Copula-Based Time Series With Filtered Nonstationarity*

Xiaohong Chen†  Zhijie Xiao‡  Bo Wang§
Yale University  Boston College  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 dis-
tribution 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 af-
fected 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 com-
plicated 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 non-
stationary 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 semi-
parametric 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 $\omega^2$. For a generic parameter, say, $\beta$, we denote the true parameter value by $\beta^*$, the pseudo-true value by $\tilde{\beta}$ and the feasible estimator by $\hat{\beta}$.

2. The Model

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

$$Z_t = X_t'\pi^* + Y_t,$$  \hspace{1cm} (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 \( \pi^* \) 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 \( \beta \) 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 \( \beta \) and study its asymptotic properties based on the filtered time series \( \{\tilde{Y}_t\}_{t=1}^n \), where
\[
\tilde{Y}_t = Z_t - X_t^{'\hat{\pi}},
\]
and \( \hat{\pi} \) denotes some nonstationary filtering estimator for \( \pi^* \). 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
\[
\begin{bmatrix}
Y_n(r) \\
X_n(r)
\end{bmatrix} \Rightarrow \begin{bmatrix}
B_Y(r) \\
X(r)
\end{bmatrix}, \quad r \in [0, 1] \quad \text{as} \quad n \to \infty,
\]
where \( B_Y(r) \) is a Brownian motion, \( X(r) \) is a vector of stochastic or deterministic functions. And
\[
D_n (\hat{\pi} - \pi^*) \Rightarrow \xi \quad \text{as} \quad n \to \infty.
\]

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, \( \xi \), 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 \( \hat{\pi} \) be the OLS estimator of \( \pi^* \),
\[
D_n (\hat{\pi} - \pi^*) \Rightarrow \xi_1,
\]
where in general \( \xi_1 \) 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 $D_n = \text{diag}(n^{1/2}, n^{3/2})$.

**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 $\tilde{\pi}$ converges at rate-$n$ to a non-normal limit: $n(\tilde{\pi} - \pi) \Rightarrow \xi_2$, 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 $\lambda = \sum_{h=1}^\infty E(Y_1Y_{1+h})$.

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

$X_1(r)$ is the limiting trending function, and $B_2(r)$ is a stochastic process. Let $D_n = \text{diag}\{D_1, n, \cdots, n\}$,

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

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

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

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

\[ G^*(y_{t-1}, y_t) = C(F^*(y_{t-1}), F^*(y_t)), \]

where the copula function \(C(\cdot, \cdot)\) is a bivariate probability distribution function with uniform marginals. Denote the corresponding copula density of \(C(\cdot, \cdot)\) by \(c(\cdot, \cdot)\), 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:** \(\{Y_t\}_{t=1}^{\infty}\) 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 \(\beta^*\).

#### 3.1. Feasible estimation of copula parameter using filtered data \(\hat{Y}_t\)

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

\[ \hat{\beta} = \arg \max_{\beta} \hat{Q}_n(\hat{F}, \beta), \text{ where } \hat{Q}_n(\hat{F}, \beta) = \frac{1}{n} \sum_{t=2}^{n} \log c(\hat{F}(\hat{Y}_{t-1}), \hat{F}(\hat{Y}_{t}), \beta). \]  

(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
We could then estimate the true marginal \( F^* (\cdot) \) by \( F (\cdot, \hat{\alpha}) \) where

\[
\hat{\alpha} = \arg \max_\alpha \sum_{t=1}^n \log f (\hat{Y}_t, \alpha),
\]  

(3.2)

and estimate the copula parameter \( \beta^* \) by the following “parametric copula estimator”:

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

### 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 \( \hat{Y}_t \). We use the so-called rescaled empirical distribution function (EDF) to estimate \( F (\cdot) \):

\[
\hat{F}_n (y) = \frac{1}{n+1} \sum_{t=1}^n 1 (\hat{Y}_t \leq y),
\]

and estimate the copula parameter \( \beta^* \) by the following “semiparametric copula estimator”:

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

### 3.2. Infeasible estimation of copula parameter using \( Y_t \)

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

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

Again, \( \tilde{\beta}_p \) denotes the parametric copula estimator using the infeasible parametric marginal estimator \( \tilde{F} = F (\cdot, \tilde{\alpha}) \), where\(^1\)

\[
\tilde{\alpha} = \arg \max_\alpha \sum_{t=1}^n \log f (Y_t, \alpha).
\]

\(^1\)Previously, Joe and Xu (1996) and Joe (2005) studied two-step parametric estimation of copula parameter \( \beta \) 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 $$\tilde{\beta}_{SP}$$ denotes the semiparametric copula estimator using the infeasible rescaled estimator for $$F^*(\cdot)$$:

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

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

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

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

In the next two subsections, we show that although the parameter estimators $$\hat{\beta}_P$$ and $$\tilde{\beta}_P$$ could have different asymptotic properties, the semiparametric estimators $$\beta_{SP}$$ and $$\tilde{\beta}_{SP}$$ 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 $$g(Y_{t-1}, Y_t, \alpha, \beta) = \log c(F(Y_{t-1}, \alpha), F(Y_t, \alpha), \beta)$$ and $$g_\beta(s_1, s_2, \alpha, \beta) = \partial g(s_1, s_2, \alpha, \beta) / \partial \beta$$. For $$i = 1, 2$$, $$j = 1, 2$$, we define

$$\frac{\partial g_\beta(s_1, s_2, \alpha, \beta)}{\partial \alpha} = g_\beta(s_1, s_2, \alpha, \beta); \quad \frac{\partial g_\beta(s_1, s_2, \alpha, \beta)}{\partial s_j} = g_\beta(s_1, s_2, \alpha, \beta).$$

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

$$\frac{\partial \ell(u, v, \beta)}{\partial \beta} = \ell_\beta(u, v, \beta); \quad \frac{\partial \ell(u, v, \beta)}{\partial u} = \ell_1(u, v, \beta); \quad \frac{\partial \ell(u, v, \beta)}{\partial v} = \ell_2(u, v, \beta).$$

For consistency in the parametric case, we make the following assumptions.
Assumption ID1: (1) \( \mathcal{A} \) and \( \mathfrak{B} \) are compact subsets of \( \mathcal{R}_k^m \) and \( \mathcal{R}_k \). (2) \( q(\alpha) = \mathbb{E}[\log f(Y_t, \alpha)] \) has a unique maximizer \( \alpha^* \in \mathcal{A} \); and \( Q(\beta) = \mathbb{E}[\ell(F(Y_{t-1}, \alpha^*), F(Y_t, \alpha^*) \beta)] \) has a unique maximizer \( \beta^* \in \mathfrak{B} \). (3) \( f(y, \alpha) \) is continuous in \( \alpha \in \mathcal{A} \), and \( g(\alpha, \beta) = \mathbb{E}[g(Y_{t-1}, Y_t, \alpha, \beta)] \) is Lipschitz continuous in \( \alpha \in \mathcal{A} \) and \( \beta \in \mathfrak{B} \).

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

\[
\sup_{\alpha \in \mathcal{A}} |\log f(y_1, \alpha) - \log f(y_2, \alpha)| < \varepsilon.
\]

Similarly, \( g(s_1, s_2, \alpha, \beta) \) is uniformly continuous in \( (s_1, s_2, \alpha) \), uniformly over \( \beta \in \mathfrak{B} \), in the sense that for any \( \varepsilon > 0 \), there exists \( \delta > 0 \), such that if \( |s'_1 - s''_1| + |s'_2 - s''_2| + |\alpha' - \alpha''| < \delta \), then

\[
\sup_{\beta \in \mathfrak{B}} |g(s'_1, s'_2, \alpha', \beta) - g(s''_1, s''_2, \alpha'', \beta)| < \varepsilon.
\]

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

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

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

and

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

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

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:

\[
P_{nj} = \frac{1}{n} \sum_{t=2}^{n} g_{\beta_j} (Y_{t-1}, Y_t, \alpha^*, \beta^*) X_{t-2+j}^{-1} D_n^{-1/2} = P_j + o_p(1), \quad j = 1, 2,
\]

\[
P_{n3} = n^{-1} \sum_{t=2}^{n} g_{\beta_0} (Y_{t-1}, Y_t, \alpha^*, \beta^*) = P_3 + o_p(1).
\]

\[
H_{n\alpha Y} = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \log f(Y_t, \alpha^*)}{\partial \alpha \partial Y} \left( X_t' D_n^{-1/2} \right) = H_{\alpha Y} + o_p(1).
\]
Theorem 2: Under Assumptions DGP, MX, ID2, M2, and X, as $n \to \infty$,

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

where

$$
\Omega^\#_\beta = \lim_n \text{Var} \left( \frac{1}{\sqrt{n}} \sum_{t=2}^n \left( \ell_\beta (F^*(Y_{t-1}), F^*(Y_t), \beta^*) + P_3 \Omega^{-1}_\alpha \frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha} \right) \right)
$$

$$
= \Omega_\beta + P_3 \Omega^{-1}_\alpha P'_3.
$$

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 $\hat{\beta}$ into three components: The first part is $N \left( 0, H^{-1}_\beta \Omega_\beta H^{-1}_\beta \right) = N \left( 0, \Omega_\beta \right)$, 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 $\alpha^*$); The second part is $N(0, H^{-1}_\beta P_3 \Omega^{-1}_\alpha P'_3 H^{-1}_\beta)$, the normal limit from the parametric estimation of marginal parameter $\alpha^*$ using $Y_t$; The third part is $H^{-1}_\beta \left( P_1 + P_2 + P_3 \Omega^{-1}_\alpha H_{\alpha Y} \right) \xi$, the effect of nonstationary filtration $\tilde{Y}_t$. The first two parts are normal random variates but the third part may not be normal. Unless $P_1 + P_2 + P_3 \Omega^{-1}_\alpha H_{\alpha Y} = o_p(1)$, the nonstationary filtration will affect the limiting distribution of the parametric copula estimator $\hat{\beta}_P$.

In particular, the filtration affects the limiting distribution of $\sqrt{n} \left( \hat{\beta}_P - \beta^* \right)$ directly through $\tilde{Y}_t$ and indirectly through $\tilde{\alpha}$. Unless $X_t$ is purely deterministic, the limiting distribution of $\sqrt{n} \left( \hat{\beta}_P - \beta^* \right)$ 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 $\hat{\beta}$, assuming $Y_t$ is observed with a completely known marginal distribution $F^*(\cdot)$, is given by

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

From the proof of Theorem 2, we have

$$
\sqrt{n} \left( \hat{\beta}_P - \beta^* \right) \Rightarrow N \left( 0, H^{-1}_\beta \Omega^\#_\beta H^{-1}_\beta \right).
$$

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

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

(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,
\[
P_{nj} \to P_j = E \int_0^1 X(r)'dr, \quad j = 1, 2,
\]
\[
H_{nj\alpha Y} \to H_{\alpha Y} = \int_0^1 X(r)'dr,
\]
we have
\[
P_1 + P_2 + P_3\alpha^{-1}H_{\alpha Y} = \eta \int_0^1 X(r)'dr,
\]
and
\[
\sqrt{n} \left( \beta_P - \beta^* \right) \Rightarrow N \left( 0, \Omega_{\beta}^{-1}\Omega_{\beta}^{\#}\beta^{-1} \right),
\]
where
\[
\Omega_{\beta}^{\#} = \lim_{n \to \infty} \text{Var} \left( \frac{1}{\sqrt{n}} \sum_{t=2}^n g_\beta (Y_{t-1}, Y_t, \alpha^*, \beta^*) + P_{n3}\alpha^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\partial \log f(Y_t, \alpha^*)}{\partial \alpha} - \eta_X \sum_t D_{n-1}x_i Y_t \right).
\]

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

**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
\[
Z_t = \hat{\pi}Z_{t-1} + \hat{\eta}_t,
\]
\[
n(\hat{\pi} - \pi^*) \Rightarrow \xi_2 = \left[ \int_0^1 B_Y(r)^2dr \right]^{-1} \left[ \int_0^1 B_Y(r)dB_Y(r) + \lambda \right]
\]
with \(\lambda = \sum^\infty_{h=1} E(Y_1Y_{1+h})\). Then,
\[
\sqrt{n} \left( \beta_P - \beta^* \right) \Rightarrow N \left( 0, \beta^{-1}\Omega_{\beta}^{\#}\beta^{-1} \right) - \eta H^{-1}\beta h(B_Y(r))
\]
where \(\eta\) is defined as (3.4), and
\[
h(B_Y(r)) = \int_0^1 B_Y(r)dr \left[ \int_0^1 B_Y(r)^2dr \right]^{-1} \left[ \int_0^1 B_Y(r)dB_Y(r) + \lambda \right].
\]

In this example, the limiting distribution \(\xi_2\) coming from nonstationary filtration is non-normal, and thus the limiting distribution of the parametric copula estimator \(\hat{\beta}_P\) 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

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

and

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

Then,

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

where

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

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 \( \Lambda_{2Y} \), the limiting distribution coming from nonstationary filtration is not normal. Thus the limiting distribution of the parametric copula estimator in this case \( \hat{\beta}_P \) 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

\[
\|F - F^*\|_w = \sup_y \left\{ \frac{|F(y) - F^*(y)|}{w(F^*(y))} \right\}.
\]

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 \( w(u) \geq \zeta [u(1-u)]^{\mu_1} \log(1/(u(1-u)))^{\mu_2} \) 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 \( b = (b_1, \ldots, b_n)' \), 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 $|b| = \max_{i} |b_i|$.

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

**Assumption M3** (1). $E[\sup_{\beta \in \mathcal{B}} \|\ell_{\beta_j}(F^*(Y_{t-1}), F^*(Y_t), \beta)\| \log(1 + \|\ell_{\beta_j}(F^*(Y_{t-1}), F^*(Y_t), \beta)\|)] < \infty$. (2). $E[\sup_{\beta \in \mathcal{B}, F \in \mathcal{F}_3} \|\ell_{\beta_j}(F^*(Y_{t-1}), F(Y_t), \beta)\| w(F^*(Y_{t-1}+j))] < \infty$, $j = 1, 2$. (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, $\hat{\beta}_{SP} = \beta^* + o_p(1)$.

