis of a unit root can not be rejected.

We then exam the relationship between these two time series based on the cointegrating regression (6.1). The Engle-Granger two-step cointegration test statistic is -4.13, rejecting the null hypothesis of no cointegration (5% critical value -3.78).

Next, we study the short term dynamics in the latent process $\{Y_t\}$ using the fitted residual series $\{\widehat{Y}_t\}$ obtained from the cointegrating regression (6.1). Figure 6.1 presents the scatter plot of the empirical cdf standardized realizations of the filtered time series $\{\widehat{Y}_t\}$. The figure indicates possibly presence of asymmetric tail dependence.

Given the small sample size of $n = 290$, to capture possibly asymmetric tail dependence we consider the Joe-Clayton copula:

$$C(u, v; \beta) = 1 - \{1 - [(1 - \bar{u}^{\beta_2})^{-\beta_1} + (1 - \overline{v}^{\beta_2})^{-\beta_1} - 1]^{-1/\beta_1}\}^{1/\beta_2}, \tag{6.2}$$

where $\bar{u} = 1 - u$, $\overline{v} = 1 - v$, $\beta = (\beta_1, \beta_2)'$ and $\beta_1 > 0$, $\beta_2 \geq 1$. This family of copulas has the lower tail dependence given by $\lambda_L = 2^{-1/\beta_1}$ and the upper tail dependence given by $\lambda_U = 2 - 2^{1/\beta_2}$. When

[3] https://fred.stlouisfed.org/https://fred.stlouisfed.org/

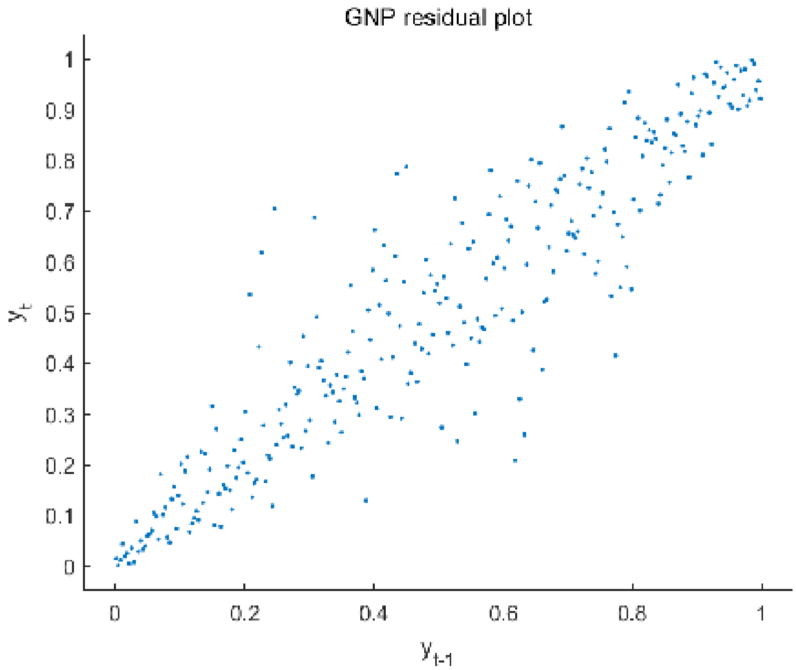

Figure 6.1: Scatter Plot of the standardized GNP residuals

$\beta_2 = 1$, the Joe-Clayton copula reduces to the Clayton copula:

$$C(u, v; \beta) = [u^{-\beta} + v^{-\beta} - 1]^{-1/\beta}, \quad \text{where} \quad \beta = \beta_1 > 0.$$

When $\beta_1 \to 0$, the Joe-Clayton copula approaches the Joe copula whose upper tail dependence increase as β_2 increases. See Joe (1997) and Patton (2006) for other properties of the Joe-Clayton copula.

The semiparametric two-step copula parameter estimates are: $\hat{\beta}_1 = 3.902$; $\hat{\beta}_2 = 2.765$. We examine tail dependence based on the copula parameter values β_1 and β_2. We first test the lower tail dependence based β_1. The estimated value of β_1 is 3.902, and the corresponding t-statistic is 5.04 (p-value $< 0.1\%$) which is significantly greater than 0, rejecting the null hypothesis of no lower tail dependence at 5% level. Next, for upper tail dependence, the estimated value of β_2 is 2.765, and the corresponding t-statistic is 5.36 (p-value $< 0.1\%$). We reject the null hypothesis $\mathrm{H}_0 : \beta_2 = 1$ at 5% level. Thus, we conclude that we find tail dependence in the short term dynamics of GNP.

6.2. An application to financial time series

The CAY time series (Lettau and Ludvigson (2001)) has been often used in macro-finance applications. Lettau and Ludvigson (2001, 2003, 2009), Chen and Ludvigson (2009) studied the role of consumption and fluctuations in the aggregate consumption–wealth ratio for predicting stock returns. They argue that investors who want to maintain a flat consumption path over time will attempt to "smooth out" transitory movements in their asset wealth arising from time variation in expected returns. When excess returns are expected to be higher in the future, forward-looking investors will react by increasing consumption out of current asset wealth and labor income, allowing consumption to rise above its common trend with those variables. When excess returns are expected to be lower in the future, these investors will react by decreasing consumption out of current asset wealth and labor income, and consumption will fall below its shared trend with these variables. In this way, investors may insulate future consumption from fluctuations in expected returns, and stationary deviations from the shared trend among consumption, asset holdings, and labor income are likely to be a predictor of excess stock returns.

We apply the copula model to capture the short term dynamics in the consumption–wealth ratio time series. Since this time series is not directly observed, Lettau and Ludvigson (2001) argue that consumption (c_t), asset holding (a_t) and labor income (y_t) are cointegrated, and that deviations from this shared trend summarize agents' expectations of future returns on the market portfolio. In particular, the residual term from a cointegrating regression of consumption (c_t) on asset holding (a_t) and labor income (y_t) is called the "CAY" time series by Lettau and Ludvigson (2001). The "CAY" time series contain important information of future returns at short horizons.

We use the dataset from the website of Martin Lettau. The time series is from 1952Q4 to 1998Q3. The unit root nonstationarity in time series c_t, a_t, y_t can be verified. In particular, the ADF t-test statistics corresponding to (c_t, a_t, y_t) are -1.233, -2.603, -0.7918, thus the unit root hypothesis can not be rejected. We then consider a cointegrating regression of consumption (c_t) on asset holding (a_t) and labor income (y_t). The Engle-Granger 2-stage cointegration test statistic is -3.93, rejecting the null hypothesis of no cointegration (the 5% level critical value is -3.81). Figure 6.2 presents the corresponding scatter plot of standardized realizations of the CAY time series. The figure indicates presence of lower tail dependence.

We again consider the Joe-Clayton copula model given by (6.2). The semiparametric two-step copula estimates are $\hat{\beta}_1 = 2.050$; $\hat{\beta}_2 = 1.356$. We test lower tail dependence based on β_1. The estimated value of this parameter is 2.05, and the corresponding t-statistic is 4.95 (p-value $< 0.1\%$). The null hypothesis of no lower tail dependence in the CAY time series is rejected at 5% level of significance and lower tail dependence is detected.

For upper tail dependence, the estimated value of β_2 is 1.356. Corresponding to the null hypothesis $\mathrm{H}_0 : \beta_2 = 1$, the t-statistic is 1.825. We reject the null at 5% level. However, the p-value corresponding to this t-statistic is 3.414%, we can not reject the null hypothesis at 1% level. Given this marginal empirical evidence for upper tail dependence, we further conduct a likelihood ratio (LR) test for

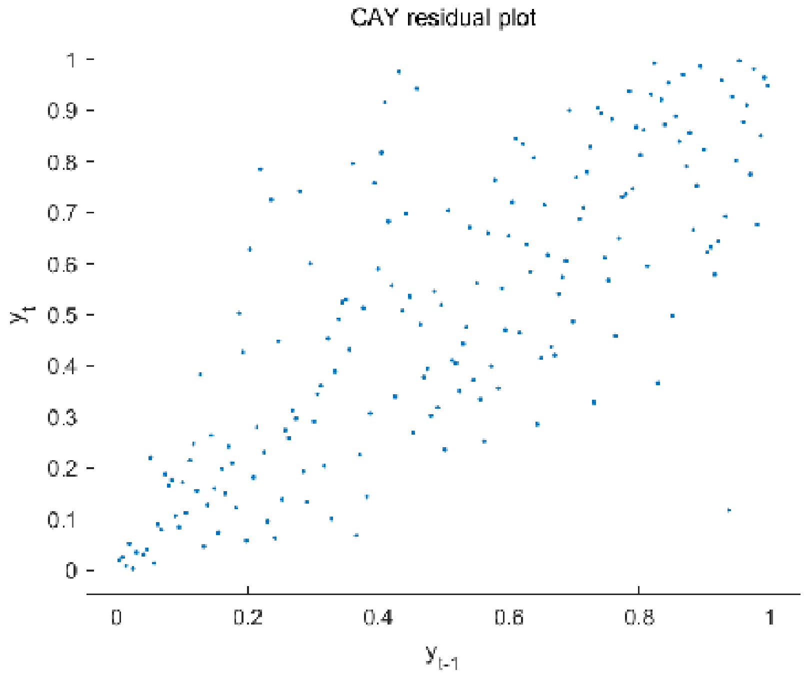

Figure 6.2: Scatter Plot of the standardized CAY residual time series

$H_0 : \beta_2 = 1$. The corresponding LR statistic equals 4, with a p-value equals 4.6%, marginally rejecting the null at 5% level, but could not reject it even at 4% level of significance. Thus, the evidence of upper tail dependence is relatively weak.

Thus, we conclude that we find significant lower tail dependence and moderate upper tail dependence in the CAY time series.

7. Conclusion

We propose a component approach to study nonstationary time series with nonlinear short term dynamics that may also exhibit tail dependence. The observed time series can be decomposed into a nonstationary part and a stationary Markov component generated via a copula. The nonstationary component can be removed by a filtration, and the copula-based Markov model is used to capture the weakly dependent nonlinear dynamics (and the tail dependence) in the filtered time series.

When the marginal distribution of the filtered time series is parametrically estimated, we show that the limiting distribution of the parametric (two-step) copula estimator can be affected by the filtration and the estimation of the marginal distribution, and may not be normal under stochastic trend filtration. However, when the marginal distribution of the filtered time series is nonparametrically estimated, we find that the limiting distribution of the semiparametric (two-step) copula estimator is not affected by the nonstationary filtration and is asymptotically normal. The surprising result for the semiparametric two-step copula estimator is also extended to models with misspecified residual copula function. Monte Carlo studies reveal that, for different kinds of nonstationarity, symmetric or asymmetric unknown marginal distributions, various copula functions with or without tail dependence, our semiparametric (two-step) copula estimator not only is robust, but also performs very similarly to the infeasible semiparametric copula estimator without filtration. The simple and robust asymptotic properties of the semiparametric copula estimators greatly simplify statistical inference on nonstationary filtered copula-based time series models. These results have many practical implications for empirical analysis of nonstationary nonlinear time series in economics and finance.

The results in this paper can be extended in many directions. First, other copula estimators, such as those in Oh and Patton (2013) and Chen, Wu and Yi (2009), can be studied. Second, notice that, given a copula function $C(u, v)$ of the latent first-order Markov process $\{Y_t\}$, differentiating $C(u, v)$ with respect to u, and evaluate at $u = F^*(x)$, $v = F^*(y)$, we obtain the conditional distribution of Y_t given $Y_{t-1} = x$. Consequently, a time series with nonlinear dynamics satisfying the specific copula can be generated based on the conditional distribution (Chen and Fan 2006a, Chen, Koenker and Xiao 2009), and thus the bootstrap approach can be studied as an alternative inference method. Finally, multivariate nonstationary filtration may be considered with the latent stationary multivariate Markov process $\{Y_t\}$ generated by contemporary and temporal copulas as in Remillard, Papageorgiou and Soustra (2012), Beare and Seo (2015) and others.

References

[1] Andrews, D. W., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica, 817-858.

[2] Beare, B., 2010, Copulas and temporal dependence, Econometrica 78, 395-410.

[3] Beare, B., and J. Seo, 2015, Vine copula specifications for stationary multivariate Markov chains. Journal of Time Series Analysis, 36(2), 228-246.

[4] Billingsley, P., 1968, Convergence of probability measures. John Wiley & Sons.

[5] Blanchard, O. J. 1981, Output, the stock market, and interest rates, The American Economic Review, 71, 132-143.

[6] Chen, X. and Y. Fan, 2006a, Estimation of Copula-Based Semiparametric Time Series Models, Journal of Econometrics 130, 307-335.

[7] Chen, X. and Y. Fan, 2006b, Estimation and Model Selection of Semiparametric Copula-based Multivariate Dynamic Models under Copula Misspecification, Journal of Econometrics 135, 125-154.

[8] Chen, X. and S. Ludvigson., 2009, Land of addicts? an empirical investigation of habit-based asset pricing models. Journal of Applied Econometrics 24, 1057-1093.

[9] Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile autoregression. The Econometrics Journal, 12, S50-S67.

[10] Chen, X., W. Wu and Y. Yi, 2009, Efficient Estimation of Copula-Based Semiparametric Markov Models. Annals of Statistics 37 (6B): 4214–53.

[11] Cherubini, U., F. Gobbi, S. Mulinacci and S. Romagnoli, 2012, Dynamic Copula Methods in Finance. Chichester: Wiley.

[12] Cochrane, J., 1994, Permanent and transitory components of GNP and stock prices. The Quarterly Journal of Economics 109, 241-265.

[13] Csorgo, M., Csorgo, S., Horváth, L., & Mason, D. M. (1986). Weighted empirical and quantile processes. The Annals of Probability, 31-85.

[14] Csörgö, M., & Horváth, L. (1993). Weighted approximations in probability and statistics. J. Wiley & Sons.

[15] Doukhan, P., 1994, Mixing: Properties and Examples. New York, Springer.

[16] Gallant, A. R., 2009, Nonlinear statistical models. John Wiley & Sons.

[17] Gallant, A. R., Rossi, P. E., & Tauchen, G., 1993, Nonlinear dynamic structures, Econometrica, 871-907.

[18] Genest, C., K. Ghoudi and L. Rivest, 1995, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82, 543–552.

[19] Granger, C.W.J., 2002, Time Series Concept for Conditional Distributions, Manuscript, UCSD.

[20] Hannan, E., 1970, Multiple Time Series, John Wiley & Sons.

[21] Ibragimov, R., 2009, Copulas-based characterizations and higher-order Markov processes. Econometric Theory 25, 819-846.

[22] Ibragimov, R. and G. Lentzas, 2017, Copulas and long memory. Probability Surveys 14, 289-327.

[23] Joe, H., 1997, Multivariate Models and Dependence Concepts, Chapman & Hall/CRC.

[24] Joe, H., 2005, Asymptotic efficiency of the two-stage estimation method for copula-based models. J. Multivariate Anal. 94, 401–419.

[25] Joe, H. and J. Xu, 1996, The Estimation Method of Inference Functions for Margins for Multivariate Models. Technical Report 166, Department of Statistics, University of British Columbia.

[26] King, R., C. Plosser, J. Stock, and M. Watson, 1991, Stochastic trends and economic fluctuations." American Economic Review 81(4).

[27] Kydland, F. and E. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica: 1345-1370.

[28] Lettau, M., and S. Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns. The Journal of Finance 56: 815-849.

[29] Lettau, M., and S. Ludvigson, 2004, Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption. American Economic Review 94: 276-299.

[30] Longla, M. and M. Peligrad, 2012, Some aspects of modeling dependence in copula-based Markov chains. Journal of Multivariate Analysis 111, 234-240.

[31] Moricz, F. 1982, A general moment inequality for the maximum of partial sums of single series. Acta Sci. Math. 44, 67-75.

[32] Oh, D., and A. Patton, 2013. Simulated Method of Moments Estimation for Copula based Multivariate Models. Journal of the American Statistical Association, 108, 689-700.

[33] Patton, A. 2006, Modelling asymmetric exchange rate dependence, International Economic Review, 47, 527-556.