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

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

**Assumption M4** (1). Let $F^*_\eta = F^* + \eta [F - F^*]$ for $\eta \in [0, 1]$ and $F \in \mathcal{F}_3$, the interchange of differentiation and integration of $\ell_\beta(F^*_\eta(Y_{t-1}), F^*_\eta(Y_t), \beta^*_\eta)$ w.r.t. $\eta \in (0, 1)$ is valid.

(2). $E\left[\sup_{\|\beta - \beta^*\| \leq \delta, F \in \mathcal{F}_3} \|\ell_\beta(F^*_\eta(Y_{t-1}), F^*_\eta(Y_t), \beta^*_\eta)\|^{2} \log(1 + \|\ell_\beta(F^*_\eta(Y_{t-1}), F^*_\eta(Y_t), \beta^*_\eta)\|)\right] < \infty$,

(3). $E\left[\sup_{\|\beta - \beta^*\| \leq \delta, F \in \mathcal{F}_3} \|\ell_{\beta_j}(F^*_\eta(Y_{t-1}), F^*_\eta(Y_t), \beta^*_\eta)\| w(F^*(Y_{t-1}+j))]^{2} < \infty$, $j = 1, 2$,

(4). $E\left[\sup_{\|\beta - \beta^*\| \leq \delta, F \in \mathcal{F}_3} \|\ell_{\beta_{ij}}(F^*_\eta(Y_{t-1}), F^*_\eta(Y_t), \beta^*_\eta) w(F^*(Y_{t+i-2})) w(F^*(Y_{t+j-2})))\right] < \infty$, $i, j = 1, 2$.

Denote

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

where

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

Let

$$
\Omega_{\beta}^+ = \lim_{n \to \infty} \text{Var}(G_n) = \Omega_{\beta} + \text{Var}(G_0(Y_t) + G_1(Y_{t-1})).
$$
Theorem 4: Under Assumptions DGP, SP, MX, ID4, M4, and X, as \( n \to \infty \),
\[
\sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) = \sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) + o_p(1) \Rightarrow N \left( 0, H_{\beta}^{-1} \Omega_{\beta} H_{\beta}^{-1} \right).
\]

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

From the proof of Theorem 4 in the Appendix, we can again decompose the distribution of the semiparametric copula estimator \( \beta_{SP} \) into three components: The first part is \( N \left( 0, H_{\beta}^{-1} \Omega_{\beta} H_{\beta}^{-1} \right) \), the normal limit of the ideal infeasible estimator \( \beta \) when \( Y_t \) is observed with a completely known marginal distribution \( F \); The second part, denoted as \( A_{n} + A_{n4} \) 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 \( Y_t \). 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 \( \sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) \) is only asymptotically affected by the first two parts. Consequently, the limiting distribution of \( \sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) \) is the same as that of \( \sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) \), which is always normal.

Remark 2. Chen and Fan (2006b) studied the following class of semiparametric copula-based multivariate dynamic models
\[
Z_t = (Z_{1,t}, ..., Z_{d,t}), \quad Z_{j,t} = \mu_{j,t}(\theta^*) + \sigma_{j,t}(\theta^*)Y_{j,t},
\]
\[
\mu_{j,t}(\theta^*) = E[Z_{j,t}|I_{t-1}], \quad \sigma_{j,t}^2(\theta^*) = VarE[Z_{j,t}|I_{t-1}],
\]
\[
Y_t = (Y_{1,t}, ..., Y_{d,t}) \text{ is independent of } I_{t-1}, \text{ and } \{Y_t\}_{t=1}^n \text{ is i.i.d. over } t
\]
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 \( \hat{Y}_t \) 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 \( \theta \). 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 \( I_{t-1} \) 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 \( \beta \) using the asymptotic results of Theorem 4.

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

\[
\sqrt{n} \left( \hat{R}\beta_{SP} - r \right) \Rightarrow N \left( 0, R\beta^* \Omega_{\beta}^+ H^{-1} R' \right),
\]

where \( H_{\beta} \) and \( \Omega_{\beta}^+ \) are defined in Theorem 4. Thus, under \( H_{01} \), as \( n \to \infty \),

\[
n \left( \hat{R}\beta_{SP} - r \right)' \left[ R\beta^* \Omega_{\beta}^+ H^{-1} R' \right]^{-1} \left( \hat{R}\beta_{SP} - r \right) \Rightarrow \chi^2_d,
\]

where \( d \) is the number of restrictions.

In order to construct the Wald test, we need to estimate \( \Omega_{\beta}^+ = \lim_{n \to \infty} \text{Var}(G_n) \), and \( H_{\beta} = -\mathbb{E}_{\beta^*}(F^*(Y_{t-1}), F^*(Y_t), \beta^*) \). We may estimate \( H_{\beta} \) by the sample analog:

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

and estimate \( \Omega_{\beta}^+ \) by a nonparametric kernel estimator:

\[
\hat{\Omega}_{\beta}^+ = \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \hat{\gamma}_n(h),
\]

with

\[
\hat{\gamma}_n(h) = \frac{1}{\sqrt{n}} \sum_{2 \leq t, t+h \leq n} S_t \left( \hat{F}_n, \hat{\beta} \right) S_{t+h} \left( \hat{F}_n, \hat{\beta} \right),
\]

where

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

\[
\hat{G}_j(\hat{Y}_{t-j}) = \int_{0}^{1} \int_{0}^{1} \left[ 1 \left( \hat{F}_n(\hat{Y}_{t-j}) \leq v_{2-j} \right) - v_{2-j} \right] \ell_{\beta,2-j}(v_1, v_{2-j}) c(v_1, v_{2-j}) dv_1 dv_{2-j}, j = 0, 1.
\]

We define the Wald test statistic as

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

We assume the following bandwidth condition for the consistency of covariance estimator for \( \Omega_{\beta}^+ \).

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) \( \hat{\Omega}_{\beta}^+ = \Omega_{\beta}^+ + o_p(1) \). (2) Under \( H_{01} \), \( W_n \Rightarrow \chi^2_d \) where \( d \) 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 $\hat{F}_n(\hat{Y}_t)$ against $\hat{F}_n(\hat{Y}_{t-1})$ 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 $\beta$ by $\hat{\beta}$ which maximizes

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

where $\hat{F}_n(Y_t)$ is the EDF of $Y_t$ estimated based on the filtered time series $\{\hat{Y}_t\}$ 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

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

The maximizer of $L_n(\beta)$ will converge to the pseudo-true value $\bar{\beta}$ 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,

$$ \bar{\beta} = \arg \min_{\beta} KLIC(c^*, c(\cdot, \beta)) $$

where following White (1982),

$$ KLIC(c^*, c(\cdot, \beta)) = E \log c^*(F^*(Y_{t-1}), F^*(Y_t)) - E \log c(F^*(Y_{t-1}), F^*(Y_t), \beta). $$
In the special case when the class of copula functions $C(u, v, \beta)$ is correctly specified, $C^*(u, v) = C(u, v, \tilde{\beta})$. 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 $\tilde{\beta}$. 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) $\tilde{\beta} \in \mathfrak{B}$, $\mathfrak{B}$ is a compact subset of $\mathbb{R}^k$. (2) $Q(\beta) = \mathbb{E}[\ell(F^*(Y_{t-1}), F^*(Y_t), \beta)]$ has a unique maximizer $\tilde{\beta}$ on $\mathfrak{B}$. (3) $Q(\beta)$ is Lipschitz continuous in $\beta$.

**Theorem 6.** Under Assumptions DGP, MX, ID5, M3, and X, $b = \beta + o_p(1)$.

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

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

Let $\vartheta_{\beta} = \lim_{n \to \infty} \text{Var}(\vartheta_n)$ where $\vartheta_n = n^{-1/2} \sum_{j=2}^n \ell_{\beta} (U_{j-1}, U_j, \tilde{\beta})$ and $U_t = F^*(Y_t)$,

\[
\ell_{\beta} (U_{j-1}, U_j, \tilde{\beta}) = \ell_{\beta} (U_{j-1}, U_j, \beta) + \sum_{i=0}^1 \mathbb{E} \left[ \ell_{\beta, 2-i} (U_{t-1}, U_t, \beta) | 1 (U_j \leq U_{t-i}) - U_{t-i} \right] | U_j |.
\]

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

\[
\sqrt{n} \left( \tilde{\beta}_{SP} - \beta \right) = \sqrt{n} \left( \tilde{\beta}_{SP} - \beta \right) + o_p(1) \Rightarrow N \left( 0, \varpi_\beta^{-1} \Omega_\beta \varpi_\beta^{-1} \right).
\]

Theorem 7 shows that, in the case of misspecified copula, the nonstationary filtration does not affect the limiting distribution of the semiparametric copula estimator $\tilde{\beta}_{SP}$ (centered at the pseudo-true parameter $\tilde{\beta}$), which is again normal, the same as that of the infeasible semiparametric copula estimator $\tilde{\beta}_{SP}$ 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 \( C_j(u_1, u_2, \beta_j), \ j = 1, 2 \). 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

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

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

\[
LR = \frac{c_2(F^*(y_{t-1}), F^*(y_t), \beta_2)}{c_1(F^*(y_{t-1}), F^*(y_t), \beta_1)}.
\]

If we consider the hypothesis \( H_0: \) Copula model \( C_1(u_1, u_2, \beta_1) \) is not worse than copula model \( C_2(u_1, u_2, \beta_2) \); vs. \( H_1: \) Copula model \( C_1(u_1, u_2, \beta_1) \) is worse than copula model \( C_2(u_1, u_2, \beta_2) \). 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 \( \hat{\beta}_j (j = 1, 2) \) be the semiparametric estimator \( \hat{\beta}_{SP} \) using the filtered time series \( \{\hat{Y}_t\}_{t=1}^n \) and based on model \( j \), we construct the following pseudo-likelihood-ratio (PLR) statistic

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

based on

\[
\hat{F}_n(\hat{Y}_t) = \frac{1}{n + 1} \sum_{j=1}^n 1(\hat{Y}_j \leq \hat{Y}_t).
\]

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

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

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

\[
F_n(Y_t) = \frac{1}{n + 1} \sum_{j=1}^n 1(Y_j \leq Y_t).
\]

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
\[
\Pr \{ (Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \beta_1) \neq c_2(F^*(Y_1), F^*(Y_2), \beta_2) \} > 0,
\]
where \( \beta_j \) are the pseudo-true values of the copula dependence parameters,
\[
\sqrt{n} \left( \hat{L}_R_n - L_{R_n} \right) = o_p(1).
\]
(ii) If \( \Pr \{ (Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \beta_1) = c_2(F^*(Y_1), F^*(Y_2), \beta_2) \} = 1, \)
\[
n \left( \hat{L}_R_n - L_{R_n} \right) = o_p(1).
\]

Theorem 8 shows that, under our assumptions, the limiting distribution of the pseudo-likelihood-ratio (PLR) test \( \hat{L}_R_n \) is the same as the infeasible PLR statistic \( L_{R_n} \) 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
\[
\Pr \{ (Y_1, Y_2) : c_1(F^*(Y_1), F^*(Y_2), \beta_1) \neq c_2(F^*(Y_1), F^*(Y_2), \beta_2) \} > 0,
\]
the null hypothesis \( H_0 \) is a composite hypothesis, and we may consider the least favorable configuration (LFC) which satisfies
\[
\mathbb{E} \left[ \log \frac{c_2(F^*(Y_{t-1}), F^*(Y_t), \beta_2)}{c_1(F^*(Y_{t-1}), F^*(Y_t), \beta_1)} \right] = 0.
\]
Thus, under the LFC and other regularity Assumptions,
\[
\sqrt{n} \hat{L}_R_n \Rightarrow N(0, \omega^2), \text{ as } n \to \infty,
\]
with
\[
\omega^2 = \lim \mathbb{V} \left[ \frac{1}{\sqrt{n}} \sum_{t=2}^{n} s(U_{t-1}, U_t, \beta_2, \beta_1) + \frac{2}{\sqrt{n}} \sum_{j=1}^{2} \left\{ g_{2j}(U_t, \beta_2) - g_{1j}(U_t, \beta_1) \right\} \right],
\]
where
\[
s(U_{t-1}, U_t, \beta_2, \beta_1) = \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)}, \quad U_t = F^*(Y_t),
\]
and
\[
g_{ij}(U_t, \beta_i) = \mathbb{E} \left\{ \frac{\partial \log c_i(U_{t-1}, U_t, \beta_i)}{\partial U_{t-2+j}} \left[ (1(U_t \leq U_{t-2+j}) - U_{t-2+j}) \right| U_t \right\}.
\]
Let $\omega^2$ be a consistent long-run variance estimator of $\omega^2$ based on

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

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

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

Then we may consider the following testing statistic

$$L_n = \frac{\sqrt{nLR_n}}{\hat{\omega}}.$$

Under the LFC and generalized non-nested case,

$$L_n \to N(0, 1), \text{ as } n \to \infty.$$
5. Monte Carlo Studies

In this section, we examine the finite sample performance of the parametric and semiparametric copula estimators based on filtered time series \( \{ \tilde{Y}_t \} \). We compare the sampling performance of the semiparametric estimator \( \hat{\beta}_{SP} \) with the parametric estimator \( \hat{\beta}_P \) under correct and incorrect specifications of the marginal distribution \( F^* \) (of the latent \( Y_t \)); in particular, \( \hat{\beta}_{P^*} \) signifies the \( \hat{\beta}_P \) under correct specification and \( \hat{\beta}_{P^1} \) signifies the \( \hat{\beta}_P \) under incorrect specification of \( F^* \). In addition, for comparison purpose, we also look at two infeasible copula estimators based on the true values of \( f_{Y_t} \): the infeasible parametric estimator \( e_P \) under correct specification of \( F \), and the infeasible semiparametric estimator \( e_{SP} \) using \( f_{Y_t} \) process (no filtration is needed).