[34] Patton, A. 2009, Copula-based models for financial time series, Handbook of Financial Time Series, Springer-Verlag.

[35] Patton, A. 2012, A review of copula models for economic time series, Journal of Multivariate Analysis 110, 4-18.

[36] Pollard, D., 1985, New ways to prove central limit theorems, Econometric Theory, 1, 295-313.

[37] Remillard, B., N. Papageorgiou, and F. Soustra, 2012. Copula-based Semiparametric Models for Multivariate Time Series. Journal of Multivariate Analysis, 110, 30-42.

[38] Shao, Qi-Man, and Hao Yu, 1996, Weak convergence for weighted empirical processes of dependent sequences. The Annals of Probability 24, 2098-2127.

[39] Sklar, A., 1959, Fonctions de r'epartition 'a n dimensionset leurs marges, Publ. Inst. Statis. Univ. Paris 8, 229-231.

[40] Viennet, Gabrielle, 1997, Inequalities for absolutely regular sequences: application to density estimation. Probability theory and related fields 107, 467-492.

[41] White, H., 1982, Maximum likelihood estimation of misspecified models, Econometrica 50, l-26.

[42] Yoshihara, K., 1976, Limiting behavior of U-statistics for stationary, absolutely regular processes. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 35, 237-252.

A. Appendix A: Monte Carlo Results

In the Monte Carlo studies, we have examed various DGPs that are different combinations of three kinds of filtering part $X_t'\pi^*$, four kinds of copula functions $C(\cdot,\cdot;\beta)$ with a range value of the copula parameter β, and two kinds of marginal distributions F^* of Y_t given in Section 5 of the paper. In each table below, the number of Monte Carlo repetition is 2000 and sample size is $n = 500$ (we also considered a larger sample size of $n = 2000$ in a few tables). The Monte Carlo bias, variance, and "Rmse" (the Ratio of MSE of an estimator over the MSE of $\widehat{\beta}_{P*}$) are reported in each table.

We investigate the finite sample performance of the semiparametric copula estimator $\widehat{\beta}_{SP}$, the parametric copula estimator with corrected specified parametric marginals $\widehat{\beta}_{P*}$; the parametric copula estimator with a normal distribution as the incorrectly specified distribution $\widehat{\beta}_{P1}$; the infeasible parametric estimator $\widetilde{\beta}_{P*}$ with corrected specified parametric marginals; and the infeasible semiparametric estimator $\widetilde{\beta}_{SP}$. Both $\widetilde{\beta}_{SP}$ and $\widetilde{\beta}_{P*}$ are computed using $\{Y_t\}$ directly, and are presented for comparison purpose.

Recall that $\widehat{\beta}_{SP}$ and $\widetilde{\beta}_{SP}$ have the same asymptotic normal distribution, which does not depend on any filtration and the specification of F^*. The infeasible $\widetilde{\beta}_{P*}$ is asymptotically normal, with the limiting distribution independent of the filtration but does depend on the parametric estimation of F^*. The limiting distributions of $\widehat{\beta}_{P*}$ and $\widehat{\beta}_{P1}$ depend on the filtration and the parametric estimation of F^* in complicated ways; they are normal under the deterministic trend filtration, but, are generally non-normal under the stochastic trend (the unit root and cointegration) filtration.

Table 1 and Table 2 report the finite sample performances of the estimators for models with deterministic trending time series. In particular, Tables 1A - 1D below summarize simulation results corresponding to the *deterministic trending model* (5.1) when the true marginal distribution is student-$t(3)$ distribution (symmetric dist.), with Table 1A for Gaussian copula, Table 1B for Frank copula, Table 1C for Clayton copula and Table 1D for Gumbel copula. Similarly, Tables 2A - 2D summarize results corresponding to the deterministic trending model (5.1) when the true marginal distribution is re-centered Chi-square with d.f. 3, again with "A to D" corresponding to Gaussian, Frank, Clayton and Gumbel copulas.

Tables 3 - 6 report the finite sample behaviors of the estimators for models with stochastic trends. In particular, Tables 3A - 3D correspond to the *unit root model* when the true marginal distribution is student-$t(3)$. Tables 4A - 4D summarize results for the *unit root model* when the true marginal distribution is re-centered Chi-square with d.f. 3. Tables 5A - 5D correspond to the *cointegrated model* when the true marginal distribution is student-$t(3)$. Tables 6A - 6D summarize results for the *cointegrated model* when the true marginal distribution is re-centered Chi-square with d.f. 3. Again, "A to D" correspond to Gaussian, Frank, Clayton and Gumbel copulas.

Table 1A: Trending Time Series, Gaussian Copula

(True marginal is student t(3))

$n = 500$						
β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	-0.0066	-0.0077	-0.0063	-0.0042	-0.0033	-0.0049
$\hat{\beta}_{SP}$ Std	0.0391	0.0438	0.0462	0.0465	0.0445	0.0401
$\hat{\beta}_{SP}$ Rmse	**1.1224**	**1.0912**	**1.0613**	**1.0389**	**1.0369**	**1.0588**
$\hat{\beta}_{P*}$ Bias	0.0004	-0.0014	-0.0035	-0.0056	-0.0076	-0.0094
$\hat{\beta}_{P*}$ Std	0.0374	0.0425	0.0452	0.0455	0.0431	0.0381
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.0046	-0.0151	-0.0193	0.0078	0.0048	-0.0067
$\hat{\beta}_{P1}$ Std	0.0721	0.0835	0.0911	0.0945	0.0871	0.0725
$\hat{\beta}_{P1}$ Rmse	**3.7261**	**3.9751**	**4.2273**	**4.2896**	**3.9660**	**3.4407**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**1.1069**	**1.0763**	**1.0508**	**1.0264**	**1.0181**	**1.0257**
$\tilde{\beta}_{P*}$ Bias	0.0002	-0.0007	-0.0014	-0.0022	-0.0030	-0.0037
$\tilde{\beta}_{P*}$ Std	0.0370	0.0423	0.0450	0.0452	0.0427	0.0375
$\tilde{\beta}_{P*}$ Rmse	**0.9758**	**0.9873**	**0.9889**	**0.9775**	**0.9569**	**0.9225**

$n = 500$						
$\tilde{\beta}_{SP}$MSE / $\hat{\beta}_{SP}$ MSE	**0.9862**	**0.9864**	**0.9901**	**0.9879**	**0.9819**	**0.9687**
$\tilde{\beta}_P$ MSE / $\hat{\beta}_{P*}$ MSE	0.9758	0.9873	0.9889	0.9775	0.9569	0.9225
$n = 2000$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9992**	**0.9981**	**0.9978**	**0.9983**	**0.9980**	**0.9935**
$\tilde{\beta}_P$ MSE / $\hat{\beta}_{P*}$ MSE	0.9977	0.9960	0.9958	0.9926	0.9859	0.9731

Table 1B: Trending Time Series, Frank Copula

(True marginal is student t(3))

$n = 500$						
β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	-0.0115	-0.0229	-0.0242	-0.0310	-0.0591	-0.1280
$\hat{\beta}_{SP}$ Std	0.4025	0.3230	0.2812	0.2812	0.3194	0.3925
$\hat{\beta}_{SP}$ Rmse	**1.2118**	**1.1066**	**1.0170**	**1.0207**	**1.1254**	**1.2741**
$\hat{\beta}_{P*}$ Bias	0.0393	0.0093	-0.0103	-0.0288	-0.0581	-0.1116
$\hat{\beta}_{P*}$ Std	0.3637	0.3077	0.2797	0.2785	0.3006	0.3483
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-1.5653	-1.3416	-0.8315	0.7674	1.2818	1.4765
$\hat{\beta}_{P1}$ Std	1.1554	1.1182	1.1144	1.1915	1.2066	1.2242
$\hat{\beta}_{P1}$ Rmse	**28.2919**	**32.1860**	**24.6847**	**25.6159**	**33.0572**	**27.5063**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**1.1879**	**1.0963**	**1.0075**	**1.0124**	**1.1010**	**1.1896**
$\tilde{\beta}_{P*}$ Bias	-0.0144	-0.0134	-0.0108	-0.0092	-0.0112	-0.0128
$\tilde{\beta}_{P*}$ Std	0.3489	0.3022	0.2776	0.2778	0.3003	0.3454
$\tilde{\beta}_{P*}$ Rmse	**0.9114**	**0.9658**	**0.9857**	**0.9854**	**0.9634**	**0.8935**

$n = 500$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9803**	**0.9907**	**0.9907**	**0.9919**	**0.9783**	**0.9336**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.9114	0.9658	0.9857	0.9854	0.9634	0.8935
$n = 2000$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9935**	**0.9985**	**0.9992**	**0.9993**	**0.9975**	**0.9875**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.9696	0.9887	0.9965	0.9951	0.9867	0.9615

Table 1C: Trending Time Series, Clayton Copula

(True marginal is student t(3))

$n = 500$						
β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	-0.0012	-0.0307	-0.1672	-0.7897	-1.8797	-3.2800
$\hat{\beta}_{SP}$ Std	0.1040	0.1989	0.4486	0.9392	1.2412	1.4254
$\hat{\beta}_{SP}$ Rmse	**1.3184**	**1.4836**	**1.4314**	**1.2141**	**1.7435**	**2.3035**
$\hat{\beta}_{P*}$ Bias	-0.0098	-0.0217	-0.0787	-0.3700	-0.9417	-1.6985
$\hat{\beta}_{P*}$ Std	0.0900	0.1638	0.3923	1.0504	1.4224	1.6333
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.0706	-0.0086	0.1218	0.1131	-0.2723	-0.9375
$\hat{\beta}_{P1}$ Std	0.4077	0.5114	0.6111	0.9539	1.3258	1.7819
$\hat{\beta}_{P1}$ Rmse	**20.8799**	**9.5796**	**2.4249**	**0.7439**	**0.6296**	**0.7301**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**1.2899**	**1.3534**	**1.3191**	**1.1583**	**1.5055**	**1.8639**
$\tilde{\beta}_{P*}$ Bias	-0.0026	-0.0069	-0.0171	-0.0257	-0.0240	-0.0160
$\tilde{\beta}_{P*}$ Std	0.0854	0.1343	0.2602	0.6389	1.1813	1.7828
$\tilde{\beta}_{P*}$ Rmse	**0.8896**	**0.6621**	**0.4246**	**0.3296**	**0.4797**	**0.5725**

$n = 500$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9784**	**0.9122**	**0.9215**	**0.9289**	**0.8635**	**0.8092**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.8896	0.6621	0.4246	0.3296	0.4797	0.5725
$n = 2000$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9948**	**0.9832**	**0.9577**	**0.9464**	**0.9520**	**0.9331**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.9051	0.7167	0.3915	0.2155	0.1923	0.2537

Table 1D: Trending Time Series, Gumbel Copula

(True marginal is student t(3))

$n = 500$						
β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0379	-0.1785	-0.4513	-0.8697	-1.4093	-2.0454
$\hat{\beta}_{SP}$ Std	0.1666	0.3793	0.5882	0.7423	0.8490	0.9330
$\hat{\beta}_{SP}$ Rmse	**1.0719**	**1.0647**	**1.1286**	**1.3556**	**1.7370**	**2.1476**
$\hat{\beta}_{P*}$ Bias	-0.0236	-0.0907	-0.2292	-0.4523	-0.7562	-1.1173
$\hat{\beta}_{P*}$ Std	0.1633	0.3960	0.6592	0.8717	0.9932	1.0512
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.1096	0.0663	-0.0742	-0.3122	-0.6547	-1.0985
$\hat{\beta}_{P1}$ Std	0.3842	0.5599	0.7989	1.0189	1.2148	1.4015
$\hat{\beta}_{P1}$ Rmse	**5.8626**	**1.9262**	**1.3218**	**1.1775**	**1.2220**	**1.3473**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**0.9732**	**0.8909**	**0.9349**	**1.1187**	**1.4204**	**1.7311**
$\tilde{\beta}_{P*}$ Bias	-0.0066	-0.0225	-0.0533	-0.0962	-0.1456	-0.1927
$\tilde{\beta}_{P*}$ Std	0.1264	0.2810	0.4848	0.7297	1.0384	1.4401
$\tilde{\beta}_{P*}$ Rmse	**0.5887**	**0.4815**	**0.4883**	**0.5618**	**0.7054**	**0.8971**

$n = 500$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9079**	**0.8368**	**0.8284**	**0.8252**	**0.8177**	**0.8061**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.5887	0.4815	0.4883	0.5618	0.7054	0.8971
$n = 2000$						
$\tilde{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE	**0.9330**	**0.8732**	**0.8819**	**0.8744**	**0.8589**	**0.8521**
$\tilde{\beta}_{P}$ MSE / $\hat{\beta}_{P*}$ MSE	0.6260	0.4710	0.4435	0.4376	0.4451	0.4496

Table 2A: Trending Time Series, Gaussian Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	-0.0062	-0.0074	-0.0059	-0.0037	-0.0028	-0.0046
$\hat{\beta}_{SP}$ Std	0.0387	0.0436	0.0463	0.0466	0.0447	0.0404
$\hat{\beta}_{SP}$ Rmse	**1.3211**	**1.0519**	**0.9589**	**0.9521**	**0.9309**	**0.9054**
$\hat{\beta}_{P*}$ Bias	-0.0053	-0.0078	-0.0068	-0.0006	0.0083	0.0147
$\hat{\beta}_{P*}$ Std	0.0337	0.0425	0.0472	0.0479	0.0456	0.0401
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.0897	0.0437	0.0079	-0.0181	-0.0344	-0.0414
$\hat{\beta}_{P1}$ Std	0.0302	0.0371	0.0431	0.0476	0.0496	0.0479
$\hat{\beta}_{P1}$ Rmse	**7.7163**	**1.7650**	**0.8457**	**1.1262**	**1.6895**	**2.1902**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**1.3371**	**1.0460**	**0.9483**	**0.9371**	**0.9077**	**0.8639**
$\tilde{\beta}_{P*}$ Bias	0.0044	0.0029	0.0000	-0.0036	-0.0063	-0.0074
$\tilde{\beta}_{P*}$ Std	0.0320	0.0400	0.0444	0.0446	0.0404	0.0324
$\tilde{\beta}_{P*}$ Rmse	**0.9013**	**0.8646**	**0.8679**	**0.8705**	**0.7763**	**0.6047**

Table 2B: Trending Time Series, Frank Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	-0.0297	-0.0344	-0.0297	-0.0296	-0.0440	-0.0851
$\hat{\beta}_{SP}$ Std	0.3970	0.3214	0.2809	0.2819	0.3222	0.4001
$\hat{\beta}_{SP}$ Rmse	**1.3150**	**1.0811**	**0.9519**	**0.9623**	**0.8341**	**0.6380**
$\hat{\beta}_{P*}$ Bias	-0.0425	-0.0523	-0.0433	0.0036	0.0988	0.2274
$\hat{\beta}_{P*}$ Std	0.3445	0.3065	0.2863	0.2889	0.3421	0.4589
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.4944	0.0962	0.0035	0.1712	0.3759	0.5257
$\hat{\beta}_{P1}$ Std	0.3021	0.2970	0.3018	0.3392	0.4140	0.5400
$\hat{\beta}_{P1}$ Rmse	**2.7855**	**1.0084**	**1.0861**	**1.7296**	**2.4664**	**2.1656**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**1.3188**	**1.0747**	**0.9411**	**0.9508**	**0.8140**	**0.6066**
$\tilde{\beta}_{P*}$ Bias	0.0033	-0.0013	-0.0065	-0.0132	-0.0208	-0.0255
$\tilde{\beta}_{P*}$ Std	0.3370	0.2967	0.2764	0.2762	0.2943	0.3336
$\tilde{\beta}_{P*}$ Rmse	**0.9423**	**0.9108**	**0.9114**	**0.9158**	**0.6866**	**0.4267**

Table 2C: Trending Time Series, Clayton Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	-0.0077	-0.0524	-0.2290	-0.9035	-1.9578	-3.2889
$\hat{\beta}_{SP}$ Std	0.1014	0.1830	0.4007	0.8853	1.2933	1.5443
$\hat{\beta}_{SP}$ Rmse	**0.8758**	**1.0248**	**1.2213**	**1.2928**	**1.2684**	**1.1733**
$\hat{\beta}_{P*}$ Bias	0.0022	-0.0198	-0.1264	-0.5526	-1.2366	-2.0305
$\hat{\beta}_{P*}$ Std	0.1086	0.1870	0.3981	0.9655	1.6767	2.6700
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.6251	0.7053	0.7347	0.6051	0.3685	-0.0129
$\hat{\beta}_{P1}$ Std	0.1651	0.2284	0.4478	1.1839	2.3474	3.5508
$\hat{\beta}_{P1}$ Rmse	**35.4067**	**15.5463**	**4.2438**	**1.4283**	**1.3008**	**1.1205**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**0.8959**	**1.0454**	**1.2109**	**1.1607**	**1.0093**	**0.9198**
$\tilde{\beta}_{P*}$ Bias	-0.0327	-0.0773	-0.2062	-0.6221	-1.2212	-1.9876
$\tilde{\beta}_{P*}$ Std	0.0851	0.1402	0.2823	0.6896	1.2753	1.8589
$\tilde{\beta}_{P*}$ Rmse	**0.7039**	**0.7254**	**0.7007**	**0.6969**	**0.7183**	**0.6582**

Table 2D: Trending Time Series, Gumbel Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0217	-0.1278	-0.3610	-0.7509	-1.2756	-1.9110
$\hat{\beta}_{SP}$ Std	0.1736	0.4040	0.6410	0.8087	0.9238	1.0039
$\hat{\beta}_{SP}$ Rmse	**0.9308**	**0.9498**	**1.0090**	**1.1762**	**1.6286**	**2.4850**
$\hat{\beta}_{P*}$ Bias	0.1061	0.2632	0.4169	0.5329	0.5779	0.5270
$\hat{\beta}_{P*}$ Std	0.1471	0.3461	0.6021	0.8668	1.0905	1.2639
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.1716	-0.2440	-0.4207	-0.7133	-1.1187	-1.6247
$\hat{\beta}_{P1}$ Std	0.2353	0.5360	0.8422	1.1149	1.3327	1.4940
$\hat{\beta}_{P1}$ Rmse	**2.5773**	**1.8340**	**1.6526**	**1.6922**	**1.9876**	**2.5980**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**0.8052**	**0.7776**	**0.8489**	**1.0421**	**1.4532**	**2.1726**
$\tilde{\beta}_{P*}$ Bias	-0.0091	-0.0234	-0.0334	-0.0305	-0.0072	0.0184
$\tilde{\beta}_{P*}$ Std	0.0758	0.1225	0.2694	0.5207	0.8738	1.2924
$\tilde{\beta}_{P*}$ Rmse	**0.1773**	**0.0822**	**0.1374**	**0.2628**	**0.5013**	**0.8909**

Table 3A: Unit Root Time Series, Gaussian Copula

(True marginal is student-t(3), $n = 500$)

β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	0.0032	-0.0015	-0.0022	-0.0010	-0.0005	-0.0020
$\hat{\beta}_{SP}$ Std	0.0413	0.0444	0.0464	0.0464	0.0443	0.0398
$\hat{\beta}_{SP}$ Rmse	**0.9609**	**1.0487**	**1.0587**	**1.0552**	**1.0651**	**1.0977**
$\hat{\beta}_{P*}$ Bias	0.0149	0.0072	0.0024	-0.0010	-0.0036	-0.0054
$\hat{\beta}_{P*}$ Std	0.0396	0.0428	0.0451	0.0452	0.0428	0.0376
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.0068	-0.0072	-0.0130	0.0132	0.0094	-0.0024
$\hat{\beta}_{P1}$ Std	0.0738	0.0844	0.0918	0.0945	0.0869	0.0720
$\hat{\beta}_{P1}$ Rmse	**3.0701**	**3.8195**	**4.2210**	**4.4582**	**4.1482**	**3.5967**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**0.8674**	**1.0368**	**1.0589**	**1.0549**	**1.0615**	**1.0943**

Table 3B: Unit Root Time Series, Frank Copula

(True marginal is student-t(3), $n = 500$)

β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	0.1320	0.0370	0.0026	-0.0118	-0.0312	-0.0746
$\hat{\beta}_{SP}$ Std	0.4599	0.3355	0.2831	0.2819	0.3205	0.3926
$\hat{\beta}_{SP}$ Rmse	**0.9452**	**1.0435**	**1.0200**	**1.0293**	**1.1367**	**1.3053**
$\hat{\beta}_{P*}$ Bias	0.2276	0.0858	0.0239	-0.0032	-0.0219	-0.0444
$\hat{\beta}_{P*}$ Std	0.4363	0.3190	0.2793	0.2781	0.3012	0.3469
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-1.3618	-1.2542	-0.7833	0.8126	1.3305	1.5537
$\hat{\beta}_{P1}$ Std	1.3053	1.2081	1.1563	1.1914	1.2061	1.2220
$\hat{\beta}_{P1}$ Rmse	**14.6941**	**27.7834**	**24.8172**	**26.8892**	**35.3614**	**31.9379**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**0.6563**	**0.9518**	**1.0039**	**1.0264**	**1.1317**	**1.3005**

Table 3C: Unit Root Time Series, Clayton Copula

(True marginal is student-t(3), $n = 500$)

β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	0.0029	-0.0238	-0.1400	-0.6490	-1.5641	-2.7850
$\hat{\beta}_{SP}$ Std	0.1032	0.1930	0.4410	1.0001	1.3963	1.6425
$\hat{\beta}_{SP}$ Rmse	**1.4129**	**1.7608**	**2.0309**	**1.7618**	**1.7485**	**2.1501**
$\hat{\beta}_{P*}$ Bias	-0.0044	-0.0137	-0.0504	-0.2014	-0.4862	-0.9244
$\hat{\beta}_{P*}$ Std	0.0868	0.1459	0.3207	0.8753	1.5092	2.0019
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.0623	0.0084	0.1702	0.2957	0.1473	-0.1913
$\hat{\beta}_{P1}$ Std	0.4181	0.5283	0.6247	0.9293	1.2528	1.6933
$\hat{\beta}_{P1}$ Rmse	**23.6719**	**12.9987**	**3.9770**	**1.1788**	**0.6329**	**0.5972**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**1.4013**	**1.7206**	**2.0036**	**1.7806**	**1.7425**	**2.1287**

Table 3D: Unit Root Time Series, Gumbel Copula

(True marginal is student-t(3), $n = 500$)

β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0294	-0.1470	-0.3747	-0.7229	-1.1864	-1.7400
$\hat{\beta}_{SP}$ Std	0.1641	0.3615	0.5748	0.7517	0.8779	0.9840
$\hat{\beta}_{SP}$ Rmse	**1.3930**	**1.4290**	**1.4408**	**1.3654**	**1.4689**	**1.6783**
$\hat{\beta}_{P*}$ Bias	-0.0148	-0.0569	-0.1378	-0.2572	-0.4252	-0.6287
$\hat{\beta}_{P*}$ Std	0.1404	0.3215	0.5548	0.8546	1.1411	1.4091
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.1259	0.1172	0.0386	-0.1034	-0.3119	-0.5863
$\hat{\beta}_{P1}$ Std	0.3842	0.5646	0.8089	1.0408	1.2631	1.4861
$\hat{\beta}_{P1}$ Rmse	**8.1965**	**3.1196**	**2.0069**	**1.3733**	**1.1414**	**1.0719**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**1.3284**	**1.3795**	**1.3933**	**1.3545**	**1.4927**	**1.7112**

Table 4A: Unit Root Time Series, Gaussian Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	0.0049	0.0010	-0.0003	0.0001	0.0001	-0.0017
$\hat{\beta}_{SP}$ Std	0.0421	0.0447	0.0462	0.0463	0.0442	0.0398
$\hat{\beta}_{SP}$ Rmse	**1.6123**	**1.1434**	**0.9912**	**0.9845**	**1.0668**	**1.2028**
$\hat{\beta}_{P*}$ Bias	0.0026	0.0004	0.0017	0.0027	0.0029	0.0029
$\hat{\beta}_{P*}$ Std	0.0333	0.0418	0.0463	0.0466	0.0427	0.0362
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.0989	0.0511	0.0137	-0.0133	-0.0301	-0.0372
$\hat{\beta}_{P1}$ Std	0.0309	0.0371	0.0429	0.0472	0.0493	0.0475
$\hat{\beta}_{P1}$ Rmse	**9.6256**	**2.2816**	**0.9414**	**1.1046**	**1.8186**	**2.7519**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**1.3922**	**1.1162**	**1.0032**	**0.9870**	**1.0666**	**1.1961**

Table 4B: Unit Root Time Series, Frank Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	0.1025	0.0325	0.0014	-0.0109	-0.0275	-0.0624
$\hat{\beta}_{SP}$ Std	0.4346	0.3291	0.2801	0.2815	0.3201	0.3923
$\hat{\beta}_{SP}$ Rmse	**1.5689**	**1.1906**	**0.9808**	**0.9860**	**1.0704**	**1.0627**
$\hat{\beta}_{P*}$ Bias	0.0513	-0.0012	0.0002	0.0144	0.0327	0.0735
$\hat{\beta}_{P*}$ Std	0.3528	0.3031	0.2828	0.2833	0.3088	0.3783
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.5930	0.1565	0.0413	0.2045	0.4112	0.5774
$\hat{\beta}_{P1}$ Std	0.5355	0.4057	0.3297	0.3397	0.4119	0.5258
$\hat{\beta}_{P1}$ Rmse	**5.0235**	**2.0582**	**1.3803**	**1.9540**	**3.5128**	**4.1070**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**1.2505**	**1.1307**	**0.9867**	**0.9866**	**1.0703**	**1.0714**

Table 4C: Unit Root Time Series, Clayton Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	0.0030	-0.0260	-0.1464	-0.6391	-1.4513	-2.5290
$\hat{\beta}_{SP}$ Std	0.1030	0.1901	0.4360	1.0528	1.7108	2.3180
$\hat{\beta}_{SP}$ Rmse	**1.1142**	**1.3112**	**1.4351**	**1.3368**	**1.2267**	**1.1781**
$\hat{\beta}_{P*}$ Bias	-0.0068	-0.0431	-0.1549	-0.5338	-1.1085	-1.8549
$\hat{\beta}_{P*}$ Std	0.0973	0.1619	0.3513	0.9218	1.6954	2.5592
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.6387	0.7224	0.7678	0.7159	0.6593	0.5805
$\hat{\beta}_{P1}$ Std	0.1603	0.2091	0.3837	1.0003	2.1043	3.3443
$\hat{\beta}_{P1}$ Rmse	**45.5370**	**20.1466**	**4.9984**	**1.3336**	**1.1852**	**1.1532**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**1.1108**	**1.3163**	**1.4329**	**1.2661**	**1.0677**	**1.0360**

Table 4D: Unit Root Time Series, Gumbel Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0243	-0.1264	-0.3328	-0.6624	-1.1074	-1.6450
$\hat{\beta}_{SP}$ Std	0.1645	0.3706	0.5923	0.7663	0.8860	0.9805
$\hat{\beta}_{SP}$ Rmse	**1.5436**	**1.7158**	**1.8271**	**1.8169**	**2.0074**	**2.3653**
$\hat{\beta}_{P*}$ Bias	0.0432	0.1260	0.2160	0.3035	0.3676	0.3911
$\hat{\beta}_{P*}$ Std	0.1266	0.2711	0.4538	0.6875	0.9310	1.1822
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.1573	-0.1898	-0.2874	-0.4533	-0.6804	-0.9590
$\hat{\beta}_{P1}$ Std	0.2221	0.5060	0.8124	1.1127	1.3962	1.6602
$\hat{\beta}_{P1}$ Rmse	**4.1361**	**3.2682**	**2.9395**	**2.5562**	**2.4076**	**2.3709**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**1.4798**	**1.6453**	**1.8024**	**1.9105**	**2.2092**	**2.6276**

Table 5A: Cointegrated Time Series, Gaussian Copula

(True marginal is student t(3), $n = 500$)

β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	-0.0066	-0.0074	-0.0058	-0.0034	-0.0023	-0.0037
$\hat{\beta}_{SP}$ Std	0.0388	0.0435	0.0462	0.0465	0.0444	0.0398
$\hat{\beta}_{SP}$ Rmse	**1.1386**	**1.0925**	**1.0611**	**1.0460**	**1.0519**	**1.0850**
$\hat{\beta}_{P*}$ Bias	0.0003	-0.0011	-0.0025	-0.0039	-0.0053	-0.0066
$\hat{\beta}_{P*}$ Std	0.0369	0.0422	0.0451	0.0454	0.0430	0.0378
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.0039	-0.0140	-0.0176	0.0102	0.0075	-0.0038
$\hat{\beta}_{P1}$ Std	0.0725	0.0838	0.0915	0.0945	0.0870	0.0722
$\hat{\beta}_{P1}$ Rmse	**3.8714**	**4.0452**	**4.2554**	**4.3448**	**4.0632**	**3.5518**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**1.1401**	**1.0916**	**1.0567**	**1.0350**	**1.0411**	**1.0730**

Table 5B: Cointegrated Time Series, Frank Copula

(True marginal is student t(3), $n = 500$)

β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	-0.0213	-0.262	-0.0233	-0.0257	-0.0470	-0.1018
$\hat{\beta}_{SP}$ Std	0.3981	0.3216	0.2811	0.2819	0.3196	0.3913
$\hat{\beta}_{SP}$ Rmse	**1.2980**	**1.1326**	**1.0182**	**1.0221**	**1.1355**	**1.3120**
$\hat{\beta}_{P*}$ Bias	0.0137	-0.0018	-0.0106	-0.0189	-0.0347	-0.0628
$\hat{\beta}_{P*}$ Std	0.3496	0.3032	0.2793	0.2793	0.3012	0.3473
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-1.5928	-1.3566	-0.8338	0.7883	1.3134	1.5319
$\hat{\beta}_{P1}$ Std	1.2267	1.1657	1.1345	1.1913	1.2069	1.2233
$\hat{\beta}_{P1}$ Rmse	**33.0116**	**34.7982**	**25.3703**	**26.0401**	**34.6178**	**30.8483**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**1.2980**	**1.1301**	**1.0099**	**1.0130**	**1.1229**	**1.2770**

Table 5C: Cointegrated Time Series, Clayton Copula

(True marginal is student t(3), $n = 500$)

β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	0.0004	-0.0280	-0.1519	-0.7054	-1.6939	-2.9915
$\hat{\beta}_{SP}$ Std	0.1032	0.1927	0.4434	0.9793	1.3301	1.5500
$\hat{\beta}_{SP}$ Rmse	**1.3655**	**1.6828**	**1.9211**	**1.7613**	**2.1061**	**2.6836**
$\hat{\beta}_{P*}$ Bias	-0.0063	-0.0149	-0.0498	-0.2098	-0.5225	-0.9808
$\hat{\beta}_{P*}$ Std	0.0881	0.1494	0.3344	0.8849	1.3890	1.8078
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.0647	0.0067	0.1725	0.3067	0.1600	-0.1894
$\hat{\beta}_{P1}$ Std	0.4123	0.5222	0.6256	0.9401	1.2729	1.7079
$\hat{\beta}_{P1}$ Rmse	**22.3337**	**12.1029**	**3.6831**	**1.1824**	**0.7473**	**0.6980**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**1.3561**	**1.6400**	**1.8475**	**1.7371**	**1.9892**	**2.4468**

Table 5D: Cointegrated Time Series, Gumbel Copula

(True marginal is student t(3), $n = 500$)

β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0349	-0.1676	-0.4205	-0.8015	-1.3003	-1.8937
$\hat{\beta}_{SP}$ Std	0.1627	0.3558	0.5579	0.7233	0.8493	0.9527
$\hat{\beta}_{SP}$ Rmse	**1.1636**	**1.2718**	**1.3916**	**1.6076**	**1.9301**	**2.2544**
$\hat{\beta}_{P*}$ Bias	-0.0140	-0.0559	-0.1443	-0.2866	-0.4859	-0.7285
$\hat{\beta}_{P*}$ Std	0.1537	0.3442	0.5743	0.8018	1.0068	1.2094
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.1251	0.1147	0.0301	-0.1249	-0.3561	-0.6626
$\hat{\beta}_{P1}$ Std	0.3855	0.5664	0.8119	1.0448	1.2625	1.4788
$\hat{\beta}_{P1}$ Rmse	**6.8989**	**2.7456**	**1.8822**	**1.5274**	**1.3769**	**1.3172**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**1.1129**	**1.2088**	**1.2984**	**1.4882**	**1.7713**	**2.0438**