DGP designs: The observed time series \( \{ Z_t \}_{t=1}^n \) is generated by \( Z_t = X_{t}^0 + Y_t \), 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

\[
G^*(y_{t-1}, y_t) = C(F^*(y_{t-1}), F^*(y_t); \beta^*).
\]

In the Monte Carlo studies, we have examined various combinations of three kinds of filtering part \( X_t^0 \), four kinds of copula functions \( C(\cdot, \cdot; \beta) \) with a range value of the copula parameter \( \beta \), and two kinds of marginal distributions \( F^* \).

Three types of \( X_t^0 \): (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 \( \beta \), and \( \Phi \) be the CDF of a univariate standard normal. The bivariate Gaussian copula is given by

\[
C(u, v; \beta) = \Phi_\beta(\Phi^{-1}(u), \Phi^{-1}(v)) = \frac{1}{2\pi\sqrt{1-\beta^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp \left\{ -\frac{(s^2 - 2\beta st + t^2)}{2(1-\beta^2)} \right\} ds dt.
\]
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:

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

(C). The Clayton copula:

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

(D) The Gumbel copula:

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

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 $\beta$. Gumbel copula has zero lower tail dependence but positive upper tail dependence $(2 - 2^{1/\beta})$ that increases with $\beta$. The overall temporal dependence in $Y_t$ measured as Kendall's tau is all increasing with copula parameter $\beta$ 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 $\hat{\beta}_{SP}, \hat{\beta}_{P}$, and $\hat{\beta}_{P1}$ use the nonstationary filtered data; and the two infeasible ones $\tilde{\beta}_{SP}$ and $\tilde{\beta}_{P}$ use the true $Y_t$ process (without filtration). Recall that $\tilde{\beta}_{SP}$ and $\tilde{\beta}_{SP}$ have the same asymptotic normal distribution, which does not depend on the filtration or the functional form of $F^*$. The infeasible $\tilde{\beta}_{P}$ 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 $\hat{\beta}_{P}$ and $\hat{\beta}_{P1}$ 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 $\hat{\beta}_{P}$ denoted by "Rmse", are reported in each table.

All the simulations reveal the following patterns. First, the semiparametric copula estimator $\hat{\beta}_{SP}$ performs well in terms of finite sample bias, variance, "Rmse" compared to the correctly specified
parametric estimator $\hat{\beta}_p$, in most situations. Second, for all the cases when there is no strong lower tail dependence, both the semiparametric copula estimator $\hat{\beta}_{SP}$ and the correctly specified parametric copula estimator $\hat{\beta}_P$ perform much better than the parametric copula estimator $\hat{\beta}_P$ using incorrectly specified parametric marginals. The parametric copula estimator for copula dependence parameter $\beta^*$ 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 $\hat{\beta}_{SP}$ 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 $\hat{\beta}_P$ and its infeasible version $\tilde{\beta}_P$ 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 $\beta^*$). In this case, the infeasible parametric copula estimator $\tilde{\beta}_P$ performs much better than the feasible parametric estimator $\hat{\beta}_P$ and the semiparametric estimators, $\hat{\beta}_{SP}$ and $\tilde{\beta}_{SP}$. The performance of $\hat{\beta}_{SP}$ 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

$$Z_t = a_0 + a_1 t + a_2 X_t + Y_t$$ \hspace{1cm} (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. 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 $\{\hat{Y}_t\}$ 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 $\{\hat{Y}_t\}$. 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:

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

where $\bar{u} = 1 - u$, $\bar{v} = 1 - v$, $\beta = (\beta_1, \beta_2)'$ 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

Figure 6.1: Scatter Plot of the standardized GNP residuals

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

When \( \beta_1 \to 0 \), the Joe-Clayton copula approaches the Joe copula whose upper tail dependence increases as \( \beta_2 \) 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 \( \beta_1 \) and \( \beta_2 \). We first test the lower tail dependence based on \( \beta_1 \). The estimated value of \( \beta_1 \) 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 \( \beta_2 \) is 2.765, and the corresponding t-statistic is 5.36 (p-value < 0.1%). We reject the null hypothesis \( H_0 : \beta_2 = 1 \) 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 \( \hat{\beta}_1 \). 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 \( \hat{\beta}_2 \) is 1.356. Corresponding to the null hypothesis \( H_0 : \beta_2 = 1 \), 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
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


A. Appendix A: Monte Carlo Results

In the Monte Carlo studies, we have examined various DGPs that are different combinations of three kinds of filtering part $X_t^t\pi^*$, four kinds of copula functions $C(\cdot, \cdot; \beta)$ with a range value of the copula parameter $\beta$, 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 $\hat{\beta}_{P^*}$) are reported in each table.

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

Recall that $\hat{\beta}_{SP}$ and $\hat{\beta}_{SP}$ have the same asymptotic normal distribution, which does not depend on any filtration and the specification of $F^*$. The infeasible $\hat{\beta}_{P^*}$ is asymptotically normal, with the limiting distribution independent of the filtration but does depend on the parametric estimation of $F^*$. The limiting distributions of $\hat{\beta}_{P^*}$ and $\hat{\beta}_{P1}$ 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 determinstic 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 for the determinstic 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-$t(3)$ distribution. 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))

<table>
<thead>
<tr>
<th></th>
<th>$n = 500$</th>
<th></th>
<th>$n = 2000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^*$</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0066</td>
<td>-0.0077</td>
<td>-0.0063</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.0391</td>
<td>0.0438</td>
<td>0.0462</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td>1.1224</td>
<td>1.0912</td>
<td>1.0613</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>0.0004</td>
<td>-0.0014</td>
<td>-0.0035</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.0374</td>
<td>0.0425</td>
<td>0.0452</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>1.1224</td>
<td>1.0912</td>
<td>1.0613</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Bias</td>
<td>-0.0046</td>
<td>-0.0151</td>
<td>-0.0193</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Std</td>
<td>0.0721</td>
<td>0.0835</td>
<td>0.0911</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Rmse</td>
<td>3.7261</td>
<td>3.9751</td>
<td>4.2273</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0065</td>
<td>-0.0071</td>
<td>-0.0053</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.0388</td>
<td>0.0436</td>
<td>0.0461</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td>1.1224</td>
<td>1.0912</td>
<td>1.0613</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>0.0002</td>
<td>-0.0007</td>
<td>-0.0014</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.0370</td>
<td>0.0423</td>
<td>0.0450</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>0.9758</td>
<td>0.9873</td>
<td>0.9889</td>
</tr>
</tbody>
</table>

$\hat{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE

$\hat{\beta}_{P^*}$ MSE / $\hat{\beta}_{P^*}$ MSE

$\hat{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE

$\hat{\beta}_{P^*}$ MSE / $\hat{\beta}_{P^*}$ MSE
Table 1B: Trending Time Series, Frank Copula

(True marginal is student t(3))

<table>
<thead>
<tr>
<th>$n$ = 500</th>
<th>$\beta^*$</th>
<th>-5</th>
<th>-3</th>
<th>-1</th>
<th>1</th>
<th>3</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0115</td>
<td>-0.0229</td>
<td>-0.0242</td>
<td>-0.0310</td>
<td>-0.0591</td>
<td>-0.1280</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.4025</td>
<td>0.3230</td>
<td>0.2812</td>
<td>0.2812</td>
<td>0.3194</td>
<td>0.3925</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td><strong>1.2118</strong></td>
<td><strong>1.1066</strong></td>
<td><strong>1.0170</strong></td>
<td><strong>1.0207</strong></td>
<td><strong>1.1254</strong></td>
<td><strong>1.2741</strong></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>0.0393</td>
<td>0.0093</td>
<td>-0.0103</td>
<td>-0.0288</td>
<td>-0.0581</td>
<td>-0.1116</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.3637</td>
<td>0.3077</td>
<td>0.2797</td>
<td>0.2785</td>
<td>0.3006</td>
<td>0.3483</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Bias</td>
<td>-1.5653</td>
<td>-1.3416</td>
<td>-0.8315</td>
<td>0.7674</td>
<td>1.2818</td>
<td>1.4765</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Std</td>
<td>1.1554</td>
<td>1.1182</td>
<td>1.1144</td>
<td>1.1915</td>
<td>1.2066</td>
<td>1.2242</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Rmse</td>
<td><strong>28.2919</strong></td>
<td><strong>32.1860</strong></td>
<td><strong>24.6847</strong></td>
<td><strong>25.6159</strong></td>
<td><strong>33.0572</strong></td>
<td><strong>27.5063</strong></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0330</td>
<td>-0.0307</td>
<td>-0.0232</td>
<td>-0.0218</td>
<td>-0.0362</td>
<td>-0.0764</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.3973</td>
<td>0.3209</td>
<td>0.2799</td>
<td>0.2809</td>
<td>0.3192</td>
<td>0.3915</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td><strong>1.1879</strong></td>
<td><strong>1.0963</strong></td>
<td><strong>1.0075</strong></td>
<td><strong>1.0124</strong></td>
<td><strong>1.1010</strong></td>
<td><strong>1.1896</strong></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>-0.0144</td>
<td>-0.0134</td>
<td>-0.0108</td>
<td>-0.0092</td>
<td>-0.0112</td>
<td>-0.0128</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.3489</td>
<td>0.3022</td>
<td>0.2776</td>
<td>0.2778</td>
<td>0.3003</td>
<td>0.3454</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td><strong>0.9114</strong></td>
<td><strong>0.9658</strong></td>
<td><strong>0.9857</strong></td>
<td><strong>0.9854</strong></td>
<td><strong>0.9634</strong></td>
<td><strong>0.8935</strong></td>
<td></td>
</tr>
</tbody>
</table>

$n = 500$

| $\hat{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE | 0.9803 | 0.9907 | 0.9907 | 0.9919 | 0.9783 | 0.9336 |
| $\hat{\beta}_{P^*}$ MSE / $\hat{\beta}_{P^*}$ MSE | 0.9114 | 0.9658 | 0.9857 | 0.9854 | 0.9634 | 0.8935 |

$n = 2000$

| $\hat{\beta}_{SP}$ MSE / $\hat{\beta}_{SP}$ MSE | 0.9935 | 0.9985 | 0.9992 | 0.9993 | 0.9975 | 0.9875 |
| $\hat{\beta}_{P^*}$ 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))

<table>
<thead>
<tr>
<th>$n = 500$</th>
<th>$\beta^*$</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_S$ Bias</td>
<td>0.0012</td>
<td>-0.0307</td>
<td>-0.1672</td>
<td>-0.7897</td>
<td>-1.8797</td>
<td>-3.2800</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_S$ Std</td>
<td>0.1040</td>
<td>0.1989</td>
<td>0.4486</td>
<td>0.9392</td>
<td>1.2412</td>
<td>1.4254</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_S$ Rmse</td>
<td>1.3184</td>
<td>1.4836</td>
<td>1.4314</td>
<td>1.2141</td>
<td>1.7435</td>
<td>2.3035</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_p$ Bias</td>
<td>-0.0098</td>
<td>-0.0217</td>
<td>-0.0787</td>
<td>-0.3700</td>
<td>-0.9417</td>
<td>-1.6985</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_p$ Std</td>
<td>0.0900</td>
<td>0.1638</td>
<td>0.3923</td>
<td>1.0504</td>
<td>1.4224</td>
<td>1.6333</td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_p$ Rmse</td>
<td>1.3184</td>
<td>1.4836</td>
<td>1.4314</td>
<td>1.2141</td>
<td>1.7435</td>
<td>2.3035</td>
<td></td>
</tr>
</tbody>
</table>

| $n = 2000$ | $\hat{\beta}_S$ Bias | -0.0026 | -0.0069  | -0.0171 | -0.0257 | -0.0240 | -0.0160 |
| $\hat{\beta}_p$ Std  | 0.0854 | 0.1343 | 0.2602 | 0.6389 | 1.1813 | 1.7828 |
| $\hat{\beta}_p$ Rmse | 0.8896 | 0.6621 | 0.4246 | 0.3296 | 0.4797 | 0.5725 |

| $\hat{\beta}_S$ MSE / $\hat{\beta}_S$ MSE | 0.9784 | 0.9122 | 0.9215 | 0.9289 | 0.8635 | 0.8092 |
| $\hat{\beta}_p$ MSE / $\hat{\beta}_p$ MSE | 0.8896 | 0.6621 | 0.4246 | 0.3296 | 0.4797 | 0.5725 |

| $n = 36$ | $\hat{\beta}_S$ Bias | -0.0026 | -0.0069 | -0.0171 | -0.0257 | -0.0240 | -0.0160 |
| $\hat{\beta}_p$ Std  | 0.0854 | 0.1343 | 0.2602 | 0.6389 | 1.1813 | 1.7828 |
| $\hat{\beta}_p$ Rmse | 0.8896 | 0.6621 | 0.4246 | 0.3296 | 0.4797 | 0.5725 |

| $\hat{\beta}_S$ MSE / $\hat{\beta}_S$ MSE | 0.9948 | 0.9832 | 0.9577 | 0.9464 | 0.9520 | 0.9331 |
| $\hat{\beta}_p$ 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))

\[ n = 500 \]