Table 6A: Cointegrated Time Series, Gaussian Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-0.5	-0.3	-0.1	0.1	0.3	0.5
$\hat{\beta}_{SP}$ Bias	-0.0063	-0.0072	-0.0056	-0.0032	-0.0021	-0.0035
$\hat{\beta}_{SP}$ Std	0.0388	0.0436	0.0463	0.0465	0.0444	0.0399
$\hat{\beta}_{SP}$ Rmse	**1.3898**	**1.1142**	**1.0103**	**0.9926**	**0.9952**	**1.0527**
$\hat{\beta}_{P*}$ Bias	-0.0013	-0.0034	-0.0040	-0.0015	0.0033	0.0073
$\hat{\beta}_{P*}$ Std	0.0333	0.0417	0.0462	0.0468	0.0444	0.0384
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.0911	0.0453	0.0097	-0.0159	-0.0318	-0.0384
$\hat{\beta}_{P1}$ Std	0.0302	0.0371	0.0431	0.0474	0.0493	0.0475
$\hat{\beta}_{P1}$ Rmse	**8.2865**	**1.9519**	**0.9062**	**1.1415**	**1.7373**	**2.4417**
$\tilde{\beta}_{SP}$ Bias	-0.0065	-0.0071	-0.0053	-0.0027	-0.0013	-0.0024
$\tilde{\beta}_{SP}$ Std	0.0388	0.0436	0.0461	0.0463	0.0442	0.0397
$\tilde{\beta}_{SP}$ Rmse	**1.3971**	**1.1118**	**1.0040**	**0.9835**	**0.9857**	**1.0339**

Table 6B: Cointegrated Time Series, Frank Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	-5	-3	-1	1	3	5
$\hat{\beta}_{SP}$ Bias	-0.0313	-0.0325	-0.0263	-0.0252	-0.0387	-0.0773
$\hat{\beta}_{SP}$ Std	0.3968	0.3213	0.2806	0.2816	0.3201	0.3937
$\hat{\beta}_{SP}$ Rmse	**1.3420**	**1.1197**	**0.9819**	**0.9849**	**0.9466**	**0.8169**
$\hat{\beta}_{P*}$ Bias	-0.0243	-0.0303	-0.0270	-0.0015	0.0548	0.1379
$\hat{\beta}_{P*}$ Std	0.3427	0.3037	0.2831	0.2849	0.3268	0.4219
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.5008	0.1040	0.0149	0.1884	0.3985	0.5604
$\hat{\beta}_{P1}$ Std	0.3628	0.3278	0.3109	0.3385	0.4141	0.5344
$\hat{\beta}_{P1}$ Rmse	**3.2402**	**1.2697**	**1.1977**	**1.8496**	**3.0082**	**3.0429**
$\tilde{\beta}_{SP}$ Bias	-0.0330	-0.0307	-0.0232	-0.0218	-0.0362	-0.0764
$\tilde{\beta}_{SP}$ Std	0.3973	0.3209	0.2799	0.2809	0.3192	0.3915
$\tilde{\beta}_{SP}$ Rmse	**1.3463**	**1.1153**	**0.9757**	**0.9782**	**0.9400**	**0.8075**

Table 6C: Cointegrated Time Series, Clayton Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	0.5	1	2	4	6	8
$\hat{\beta}_{SP}$ Bias	-0.0034	-0.0399	-0.1888	-0.7936	-1.7964	-3.0777
$\hat{\beta}_{SP}$ Std	0.1025	0.1872	0.4119	0.9159	1.3067	1.5658
$\hat{\beta}_{SP}$ Rmse	**0.9985**	**1.1506**	**1.3626**	**1.4072**	**1.4238**	**1.4918**
$\hat{\beta}_{P*}$ Bias	-0.0091	-0.0403	-0.1571	-0.5909	-1.2861	-2.1973
$\hat{\beta}_{P*}$ Std	0.1022	0.1739	0.3550	0.8333	1.3460	1.7789
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	0.6315	0.7141	0.7526	0.6658	0.4923	0.1799
$\hat{\beta}_{P1}$ Std	0.1626	0.2150	0.3894	0.9684	1.8165	2.6612
$\hat{\beta}_{P1}$ Rmse	**40.3787**	**17.4608**	**4.7656**	**1.3233**	**1.0220**	**0.8901**
$\tilde{\beta}_{SP}$ Bias	0.0016	-0.0256	-0.1415	-0.6389	-1.5373	-2.7485
$\tilde{\beta}_{SP}$ Std	0.1028	0.1905	0.4373	1.0141	1.4205	1.6720
$\tilde{\beta}_{SP}$ Rmse	**1.0042**	**1.1603**	**1.4019**	**1.3764**	**1.2641**	**1.2949**

Table 6D: Cointegrated Time Series, Gumbel Copula

(True marginal is re-centered Chi-square with d.f. 3, $n = 500$)

β^*	2	3	4	5	6	7
$\hat{\beta}_{SP}$ Bias	-0.0264	-0.1393	-0.3687	-0.7297	-1.2112	-1.7912
$\hat{\beta}_{SP}$ Std	0.1646	0.3676	0.5754	0.7426	0.8632	0.9660
$\hat{\beta}_{SP}$ Rmse	**1.4518**	**1.5389**	**1.7695**	**2.0905**	**2.5928**	**3.3765**
$\hat{\beta}_{P*}$ Bias	0.0663	0.1697	0.2678	0.3417	0.3741	0.3457
$\hat{\beta}_{P*}$ Std	0.1214	0.2676	0.4385	0.6338	0.8445	1.0522
$\hat{\beta}_{P*}$ Rmse	**1**	**1**	**1**	**1**	**1**	**1**
$\hat{\beta}_{P1}$ Bias	-0.1548	-0.1821	-0.2766	-0.4411	-0.6698	-0.9527
$\hat{\beta}_{P1}$ Std	0.2238	0.5124	0.8083	1.0926	1.3600	1.6112
$\hat{\beta}_{P1}$ Rmse	**3.8690**	**2.9455**	**2.7646**	**2.6779**	**2.6937**	**2.8563**
$\tilde{\beta}_{SP}$ Bias	-0.0321	-0.1540	-0.3861	-0.7354	-1.1963	-1.7464
$\tilde{\beta}_{SP}$ Std	0.1596	0.3512	0.5534	0.7335	0.8846	1.0121
$\tilde{\beta}_{SP}$ Rmse	**1.3843**	**1.4646**	**1.7249**	**2.0810**	**2.5945**	**3.3213**

B. Appendix B: Proofs

B.1. The Parametric Models

We first introduce a useful inequality of absolutely regular process given by Yoshihara (1976).

Lemma A. Let $x_{t_1}, x_{t_2}, \ldots, x_{t_k}$ (with $t_1 < t_2 < \cdots < t_k$) be absolutely regular random vectors with mixing coefficients $\beta(t)$. Let $h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$ be a Borel measurable function and let there be a $\delta > 0$ such that

$$P = \max\{M_1, M_2\} < \infty$$

where

$$M_1 = \sup_{t_1, t_2, \ldots, t_k} \int |h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})|^{1+\delta} \, dF(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$$
$$M_2 = \sup_{t_1, t_2, \ldots, t_k} \int |h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})|^{1+\delta} \, dF(x_{t_1}, \ldots, x_{t_j}) dF(x_{t_{j+1}}, \ldots, x_{t_k}).$$

Then

$$\left| \int h(x_{t_1}, \ldots, x_{t_k}) dF(x_{t_1}, \ldots, x_{t_k}) - h(x_{t_1}, \ldots, x_{t_k}) dF(x_{t_1}, \ldots, x_{t_j}) dF(x_{t_{j+1}}, \ldots, x_{t_k}) \right|$$
$$\leq 4P^{\frac{1}{1+\delta}} \beta\left(t_{j+1} - t_j\right)^{\frac{\delta}{1+\delta}}$$

for all j.

B.1.1. Consistency of $\widehat{\beta}_P$

For the first step estimator, $\widehat{\alpha} = \arg\max_{\alpha \in \mathcal{A}} \sum_{t=1}^{n} \log f(\widehat{Y}_t, \alpha)$, let $q(\alpha)$ =E[$\log f(Y_t, \alpha)$], we need to verify that

$$\sup_{\alpha \in \mathcal{A}} \left| \frac{1}{n} \sum_{t=1}^{n} \log f(\widehat{Y}_t, \alpha) - q(\alpha) \right| = o_p\left(1\right),$$

where $q(\alpha) = E\left[\log f(Y_t, \alpha)\right]$. By (1) Assumption ID1(1): compactness of $\mathcal{A}$; (2) Assumption MX: weak dependence of Y_t; (3) Assumption ID1(3): $f(y, \alpha)$ is continuous in $\alpha \in \mathcal{A}$; and (4) Assumption M1(1): E[$\sup_{\alpha \in \mathcal{A}} |\log f(Y_t, \alpha)|$] $< \infty$, we can show that

$$\sup_{\alpha \in \mathcal{A}} \left| \frac{1}{n} \sum_{t=1}^{n} \log f(Y_t, \alpha) - q(\alpha) \right| = o_p\left(1\right).$$

Thus, we only need to show that

$$\sup_{\alpha \in \mathcal{A}} \left| \frac{1}{n} \sum_{t=1}^{n} \left[\log f(\widehat{Y}_t, \alpha) - \log f(Y_t, \alpha) \right] \right| = o_p\left(1\right).$$

Denote the re-standardized X_t by $\underline{X}_t$, i.e. $\underline{X}_t = n^{1/2} D_n^{-1} X_t$, and define $q_t(\eta, \alpha) = \log f(Y_t - \underline{X}_t' \eta, \alpha)$. Under Assumption M1(2), we have, for all sequences of positive numbers $\{\epsilon_n\}$ with $\epsilon_n = o(1)$,

$$\sup_{\alpha \in \mathcal{A}, \|\eta\| \leq \epsilon_n} \left| \frac{1}{n} \sum_{t=1}^{n} \left[q_t(\eta, \alpha) - q_t(0, \alpha) \right] \right| = o_p(1).$$

Thus

$$\begin{aligned} & \sup_{\alpha \in \mathcal{A}} \left| \frac{1}{n} \sum_{t=1}^{n} \left[\log f(\widehat{Y}_t, \alpha) - \log f(Y_t, \alpha) \right] \right| \\ \leq & \sup_{\alpha \in \mathcal{A}, \|\eta\| \leq \epsilon_n} \left| \frac{1}{n} \sum_{t=1}^{n} \left[q_t(\eta, \alpha) - q_t(0, \alpha) \right] \right| = o_p(1). \end{aligned}$$

Together with Assumption ID1(2), we obtain consistency of $\widehat{\alpha}$.

For the second step estimation, we need to verify that $\sup_{\beta \in \mathcal{B}} \left\| \widehat{Q}_n(\beta) - Q(\beta) \right\| = o_p(1)$, where

$$\widehat{Q}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} g(\widehat{Y}_{t-1}, \widehat{Y}_t, \widehat{\alpha}, \beta), \quad Q(\beta) = \mathrm{E}\left[g(Y_{t-1}, Y_t, \alpha^*, \beta) \right].$$

Denote

$$Q_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} g(Y_{t-1}, Y_t, \alpha^*, \beta),$$

similarly, by: (1) Assumption ID1(1): compactness of $\mathcal{B}$; (2) Assumption MX: weak dependence of Y_t; (3) Assumption ID(3): $g(\cdot)$ is continuous in β; (4) Assumption M1(1): $\mathrm{E}\left[\sup_{\beta \in \mathfrak{B}, \alpha \in \mathcal{A}_\delta} |g(Y_{t-1}, Y_t, \alpha, \beta)| \right] < \infty$, we have

$$\sup_{\beta \in \mathcal{B}} |Q_n(\beta) - Q(\beta)| = o_p(1).$$

Thus, it suffice to show that

$$\sup_{\beta \in \mathcal{B}} \left| \widehat{Q}_n(\beta) - Q_n(\beta) \right| = o_p(1).$$

Notice that $\widehat{Y}_t = Y_t - X_t'(\widehat{\pi} - \pi^*) = Y_t - n^{-1/2} \left(X_t' n^{1/2} D_n^{-1} \right) D_n (\widehat{\pi} - \pi^*)$, let

$$D_n(\widehat{\pi} - \pi^*) = \delta_n, \ \sqrt{n}(\widehat{\alpha} - \alpha^*) = \Delta_{1n},$$

then we may write

$$\widehat{Q}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} g\left(Y_{t-1} - n^{-1/2} \left(X_{t-1}' n^{1/2} D_n^{-1} \right) \delta_n, Y_t - n^{-1/2} \left(X_t' n^{1/2} D_n^{-1} \right) \delta_n, \alpha^* + n^{-1/2} \Delta_{1n}, \beta \right).$$

Recall $\underline{X}_t = n^{1/2} D_n^{-1} X_t$, we define

$$m_t(\eta, \alpha, \beta) = g\left(Y_{t-1} - \underline{X}_{t-1}' \eta, Y_t - \underline{X}_t' \eta, \alpha, \beta \right).$$

Under the Assumption M1(2) that $g(s_1, s_2, \alpha, \beta)$ is uniformly continuous in (s_1, s_2, α), uniformly over $\beta \in \mathcal{B}$, thus we can show that, for all sequences $\{\epsilon_n\}$ with $\epsilon_n = o(1)$,

$$\sup_{\beta\in\mathcal{B},\|\alpha-\alpha^*\|+\|\eta\|\leq\epsilon_n}\left|\frac{1}{n}\sum_{t=2}^{n}\left[m_t(\eta,\alpha,\beta)-m_t(0,\alpha^*,\beta)\right]\right| = o_p(1).$$

Let $\widehat{\eta} = n^{-1/2}\delta_n$, then

$$\widehat{Q}_n(\beta) - Q_n(\beta) = \frac{1}{n}\sum_{t=2}^{n}\left[m_t(\widehat{\eta},\widehat{\alpha},\beta)-m_t(0,\alpha^*,\beta)\right]$$

Notice that

$$\begin{aligned}
&\sup_{\beta\in\mathcal{B}}\left|\widehat{Q}_n(\beta)-Q_n(\beta)\right| \\
\leq\ & \sup_{\beta\in\mathcal{B},\|\alpha-\alpha^*\|+\|\eta\|\leq\epsilon_n}\left|\frac{1}{n}\sum_{t=2}^{n}\left[g(Y_{t-1}-\underline{X}'_{t-1}\eta, Y_t-\underline{X}'_t\eta,\alpha,\beta)-g(Y_{t-1},Y_t,\alpha^*,\beta)\right]\right| \\
=\ & o_p(1).
\end{aligned}$$

Thus, $\sup_{\beta\in\mathcal{B}}\left|\widehat{Q}_n(\beta)-Q_n(\beta)\right| = o_p(1)$. In addition with Assumption ID1, Theorem 1 is proved.

B.1.2. Limiting Distribution of $\widehat{\beta}_P$

Let $g\left(\widehat{Y}_{t-1},\widehat{Y}_t,\widehat{\alpha},\beta\right) = \log c(F(\widehat{Y}_{t-1},\widehat{\alpha}),F(\widehat{Y}_t,\widehat{\alpha}),\beta)$, then the likelihood function is given by

$$\widehat{Q}_n(\beta) = \frac{1}{n}\sum_{t=2}^{n} g\left(\widehat{Y}_{t-1},\widehat{Y}_t,\widehat{\alpha},\beta\right).$$

Let $\sqrt{n}(\beta-\beta^*) = \Delta_2$, and $D_n(\widehat{\pi}-\pi^*) = \delta_n$, $\sqrt{n}(\widehat{\alpha}-\alpha^*) = \Delta_{1n}$, $\sqrt{n}\left(\widehat{\beta}-\beta^*\right) = \Delta_{2n}$, then, we may re-write the criterion function $\widehat{Q}_n(\beta)$ as

$$\begin{aligned}
& V_n(\Delta_2) \\
=\ & \frac{1}{n}\sum_{t=2}^{n} g\left(Y_{t-1}-n^{-1/2}\left(X'_{t-1}n^{1/2}D_n^{-1}\right)\delta_n, Y_t-n^{-1/2}\left(X'_t n^{1/2}D_n^{-1}\right)\delta_n, \alpha^*+n^{-1/2}\Delta_{1n}, \beta^*+n^{-1/2}\Delta_2\right).
\end{aligned}$$

and $\min_\beta \widehat{Q}_n(\beta)$ is equivalent to $\min_{\Delta_2} V_n(\Delta_2)$.

The FOC corresponding to minimize $V_n(\Delta_2)$ w.r.t. Δ_2 is given by

$$\left.\frac{\partial V_n(\Delta_2)}{\partial\Delta_2}\right|_{\Delta_2=\Delta_{2n}} = 0.$$

Expanding $\left.\frac{\partial V_n(\Delta_2)}{\partial\Delta_2}\right|_{\Delta_2=\Delta_{2n}}$ around $\Delta_2 = 0$, we have

$$\begin{aligned}
0 &= \left.\frac{\partial V_n(\Delta_2)}{\partial\Delta_2}\right|_{\Delta_2=\Delta_{2n}} \\
&= \frac{1}{n}\sum_{t=2}^{n} g_\beta\left(\widehat{Y}_{t-1},\widehat{Y}_t,\widehat{\alpha},\beta^*\right) + n^{-1/2}\left[\frac{1}{n}\sum_{t=2}^{n} g_{\beta\beta}\left(\widehat{Y}_{t-1},\widehat{Y}_t,\widehat{\alpha},\beta^{\#}\right)\right]\Delta_{2n}
\end{aligned}$$

where $\beta^{\#}$ is the middle value between β^* and $\widehat{\beta}$.

Let $\widehat{H}_{n\beta} = -n^{-1}\sum_{t=2}^{n} g_{\beta\beta}\left(\widehat{Y}_{t-1}, \widehat{Y}_t, \widehat{\alpha}, \beta^{\#}\right)$, $\widehat{S}_{n\beta} = n^{-1/2}\sum_{t=2}^{n} g_{\beta}\left(\widehat{Y}_{t-1}, \widehat{Y}_t, \widehat{\alpha}, \beta^*\right)$. First, denote $\eta = (\eta_1', \eta_2', \eta_3')'$, by consistency of $\widehat{\beta}$, Assumption X, and Assumption M2, we can show that, for any sequence $\{\epsilon_n\}$ with $\epsilon_n = o(1)$,

$$\sup_{\|\eta\|\le\epsilon_n} \frac{1}{n}\sum_{t=2}^{n}\left\|g_{\beta\beta}\left(Y_{t-1} + \underline{X}_{t-1}'\eta_1, Y_t + \underline{X}_t'\eta_1, \alpha^* + \eta_2, \beta^* + \eta_3\right) - g_{\beta\beta}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\right\| = o_p(1)$$

$$\sup_{\|\eta\|\le\epsilon_n} \frac{1}{n}\sum_{t=2}^{n}\left\|g_{\beta\alpha}\left(Y_{t-1} + \underline{X}_{t-1}'\eta_1, Y_t + \underline{X}_t'\eta_1, \alpha^* + \eta_2, \beta^* + \eta_3\right) - g_{\beta\beta}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\right\| = o_p(1)$$

$$\begin{aligned}\sup_{\|\eta\|\le\epsilon_n} \frac{1}{n}\sum_{t=2}^{n}\left\|g_{\beta j}\left(Y_{t-1} + \underline{X}_{t-1}'\eta_1, Y_t + \underline{X}_t'\eta_1, \alpha^* + \eta_2, \beta^* + \eta_3\right) - g_{\beta\beta}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\right\| &= o_p(1), \\ j &= 1, 2\end{aligned}$$

we have

$$\widehat{H}_{n\beta} = H_{n\beta} + o_p(1),$$

where

$$H_{n\beta} = -\frac{1}{n}\sum_{t=2}^{n} g_{\beta\beta}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right).$$

Denote

$$S_{n\beta} = \frac{1}{\sqrt{n}}\sum_{t=2}^{n} g_{\beta}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right),$$

and expanding $g_\beta\left(\widehat{Y}_{t-1}, \widehat{Y}_t, \widehat{\alpha}, \beta^*\right)$ around (Y_{t-1}, Y_t, α^*), Using a similar argument as the previous term, we can show that

$$\begin{aligned}\widehat{S}_{n\beta} &= S_{n\beta} + n^{-1}\sum_{t=2}^{n} g_{\beta 1}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right) X_{t-1}' n^{1/2} D_n^{-1}\delta_n \\ &\quad + n^{-1}\sum_{t=2}^{n} g_{\beta 2}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\left(X_t' n^{1/2} D_n^{-1}\right)\delta_n + n^{-1}\sum_{t=2}^{n} g_{\beta\alpha}\left(Y_{t-1}, Y_t, \alpha^*, \beta^*\right)\Delta_{1n} + o_p(1)\end{aligned}$$

Thus,

$$\begin{aligned}&\sqrt{n}\left(\widehat{\beta} - \beta^*\right) \\ = \; & H_{n\beta}^{-1} S_{n\beta} - H_{n\beta}^{-1}\left(P_{n1} + P_{n2}\right) D_n\left(\widehat{\pi} - \pi^*\right) + H_{n\beta}^{-1} P_{n3}\sqrt{n}\left(\widehat{\alpha} - \alpha^*\right) + o_p(1) \\ = \; & H_{\beta}^{-1} N\left(0, \Omega_\beta\right) - H_{\beta}^{-1}\left(P_1 + P_2\right) D_n\left(\widehat{\pi} - \pi^*\right) + H_{\beta}^{-1} P_3\sqrt{n}\left(\widehat{\alpha} - \alpha^*\right) + o_p(1) \\ = \; & H_{\beta}^{-1} N\left(0, \Omega_\beta\right) - H_{\beta}^{-1}\left(P_1 + P_2 + P_3\Omega_\alpha^{-1} H_{\alpha Y}\right) D_n\left(\widehat{\pi} - \pi^*\right) + H_{\beta}^{-1} P_3\sqrt{n}\left(\widetilde{\alpha} - \alpha^*\right) + o_p(1)\end{aligned}$$

Notice that $\sqrt{n}\left(\widetilde{\alpha} - \alpha^*\right) = H_{n\alpha}^{-1} S_{n\alpha} + o_p(1)$, where

$$H_{n\alpha} = -\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial\alpha\partial\alpha'};\; S_{n\alpha} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \log f(Y_t, \alpha^*)}{\partial\alpha},$$

thus,

$$\sqrt{n}\left(\widehat{\beta}-\beta^*\right)=H_{n\beta}^{-1}\left[S_{n\beta}+P_{n3}H_{n\alpha}^{-1}S_{n\alpha}\right]-H_{\beta}^{-1}\left(P_1+P_2+P_3\Omega_{\alpha}^{-1}H_{\alpha Y}\right)D_n\left(\widehat{\pi}-\pi^*\right)+o_p(1).$$

B.2. The Semiparametric Copula Model

We use ζ and $\eta\in(0,1)$ to signify generic constants whose value may vary throughout the paper.

Recall that we denote the true values of F and β by F^* and β^*. We first restate the important Lemma 1 from the main text. Consider $b=(b_1,\cdots,b_n)'$, let

$$Z_n(y,b)=\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y+n^{-1/2}b_t\right)-F^*(y+n^{-1/2}b_t)\right]$$

and denote $|b|=\max_t|b_t|$.

Lemma 1. Under Assumptions DGP, MX, SP, and X, for any given $B>0$,

$$\sup_{|b|\leq B}\sup_{y}\left|\frac{Z_n(y,b)-Z_n(y,0)}{w(F^*(y))}\right|=o_p(1),$$

Proof of Lemma 1.

Following the argument of Csörgö, Csörgö, Horvath and Mason (1986), Csörgö and Horvath (1993), Shao and Yu (1996), we only need to show that, for any $\epsilon>0$,

$$\lim_{L\to\infty}\limsup_{n\to\infty}\Pr\left[\sup_{y\leq -L}\left|\frac{Z_n(y,b)-Z_n(y,0)}{w(F^*(y))}\right|\geq\epsilon\right]=0, \tag{B.1}$$

and

$$\lim_{L\to\infty}\limsup_{n\to\infty}\Pr\left[\sup_{y\geq L}\left|\frac{Z_n(y,b)-Z_n(y,0)}{w(F^*(y))}\right|\geq\epsilon\right]=0. \tag{B.2}$$

We show (B.1), (B.2) can be proved in the same way. For a large L, partition $(-\infty,-L]$ into $\cup_{j=1}^{\infty}(y_j,y_{j-1}]$, with $F^*(y_j)=2^{-j}\delta$, where $\delta=\delta_L=F^*(-L)$, then

$$\Pr\left[\sup_{y\leq -L}\left|\frac{Z_n(y,b)-Z_n(y,0)}{w(F^*(y))}\right|\geq\epsilon\right]\leq\sum_{j=1}^{\infty}\Pr\left[\sup_{y_j<y\leq y_{j-1}}\left|\frac{Z_n(y,b)-Z_n(y,0)}{w(2^{-j}\delta)}\right|\geq\epsilon\right].$$

Thus, we need to show that

$$\lim_{L\to\infty}\limsup_{n\to\infty}\sum_{j=1}^{\infty}\Pr\left[\sup_{y_j<y\leq y_{j-1}}|Z_n(y,b)-Z_n(y,0)|\geq\epsilon w(2^{-j}\delta)\right]=0.$$

By monotonicity of the indicator function and the distribution function, we have

$$
\begin{aligned}
& \sup_{y_j<y\leq y_{j-1}} |Z_n(y,b)-Z_n(y,0)| \\
\leq\ & |Z_n(y_j,b)-Z_n(y_j,0)|+|Z_n(y_{j-1},b)-Z_n(y_{j-1},0)| \\
& +\sup_{y_j<y\leq y_{j-1}} |Z_n(y_{j-1},0)-Z_n(y,0)|+\sup_{y_j<y\leq y_{j-1}} |Z_n(y_j,0)-Z_n(y,0)| \\
& +\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[F^*(y_{j-1}+n^{-1/2}b_t)-F^*(y_j+n^{-1/2}b_t)\right] \\
& +\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[F(y_{j-1})-F(y_j)\right]
\end{aligned}
$$

Notice that $F^*(y_j)=2^{-j}\delta$, and, under Assumption SP, for large enough n,

$$
\begin{aligned}
& \Pr\left[\sup_{y_j<y\leq y_{j-1}} |Z_n(y,b)-Z_n(y,0)|\geq \epsilon w(2^{-j}\delta)\right] \\
\leq\ & \Pr\{|Z_n(y_j,b)-Z_n(y_j,0)|+|Z_n(y_{j-1},b)-Z_n(y_{j-1},0)| \\
& +\sup_{y_j<y\leq y_{j-1}} |Z_n(y_{j-1},0)-Z_n(y,0)|+\sup_{y_j<y\leq y_{j-1}} |Z_n(y_j,0)-Z_n(y,0)| \\
& +C^*\sqrt{n}2^{-j}\delta\geq \epsilon w(2^{-j}\delta)\Big\}.
\end{aligned}
$$

We first consider the case when $n^{1/2}2^{-j}\delta C^*\leq \epsilon w(2^{-j}\delta)/2$, $C^*=8$. Let

$$
S_1=\left\{j: n^{1/2}2^{-j}\delta C\leq \epsilon w(2^{-j}\delta)/2\right\},
$$

if $j\in S_1$,

$$
\begin{aligned}
& \Pr\left[\sup_{y_j<y\leq y_{j-1}} |Z_n(y,b)-Z_n(y,0)|\geq \epsilon w(2^{-j}\delta)\right] \\
\leq\ & \Pr\left[|Z_n(y_j,b)-Z_n(y_j,0)|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right] \\
& +\Pr\left[|Z_n(y_{j-1},b)-Z_n(y_{j-1},0)|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right] \\
& +\Pr\left[\sup_{y_j<y\leq y_{j-1}}\left|\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y_j\right)-F(y_j)-1\left(Y_t\leq y\right)+F(y)\right]\right|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right] \\
& +\Pr\left[\sup_{y_j<y\leq y_{j-1}}\left|\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y_{j-1}\right)-F(y_{j-1})-1\left(Y_t\leq y\right)+F(y)\right]\right|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right]
\end{aligned}
$$

We consider each of these terms. In particular, we show that

$$
\lim_{L\to\infty}\limsup_{n\to\infty}\sum_{j\in S_1}\Pr\left[|Z_n(y_j,b)-Z_n(y_j,0)|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right] = 0, \tag{B.3}
$$

$$
\lim_{L\to\infty}\limsup_{n\to\infty}\sum_{j\in S_1}\Pr\left[\sup_{y_j<y\leq y_{j-1}}|Z_n(y_j,0)-Z_n(y,0)|\geq \frac{\epsilon w(2^{-j}\delta)}{8}\right] = 0. \tag{B.4}
$$

and analysis of the other two terms are similar.

For the first term (B.3), by Chebyshev inequality,

$$\Pr\left[|Z_n(y_j,b)-Z_n(y_j,0)|\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]\leq\frac{2^6\mathrm{E}\,|Z_n(y_j,b)-Z_n(y_j,0)|^2}{\epsilon^2 w(2^{-j}\delta)^2}.$$

Under weak dependence of Y_t, by definition of y_j, Assumption SP, and by the inequality of Yoshihara (1976), we have:

$$\mathrm{E}\,|Z_n(y_j,b)-Z_n(y_j,0)|^2\leq\zeta\left|2^{-j+1}\delta\right|^{1/q},$$

for $\zeta>0$, $q>1$. Thus, for $1/(2q)>\mu$,

$$\sum_{j\in S_1}\Pr\left[|Z_n(y_j,b)-Z_n(y_j,0)|\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]\leq\frac{\zeta}{\epsilon^2}\left[\sum_{j=1}^{\infty}2^{-j(1/q-2\mu)}\right]\delta^{1/q-2\mu}\to 0,\text{ as }\delta\to 0.$$

Thus, under our assumptions,

$$\lim_{L\to\infty}\limsup_{n\to\infty}\sum_{j\in S_1}\Pr\left[|Z_n(y_j,b)-Z_n(y_j,0)|\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]=0$$

For the second term (B.4), using Billingsley (1968, eq.(22.17)),

$$\begin{aligned}&\Pr\left[\sup_{y_j<y\leq y_{j-1}}\left|\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y_j\right)-F(y_j)-1\left(Y_t\leq y\right)+F(y)\right]\right|\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]\\ \leq\;&\Pr\left[\left|\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y_j\right)-F(y_j)-1\left(Y_t\leq y_{j-1}\right)+F(y_{j-1})\right]\right|+\sqrt{n}2^{-j}\delta\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]\end{aligned}$$

Notice that $n^{1/2}2^{-j}\delta\leq\epsilon w(2^{-j}\delta)/16$, using (1) weak dependence of Y_t, (2) the Cauchy-Schwarz inequality, and (3) Yoshihara (1976), we have

$$\begin{aligned}&\Pr\left[\sup_{y_j<y\leq y_{j-1}}\left|\frac{1}{\sqrt{n+1}}\sum_{t=1}^{n}\left[1\left(Y_t\leq y_j\right)-F(y_j)-1\left(Y_t\leq y\right)+F(y)\right]\right|\geq\frac{\epsilon w(2^{-j}\delta)}{8}\right]\\ \leq\;&\frac{\zeta\left[2^{-j}\delta\right]^{1/q}}{\left[\epsilon w(2^{-j}\delta)\right]^2},\end{aligned}$$

and (B.4) can be proved by a similar argument as the proof of (B.3).