<table>
<thead>
<tr>
<th>( \hat{\beta}_{SP} ) Bias</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0379</td>
<td>-0.1785</td>
<td>-0.4513</td>
<td>-0.8697</td>
<td>-1.4093</td>
<td>-2.0454</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1666</td>
<td>0.3793</td>
<td>0.5882</td>
<td>0.7423</td>
<td>0.8490</td>
<td>0.9330</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.0719</td>
<td>1.0647</td>
<td>1.1286</td>
<td>1.3556</td>
<td>1.7370</td>
<td>2.1476</td>
</tr>
<tr>
<td>( \hat{\beta}_{p*} ) Bias</td>
<td>-0.0236</td>
<td>-0.0907</td>
<td>-0.2292</td>
<td>-0.4523</td>
<td>-0.7562</td>
<td>-1.1173</td>
</tr>
<tr>
<td>( \hat{\beta}_{p*} ) Std</td>
<td>0.1633</td>
<td>0.3960</td>
<td>0.6592</td>
<td>0.8717</td>
<td>0.9932</td>
<td>1.0512</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Rmse</td>
<td>1.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>-0.0321</td>
<td>-0.1540</td>
<td>-0.3861</td>
<td>-0.7354</td>
<td>-1.1963</td>
<td>-1.7464</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1596</td>
<td>0.3512</td>
<td>0.5534</td>
<td>0.7335</td>
<td>0.8846</td>
<td>1.0121</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>0.9732</td>
<td>0.8909</td>
<td>0.9349</td>
<td>1.1187</td>
<td>1.4204</td>
<td>1.7311</td>
</tr>
<tr>
<td>( \hat{\beta}_{p*} ) Bias</td>
<td>-0.0066</td>
<td>-0.0225</td>
<td>-0.0533</td>
<td>-0.0962</td>
<td>-0.1456</td>
<td>-0.1927</td>
</tr>
<tr>
<td>( \hat{\beta}_{p*} ) Std</td>
<td>0.1264</td>
<td>0.2810</td>
<td>0.4848</td>
<td>0.7297</td>
<td>1.0384</td>
<td>1.4401</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Rmse</td>
<td>0.5887</td>
<td>0.4815</td>
<td>0.4883</td>
<td>0.5618</td>
<td>0.7054</td>
<td>0.8971</td>
</tr>
</tbody>
</table>

\[ n = 500 \]

\[ n = 2000 \]
Table 2A: Trending Time Series, Gaussian Copula

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

<table>
<thead>
<tr>
<th>$\beta^*$</th>
<th>-0.5</th>
<th>-0.3</th>
<th>-0.1</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0062</td>
<td>-0.0074</td>
<td>-0.0059</td>
<td>-0.0037</td>
<td>-0.0028</td>
<td>-0.0046</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.0387</td>
<td>0.0436</td>
<td>0.0463</td>
<td>0.0466</td>
<td>0.0447</td>
<td>0.0404</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td>1.3211</td>
<td>1.0519</td>
<td>0.9589</td>
<td>0.9521</td>
<td>0.9309</td>
<td>0.9054</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>-0.0053</td>
<td>-0.0078</td>
<td>-0.0068</td>
<td>-0.0006</td>
<td>0.0083</td>
<td>0.0147</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.0337</td>
<td>0.0425</td>
<td>0.0472</td>
<td>0.0479</td>
<td>0.0456</td>
<td>0.0401</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Bias</td>
<td>0.0897</td>
<td>0.0437</td>
<td>0.0079</td>
<td>-0.0181</td>
<td>-0.0344</td>
<td>-0.0414</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Std</td>
<td>0.0302</td>
<td>0.0371</td>
<td>0.0431</td>
<td>0.0476</td>
<td>0.0496</td>
<td>0.0479</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Rmse</td>
<td>7.7163</td>
<td>1.7650</td>
<td>0.8457</td>
<td>1.1262</td>
<td>1.6895</td>
<td>2.1902</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0065</td>
<td>-0.0071</td>
<td>-0.0053</td>
<td>-0.0027</td>
<td>-0.0013</td>
<td>-0.0024</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.0388</td>
<td>0.0436</td>
<td>0.0461</td>
<td>0.0463</td>
<td>0.0442</td>
<td>0.0397</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td>1.3371</td>
<td>1.0460</td>
<td>0.9483</td>
<td>0.9371</td>
<td>0.9077</td>
<td>0.8639</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>0.0044</td>
<td>0.0029</td>
<td>0.0000</td>
<td>-0.0036</td>
<td>-0.0063</td>
<td>-0.0074</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.0320</td>
<td>0.0400</td>
<td>0.0444</td>
<td>0.0446</td>
<td>0.0404</td>
<td>0.0324</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>0.9013</td>
<td>0.8646</td>
<td>0.8679</td>
<td>0.8705</td>
<td>0.7763</td>
<td>0.6047</td>
</tr>
</tbody>
</table>
Table 2B: Trending Time Series, Frank Copula

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

\[
\begin{array}{lcccccc}
\beta^* & -5 & -3 & -1 & 1 & 3 & 5 \\
\hline
\hat{\beta}_{SP} \text{ Bias} & -0.0297 & -0.0344 & -0.0297 & -0.0296 & -0.0440 & -0.0851 \\
\hat{\beta}_{SP} \text{ Std} & 0.3970 & 0.3214 & 0.2809 & 0.2819 & 0.3222 & 0.4001 \\
\hat{\beta}_{SP} \text{ Rmse} & 1.3150 & 1.0811 & 0.9519 & 0.9623 & 0.8341 & 0.6380 \\
\hline
\hat{\beta}_{P^*} \text{ Bias} & -0.0425 & -0.0523 & -0.0433 & 0.0036 & 0.0988 & 0.2274 \\
\hat{\beta}_{P^*} \text{ Std} & 0.3445 & 0.3065 & 0.2863 & 0.2889 & 0.3421 & 0.4589 \\
\hat{\beta}_{P^*} \text{ Rmse} & 1 & 1 & 1 & 1 & 1 & 1 \\
\hline
\hat{\beta}_{P_1} \text{ Bias} & 0.4944 & 0.0962 & 0.0035 & 0.1712 & 0.3759 & 0.5257 \\
\hat{\beta}_{P_1} \text{ Std} & 0.3021 & 0.2970 & 0.3018 & 0.3392 & 0.4140 & 0.5400 \\
\hat{\beta}_{P_1} \text{ Rmse} & 2.7855 & 1.0084 & 1.0861 & 1.7296 & 2.4664 & 2.1656 \\
\hline
\hat{\beta}_{SP} \text{ Bias} & -0.0330 & -0.0307 & -0.0232 & -0.0218 & -0.0362 & -0.0764 \\
\hat{\beta}_{SP} \text{ Std} & 0.3973 & 0.3209 & 0.2799 & 0.2809 & 0.3192 & 0.3915 \\
\hat{\beta}_{SP} \text{ Rmse} & 1.3188 & 1.0747 & 0.9411 & 0.9508 & 0.8140 & 0.6066 \\
\hline
\hat{\beta}_{P^*} \text{ Bias} & 0.0033 & -0.0013 & -0.0065 & -0.0132 & -0.0208 & -0.0255 \\
\hat{\beta}_{P^*} \text{ Std} & 0.3370 & 0.2967 & 0.2764 & 0.2762 & 0.2943 & 0.3336 \\
\hat{\beta}_{P^*} \text{ Rmse} & 0.9423 & 0.9108 & 0.9114 & 0.9158 & 0.6866 & 0.4267 \\
\end{array}
\]
Table 2C: Trending Time Series, Clayton Copula

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

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{SP} ) Bias</td>
<td>-0.0077</td>
<td>-0.0524</td>
<td>-0.2290</td>
<td>-0.9035</td>
<td>-1.9578</td>
<td>-3.2889</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
<td>0.1014</td>
<td>0.1830</td>
<td>0.4007</td>
<td>0.8853</td>
<td>1.2933</td>
<td>1.5443</td>
</tr>
<tr>
<td>( \beta_{SP} ) Rmse</td>
<td><strong>0.8758</strong></td>
<td><strong>1.0248</strong></td>
<td><strong>1.2213</strong></td>
<td><strong>1.2928</strong></td>
<td><strong>1.2684</strong></td>
<td><strong>1.1733</strong></td>
</tr>
<tr>
<td>( \beta_{P*} ) Bias</td>
<td>0.0022</td>
<td>-0.0198</td>
<td>-0.1264</td>
<td>-0.5526</td>
<td>-1.2366</td>
<td>-2.0305</td>
</tr>
<tr>
<td>( \beta_{P*} ) Std</td>
<td>0.1086</td>
<td>0.1870</td>
<td>0.3981</td>
<td>0.9655</td>
<td>1.6767</td>
<td>2.6700</td>
</tr>
<tr>
<td>( \beta_{P*} ) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \beta_{P1} ) Bias</td>
<td>0.6251</td>
<td>0.7053</td>
<td>0.7347</td>
<td>0.6051</td>
<td>0.3685</td>
<td>-0.0129</td>
</tr>
<tr>
<td>( \beta_{P1} ) Std</td>
<td>0.1651</td>
<td>0.2284</td>
<td>0.4478</td>
<td>1.1839</td>
<td>2.3474</td>
<td>3.5508</td>
</tr>
<tr>
<td>( \beta_{P1} ) Rmse</td>
<td><strong>35.4067</strong></td>
<td><strong>15.5463</strong></td>
<td><strong>4.2483</strong></td>
<td><strong>1.4283</strong></td>
<td><strong>1.3008</strong></td>
<td><strong>1.1205</strong></td>
</tr>
<tr>
<td>( \beta_{SP} ) Bias</td>
<td>0.0016</td>
<td>-0.0256</td>
<td>-0.1415</td>
<td>-0.6389</td>
<td>-1.5373</td>
<td>-2.7485</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
<td>0.1028</td>
<td>0.1905</td>
<td>0.4373</td>
<td>1.0141</td>
<td>1.4205</td>
<td>1.6720</td>
</tr>
<tr>
<td>( \beta_{SP} ) Rmse</td>
<td><strong>0.8959</strong></td>
<td><strong>1.0454</strong></td>
<td><strong>1.2109</strong></td>
<td><strong>1.1607</strong></td>
<td><strong>1.0093</strong></td>
<td><strong>0.9198</strong></td>
</tr>
<tr>
<td>( \beta_{P*} ) Bias</td>
<td>-0.0327</td>
<td>-0.0773</td>
<td>-0.2062</td>
<td>-0.6221</td>
<td>-1.2212</td>
<td>-1.9876</td>
</tr>
<tr>
<td>( \beta_{P*} ) Std</td>
<td>0.0851</td>
<td>0.1402</td>
<td>0.2823</td>
<td>0.6896</td>
<td>1.2753</td>
<td>1.8589</td>
</tr>
<tr>
<td>( \beta_{P*} ) Rmse</td>
<td><strong>0.7039</strong></td>
<td><strong>0.7254</strong></td>
<td><strong>0.7007</strong></td>
<td><strong>0.6969</strong></td>
<td><strong>0.7183</strong></td>
<td><strong>0.6582</strong></td>
</tr>
<tr>
<td>$\beta^*$</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0217</td>
<td>-0.1278</td>
<td>-0.3610</td>
<td>-0.7509</td>
<td>-1.2756</td>
<td>-1.9110</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.1736</td>
<td>0.4040</td>
<td>0.6410</td>
<td>0.8087</td>
<td>0.9238</td>
<td>1.0039</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td><strong>0.9308</strong></td>
<td><strong>0.9498</strong></td>
<td><strong>1.0090</strong></td>
<td><strong>1.1762</strong></td>
<td><strong>1.6286</strong></td>
<td><strong>2.4850</strong></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>0.1061</td>
<td>0.2632</td>
<td>0.4169</td>
<td>0.5329</td>
<td>0.5779</td>
<td>0.5270</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.1471</td>
<td>0.3461</td>
<td>0.6021</td>
<td>0.8668</td>
<td>1.0905</td>
<td>1.2639</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Bias</td>
<td>-0.1716</td>
<td>-0.2440</td>
<td>-0.4207</td>
<td>-0.7133</td>
<td>-1.1187</td>
<td>-1.6247</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Std</td>
<td>0.2353</td>
<td>0.5360</td>
<td>0.8422</td>
<td>1.1149</td>
<td>1.3327</td>
<td>1.4940</td>
</tr>
<tr>
<td>$\hat{\beta}_{P_1}$ Rmse</td>
<td><strong>2.5773</strong></td>
<td><strong>1.8340</strong></td>
<td><strong>1.6526</strong></td>
<td><strong>1.6922</strong></td>
<td><strong>1.9876</strong></td>
<td><strong>2.5980</strong></td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Bias</td>
<td>-0.0321</td>
<td>-0.1540</td>
<td>-0.3861</td>
<td>-0.7354</td>
<td>-1.1963</td>
<td>-1.7464</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Std</td>
<td>0.1596</td>
<td>0.3512</td>
<td>0.5534</td>
<td>0.7335</td>
<td>0.8846</td>
<td>1.0121</td>
</tr>
<tr>
<td>$\hat{\beta}_{SP}$ Rmse</td>
<td><strong>0.8052</strong></td>
<td><strong>0.7776</strong></td>
<td><strong>0.8489</strong></td>
<td><strong>1.0421</strong></td>
<td><strong>1.4532</strong></td>
<td><strong>2.1726</strong></td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Bias</td>
<td>-0.0091</td>
<td>-0.0234</td>
<td>-0.0334</td>
<td>-0.0305</td>
<td>-0.0072</td>
<td>0.0184</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Std</td>
<td>0.0758</td>
<td>0.1225</td>
<td>0.2694</td>
<td>0.5207</td>
<td>0.8738</td>
<td>1.2924</td>
</tr>
<tr>
<td>$\hat{\beta}_{P^*}$ Rmse</td>
<td><strong>0.1773</strong></td>
<td><strong>0.0822</strong></td>
<td><strong>0.1374</strong></td>
<td><strong>0.2628</strong></td>
<td><strong>0.5013</strong></td>
<td><strong>0.8909</strong></td>
</tr>
</tbody>
</table>
Table 3A: Unit Root Time Series, Gaussian Copula
(True marginal is student-t(3), n = 500)