Next we consider the case $n^{1/2}2^{-j}\delta\zeta^*\geq\epsilon w(2^{-j}\delta)/2$. Let

$$S_2=\left\{j:n^{1/2}2^{-j}\delta\zeta^*\geq\epsilon w(2^{-j}\delta)/2\right\},$$

and

$$\Delta_{n,j}=\frac{1}{8n^{1/2}}\epsilon w(2^{-j}\delta),$$

we divide the interval $(-\infty, y_{j-1}]$ into $\cup_i (y_{j,i}, y_{j,i+1}]$, where $F(y_{j,i}) = i\Delta_{n,j}$, $0 \le i \le F(y_{j-1})/\Delta_{n,j} = 2^{-j+1}\delta/\Delta_{n,j}$, then

$$\begin{aligned} & \Pr\left[\sup_{y_j < y \le y_{j-1}} |Z_n(y,b) - Z_n(y,0)| \ge \epsilon w(2^{-j}\delta)\right] \\ \le\ & \Pr\left[\max_{0 \le i \le F(y_{j-1})/\Delta_{n,j}} \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y,b) - Z_n(y,0)| \ge \epsilon w(2^{-j}\delta)\right]. \end{aligned}$$

Notice that

$$\begin{aligned} & \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y,b) - Z_n(y,0)| \\ \le\ & |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)| + |Z_n(y_{j,i+1},b) - Z_n(y_{j,i+1},0)| \\ & + \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i},0) - Z_n(y,0)| + \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i+1},0) - Z_n(y,0)| \\ & + \frac{1}{\sqrt{n+1}} \sum_{t=1}^{n} \left[F^*(y_{j,i+1} + n^{-1/2}b_t) - F^*(y_{j,i} + n^{-1/2}b_t)\right] \\ & + \frac{1}{\sqrt{n+1}} \sum_{t=1}^{n} \left[F(y_{j,i+1}) - F(y_{j,i})\right], \end{aligned}$$

by definition $F(y_{j,i}) = i\Delta_{n,j}$, under Assumption SP, for large n,

$$\begin{aligned} & \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y,b) - Z_n(y,0)| \\ \le\ & |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)| + |Z_n(y_{j,i+1},b) - Z_n(y_{j,i+1},0)| \\ & + \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i},0) - Z_n(y,0)| + \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i+1},0) - Z_n(y,0)| \\ & + \frac{1}{4}\epsilon w(2^{-j}\delta) \end{aligned}$$

and thus

$$\begin{aligned} & \Pr\left[\sup_{y_j < y \le y_{j-1}} |Z_n(y,b) - Z_n(y,0)| \ge \epsilon w(2^{-j}\delta)\right] \\ \le\ & \Pr\left[\max_{0 \le i \le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)| \ge \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\ & + \Pr\left[\max_{0 \le i \le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},b) - Z_n(y_{j,i+1},0)| \ge \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\ & + \Pr\left[\max_{0 \le i \le F(y_{j-1})/\Delta_{n,j}} \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i},0) - Z_n(y,0)| \ge \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\ & + \Pr\left[\max_{0 \le i \le F(y_{j-1})/\Delta_{n,j}} \sup_{y_{j,i} < y \le y_{j,i+1}} |Z_n(y_{j,i+1},0) - Z_n(y,0)| \ge \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \end{aligned}$$

By Billingsley (1968, eq.(22.17)) again,

$$\sup_{y_{j,i}<y\leq y_{j,i+1}} |Z_n(y_{j,i},0) - Z_n(y,0)| \leq |Z_n(y_{j,i+1},0) - Z_n(y_{j,i},0)| + \frac{1}{8}\epsilon w(2^{-j}\delta),$$

thus

$$\begin{aligned}
& \Pr\left[\sup_{y_j<y\leq y_{j-1}} |Z_n(y,b) - Z_n(y,0)| \geq \epsilon w(2^{-j}\delta)\right] \\
\leq\ & \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)| \geq \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\
& + \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},b) - Z_n(y_{j,i+1},0)| \geq \frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\
& + \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},0) - Z_n(y_{j,i},0)| \geq \frac{\epsilon w(2^{-j}\delta)}{16}\right] \\
& + \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},0) - Z_n(y_{j,i},0)| \geq \frac{\epsilon w(2^{-j}\delta)}{16}\right]
\end{aligned}$$

We next show that

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sum_{j\in S_2} \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)| \geq \frac{3\epsilon w(2^{-j}\delta)}{16}\right] = 0$$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sum_{j\in S_2} \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},b) - Z_n(y_{j,i+1},0)| \geq \frac{3\epsilon w(2^{-j}\delta)}{16}\right] = 0$$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sum_{j\in S_2} \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},0) - Z_n(y_{j,i},0)| \geq \frac{\epsilon w(2^{-j}\delta)}{16}\right] = 0$$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sum_{j\in S_2} \Pr\left[\max_{0\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},0) - Z_n(y_{j,i},0)| \geq \frac{\epsilon w(2^{-j}\delta)}{16}\right] = 0$$

We use the maximum inequality of Moricz (1982) to bound

$$\mathrm{E} \max_{1\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b) - Z_n(y_{j,i},0)|^p,$$

and $\mathrm{E}\max_{1\leq i\leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},0)|^p$. First,

$$\mathrm{E}\,|Z_n(y_{j,k},b) - Z_n(y_{j,k},0) - Z_n(y_{j,i},b) - Z_n(y_{j,i},0)|^2 \leq \zeta(k-i)\Delta_{n,j}.$$

Next, by Viennet (1997), we obtain a Rosenthal-type inequality for

$$\mathrm{E}\,|Z_n(y_{j,k},b) - Z_n(y_{j,k},0) - Z_n(y_{j,i},b) - Z_n(y_{j,i},0)|^p.$$

For $0 \leq i < k \leq 2^{-j+1}\delta/\Delta_{n,j}$, let

$$\begin{aligned}
& \psi_t(j,k,i) \\
=\ & 1\left(Y_t \leq y_{j,k} + n^{-1/2}b_t\right) - 1\left(Y_t \leq y_{j,k}\right) + F^*(y_{j,k}) - F^*(y_{j,k} + n^{-1/2}b_t) \\
& -1\left(Y_t \leq y_{j,i} + n^{-1/2}b_t\right) + 1\left(Y_t \leq y_{j,i}\right) - F^*(y_{j,i}) + F^*(y_{j,i} + n^{-1/2}b_t).
\end{aligned}$$

Notice that $\psi_t(j,k,i)$ is a bounded function, by Theorem 2 of Viennet (1997), and application of Moricz (1982), we have

$$\begin{aligned} & \mathrm{E}\left[\max_{1\le i\le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b)-Z_n(y_{j,i},0)|\right]^p \\ \le\ & \zeta_3\left(2^{-j}\delta\right)^{p_1}+\zeta_4 n^{-p_2/2}2^{-j}\delta\log^p(2^{-j+2}\delta/\Delta_{n,j}). \end{aligned}$$

where $p_1=p/2$, $p_2=p-2$, and thus

$$\begin{aligned} & \Pr\left[\max_{0\le i\le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b)-Z_n(y_{j,i},0)|\ge\frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\ \le\ & \frac{\zeta_3\left(2^{-j}\delta\right)^{p_1}+\zeta_4 n^{-p_2/2}2^{-j}\delta\log^p(2^{-j+2}\delta/\Delta_{n,j})}{[\epsilon w(2^{-j}\delta)]^p}. \end{aligned}$$

Notice that $\Delta_{n,j}=2^{-3}n^{-1/2}\epsilon w(2^{-j}\delta)$, and $n^{1/2}2^{-j}\delta\zeta^*\ge\epsilon w(2^{-j}\delta)/2$,

$$\begin{aligned} & \Pr\left[\max_{0\le i\le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b)-Z_n(y_{j,i},0)|\ge\frac{3\epsilon w(2^{-j}\delta)}{16}\right] \\ \le\ & \zeta\left[\epsilon w(2^{-j}\delta)/8\right]^{-p}\left[\left(2^{-j}\delta\right)^{p_1}+\left(\epsilon w(2^{-j}\delta)\right)^{-p_2}\left(\delta 2^{-j}\right)^{(1+p_2)}\log^p(\frac{n^{1/2}\cdot 2^{-j+5}\delta}{\epsilon w(2^{-j}\delta)})\right] \end{aligned}$$

Under Assumption SP, we have

$$\lim_{L\to\infty}\limsup_{n\to\infty}\sum_{j\in S_2}\Pr\left[\max_{0\le i\le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i},b)-Z_n(y_{j,i},0)|\ge\frac{3\epsilon w(2^{-j}\delta)}{16}\right]=0.$$

Notice that,

$$\begin{aligned} & \Pr\left[\max_{0\le i\le F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1},0)-Z_n(y_{j,i},0)|\ge\frac{\epsilon w(2^{-j}\delta)}{16}\right] \\ \le\ & \zeta\frac{\mathrm{E}\max_{1\le i\le F(y_{j-1})/\Delta_{n,j}}|Z_n(y_{j,i},0)|^p}{[\epsilon w(2^{-j}\delta)]^p}. \end{aligned}$$

The analysis of other terms are similar. ■

B.2.1. Theorem 3.

Notice that

$$\sqrt{n+1}\left(\widehat{F}_n(y)-F^*(y)\right)=\sqrt{n+1}\left(\widehat{F}_n(y)-F_n(y)\right)+\sqrt{n+1}\left(F_n(y)-F^*(y)\right)$$

The first term, $\sqrt{n+1}\left(\widehat{F}_n(y)-F_n(y)\right)$, captures the preliminary filtering effect, and the second term, $\sqrt{n+1}\left(F_n(y)-F^*(y)\right)$, captures the effect of marginal estimation.

Let $Y_t(\gamma) = Y_t - n^{-1/2}\left(X_t' D_n^{-1} n^{1/2}\right)\gamma$, and

$$F_{n,\gamma}(y) = \frac{1}{n+1}\sum_{t=1}^{n} 1\left(Y_t(\gamma) \le y\right),$$

By Lemma 1 and differentiability (and a Taylor expansion) of F^*, we have that, for γ in an arbitrary compact set Γ of R^k,

$$\sup_{\gamma\in\Gamma}\sup_{y}\left|\left\{\sqrt{n+1}\left(F_{n,\gamma}(y) - F_n(y)\right) - f(y)\left[\frac{1}{n}\sum_{t=1}^{n} X_t' D_n^{-1} n^{1/2}\right]\gamma\right\}\Big/ w(F^*(y))\right| = o_p(1). \tag{B.5}$$

Notice that $\widehat{\gamma} = D_n\left(\widehat{\pi} - \pi^*\right)$, then $\widehat{F}_n(y)$ can be written as

$$\widehat{F}_n(y) = F_{n,\widehat{\gamma}}(y) = \frac{1}{n+1}\sum_{t=1}^{n} 1\left(Y_t(\widehat{\gamma}) \le y\right).$$

By (B.5), we have

$$\sup_{y}\left|\left\{\sqrt{n+1}\left(\widehat{F}_n(y) - F_n(y)\right) - f(y)\left[\frac{1}{n}\sum_{t=1}^{n} X_t' D_n^{-1} n^{1/2}\right] D_n\left(\widehat{\pi} - \pi^*\right)\right\}\Big/ w(F^*(y))\right| = o_p(1). \tag{B.6}$$

Let

$$s(F,\beta) = \mathrm{E}\left[\frac{\partial \log c(F(Y_{t-1}), F(Y_t), \beta)}{\partial \beta}\right],$$

Under our assumptions, the consistency of $\widetilde{\beta}$ can be obtained if

$$\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \beta)}{\partial \beta'} - s(F^*,\beta)\right\| = o_p\left(1\right)$$

By triangular inequality,

$$\begin{aligned}
&\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \beta)}{\partial \beta'} - s(F^*,\beta)\right\| \\
\le\ &\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\left[\frac{\partial \log c(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \beta)}{\partial \beta'} - \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta'}\right]\right\| \\
&+\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta'} - s(F^*,\beta)\right\|.
\end{aligned}$$

By Chen and Fan (2006a),

$$\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta'} - s(F^*,\beta)\right\| = o_p\left(1\right).$$

Next we verify that

$$\sup_{\beta\in\Theta}\left\|\frac{1}{n}\sum_{t=2}^{n}\left[\frac{\partial\log c(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta)}{\partial\beta'}-\frac{\partial\log c(F^*(Y_{t-1}),F^*(Y_t),\beta)}{\partial\beta'}\right]\right\|=o_p(1)$$

Note that

$$\begin{aligned}
&\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\left[\frac{\partial\log c(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta)}{\partial\beta'}-\frac{\partial\log c(F^*(Y_{t-1}),F^*(Y_t),\beta)}{\partial\beta'}\right]\right\| \\
\leq\ &\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta1}(F_{t-1}^{\eta},F_t^{\eta},\beta)\left(\widehat{F}_n(\widehat{Y}_{t-1})-F_n(Y_{t-1})\right)\right\| \\
&+\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta2}(F_{t-1}^{\eta},F_t^{\eta},\beta)\left(\widehat{F}_n(\widehat{Y}_t)-F_n(Y_t)\right)\right\| \\
&+\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta1}(F_{t-1}^{\eta},F_t^{\eta},\beta)\left(F_n(Y_{t-1})-F^*(Y_{t-1})\right)\right\| \\
&+\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta2}(F_{t-1}^{\eta},F_t^{\eta},\beta)\left(F_n(Y_t)-F^*(Y_t)\right)\right\|
\end{aligned}$$

where $F_s^{\eta}=\eta\widehat{F}_n(\widehat{Y}_s)+(1-\eta)F^*(Y_s)$, $s=t-1$ or t, $\eta\in(0,1)$.

We can show that the third and fourth terms are $o_p(1)$ using a similar argument as Chen and Fan (2006a). We next show that the first two terms are $o_p(1)$. Notice that

$$\begin{aligned}
&\sup_{\beta\in\mathfrak{B}}\left\|\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta2}\left(F_{t-1}^{\eta},F_t^{\eta},\beta\right)\left[\widehat{F}_n(\widehat{Y}_t)-F_n(Y_t)\right]\right\| \\
\leq\ &\frac{1}{n}\sum_{t=2}^{n}\sup_{\beta\in\mathfrak{B},F\in\mathcal{F}_\delta}\left|\ell_{\beta2}\left(F(Y_{t-1}),F(Y_t),\beta\right)w(F^*(Y_t))\right|\sup_t\left|\frac{\widehat{F}_n(\widehat{Y}_t)-F_n(Y_t)}{w(F^*(Y_t))}\right|
\end{aligned}$$

By (B.6), we have

$$\sup_t\left|\frac{\widehat{F}_n(\widehat{Y}_t)-F_n(Y_t)}{w(F^*(Y_t))}\right|=O_p\left(n^{-1/2}\right),$$

together with Assumption M4, we obtain

$$\sup_{\beta\in\Theta}\left\|\frac{1}{n}\sum_{t=2}^{n}\left[\frac{\partial\log c(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta)}{\partial\beta'}-\frac{\partial\log c(F^*(Y_{t-1}),F^*(Y_t),\beta)}{\partial\beta'}\right]\right\|=o_p(1).$$

B.2.2. Theorem 4.

A Taylor expansion of $\ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_{SP}\right)$ w.r.t β around β^* gives

$$\begin{aligned} 0 &= \frac{1}{n}\sum_{t=2}^{n}\ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_{SP}\right) \\ &= \frac{1}{n}\sum_{t=2}^{n}\ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta^*\right)+\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta\beta}\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\mathring{\beta}\right)\left(\widehat{\beta}_{SP}-\beta^*\right), \end{aligned}$$

where $\mathring{\beta}$ is a middle value between $\widehat{\beta}_{SP}$ and β^*, and $\widehat{\beta}_{SP}$ is a consistent estimator of β^*.

Expanding $\ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta^*\right)$ around $(F^*(Y_{t-1}),F^*(Y_t))$, we have

$$\begin{aligned} &\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\ell_\beta\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\beta^*\right) \\ =\ &\frac{1}{\sqrt{n}}\sum_{t=2}^{n}\ell_\beta\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right) \\ &+\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta 1}\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1})-F^*(Y_{t-1})\right) \\ &+\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta 2}\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t)-F^*(Y_t)\right) \\ &+\frac{1}{n^{3/2}}\sum_{i,j=1}^{2}\sum_{t=2}^{n}\ell_{\beta ij}\left(F^\eta_{t-1},F^\eta_t,\beta^*\right)\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+i-2})-F^*(Y_{t+i-2})\right)\right]\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+j-2})-F^*(Y_{t+j-2})\right)\right] \end{aligned}$$

where $F^\eta_s=\eta\widehat{F}_n(\widehat{Y}_s)+(1-\eta)F^*(Y_s)$, $\eta\in(0,1)$.

First, for $i=1,2,\ j=1,2,$

$$\frac{1}{n^{3/2}}\sum_{t=2}^{n}\ell_{\beta ij}\left(F^\eta_{t-1},F^\eta_t,\beta^*\right)\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+i-2})-F^*(Y_{t+i-2})\right)\right]\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+j-2})-F^*(Y_{t+j-2})\right)\right]=o_p(1).$$

Consider, for example, the case $i=1,\ j=2,$

$$\begin{aligned} &\left|\frac{1}{n^{3/2}}\sum_{t=2}^{n}\ell_{\beta 12}\left(F^\eta_{t-1},F^\eta_t,\beta^*\right)\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1})-F^*(Y_{t-1})\right)\right]\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t)-F^*(Y_t)\right)\right]\right| \\ \leq\ &\frac{1}{n^{3/2}}\sum_{t=2}^{n}\sup_{\|\beta-\beta^*\|\leq\delta,F\in\mathcal{F}_\delta}\left|\ell_{\beta 12}\left(F(Y_{t-1}),F(Y_t),\beta^*\right)w(F^*(Y_{t-1}))w(F^*(Y_t))\right| \\ &\times\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1})-F^*(Y_{t-1})\right)}{w(F^*(Y_{t-1}))}\right|\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t)-F^*(Y_t)\right)}{w(F^*(Y_t))}\right| \end{aligned}$$

Under Assumption M4,

$$\frac{1}{n^{3/2}}\sum_{t=2}^{n}\sup_{\|\beta-\beta^*\|\leq\delta,F\in\mathcal{F}_\delta}\left|\ell_{\beta 12}\left(F(Y_{t-1}),F(Y_t),\beta^*\right)w(F^*(Y_{t-1}))w(F^*(Y_t))\right|=o_p(1),$$

and by application of Lemma 1,

$$\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1})-F^*(Y_{t-1})\right)}{w(F^*(Y_{t-1}))}\right|=O_p(1),\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t)-F^*(Y_t)\right)}{w(F^*(Y_t))}\right|=O_p(1),$$

thus

$$\left|\frac{1}{n^{3/2}}\sum_{t=2}^{n}\ell_{\beta 12}\left(F_{t-1}^{\eta},F_t^{\eta},\beta^*\right)\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1})-F^*(Y_{t-1})\right)\right]\left[\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t)-F^*(Y_t)\right)\right]\right|=o_p\left(1\right).$$

Second, by Taylor expansion,

$$\begin{aligned}
&\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta\beta}\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\mathring{\beta}\right)\\
=\;&\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta\beta}\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right)\\
&+\frac{1}{n^{3/2}}\sum_{j=1}^{2}\sum_{t=2}^{n}\ell_{\beta\beta j}\left(F_{t-1}^{\eta},F_t^{\eta},\overline{\overline{\beta}}\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+j-2})-F^*(Y_{t+j-2})\right)\\
&+\frac{1}{n^{3/2}}\sum_{t=2}^{n}\ell_{\beta\beta\beta}\left(F_{t-1}^{\eta},F_t^{\eta},\overline{\overline{\beta}}\right)\sqrt{n}(\overline{\beta}-\beta),
\end{aligned}$$

where $\overline{\overline{\beta}}=\eta\beta^*+(1-\eta)\mathring{\beta}$. Thus, by Assumptions M4, ST, and Lemma 1,

$$\begin{aligned}
&\left\|\frac{1}{n}\sum_{t=2}^{n}\left[\ell_{\beta\beta}\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\mathring{\beta}\right)-\ell_{\beta\beta}\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right)\right]\right\|\\
\leq\;&\frac{1}{n^{3/2}}\sum_{j=1}^{2}\sum_{t=2}^{n}\sup_{\|\beta-\beta^*\|\leq\delta,F\in\mathcal{F}_\delta}\left\|\ell_{\beta\beta j}\left(F(Y_{t-1}),F(Y_t),\beta\right)w(F^*(Y_{t+j-2}))\right\|\\
&\times\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+j-2})-F^*(Y_{t+j-2})\right)}{w(F^*(Y_{t+j-2}))}\right|\\
&+\frac{1}{n^{3/2}}\sum_{t=2}^{n}\sup_{\|\beta-\beta^*\|\leq\delta,F\in\mathcal{F}_\delta}\left\|\ell_{\beta\beta\beta}\left(F(Y_{t-1}),F(Y_t),\beta\right)\right\|\left\|\sqrt{n}(\mathring{\beta}-\beta^*)\right\|\\
=\;&o_p\left(1\right).
\end{aligned}$$

Thus,

$$\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta\beta}\left(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\overline{\beta}\right)=\frac{1}{n}\sum_{t=2}^{n}\ell_{\beta\beta}\left(F^*(Y_{t-1}),F^*(Y_t),\beta^*\right)+o_p\left(1\right),$$

Let

$$A_{n1} = \frac{1}{n}\sum_{t=2}^{n} \ell_{\beta 1}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t-1}) - F_n(Y_{t-1})\right),$$

$$A_{n2} = \frac{1}{n}\sum_{t=2}^{n} \ell_{\beta 1}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\sqrt{n}\left(F_n(Y_{t-1}) - F^*(Y_{t-1})\right),$$

$$A_{n3} = \frac{1}{n}\sum_{t=2}^{n} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t) - F_n(Y_t)\right),$$

$$A_{n4} = \frac{1}{n}\sum_{t=2}^{n} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\sqrt{n}\left(F_n(Y_t) - F^*(Y_t)\right),$$

and

$$\Sigma_n = -\left[\frac{1}{n}\sum_{t=2}^{n} \ell_{\beta\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\right],\ S_n = \frac{1}{\sqrt{n}}\sum_{t=2}^{n} \ell_{\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right),$$

then we have

$$\Sigma_n\sqrt{n}\left(\widehat{\beta}_{SP} - \beta^*\right) = S_n + A_{n1} + A_{n2} + A_{n3} + A_{n4} + o_p(1),$$

where $A_{n2}+A_{n4}$ is the effect of estimating $F^*(\cdot)$ based on Y_t (unobserved), and $A_{n1}+A_{n3}$ is the effect of filtration. Thus, the first part

$$S_n + A_{n2} + A_{n4}$$

is the leading part of the *infeasible* estimator based on knowledge of $Y_t's$, and the effect of filtration is captured by A_{n1} and A_{n3}.

The analysis of A_{n1} and A_{n3} are similar, we illustrate our proof for A_{n3}. Notice that

$$\begin{aligned} A_{n3} &= \frac{1}{n}\sum_{t=2}^{n} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right)\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_t) - F_n(Y_t)\right) \\ &= -\frac{1}{n^2}\sum_{t=2}^{n}\sum_{j=2}^{n} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\left[(X_j - X_t)' D_n^{-1} n^{1/2}\right] D_n(\widehat{\pi} - \pi^*) + o_p(1). \end{aligned}$$

and

$$\begin{aligned} &\frac{1}{n^2}\sum_{t=2}^{n}\sum_{j=2}^{n} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\left[(X_j - X_t)' D_n^{-1} n^{1/2}\right] \\ = &\frac{1}{n^2}\sum\sum_{t>j} \ell_{\beta 2}\left(F^*(Y_{j-1}), F^*(Y_j), \beta^*\right) f(Y_j)\left[X_t' D_n^{-1} n^{1/2}\right] \\ &+\frac{1}{n^2}\sum\sum_{t>j} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\left[X_j' D_n^{-1} n^{1/2}\right] \\ &-\frac{1}{n^2}\sum\sum_{t>j} \ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\left[X_t' D_n^{-1} n^{1/2}\right] \\ &-\frac{1}{n^2}\sum\sum_{t>j} \ell_{\beta 2}\left(F^*(Y_{j-1}), F^*(Y_j), \beta^*\right) f(Y_j)\left[X_j' D_n^{-1} n^{1/2}\right] \\ = &H_{1n} + H_{2n} - H_{3n} - H_{4n}. \end{aligned}$$

We investigate the behavior of each of the above terms and show that

$$\begin{aligned}
H_{1n} &\rightarrow \left[\int_0^1 rX(r)dr\right] \mathrm{E}\left[\ell_{\beta 2}\left(F^*(Y_{j-1}), F^*(Y_j), \beta^*\right) f(Y_j)\right], \\
H_{2n} &\rightarrow \int_0^1 \int_0^r X(s)dsdr \mathrm{E}\left[\ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f\left(Y_t\right)\right], \\
H_{3n} &\rightarrow \left[\int_0^1 rX(r)dr\right] \mathrm{E}\left\{\ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\right\}, \\
H_{4n} &\rightarrow \int_0^1 \int_0^r X(s)dsdr \mathrm{E}\left\{\ell_{\beta 2}\left(F^*(Y_{t-1}), F^*(Y_t), \beta^*\right) f(Y_t)\right\}.
\end{aligned}$$

Thus $A_{n3} = o_p(1)$. Similarly, $A_{n1} = o_p(1)$. The semiparametric copula estimator of β based on filtered data is asymptotically equivalent to the infeasible semiparametric copula estimator of β based on the unobserved data Y_t,

$$\Sigma_n \sqrt{n}\left(\widehat{\beta}_{SP} - \beta^*\right) = \Sigma_n \sqrt{n}\left(\widetilde{\beta}_{SP} - \beta^*\right) + o_p(1) = S_n + A_{n2} + A_{n4} + o_p(1).$$

By Chen and Fan (2006a), we can then obtain the result of Theorem 4.

B.2.3. Theorem 5

We may re-write the variance estimator $\widehat{\Omega}_\beta^+$ as:

$$\widehat{\Omega}_\beta^+ = \sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\gamma_n(h) + \sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\left[\gamma_{n1}(h) - \gamma_n(h)\right] + \sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\left[\widehat{\gamma}_n(h) - \gamma_{n1}(h)\right]$$

where

$$\gamma_n(h) = \frac{1}{n}\sum_{\substack{t=2 \\ 2\le t, t+h\le n}}^{n} S_t(F,\beta)\, S_{t+h}(F,\beta),$$

and

$$\gamma_{n1}(h) = \frac{1}{n}\sum_{\substack{t=2 \\ 2\le t, t+h\le n}}^{n} S_t\left(F,\widehat{\beta}\right) S_{t+h}\left(F,\widehat{\beta}\right).$$

The first part,

$$\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\gamma_n(h)$$

is the conventional long-run variance (spectral density) estimator, which converges to Ω_β by the standard arguments as Hannan (1970).

The second part,

$$\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\left[\gamma_{n1}(h) - \gamma_n(h)\right],$$

contains the effect of copula estimation error $(\widehat{\beta}-\beta)$, this term converges to 0 following a similar argument as Andrews (1991, p852).

We now consider the third term,

$$\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\left[\widehat{\gamma}_n(h)-\gamma_{n1}(h)\right],$$

which contains the estimation error from the filtration and the estimation of marginal. Notice that

$$\begin{aligned}
\widehat{\gamma}_n(h) &= \frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n} S_t\left(\widehat{F}_n,\widehat{\beta}\right)S_{t+h}\left(\widehat{F}_n,\widehat{\beta}\right)\\
&= \frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n}\left[S_t\left(F,\widehat{\beta}\right)+S_t\left(\widehat{F}_n,\widehat{\beta}\right)-S_t\left(F,\widehat{\beta}\right)\right]\left[S_{t+h}\left(F,\widehat{\beta}\right)+S_{t+h}\left(\widehat{F}_n,\widehat{\beta}\right)-S_{t+h}\left(F,\widehat{\beta}\right)\right]
\end{aligned}$$

thus

$$\begin{aligned}
&\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\left[\widehat{\gamma}_n(h)-\gamma_{n1}(h)\right]\\
=\;&\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n}\left[S_t\left(\widehat{F}_n,\widehat{\beta}\right)-S_t\left(F,\widehat{\beta}\right)\right]\left[S_{t+h}\left(F,\widehat{\beta}\right)\right]\\
&+\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n}\left[S_t\left(F,\widehat{\beta}\right)\right]\left[S_{t+h}\left(\widehat{F}_n,\widehat{\beta}\right)-S_{t+h}\left(F,\widehat{\beta}\right)\right]\\
&+\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n}\left[S_t\left(\widehat{F}_n,\widehat{\beta}\right)-S_t\left(F,\widehat{\beta}\right)\right]\left[S_{t+h}\left(\widehat{F}_n,\widehat{\beta}\right)-S_{t+h}\left(F,\widehat{\beta}\right)\right]
\end{aligned}$$

We can verify the order of magnitude for each of these terms. For example, consider the second term

$$\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\2\leq t,t+h\leq n}}^{n}\left[S_t\left(F,\widehat{\beta}\right)\right]\left[S_{t+h}\left(\widehat{F}_n,\widehat{\beta}\right)-S_{t+h}\left(F,\widehat{\beta}\right)\right],$$

notice that

$$\begin{aligned}
&\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\ 2\le t,t+h\le n}}^{n} S_t\left(F,\widehat{\beta}\right)\left[S_{t+i}\left(\widehat{F}_n,\widehat{\beta}\right)-S_{t+i}\left(F_n,\widehat{\beta}\right)\right]\\
=\;&\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_t S_t\left(F,\widehat{\beta}\right)\sum_{j=1}^{2}\ell_{\beta j}\left(F^{\eta}_{n,t+i-1},F^{\eta}_{n,t+i},\widehat{\beta}\right)\left(\widehat{F}_n(\widehat{Y}_{t+i+j-2})-F(Y_{t+i+j-2})\right)\\
&-\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_t S_t\left(F,\widehat{\beta}\right)\int_0^1 \ell_{\beta,2}\left(v_1,F^{\eta}_{n,t+i};\widehat{\beta}\right)c\left(v_1,F^{\eta}_{n,t+i};\widehat{\beta}\right)dv_1\left(\widehat{F}_n(\widehat{Y}_{t+i})-F(Y_{t+i})\right)\\
&-\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_t S_t\left(F,\widehat{\beta}\right)\int_0^1 \ell_{\beta,1}\left(F^{\eta}_{n,t+i-1},v_2;\widehat{\beta}\right)c\left(F^{\eta}_{n,t+i-1},v_2;\widehat{\beta}\right)dv_2\left(\widehat{F}_n(\widehat{Y}_{t+i-1})-F(Y_{t+i-1})\right)
\end{aligned}$$

where $F^{\eta}_{n,s}=F(Y_s)+\eta\left[\widehat{F}_n(\widehat{Y}_s)-F(Y_s)\right]$, $\eta\in[0,1]$, denotes a (generic) middle value between $\widehat{F}_n(\widehat{Y}_s)$ and $F(Y_s)$. Under our regularity assumptions, the order of magnitude for each of these terms are $o_p(1)$. For example

$$\begin{aligned}
&\left|\sum_{h=-M}^{M} K\left(\frac{h}{M}\right)\frac{1}{n}\sum_{\substack{t=2\\ 2\le t,t+h\le n}}^{n} S_t\left(F,\widehat{\beta}\right)\sum_{j=1}^{2}\ell_{\beta j}\left(F^{\eta}_{n,t+h-1},F^{\eta}_{n,t+h},\widehat{\beta}\right)\left(\widehat{F}_n(\widehat{Y}_{t+h+j-2})-F(Y_{t+h+j-2})\right)\right|\\
\le\;&\frac{1}{\sqrt{n}}\frac{1}{n}\sum_{h=-M}^{M}\sum_{\substack{t=2\\ 2\le t,t+h\le n}}^{n}\left|K\left(\frac{h}{M}\right)\right|\sum_{j=1}^{2}\sup_{F\in\mathcal{F}_\delta}\left|S_t\left(F^*,\widehat{\beta}\right)w\left(F^*(Y_{t+h+j-2})\right)\ell_{\beta j}\left(F(Y_{t+h-1}),F(Y_{t+h}),\widehat{\beta}\right)\right|\\
&\times\left|\frac{\sqrt{n}\left(\widehat{F}_n(\widehat{Y}_{t+h+j-2})-F^*(Y_{t+h+j-2})\right)}{w\left(F^*(Y_{t+h+j-2})\right)}\right|
\end{aligned}$$

under our regularity assumptions and the bandwidth condition, the above term is $o_p(1)$.