<table>
<thead>
<tr>
<th>(\beta^*)</th>
<th>-0.5</th>
<th>-0.3</th>
<th>-0.1</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>0.0032</td>
<td>-0.0015</td>
<td>-0.0022</td>
<td>-0.0010</td>
<td>-0.0005</td>
<td>-0.0020</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.0413</td>
<td>0.0444</td>
<td>0.0464</td>
<td>0.0464</td>
<td>0.0443</td>
<td>0.0398</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td>0.9609</td>
<td>1.0487</td>
<td>1.0587</td>
<td>1.0552</td>
<td>1.0651</td>
<td>1.0977</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Bias</td>
<td>0.0149</td>
<td>0.0072</td>
<td>0.0024</td>
<td>-0.0010</td>
<td>-0.0036</td>
<td>-0.0054</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Std</td>
<td>0.0396</td>
<td>0.0428</td>
<td>0.0451</td>
<td>0.0452</td>
<td>0.0428</td>
<td>0.0376</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Bias</td>
<td>0.0068</td>
<td>-0.0072</td>
<td>-0.0130</td>
<td>0.0132</td>
<td>0.0094</td>
<td>-0.0024</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Std</td>
<td>0.0738</td>
<td>0.0844</td>
<td>0.0918</td>
<td>0.0945</td>
<td>0.0869</td>
<td>0.0720</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Rmse</td>
<td>3.0701</td>
<td>3.8195</td>
<td>4.2210</td>
<td>4.4582</td>
<td>4.1482</td>
<td>3.5967</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>-0.0065</td>
<td>-0.0071</td>
<td>-0.0053</td>
<td>-0.0027</td>
<td>-0.0013</td>
<td>-0.0024</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.0388</td>
<td>0.0436</td>
<td>0.0461</td>
<td>0.0463</td>
<td>0.0442</td>
<td>0.0397</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td>0.8674</td>
<td>1.0368</td>
<td>1.0589</td>
<td>1.0549</td>
<td>1.0615</td>
<td>1.0943</td>
</tr>
</tbody>
</table>

Table 3B: Unit Root Time Series, Frank Copula
(True marginal is student-t(3), n = 500)

<table>
<thead>
<tr>
<th>(\beta^*)</th>
<th>-5</th>
<th>-3</th>
<th>-1</th>
<th>1</th>
<th>3</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>0.1320</td>
<td>0.0370</td>
<td>0.0026</td>
<td>-0.0118</td>
<td>-0.0312</td>
<td>-0.0746</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.4599</td>
<td>0.3355</td>
<td>0.2831</td>
<td>0.2819</td>
<td>0.3205</td>
<td>0.3926</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td>0.9452</td>
<td>1.0435</td>
<td>1.0200</td>
<td>1.0293</td>
<td>1.1367</td>
<td>1.3053</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Bias</td>
<td>0.2276</td>
<td>0.0858</td>
<td>0.0239</td>
<td>-0.0032</td>
<td>-0.0219</td>
<td>-0.0444</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Std</td>
<td>0.4363</td>
<td>0.3190</td>
<td>0.2793</td>
<td>0.2781</td>
<td>0.3012</td>
<td>0.3469</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Bias</td>
<td>-1.3618</td>
<td>-1.2542</td>
<td>-0.7833</td>
<td>0.8126</td>
<td>1.3305</td>
<td>1.5537</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Std</td>
<td>1.3053</td>
<td>1.2081</td>
<td>1.1563</td>
<td>1.1914</td>
<td>1.2061</td>
<td>1.2220</td>
</tr>
<tr>
<td>(\hat{\beta}_{P_1}) Rmse</td>
<td>14.6941</td>
<td>27.7834</td>
<td>24.8172</td>
<td>26.8892</td>
<td>35.3614</td>
<td>31.9379</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>-0.0330</td>
<td>-0.0307</td>
<td>-0.0232</td>
<td>-0.0218</td>
<td>-0.0362</td>
<td>-0.0764</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.3973</td>
<td>0.3209</td>
<td>0.2799</td>
<td>0.2809</td>
<td>0.3192</td>
<td>0.3915</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td>0.6563</td>
<td>0.9518</td>
<td>1.0039</td>
<td>1.0264</td>
<td>1.1317</td>
<td>1.3005</td>
</tr>
</tbody>
</table>
### Table 3C: Unit Root Time Series, Clayton Copula

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

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>0.0029</td>
<td>-0.0238</td>
<td>-0.1400</td>
<td>-0.6490</td>
<td>-1.5641</td>
<td>-2.7850</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1032</td>
<td>0.1930</td>
<td>0.4410</td>
<td>1.0001</td>
<td>1.3963</td>
<td>1.6425</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.4129</td>
<td>1.7608</td>
<td>2.0309</td>
<td>1.7618</td>
<td>1.7485</td>
<td>2.1501</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Bias</td>
<td>-0.0044</td>
<td>-0.0137</td>
<td>-0.0504</td>
<td>-0.2014</td>
<td>-0.4862</td>
<td>-0.9244</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Std</td>
<td>0.0868</td>
<td>0.1459</td>
<td>0.3207</td>
<td>0.8753</td>
<td>1.5092</td>
<td>2.0019</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Rmse</td>
<td>1 1 1 1 1 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Bias</td>
<td>-0.0623</td>
<td>0.0084</td>
<td>0.1702</td>
<td>0.2957</td>
<td>0.1473</td>
<td>-0.1913</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Std</td>
<td>0.4181</td>
<td>0.5283</td>
<td>0.6247</td>
<td>0.9293</td>
<td>1.2528</td>
<td>1.6933</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Rmse</td>
<td>23.6719</td>
<td>12.9987</td>
<td>3.9770</td>
<td>1.1788</td>
<td>0.6329</td>
<td>0.5972</td>
</tr>
</tbody>
</table>

### Table 3D: Unit Root Time Series, Gumbel Copula

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

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>-0.0294</td>
<td>-0.1470</td>
<td>-0.3747</td>
<td>-0.7229</td>
<td>-1.1864</td>
<td>-1.7400</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1641</td>
<td>0.3615</td>
<td>0.5748</td>
<td>0.7517</td>
<td>0.8779</td>
<td>0.9840</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.3930</td>
<td>1.4290</td>
<td>1.4408</td>
<td>1.3654</td>
<td>1.4689</td>
<td>1.6783</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Bias</td>
<td>-0.0174</td>
<td>-0.0569</td>
<td>-0.1378</td>
<td>-0.2572</td>
<td>-0.4252</td>
<td>-0.6287</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Std</td>
<td>0.1404</td>
<td>0.3215</td>
<td>0.5548</td>
<td>0.8546</td>
<td>1.1411</td>
<td>1.4091</td>
</tr>
<tr>
<td>( \hat{\beta}_p ) Rmse</td>
<td>1 1 1 1 1 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Bias</td>
<td>0.1259</td>
<td>0.1172</td>
<td>0.0386</td>
<td>-0.1034</td>
<td>-0.3119</td>
<td>-0.5863</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Std</td>
<td>0.3842</td>
<td>0.5646</td>
<td>0.8089</td>
<td>1.0408</td>
<td>1.2631</td>
<td>1.4861</td>
</tr>
<tr>
<td>( \hat{\beta}_{p1} ) Rmse</td>
<td>8.1965</td>
<td>3.1196</td>
<td>2.0069</td>
<td>1.3733</td>
<td>1.1414</td>
<td>1.0719</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>-0.0321</td>
<td>-0.1540</td>
<td>-0.3861</td>
<td>-0.7354</td>
<td>-1.1963</td>
<td>-1.7464</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1596</td>
<td>0.3512</td>
<td>0.5534</td>
<td>0.7335</td>
<td>0.8846</td>
<td>1.0121</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.3284</td>
<td>1.3795</td>
<td>1.3933</td>
<td>1.3545</td>
<td>1.4927</td>
<td>1.7112</td>
</tr>
</tbody>
</table>

43
Table 4A: Unit Root Time Series, Gaussian Copula

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

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>-0.5</th>
<th>-0.3</th>
<th>-0.1</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_S ) Bias</td>
<td>0.0049</td>
<td>0.0010</td>
<td>-0.0003</td>
<td>0.0001</td>
<td>0.0001</td>
<td>-0.0017</td>
</tr>
<tr>
<td>( \hat{\beta}_S ) Std</td>
<td>0.0421</td>
<td>0.0447</td>
<td>0.0462</td>
<td>0.0463</td>
<td>0.0442</td>
<td>0.0398</td>
</tr>
<tr>
<td>( \hat{\beta}_S ) Rmse</td>
<td>1.6123</td>
<td>1.1434</td>
<td>0.9912</td>
<td>0.9845</td>
<td>1.0668</td>
<td>1.2028</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Bias</td>
<td>0.0026</td>
<td>0.0004</td>
<td>0.0017</td>
<td>0.0027</td>
<td>0.0029</td>
<td>0.0029</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Std</td>
<td>0.0333</td>
<td>0.0418</td>
<td>0.0463</td>
<td>0.0466</td>
<td>0.0427</td>
<td>0.0362</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

| \( \hat{\beta}_{P1} \) 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 |

Table 4B: Unit Root Time Series, Frank Copula

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

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>-5</th>
<th>-3</th>
<th>-1</th>
<th>1</th>
<th>3</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_S ) Bias</td>
<td>0.1025</td>
<td>0.0325</td>
<td>0.0014</td>
<td>-0.0109</td>
<td>-0.0275</td>
<td>-0.0624</td>
</tr>
<tr>
<td>( \hat{\beta}_S ) Std</td>
<td>0.4346</td>
<td>0.3291</td>
<td>0.2801</td>
<td>0.2815</td>
<td>0.3201</td>
<td>0.3923</td>
</tr>
<tr>
<td>( \hat{\beta}_S ) Rmse</td>
<td>1.5689</td>
<td>1.1906</td>
<td>0.9808</td>
<td>0.9860</td>
<td>1.0704</td>
<td>1.0627</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Bias</td>
<td>0.0513</td>
<td>-0.0012</td>
<td>0.0002</td>
<td>0.0144</td>
<td>0.0327</td>
<td>0.0735</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Std</td>
<td>0.3528</td>
<td>0.3031</td>
<td>0.2828</td>
<td>0.2833</td>
<td>0.3088</td>
<td>0.3783</td>
</tr>
<tr>
<td>( \hat{\beta}_P ) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Bias</td>
<td>0.5930</td>
<td>0.1565</td>
<td>0.0413</td>
<td>0.2045</td>
<td>0.4112</td>
<td>0.5774</td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Std</td>
<td>0.5355</td>
<td>0.4057</td>
<td>0.3297</td>
<td>0.3397</td>
<td>0.4119</td>
<td>0.5258</td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Rmse</td>
<td>5.0235</td>
<td>2.0582</td>
<td>1.3803</td>
<td>1.9540</td>
<td>3.5128</td>
<td>4.1070</td>
</tr>
</tbody>
</table>

| \( \hat{\beta}_S \) Bias | -0.0330 | -0.0307 | -0.0232 | -0.0218 | -0.0362 | -0.0764 |
| \( \hat{\beta}_S \) Std | 0.3973 | 0.3209 | 0.2799 | 0.2809 | 0.3192 | 0.3915 |
| \( \hat{\beta}_S \) 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, \( n = 500 \))

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>0.0030</td>
<td>-0.0260</td>
<td>-0.1464</td>
<td>-0.6391</td>
<td>-1.4513</td>
<td>-2.5290</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1030</td>
<td>0.1901</td>
<td>0.4360</td>
<td>1.0528</td>
<td>1.7108</td>
<td>2.3180</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.1142</td>
<td>1.3112</td>
<td>1.4351</td>
<td>1.3368</td>
<td>1.2267</td>
<td>1.1781</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Bias</td>
<td>-0.0068</td>
<td>-0.0431</td>
<td>-0.1549</td>
<td>-0.5338</td>
<td>-1.1085</td>
<td>-1.8549</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Std</td>
<td>0.0973</td>
<td>0.1619</td>
<td>0.3513</td>
<td>0.9218</td>
<td>1.6954</td>
<td>2.5592</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Rmse</td>
<td>1.1142</td>
<td>1.3112</td>
<td>1.4351</td>
<td>1.3368</td>
<td>1.2267</td>
<td>1.1781</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Bias</td>
<td>0.6387</td>
<td>0.7224</td>
<td>0.7678</td>
<td>0.7159</td>
<td>0.6593</td>
<td>0.5805</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Std</td>
<td>0.1603</td>
<td>0.2091</td>
<td>0.3837</td>
<td>1.0003</td>
<td>2.1043</td>
<td>3.3443</td>
<td></td>
</tr>
<tr>
<td>( \hat{\beta}_{P1} ) Rmse</td>
<td>45.5370</td>
<td>20.1466</td>
<td>4.9984</td>
<td>1.3336</td>
<td>1.1852</td>
<td>1.1532</td>
<td></td>
</tr>
</tbody>
</table>

Table 4D: Unit Root Time Series, Gumbel Copula
(True marginal is re-centered Chi-square with d.f. 3, \( n = 500 \))

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>-0.0243</td>
<td>-0.1264</td>
<td>-0.3328</td>
<td>-0.6624</td>
<td>-1.1074</td>
<td>-1.6450</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1645</td>
<td>0.3706</td>
<td>0.5923</td>
<td>0.7663</td>
<td>0.8860</td>
<td>0.9805</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>1.5436</td>
<td>1.7158</td>
<td>1.8271</td>
<td>1.8169</td>
<td>2.0074</td>
<td>2.3653</td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Bias</td>
<td>0.0432</td>
<td>0.1260</td>
<td>0.2160</td>
<td>0.3035</td>
<td>0.3676</td>
<td>0.3911</td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Std</td>
<td>0.1266</td>
<td>0.2711</td>
<td>0.4538</td>
<td>0.6875</td>
<td>0.9310</td>
<td>1.1822</td>
</tr>
<tr>
<td>( \hat{\beta}_{P} ) Rmse</td>
<td>1.1108</td>
<td>1.3163</td>
<td>1.4329</td>
<td>1.2661</td>
<td>1.0677</td>
<td>1.0360</td>
</tr>
</tbody>
</table>
Table 5A: Cointegrated Time Series, Gaussian Copula

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

<table>
<thead>
<tr>
<th>β*</th>
<th>-0.5</th>
<th>-0.3</th>
<th>-0.1</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>-0.0066</td>
<td>-0.0074</td>
<td>-0.0058</td>
<td>-0.0034</td>
<td>-0.0023</td>
<td>-0.0037</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.0388</td>
<td>0.0435</td>
<td>0.0462</td>
<td>0.0465</td>
<td>0.0444</td>
<td>0.0398</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td><strong>1.1386</strong></td>
<td><strong>1.0925</strong></td>
<td>1.0611</td>
<td><strong>1.0460</strong></td>
<td><strong>1.0519</strong></td>
<td><strong>1.0850</strong></td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Bias</td>
<td>0.0003</td>
<td>-0.0011</td>
<td>-0.0025</td>
<td>-0.0039</td>
<td>-0.0053</td>
<td>-0.0066</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Std</td>
<td>0.0369</td>
<td>0.0422</td>
<td>0.0451</td>
<td>0.0454</td>
<td>0.0430</td>
<td>0.0378</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5B: Cointegrated Time Series, Frank Copula

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

<table>
<thead>
<tr>
<th>β*</th>
<th>5</th>
<th>3</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\beta}_{SP}) Bias</td>
<td>-0.0213</td>
<td>-0.262</td>
<td>-0.0233</td>
<td>-0.0257</td>
<td>-0.0470</td>
<td>-0.1018</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Std</td>
<td>0.3981</td>
<td>0.3216</td>
<td>0.2811</td>
<td>0.2819</td>
<td>0.3196</td>
<td>0.3913</td>
</tr>
<tr>
<td>(\hat{\beta}_{SP}) Rmse</td>
<td><strong>1.2980</strong></td>
<td><strong>1.1326</strong></td>
<td>1.0182</td>
<td><strong>1.0221</strong></td>
<td><strong>1.1355</strong></td>
<td><strong>1.3120</strong></td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Bias</td>
<td>0.0137</td>
<td>-0.0018</td>
<td>-0.0106</td>
<td>-0.0189</td>
<td>-0.0347</td>
<td>-0.0628</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Std</td>
<td>0.3496</td>
<td>0.3032</td>
<td>0.2793</td>
<td>0.2793</td>
<td>0.3012</td>
<td>0.3473</td>
</tr>
<tr>
<td>(\hat{\beta}_{P}) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 5C: Cointegrated Time Series, Clayton Copula

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

<table>
<thead>
<tr>
<th>β*</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>β_{SP} Bias</td>
<td>0.0004</td>
<td>-0.0280</td>
<td>-0.1519</td>
<td>-0.7054</td>
<td>-1.6939</td>
<td>-2.9915</td>
</tr>
<tr>
<td>β_{SP} Std</td>
<td>0.1032</td>
<td>0.1927</td>
<td>0.4434</td>
<td>0.9793</td>
<td>1.3301</td>
<td>1.5500</td>
</tr>
<tr>
<td>β_{SP} Rmse</td>
<td>1.3655</td>
<td>1.6828</td>
<td>1.9211</td>
<td>1.7613</td>
<td>2.1061</td>
<td>2.6836</td>
</tr>
<tr>
<td>β_{P} Bias</td>
<td>-0.0063</td>
<td>-0.0149</td>
<td>-0.0498</td>
<td>-0.2098</td>
<td>-0.5225</td>
<td>-0.9808</td>
</tr>
<tr>
<td>β_{P} Std</td>
<td>0.0881</td>
<td>0.1494</td>
<td>0.3344</td>
<td>0.8849</td>
<td>1.3890</td>
<td>1.8078</td>
</tr>
<tr>
<td>β_{P} Rmse</td>
<td>1.3561</td>
<td>1.6400</td>
<td>1.8475</td>
<td>1.7371</td>
<td>1.9892</td>
<td>2.4468</td>
</tr>
</tbody>
</table>

Table 5D: Cointegrated Time Series, Gumbel Copula

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

<table>
<thead>
<tr>
<th>β*</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>β_{SP} Bias</td>
<td>-0.0349</td>
<td>-0.1676</td>
<td>-0.4205</td>
<td>-0.8015</td>
<td>-1.3003</td>
<td>-1.8937</td>
</tr>
<tr>
<td>β_{SP} Std</td>
<td>0.1627</td>
<td>0.3558</td>
<td>0.5579</td>
<td>0.7233</td>
<td>0.8493</td>
<td>0.9527</td>
</tr>
<tr>
<td>β_{SP} Rmse</td>
<td>1.1636</td>
<td>1.2718</td>
<td>1.3916</td>
<td>1.6076</td>
<td>1.9301</td>
<td>2.2544</td>
</tr>
<tr>
<td>β_{P} Bias</td>
<td>-0.0140</td>
<td>-0.0559</td>
<td>-0.1443</td>
<td>-0.2866</td>
<td>-0.4859</td>
<td>-0.7285</td>
</tr>
<tr>
<td>β_{P} Std</td>
<td>0.1537</td>
<td>0.3442</td>
<td>0.5743</td>
<td>0.8018</td>
<td>1.0068</td>
<td>1.2094</td>
</tr>
<tr>
<td>β_{P} Rmse</td>
<td>0.1251</td>
<td>0.1147</td>
<td>0.0301</td>
<td>-0.1249</td>
<td>-0.3561</td>
<td>-0.6626</td>
</tr>
<tr>
<td>β_{P1} Std</td>
<td>0.3855</td>
<td>0.5664</td>
<td>0.8119</td>
<td>1.0448</td>
<td>1.2625</td>
<td>1.4788</td>
</tr>
<tr>
<td>β_{P1} Rmse</td>
<td>6.8989</td>
<td>2.7456</td>
<td>1.8822</td>
<td>1.5274</td>
<td>1.3769</td>
<td>1.3172</td>
</tr>
<tr>
<td>β_{SP} Bias</td>
<td>-0.0321</td>
<td>-0.1540</td>
<td>-0.3861</td>
<td>-0.7354</td>
<td>-1.1963</td>
<td>-1.7464</td>
</tr>
<tr>
<td>β_{SP} Std</td>
<td>0.0966</td>
<td>0.3512</td>
<td>0.5534</td>
<td>0.7335</td>
<td>0.8846</td>
<td>1.0121</td>
</tr>
<tr>
<td>β_{SP} Rmse</td>
<td>1.1129</td>
<td>1.2088</td>
<td>1.2984</td>
<td>1.4882</td>
<td>1.7713</td>
<td>2.0438</td>
</tr>
<tr>
<td>Table 6A: Cointegrated Time Series, Gaussian Copula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(True marginal is re-centered Chi-square with d.f. 3, ( n = 500 ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta^* )</td>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>( \beta_{SP} ) Bias</td>
<td>-0.0063</td>
<td>-0.0072</td>
<td>-0.0056</td>
<td>-0.0032</td>
<td>-0.0021</td>
<td>-0.0035</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
<td>0.0388</td>
<td>0.0436</td>
<td>0.0463</td>
<td>0.0465</td>
<td>0.0444</td>
<td>0.0399</td>
</tr>
<tr>
<td>( \tilde{\beta}_{SP} ) Rmse</td>
<td>1.3898</td>
<td>1.1142</td>
<td>1.0103</td>
<td>0.9926</td>
<td>0.9952</td>
<td>1.0527</td>
</tr>
<tr>
<td>( \beta\hat{p}^* ) Bias</td>
<td>-0.0013</td>
<td>-0.0034</td>
<td>-0.0040</td>
<td>-0.0015</td>
<td>0.0033</td>
<td>0.0073</td>
</tr>
<tr>
<td>( \beta\hat{p}^* ) Std</td>
<td>0.0333</td>
<td>0.0417</td>
<td>0.0462</td>
<td>0.0468</td>
<td>0.0444</td>
<td>0.0384</td>
</tr>
<tr>
<td>( \tilde{\beta}_{P1} ) Rmse</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \beta_{P1} ) Bias</td>
<td>0.0911</td>
<td>0.0453</td>
<td>0.0097</td>
<td>-0.0159</td>
<td>-0.0318</td>
<td>-0.0384</td>
</tr>
<tr>
<td>( \beta_{P1} ) Std</td>
<td>0.0302</td>
<td>0.0371</td>
<td>0.0431</td>
<td>0.0474</td>
<td>0.0493</td>
<td>0.0475</td>
</tr>
<tr>
<td>( \tilde{\beta}_{P1} ) Rmse</td>
<td>8.2865</td>
<td>1.9519</td>
<td>0.9062</td>
<td>1.1415</td>
<td>1.7373</td>
<td>2.4417</td>
</tr>
<tr>
<td>( \beta_{SP} ) Bias</td>
<td>-0.0065</td>
<td>-0.0071</td>
<td>-0.0053</td>
<td>-0.0027</td>
<td>-0.0013</td>
<td>-0.0024</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
<td>0.0388</td>
<td>0.0436</td>
<td>0.0461</td>
<td>0.0463</td>
<td>0.0442</td>
<td>0.0397</td>
</tr>
<tr>
<td>( \tilde{\beta}_{SP} ) Rmse</td>
<td>1.3971</td>
<td>1.1118</td>
<td>1.0040</td>
<td>0.9835</td>
<td>0.9857</td>
<td>1.0339</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6B: Cointegrated Time Series, Frank Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>(True marginal is re-centered Chi-square with d.f. 3, ( n = 500 ))</td>
</tr>
<tr>
<td>( \beta^* )</td>
</tr>
<tr>
<td>( \beta_{SP} ) Bias</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
</tr>
<tr>
<td>( \tilde{\beta}_{SP} ) Rmse</td>
</tr>
<tr>
<td>( \beta\hat{p}^* ) Bias</td>
</tr>
<tr>
<td>( \beta\hat{p}^* ) Std</td>
</tr>
<tr>
<td>( \tilde{\beta}_{P1} ) Rmse</td>
</tr>
<tr>
<td>( \beta_{P1} ) Bias</td>
</tr>
<tr>
<td>( \beta_{P1} ) Std</td>
</tr>
<tr>
<td>( \tilde{\beta}_{P1} ) Rmse</td>
</tr>
<tr>
<td>( \beta_{SP} ) Bias</td>
</tr>
<tr>
<td>( \beta_{SP} ) Std</td>
</tr>
<tr>
<td>( \tilde{\beta}_{SP} ) Rmse</td>
</tr>
</tbody>
</table>
Table 6C: Cointegrated Time Series, Clayton Copula
(True marginal is re-centered Chi-square with d.f. 3, \( n = 500 \))

<table>
<thead>
<tr>
<th>( \beta^* )</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta}_{SP} ) Bias</td>
<td>-0.0034</td>
<td>-0.0399</td>
<td>-0.1888</td>
<td>-0.7936</td>
<td>-1.7964</td>
<td>-3.0777</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Std</td>
<td>0.1025</td>
<td>0.1872</td>
<td>0.4119</td>
<td>0.9159</td>
<td>1.3067</td>
<td>1.5658</td>
</tr>
<tr>
<td>( \hat{\beta}_{SP} ) Rmse</td>
<td>0.9985</td>
<td>1.1506</td>
<td>1.3626</td>
<td>1.4072</td>
<td>1.4238</td>
<td>1.4918</td>
</tr>
<tr>
<td>( \hat{\beta}_{P^*} ) Bias</td>
<td>-0.0091</td>
<td>-0.0403</td>
<td>-0.1571</td>
<td>-0.5909</td>
<td>-1.2861</td>
<td>-2.1973</td>
</tr>
<tr>
<td>( \hat{\beta}_{P^*} ) Std</td>
<td>0.1022</td>
<td>0.1739</td>
<td>0.3550</td>
<td>0.8333</td>
<td>1.3460</td>
<td>1.7789</td>
</tr>
<tr>
<td>( \hat{\beta}_{P^*} ) Rmse</td>
<td>0.9985</td>
<td>1.1506</td>
<td>1.3626</td>
<td>1.4072</td>
<td>1.4238</td>
<td>1.4918</td>
</tr>
</tbody>
</table>

Table 6D: Cointegrated Time Series, Gumbel Copula
(True marginal is re-centered Chi-square with d.f. 3, \( n = 500 \))

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

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

for all \( j \).

B.1.1. Consistency of \( \hat{\beta}_P \)

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

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

where \( q(\alpha) = \mathbb{E}[\log f(Y_t, \alpha)] \). 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): \( \mathbb{E}[\sup_{\alpha \in \mathcal{A}} |\log f(Y_t, \alpha)|] < \infty \), we can show that

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

Thus, we only need to show that

\[
\sup_{\alpha \in \mathcal{A}} \left| \frac{1}{n} \sum_{t=1}^{n} \left[ \log f(Y_t, \alpha) - \log f(Y_t, \alpha) \right] \right| = o_p(1).
\]
Denote the re-standardized $X_t$ by $X_t$, i.e. $X_t = n^{1/2}D_n^{-1}X_t$, and define $q_t(\eta, \alpha) = \log f(Y_t - X_t^\prime \eta, \alpha)$. Under Assumption M1(2), we have, for all sequences of positive numbers $\{\epsilon_n\}$ with $\epsilon_n = o(1)$,

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

Thus

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

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

For the second step estimation, we need to verify that

$$
\sup_{\beta \in \mathcal{B}} \| \hat{Q}_n(\beta) - Q(\beta) \| = o_p(1),
$$

where

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

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 $\beta$; (4) Assumption M1(1): $\mathbb{E}[\sup_{\beta \in \mathcal{B}, \alpha \in \mathcal{A}} g(Y_{t-1}, Y_t, \alpha, \beta)] < \infty$, we have

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

Thus, it suffice to show that

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

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

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

then we may write

$$
\hat{Q}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} g \left( Y_{t-1} - n^{-1/2} \left( X_{t-1}^\prime n^{1/2} D_n^{-1} \right) \delta_n, Y_t - n^{-1/2} \left( X_t^\prime n^{1/2} D_n^{-1} \right) \delta_n, \alpha^* + n^{-1/2} \Delta_{1n}, \beta \right).
$$

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

$$
m_t(\eta, \alpha, \beta) = g \left( Y_{t-1} - X_{t-1}^\prime \eta, Y_t - X_t^\prime \eta, \alpha, \beta \right).
$$
Under the Assumption M1(2) that \( g(s_1, s_2, \alpha, \beta) \) is uniformly continuous in \( (s_1, s_2, \alpha, \beta) \), 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 \( \tilde{\eta} = n^{-1/2} \delta_n \), then

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

Notice that

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

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

**B.1.2. Limiting Distribution of \( \tilde{\beta}_p \)**

Let \( g(\hat{\beta}_{t-1}, \hat{\gamma}_t, \hat{\alpha}, \beta) = \log c(F(\hat{\beta}_{t-1}, \hat{\alpha}), F(\hat{\gamma}_t, \hat{\alpha}), \beta) \), then the likelihood function is given by

\[
\tilde{Q}_n(\beta) = \frac{1}{n} \sum_{t=2}^{n} g(\hat{\beta}_{t-1}, \hat{\gamma}_t, \hat{\alpha}, \beta).
\]

Let \( \sqrt{n}(\beta - \beta^*) = \Delta_2 \), and \( D_n(\bar{\pi} - \pi^*) = \delta_n, \sqrt{n}(\bar{\alpha} - \alpha^*) = \Delta_1n, \sqrt{n}(\bar{\beta} - \beta^*) = \Delta_{2n} \), then, we may re-write the criterion function \( \tilde{Q}_n(\beta) \) as

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

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

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

\[
\frac{\partial V_n(\Delta_2)}{\partial \Delta_2} \bigg|_{\Delta_2 = \Delta_{2n}} = 0.
\]

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

\[
0 = \frac{\partial V_n(\Delta_2)}{\partial \Delta_2} \bigg|_{\Delta_2 = \Delta_{2n}} = \frac{1}{n} \sum_{t=2}^{n} g(\hat{\beta}_{t-1}, \hat{\gamma}_t, \hat{\alpha}, \beta^*) + n^{-1/2} \left[ \frac{1}{n} \sum_{t=2}^{n} g(\hat{\beta}_{t-1}, \hat{\gamma}_t, \hat{\alpha}, \beta^*) \right] \Delta_{2n}.
\]
where $\beta^\#$ is the middle value between $\beta^*$ and $\hat{\beta}$.