Other terms can be verified to be $o_p(1)$ using similar arguments.

B.2.4. Theorem 8

We show that the filtration does not affect the limiting distribution. Expanding $\log c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\overline{\beta}_2)$ around $\widehat{\beta}_2$, and notice that the FOC corresponding to $\widehat{\beta}_2$ implies

$$\sum_t \frac{\partial\log c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_2)}{\partial\beta}=0,$$

in the non-nested case,

$$\begin{aligned}
\Pr\left[\log\frac{c_2(U_{t-1},U_t,\overline{\beta}_2)}{c_1(U_{t-1},U_t,\overline{\beta}_1)}\neq \mathrm{E}\left[\log\frac{c_2(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_2)}{c_1(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_1)}\right]\right] &> 0\\
\Pr\left[\frac{\partial\log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}\neq\frac{\partial\log c_1(U_{t-1},U_t,\overline{\beta}_1)}{\partial U_{t-2+j}}\right] &> 0
\end{aligned}$$

we have

$$\begin{aligned}
&\frac{1}{n}\sum_{t=2}^{n}\log c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_2)\\
=&\ \frac{1}{n}\sum_{t=2}^{n}\log c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\overline{\beta}_2)-\frac{1}{2n}\sum_{t=2}^{n}\left(\overline{\beta}_2-\widehat{\beta}_2\right)\frac{\partial^2\log c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\overline{\overline{\beta}}_2)}{\partial\beta\partial\beta'}\left(\overline{\beta}_2-\widehat{\beta}_2\right)\\
=&\ \frac{1}{n}\sum_{t=2}^{n}\log c_2(U_{t-1},U_t,\overline{\beta}_2)+\frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\frac{\partial\log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}\left[\widehat{F}_n(\widehat{Y}_{t-2+j})-F(Y_{t-2+j})\right]+o_p\left(n^{-1/2}\right)
\end{aligned}$$

and

$$\begin{aligned}
&\widehat{LR}_n\\
=&\ \frac{1}{n}\sum_{t=2}^{n}\log\frac{c_2(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_2)}{c_1(\widehat{F}_n(\widehat{Y}_{t-1}),\widehat{F}_n(\widehat{Y}_t),\widehat{\beta}_1)}\\
=&\ \frac{1}{n}\sum_{t=2}^{n}\log\frac{c_2(U_{t-1},U_t,\overline{\beta}_2)}{c_1(U_{t-1},U_t,\overline{\beta}_1)}+\frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\left\{\frac{\partial\log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}-\frac{\partial\log c_1(U_{t-1},U_t,\overline{\beta}_1)}{\partial U_{t-2+j}}\right\}\left[\widehat{F}_n(\widehat{Y}_{t-2+j})-F(Y_{t-2+j})\right]\\
&+o_p\left(n^{-1/2}\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\widehat{LR}_n-\mathrm{E}\left[\log\frac{c_2(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_2)}{c_1(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_1)}\right]\\
=&\ \frac{1}{n}\sum_{t=2}^{n}\left[\log\frac{c_2(U_{t-1},U_t,\overline{\beta}_2)}{c_1(U_{t-1},U_t,\overline{\beta}_1)}-\mathrm{E}\left[\log\frac{c_2(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_2)}{c_1(F\left(Y_{t-1}\right),F\left(Y_t\right),\overline{\beta}_1)}\right]\right]\\
&+\frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\left\{\frac{\partial\log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}-\frac{\partial\log c_1(U_{t-1},U_t,\overline{\beta}_1)}{\partial U_{t-2+j}}\right\}\left[\widehat{F}_n(\widehat{Y}_{t-2+j})-F(Y_{t-2+j})\right]\\
&+o_p\left(n^{-1/2}\right)
\end{aligned}$$

$$
\begin{aligned}
& \frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}\left[\widehat{F}_n(\widehat{Y}_{t-2+j})-F(Y_{t-2+j})\right] \\
= & \frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-1}}\left[\widehat{F}_n(\widehat{Y}_{t-1})-F_n(Y_{t-1})\right] \\
& +\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}\left[\widehat{F}_n(\widehat{Y}_{t})-F_n(Y_{t})\right] \\
& +\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-1}}\left[F_n(Y_{t-1})-F(Y_{t-1})\right] \\
& +\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}\left[F_n(Y_{t})-F(Y_{t})\right]
\end{aligned}
$$

Using similar argument as in the previous Sections, we can show

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}\left[\widehat{F}_n(\widehat{Y}_{t})-F_n(Y_{t})\right] \\
= & \frac{1}{\sqrt{n}}\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}f^*\left(Y_t\right)\frac{1}{n}\sum_{j=1}^{n}\left[\left(X_j'-X_t'\right)D_n^{-1}n^{1/2}\right]D_n\left(\widehat{\pi}-\pi^*\right)+o_p\left(n^{-1/2}\right) \\
= & \frac{1}{\sqrt{n}}\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}f^*\left(Y_t\right)\frac{1}{n}\sum_{j=1}^{n}\left[X_j'D_n^{-1}n^{1/2}\right]D_n\left(\widehat{\pi}-\pi^*\right) \\
& -\frac{1}{\sqrt{n}}\frac{1}{n}\sum_{t=2}^{n}\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t}}f^*\left(Y_t\right)\frac{1}{n}\sum_{j=1}^{n}\left[X_t'D_n^{-1}n^{1/2}\right]D_n\left(\widehat{\pi}-\pi^*\right)+o_p\left(n^{-1/2}\right) \\
= & o_p\left(n^{-1/2}\right)
\end{aligned}
$$

and thus

$$
\begin{aligned}
& \frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\left\{\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}-\frac{\partial \log c_1(U_{t-1},U_t,\overline{\beta}_1)}{\partial U_{t-2+j}}\right\}\left[\widehat{F}_n(\widehat{Y}_{t-2+j})-F(Y_{t-2+j})\right] \\
= & \frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\left\{\frac{\partial \log c_2(U_{t-1},U_t,\overline{\beta}_2)}{\partial U_{t-2+j}}-\frac{\partial \log c_1(U_{t-1},U_t,\overline{\beta}_1)}{\partial U_{t-2+j}}\right\}\left[F_n(Y_{t-2+j})-F(Y_{t-2+j})\right]+o_p\left(n^{-1/2}\right).
\end{aligned}
$$

Let

$$
g_{ij}\left(U_l,\overline{\beta}_i\right)=\mathrm{E}\left\{\left[\frac{\partial \log c_i(U_{t-1},U_t,\overline{\beta}_i)}{\partial U_{t-2+j}}\right]\left[\left(1(U_l\le U_{t-2+j})-U_{t-2+j}\right)\right]\middle| U_l\right\},
$$

$$
\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{j=1}^{2} \sum_{t=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \overline{\beta}_1)}{\partial U_{t-2+j}} \right\} [F_n(Y_{t-2+j}) - F(Y_{t-2+j})] \\
= \; & \frac{1}{\sqrt{n}} \sum_{j=1}^{2} \sum_{l=2}^{n} \mathrm{E} \left\{ \left[\frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \overline{\beta}_1)}{\partial U_{t-2+j}} \right] [(1(U_l \le U_{t-2+j}) - U_{t-2+j})] \middle| U_l \right\} \\
= \; & \sum_{j=1}^{2} \left[\frac{1}{\sqrt{n}} \sum_{l=2}^{n} \mathrm{E} \left\{ g_{2j} \left(U_l, \overline{\beta}_2 \right) - g_{1j} \left(U_l, \overline{\beta}_1 \right) \right\} \right],
\end{aligned}
$$

we have

$$
\begin{aligned}
& \sqrt{n} \left(LR_n - \mathrm{E} \left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)} \right] \right) \\
= \; & \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \left[\log \frac{c_2(U_{t-1}, U_t, \overline{\beta}_2)}{c_1(U_{t-1}, U_t, \overline{\beta}_1)} - \mathrm{E} \left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)} \right] \right] \\
& + \frac{1}{\sqrt{n}} \sum_{j=1}^{2} \sum_{t=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \overline{\beta}_1)}{\partial U_{t-2+j}} \right\} [F_n(Y_{t-2+j}) - F(Y_{t-2+j})] + o_p(1) \\
= \; & \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \left[\log \frac{c_2(U_{t-1}, U_t, \overline{\beta}_2)}{c_1(U_{t-1}, U_t, \overline{\beta}_1)} - \mathrm{E} \left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)} \right] \right] \\
& + \sum_{j=1}^{2} \left[\frac{1}{\sqrt{n}} \sum_{l=2}^{n} \mathrm{E} \left\{ g_{2j} \left(U_l, \overline{\beta}_2 \right) - g_{1j} \left(U_l, \overline{\beta}_1 \right) \right\} \right] + o_p(1) \\
\Rightarrow \; & N \left(0, \omega^2 \right)
\end{aligned}
$$

In the generalized nested case, denote

$$
H_{jn} = -\frac{1}{n} \sum_{t=2}^{n} \frac{\partial^2 \log c_j(\widehat{F}_n(\widehat{Y}_{t-1}), \widehat{F}_n(\widehat{Y}_t), \overline{\beta}_j)}{\partial \beta \partial \beta'} \rightarrow \; \overline{H}_{j,\beta},
$$

Notice that

$$
\Pr \left[c_2(U_{t-1}, U_t, \overline{\beta}_2) = c_1(U_{t-1}, U_t, \overline{\beta}_1) \right] = 1
$$

thus

$$
\begin{aligned}
\Pr \left[\log \frac{c_2(U_{t-1}, U_t, \overline{\beta}_2)}{c_1(U_{t-1}, U_t, \overline{\beta}_1)} = 0 = \mathrm{E} \left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)} \right] \right] &= 1 \\
\Pr \left[\frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} = \frac{\partial \log c_1(U_{t-1}, U_t, \overline{\beta}_1)}{\partial U_{t-2+j}} \right] &= 1
\end{aligned}
$$

thus,

$$
\begin{aligned}
& LR_n - \mathrm{E}\left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)}\right] \\
= & \frac{1}{n}\sum_{t=2}^{n}\left[\log \frac{c_2(U_{t-1}, U_t, \overline{\beta}_2)}{c_1(U_{t-1}, U_t, \overline{\beta}_1)} - \mathrm{E}\left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)}\right]\right] \\
& + \frac{1}{n}\sum_{j=1}^{2}\sum_{t=2}^{n}\left\{\frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \overline{\beta}_1)}{\partial U_{t-2+j}}\right\}\left[\widehat{F}_n(\widehat{Y}_{t-2+j}) - F(Y_{t-2+j})\right] \\
& + \frac{1}{2}\left(\overline{\beta}_2 - \widehat{\beta}_2\right) H_{2n}\left(\overline{\beta}_2 - \widehat{\beta}_2\right) - \frac{1}{2}\left(\overline{\beta}_1 - \widehat{\beta}_1\right) H_{1n}\left(\overline{\beta}_1 - \widehat{\beta}_1\right) + \\
= & \frac{1}{2}\left(\overline{\beta}_2 - \widehat{\beta}_2\right) H_{2n}\left(\overline{\beta}_2 - \widehat{\beta}_2\right) - \frac{1}{2}\left(\overline{\beta}_1 - \widehat{\beta}_1\right) H_{1n}\left(\overline{\beta}_1 - \widehat{\beta}_1\right) + o_p\left(\frac{1}{n}\right)
\end{aligned}
$$

Let $U_t = F^*(Y_t)$, and $\overline{\mathcal{G}}_{j,n} = n^{-1/2}\sum_{j=2}^{n}\overline{\ell}_{j,\beta}\left(U_{j-1}, U_j, \overline{\beta}_j\right)$, $j = 1, 2$, where

$$
\overline{\ell}_{j,\beta}\left(U_{j-1}, U_j, \overline{\beta}_j\right) = \frac{\partial \log c_j\left(U_{j-1}, U_j, \overline{\beta}\right)}{\partial \beta} + \sum_{i=0}^{1}\mathrm{E}\left[\frac{\partial^2 \log c_j\left(U_{t-1}, U_t, \overline{\beta}\right)}{\partial \beta \partial U_{t-i}}\left[1\left(U_j \le U_{t-i}\right) - U_{t-i}\right]\middle| U_j\right]
$$

$$
\overline{\Omega}_{j,\beta} = \lim_{n\to\infty} Var\left(\overline{\mathcal{G}}_{j,n}\right), \; \overline{H}_{j,\beta} = -E\ell_{j,\beta\beta}\left(F^*(Y_{t-1}), F^*(Y_t), \overline{\beta}_j\right)
$$

Using the results of Section 4,

$$
\sqrt{n}\left(\widehat{\beta}_j - \overline{\beta}_j\right) \Rightarrow N\left(0, \overline{H}_{j,\beta}^{-1}\overline{\Omega}_{j,\beta}\overline{H}_{j,\beta}^{-1}\right).
$$

and

$$
\begin{aligned}
& n\left[LR_n - \mathrm{E}\left[\log \frac{c_2(F(Y_{t-1}), F(Y_t), \overline{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \overline{\beta}_1)}\right]\right] \\
= & \frac{1}{2}n\left(\overline{\beta}_2 - \widehat{\beta}_2\right) H_{2n}\left(\overline{\beta}_2 - \widehat{\beta}_2\right) - \frac{1}{2}n\left(\overline{\beta}_1 - \widehat{\beta}_1\right) H_{1n}\left(\overline{\beta}_1 - \widehat{\beta}_1\right) \\
& + o_p(1) \\
= & \frac{1}{2}\overline{\mathcal{G}}_{2,n}'\overline{H}_{2,\beta}^{-1}\left(H_{2n}\right)\overline{H}_{2,\beta}^{-1}\overline{\mathcal{G}}_{2,n} - \frac{1}{2}\overline{\mathcal{G}}_{1,n}'\overline{H}_{1,\beta}^{-1}\left(H_{1n}\right)\overline{H}_{1,\beta}^{-1}\overline{\mathcal{G}}_{1,n} + o_p(1) \\
= & \frac{1}{2}\begin{bmatrix} \overline{\mathcal{G}}_{2,n}' & \overline{\mathcal{G}}_{1,n}' \end{bmatrix}\begin{bmatrix} \overline{H}_{2,\beta}^{-1} & 0 \\ 0 & -\overline{H}_{1,\beta}^{-1} \end{bmatrix}\begin{bmatrix} \overline{\mathcal{G}}_{2,n} \\ \overline{\mathcal{G}}_{1,n} \end{bmatrix} + o_p(1)
\end{aligned}
$$

where

$$
\begin{bmatrix} \overline{\mathcal{G}}_{2,n} \\ \overline{\mathcal{G}}_{1,n} \end{bmatrix} \Rightarrow N\left(0, \begin{bmatrix} \overline{\Omega}_{2,\beta} & \overline{\Omega}_{2,1} \\ \overline{\Omega}_{2,1}' & \overline{\Omega}_{1,\beta} \end{bmatrix}\right)
$$

Thus, under the null, $2nLR_n$ converges to a weighted sum of independent χ_1^2 random variables in which the weights $(\lambda_1, \cdots, \lambda_{k_1+k_2})$ is the vector of eigenvalues of the following matrix

$$
\begin{bmatrix} \overline{\Omega}_{2,\beta}\overline{H}_{2,\beta}^{-1} & -\overline{\Omega}_{2,1}\overline{H}_{1,\beta}^{-1} \\ \overline{\Omega}_{2,1}'\overline{H}_{2,\beta}^{-1} & -\overline{\Omega}_{1,\beta}\overline{H}_{1,\beta}^{-1} \end{bmatrix} = \begin{bmatrix} \overline{\Omega}_{2,\beta} & \overline{\Omega}_{2,1} \\ \overline{\Omega}_{2,1}' & \overline{\Omega}_{1,\beta} \end{bmatrix}\begin{bmatrix} \overline{H}_{2,\beta}^{-1} & \\ & -\overline{H}_{1,\beta}^{-1} \end{bmatrix}.
$$