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

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

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

$$
\sup_{\|\eta\| \leq \epsilon_n} \frac{1}{n} \sum_{t=2}^{n} \left| g_{\beta} \left( Y_{t-1} + X_{t-1}' \eta_1, Y_t + X_t' \eta_1, \alpha^* + \eta_2, \beta^* + \eta_3 \right) - g_{\beta} \left( Y_{t-1}, Y_t, \alpha^*, \beta^* \right) \right| = o_p(1),
\quad j = 1, 2
$$

we have

$$
\hat{H}_n = H_n + o_p(1),
$$

where

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

Denote

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

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

$$
\hat{S}_n = S_n + n^{-1} \sum_{t=2}^{n} g_{\beta_1} \left( Y_{t-1}, Y_t, \alpha^*, \beta^* \right) X_{t-1}' n^{1/2} D_n^{-1} \Delta_n + o_p(1)
$$

Thus,

$$
\sqrt{n} (\hat{\beta} - \beta^*)
$$

$$
= H_n S_n - H_n \left( P_{n1} + P_{n2} \right) D_n \left( \hat{\pi} - \pi^* \right) + H_n \left( P_{n3} \sqrt{n} \left( \hat{\alpha} - \alpha^* \right) \right) + o_p(1)
$$

$$
= H_n \left( 0, \Omega_\beta \right) - H_n \left( P_1 + P_2 \right) D_n \left( \hat{\pi} - \pi^* \right) + H_n \left( P_3 \sqrt{n} \left( \hat{\alpha} - \alpha^* \right) \right) + o_p(1)
$$

$$
= H_n \left( 0, \Omega_\beta \right) - H_n \left( P_1 + P_2 + P_3 \Omega_\alpha^{-1} H\alpha \right) D_n \left( \hat{\pi} - \pi^* \right) + H_n \left( P_3 \sqrt{n} \left( \hat{\alpha} - \alpha^* \right) \right) + o_p(1)
$$

Notice that $\sqrt{n} (\hat{\alpha} - \alpha^*) = H_n \left( \alpha \right) S_n + o_p(1)$, where

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

53
thus,
\[
\sqrt{n} (\hat{\beta} - \beta^*) = H_{n\beta}^{-1} \left[ S_{n\beta} + P_{n\beta} H_{n\alpha} S_{n\alpha} \right] - H_{n\beta}^{-1} \left( P_1 + P_2 + P_3 \Omega_n^{-1} H \alpha \right) D_n (\hat{\pi} - \pi^*) + o_p(1).
\]

B.2. The Semiparametric Copula Model

We use $\zeta$ 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 $\beta$ by $F^*$ and $\beta^*$. We first restate the important Lemma 1 from the main text. Consider $b = (b_1, \ldots, b_n)'$, 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 $\epsilon > 0$,

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

and

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

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

\[
\lim_{L \to \infty} \limsup_{n \to \infty} \sum_{j=1}^{\infty} \Pr \left[ \sup_{y_j < y \leq y_{j-1}} \left| Z_n(y, b) - Z_n(y, 0) \right| \geq \epsilon w(2^{-j} \delta) \right] = 0.
\]
By monotonicity of the indicator function and the distribution function, we have

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

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

\[
\Pr \left[ \sup_{y_j < y \leq y_{j-1}} |Z_n(y, b) - Z_n(y, 0)| \geq \epsilon w(2^{-j}\delta) \right] \\
\leq \Pr \left\{ |Z_n(y_j, b) - Z_n(y_j, 0)| + |Z_n(y_{j-1}, b) - Z_n(y_{j-1}, 0)| \\
+ \sup_{y_j < y \leq y_{j-1}} |Z_n(y_{j-1}, 0) - Z_n(y, 0)| + \sup_{y_j < y \leq y_{j-1}} |Z_n(y_j, 0) - Z_n(y, 0)| \\
+ C^* \sqrt{n2^{-j}\delta} \geq \epsilon w(2^{-j}\delta) \right\}.
\]

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 \( j \in S_1 \),

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

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

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

\[
\lim_{L \to \infty} \lim_{\delta \to 0} \sup_{j \in S_1} \sum \Pr \left[ \sup_{y_j < y \leq y_{j-1}} |Z_n(y_j, 0) - Z_n(y, 0)| \geq \frac{\epsilon w(2^{-j}\delta)}{8} \right] = 0. \quad (B.4)
\]
and analysis of the other two terms are similar.

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

$$\Pr \left[ |Z_n(y_j, b) - Z_n(y_j, 0)| \geq \frac{\epsilon w(2^{-j}\delta)}{8} \right] \leq \frac{2^5 \mathbb{E} |Z_n(y_j, b) - Z_n(y_j, 0)|^2}{\epsilon^2 w(2^{-j}\delta)^2}.$$  

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

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

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,

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

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

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

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

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

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 \(\cup(i(y_j, i, y_{i+1}]\), where \(F(y_{j, i}) = i\Delta_{n,j}, 0 \leq i \leq F(y_{j-1})/\Delta_{n,j} = 2^{-j+1}\delta/\Delta_{n,j}, \]

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

Notice that

\[
\sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y, b) - Z_n(y, 0)| \\
\leq |Z_n(y_j, i, b) - Z_n(y_j, i, 0)| + |Z_n(y_j, i+1, b) - Z_n(y_j, i+1, 0)| \\
+ \sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y_j, i, 0) - Z_n(y, 0)| + \sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y_j, i+1, 0) - Z_n(y, 0)| \\
+ 1 \frac{1}{\sqrt{n+1}} \sum_{t=1}^{n} \left[ F^*(y_j, i+1) + n^{-1/2}b_t \right] - F^*(y_j, i + n^{-1/2}b_t) \\
+ \frac{1}{4} \epsilon w(2^{-j}\delta)
\]

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

\[
\sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y, b) - Z_n(y, 0)| \\
\leq \left| \sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y_j, i, b) - Z_n(y_j, i, 0)| + |Z_n(y_j, i+1, b) - Z_n(y_j, i+1, 0)| \\
+ \sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y_j, i, 0) - Z_n(y, 0)| + \sup_{y_j, i < y \leq y_{j+i+1}} |Z_n(y_j, i+1, 0) - Z_n(y, 0)| \\
+ \frac{1}{4} \epsilon w(2^{-j}\delta) \right|
\]

and thus

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

57
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

$$\Pr \left[ \sup_{y_{j,i} < y \leq y_{j,i+1}} |Z_n(y, b) - Z_n(y, 0)| \geq \epsilon w(2^{-j}\delta) \right]$$

$$\leq \Pr \left[ \max_{0 \leq i \leq F(y_{j,i})/\Delta_{n,j}} |Z_n(y_{j,i}, b) - Z_n(y_{j,i}, 0)| \geq \frac{3\epsilon w(2^{-j}\delta)}{16} \right]$$

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

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

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

We next show that

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

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

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

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

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

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

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

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

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

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

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

$$\psi_t(j, k, i) = 1 \left( Y_t \leq y_{j,k} + n^{-1/2}b_t \right) - 1 \left( Y_t \leq y_{j,k} \right) + F^*(y_{j,k}) - F^*(y_{j,k} + n^{-1/2}b_t) - 1 \left( Y_t \leq y_{j,i} + n^{-1/2}b_t \right) + 1 \left( Y_t \leq y_{j,i} \right) - F^*(y_{j,i}) - F^*(y_{j,i} + n^{-1/2}b_t).$$
Notice that $\psi_j(k, i)$ is a bounded function, by Theorem 2 of Viennet (1997), and application of Moricz (1982), we have
\[
E \left[ \max_{1 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i}, b) - Z_n(y_{j,i}, 0)| \right]^p \\
\leq \zeta_3 (2^j \delta)^{p_1} + \zeta_4 n^{-p_2/2} 2^{-j} \delta \log^p (2^{-j+2} \delta / \Delta_{n,j}).
\]
where $p_1 = p/2$, $p_2 = p - 2$, and thus
\[
\Pr \left[ \max_{0 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i}, b) - Z_n(y_{j,i}, 0)| \geq \frac{3 \epsilon w(2^{-j} \delta)}{16} \right] \\
\leq \zeta_3 (2^j \delta)^{p_1} + \zeta_4 n^{-p_2/2} 2^{-j} \delta \log^p (2^{-j+2} \delta / \Delta_{n,j})
\]
Notice that $\Delta_{n,j} = 2^{-3} n^{-1/2} \epsilon w(2^{-j} \delta)$, and $n^{1/2} 2^{-j} \delta \epsilon \geq \epsilon w(2^{-j} \delta)/2$,
\[
\Pr \left[ \max_{0 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i}, b) - Z_n(y_{j,i}, 0)| \geq \frac{3 \epsilon w(2^{-j} \delta)}{16} \right] \\
\leq \epsilon \left[ \frac{w(2^{-j} \delta)}{8} \right]^{-p} \left(2^{-j} \delta^{p_1} + \epsilon w(2^{-j} \delta)^{-p_2} (2^{-j})^{(1+p_2)} \log^p (n^{1/2} 2^{-j+5} \delta) \right)
\]
Under Assumption SP, we have
\[
\lim_{L \to \infty} \lim_{n \to \infty} \sup_{j \in C_2} \sum_{j \leq i \leq n} \Pr \left[ \max_{0 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i}, b) - Z_n(y_{j,i}, 0)| \geq \frac{3 \epsilon w(2^{-j} \delta)}{16} \right] = 0.
\]
Notice that,
\[
\Pr \left[ \max_{0 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i+1}, 0) - Z_n(y_{j,i}, 0)| \geq \frac{\epsilon w(2^{-j} \delta)}{16} \right] \\
\leq \epsilon \frac{E \max_{1 \leq i \leq F(y_{j-1})/\Delta_{n,j}} |Z_n(y_{j,i}, 0)|^p}{\epsilon w(2^{-j} \delta)^p}.
\]
The analysis of other terms are similar. □

B.2.1. Theorem 3.

Notice that
\[
\sqrt{n+1} \left( \hat{F}_n(y) - F^*(y) \right) = \sqrt{n+1} \left( \hat{F}_n(y) - F_n(y) \right) + \sqrt{n+1} \left( F_n(y) - F^*(y) \right)
\]
The first term, $\sqrt{n+1} \left( \hat{F}_n(y) - F_n(y) \right)$, 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 \( Y_t = Y_t - n^{-1/2} \left( X_t' D_n^{-1} n^{1/2} \right) \gamma \), and

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

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

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

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

\[
\hat{F}_n(y) = F_{n, \hat{\gamma}}(y) = \frac{1}{n+1} \sum_{t=1}^{n} 1 \left( Y_t(\hat{\gamma}) \leq y \right).
\]

By (B.5), we have

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

Let

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

Under our assumptions, the consistency of \( \hat{\beta} \) can be obtained if

\[
\sup_{\beta \in \mathcal{B}} \left\| \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c(\hat{F}_n(Y_t), \hat{F}_n(Y_t), \beta)}{\partial \beta} - s(F^*, \beta) \right\| = o_p(1)
\]

By triangular inequality,

\[
\sup_{\beta \in \mathcal{B}} \left\| \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c(\hat{F}_n(Y_t-1), \hat{F}_n(Y_t), \beta)}{\partial \beta} - s(F^*, \beta) \right\| \\
\leq \sup_{\beta \in \mathcal{B}} \left\| \frac{1}{n} \sum_{t=2}^{n} \left[ \frac{\partial \log c(\hat{F}_n(Y_t-1), \hat{F}_n(Y_t), \beta)}{\partial \beta} - \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} \right] \right\| \\
+ \sup_{\beta \in \mathcal{B}} \left\| \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} - s(F^*, \beta) \right\|.
\]

By Chen and Fan (2006a),

\[
\sup_{\beta \in \mathcal{B}} \left\| \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} - s(F^*, \beta) \right\| = o_p(1).
\]

Next we verify that
We next show that the first two terms are together with Assumption M4, we obtain
\[ \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \left[ \frac{\partial \log c(\tilde{F}_n(\tilde{Y}_{t-1}), \tilde{F}_n(\tilde{Y}_t), \beta)}{\partial \beta} - \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} \right] \right\| = o_p(1) \]

Note that
\[ \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \left[ \frac{\partial \log c(\tilde{F}_n(\tilde{Y}_{t-1}), \tilde{F}_n(\tilde{Y}_t), \beta)}{\partial \beta} - \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} \right] \right\| \leq \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta 1}(F_{t-1}^n, F_t^n, \beta) \left( \tilde{F}_n(\tilde{Y}_{t-1}) - F_n(Y_{t-1}) \right) \right\| + \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta 2}(F_{t-1}^n, F_t^n, \beta) \left( \tilde{F}_n(\tilde{Y}_t) - F_n(Y_t) \right) \right\| + \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta 1}(F_{t-1}^n, F_t^n, \beta) (F_n(Y_{t-1}) - F^*(Y_{t-1})) \right\| + \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta 2}(F_{t-1}^n, F_t^n, \beta) (F_n(Y_t) - F^*(Y_t)) \right\| \]

where \( F_{n}^{\eta} = \eta \tilde{F}_n(\tilde{Y}_n) + (1 - \eta)F^*(Y_n), s = t - 1 \) 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
\[ \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta 2}(F_{t-1}^n, F_t^n, \beta) \left[ \tilde{F}_n(\tilde{Y}_t) - F_n(Y_t) \right] \right\| \leq \frac{1}{n} \sum_{t=2}^{n} \sup_{\beta \in \Theta} \sup_{F \in F_S} |\ell_{\beta 2}(F(Y_{t-1}), F(Y_t), \beta) w(F^*(Y_t))| \sup_t \frac{\tilde{F}_n(\tilde{Y}_t) - F_n(Y_t)}{w(F^*(Y_t))} \]

By (B.6), we have
\[ \sup_t \left| \frac{\tilde{F}_n(\tilde{Y}_t) - F_n(Y_t)}{w(F^*(\tilde{Y}_t))} \right| = O_p \left( n^{-1/2} \right), \]

together with Assumption M4, we obtain
\[ \sup_{\beta \in \Theta} \left\| \frac{1}{n} \sum_{t=2}^{n} \left[ \frac{\partial \log c(\tilde{F}_n(\tilde{Y}_{t-1}), \tilde{F}_n(\tilde{Y}_t), \beta)}{\partial \beta} - \frac{\partial \log c(F^*(Y_{t-1}), F^*(Y_t), \beta)}{\partial \beta} \right] \right\| = o_p(1). \]
B.2.2. Theorem 4.

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

$$0 = \frac{1}{n} \sum_{t=2}^{n} \ell_\beta \left( \hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\hat{Y}_t), \hat{\beta}_{SP} \right)$$

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

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

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

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

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

$$+ \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \ell_{\beta_1} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t-1}) - F^*(Y_{t-1}) \right)$$

$$+ \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( \hat{F}_n(\hat{Y}_t) - F^*(Y_t) \right)$$

$$+ \frac{1}{n^{3/2}} \sum_{i,j=1}^{n} \ell_{\beta_{ij}} (F^n_{i-1}, F^n_t, \beta^*) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+i-2}) - F^*(Y_{t+i-2}) \right) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+j-2}) - F^*(Y_{t+j-2}) \right),$$

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

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

$$\frac{1}{n^{3/2}} \sum_{t=2}^{n} \ell_{\beta_{ij}} (F^n_{i-1}, F^n_t, \beta^*) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+i-2}) - F^*(Y_{t+i-2}) \right) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+j-2}) - F^*(Y_{t+j-2}) \right) = o_p(1).$$

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

$$\left| \frac{1}{n^{3/2}} \sum_{t=2}^{n} \ell_{\beta_{12}} (F^n_{i-1}, F^n_t, \beta^*) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t-1}) - F^*(Y_{t-1}) \right) \sqrt{n} \left( \hat{F}_n(\hat{Y}_t) - F^*(Y_t) \right) \right|$$

$$\leq \frac{1}{n^{3/2}} \sum_{t=2}^{n} \sup_{\|\beta - \beta^*\| \leq \delta, F \in F_\delta} |\ell_{\beta_{12}} (F(Y_{t-1}), F(Y_t), \beta^*) w(F^*(Y_{t-1}))| \times \frac{\sqrt{n} \left( \hat{F}_n(\hat{Y}_{t-1}) - F^*(Y_{t-1}) \right) \sqrt{n} \left( \hat{F}_n(\hat{Y}_t) - F^*(Y_t) \right)}{w(F^*(Y_{t-1}))}$$

Under Assumption M4,

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

62
and by application of Lemma 1,
\[
\frac{\sqrt{n} \left( \hat{F}_n(\hat{Y}_{t-1}) - F^*(Y_{t-1}) \right)}{w(F^*(Y_{t-1}))} = O_p(1), \quad \frac{\sqrt{n} \left( \hat{F}_n(\bar{Y}_t) - F^*(Y_t) \right)}{w(F^*(Y_t))} = O_p(1),
\]

thus
\[
\left| \frac{1}{n^{3/2}} \sum_{t=2}^{n} \ell_{\beta \beta} \left( \hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\bar{Y}_t), \hat{\beta} \right) \right| = o_p(1).
\]

Second, by Taylor expansion,
\[
\frac{1}{n} \sum_{t=2}^{n} \ell_{\beta \beta} \left( \hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\bar{Y}_t), \hat{\beta} \right) = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta \beta} \left( F^*(Y_{t-1}), F^*(Y_t), \beta^* \right)
\]
\[
+ \frac{1}{n^{3/2}} \sum_{j=1}^{2} \sum_{t=2}^{n} \ell_{\beta \beta} \left( F_{t-1}^\eta, F_t^\eta, \beta \right) \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+j-2}) - F^*(Y_{t+j-2}) \right)
\]
\[
+ \frac{1}{n^{3/2}} \sum_{t=2}^{n} \ell_{\beta \beta \beta} \left( F_{t-1}^\eta, F_t^\eta, \beta \right) \sqrt{n}(\hat{\beta} - \beta),
\]

where \( \hat{\beta} = \eta \beta^* + (1 - \eta) \hat{\beta} \). Thus, by Assumptions M4, ST, and Lemma 1,
\[
\left\| \frac{1}{n} \sum_{t=2}^{n} \left[ \ell_{\beta \beta} \left( \hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\bar{Y}_t), \hat{\beta} \right) - \ell_{\beta \beta} \left( F^*(Y_{t-1}), F^*(Y_t), \beta^* \right) \right] \right\|
\]
\[
\leq \frac{1}{n^{3/2}} \sum_{j=1}^{2} \sum_{t=2}^{n} \sup_{\|\beta - \beta^*\| \leq \delta, F \in \mathcal{F}_\delta} \|\ell_{\beta \beta} \left( F(Y_{t-1}), F(Y_t), \beta \right) w(F^*(Y_{t+j-2}))\|
\]
\[
\times \left| \frac{\sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+j-2}) - F^*(Y_{t+j-2}) \right)}{w(F^*(Y_{t+j-2}))} \right|
\]
\[
+ \frac{1}{n^{3/2}} \sum_{t=2}^{n} \sup_{\|\beta - \beta^*\| \leq \delta, F \in \mathcal{F}_\delta} \|\ell_{\beta \beta \beta} \left( F(Y_{t-1}), F(Y_t), \beta \right) \| \sqrt{n}(\hat{\beta} - \beta^*)\|
\]
\[
= o_p(1).
\]

Thus,
\[
\frac{1}{n} \sum_{t=2}^{n} \ell_{\beta \beta} \left( \hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\bar{Y}_t), \hat{\beta} \right) = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta \beta} \left( F^*(Y_{t-1}), F^*(Y_t), \beta^* \right) + o_p(1),
\]
Let
\[
A_{n1} = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_1} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( \hat{F}_n(Y_{t-1}) - F_n(Y_{t-1}) \right),
\]
\[
A_{n2} = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_1} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( F_n(Y_{t-1}) - F^*(Y_{t-1}) \right),
\]
\[
A_{n3} = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( \hat{F}_n(Y_t) - F_n(Y_t) \right),
\]
\[
A_{n4} = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( F_n(Y_t) - F^*(Y_t) \right),
\]
and
\[
\Sigma_n = - \left[ \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_3} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \right],
\]
\[
S_n = \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \ell_{\beta} (F^*(Y_{t-1}), F^*(Y_t), \beta^*),
\]
then we have
\[
\Sigma_n \sqrt{n} \left( \beta_{SP} - \beta^* \right) = S_n + A_{n1} + A_{n2} + A_{n3} + A_{n4} + o_p(1),
\]
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 \textit{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
\[
A_{n3} = \frac{1}{n} \sum_{t=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) \sqrt{n} \left( \hat{F}_n(Y_t) - F_n(Y_t) \right)
= - \frac{1}{n^2} \sum_{t=2}^{n} \sum_{j=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ (X_j - X_t)^t D_n^{-1} n^{1/2} \right] D_n(\tilde{\pi} - \pi^*) + o_p(1),
\]
and
\[
\frac{1}{n^2} \sum_{t=2}^{n} \sum_{j=2}^{n} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ (X_j - X_t)^t D_n^{-1} n^{1/2} \right] = \frac{1}{n^2} \sum_{t>j} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ X_j^t D_n^{-1} n^{1/2} \right]
+ \frac{1}{n^2} \sum_{t>j} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ X_t^t D_n^{-1} n^{1/2} \right]
- \frac{1}{n^2} \sum_{t>j} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ X_j^t D_n^{-1} n^{1/2} \right]
- \frac{1}{n^2} \sum_{t>j} \ell_{\beta_2} (F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t) \left[ X_j^t D_n^{-1} n^{1/2} \right] = H_{1n} + H_{2n} - H_{3n} - H_{4n}.
\]
We investigate the behavior of each of the above terms and show that

\[
H_{1n} \to \left[ \int_0^1 rX(r)dr \right] E\{\ell_{\beta_2}(F^*(Y_{j-1}), F^*(Y_j), \beta^*) f(Y_j)\},
\]

\[
H_{2n} \to \int_0^1 \int_0^r X(s)dsdr E\{\ell_{\beta_2}(F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t)\},
\]

\[
H_{3n} \to \left[ \int_0^1 rX(r)dr \right] E\{\ell_{\beta_2}(F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t)\},
\]

\[
H_{4n} \to \int_0^1 \int_0^r X(s)dsdr E\{\ell_{\beta_2}(F^*(Y_{t-1}), F^*(Y_t), \beta^*) f(Y_t)\}.
\]

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

\[
\Sigma_n \sqrt{n} \left( \hat{\beta}_{SP} - \beta^* \right) = \Sigma_n \sqrt{n} \left( \tilde{\beta}_{SP} - \beta^* \right) + o_p(1) = S_n + A_{n2} + A_{n4} + o_p(1).
\]

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 \(\hat{\Omega}_{\beta}^+\) as:

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

where

\[
\gamma_n(h) = \frac{1}{n} \sum_{2 \leq t, t+h \leq n} S_t(F, \beta) S_{t+h}(F, \beta),
\]

and

\[
\gamma_{n1}(h) = \frac{1}{n} \sum_{2 \leq t, t+h \leq n} S_t \left( F, \tilde{\beta} \right) S_{t+h} \left( F, \tilde{\beta} \right).
\]

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 \(\Omega_\beta\) by the standard arguments as Hannan (1970).

The second part,

\[
\sum_{h=-M}^M K \left( \frac{h}{M} \right) [\gamma_{n1}(h) - \gamma_n(h)],
\]
contains the effect of copula estimation error ($\hat{\beta} - \beta$), this term converges to 0 following a similar argument as Andrews (1991, p852).

We now consider the third term,

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

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

$$\hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=2}^{n} \sum_{2 \leq t, t+h \leq n} \left[ S_t \left( \hat{\gamma}_n(h) - S_t \left( \hat{\gamma}_n(h) \right) \right) \right]$$

thus

$$\sum_{h=-M}^{M} K \left( \frac{h}{M} \right) [\hat{\gamma}_n(h) - \gamma_n(h)]$$

$$= \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t=2}^{n} \sum_{2 \leq t, t+h \leq n} \left[ S_t \left( \hat{\gamma}_n(h) - S_t \left( \hat{\gamma}_n(h) \right) \right) \right]$$

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

$$\sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t=2}^{n} \sum_{2 \leq t, t+h \leq n} \left[ S_t \left( \hat{\gamma}_n(h) - S_t \left( \hat{\gamma}_n(h) \right) \right) \right],$$
notice that

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

\[
= \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t} S_t \left( F, \hat{\beta} \right) \sum_{j=1}^{2} \ell_{\beta_j} \left( F_{n,t+i-1}, F_{n,t+i}, \hat{\beta} \right) \left( \hat{F}_n(\hat{Y}_{t+i+j-2}) - F(Y_{t+i+j-2}) \right)
\]

\[
- \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t} S_t \left( F, \hat{\beta} \right) \int_{0}^{1} \ell_{\beta,2} \left( v_1, F_{n,t+i}, \hat{\beta} \right) c \left( v_1, F_{n,t+i}, \hat{\beta} \right) dv_1 \left( \hat{F}_n(\hat{Y}_{t+i}) - F(Y_{t+i}) \right)
\]

\[
- \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t} S_t \left( F, \hat{\beta} \right) \int_{0}^{1} \ell_{\beta,1} \left( F_{n,t+i-1}, v_2, \hat{\beta} \right) c \left( F_{n,t+i-1}, v_2, \hat{\beta} \right) dv_2 \left( \hat{F}_n(\hat{Y}_{t+i-1}) - F(Y_{t+i-1}) \right)
\]

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

\[
\left| \sum_{h=-M}^{M} K \left( \frac{h}{M} \right) \frac{1}{n} \sum_{t=2}^{n} S_t \left( F, \hat{\beta} \right) \sum_{j=1}^{2} \ell_{\beta_j} \left( F_{n,t+h-1}, F_{n,t+h}, \hat{\beta} \right) \left( \hat{F}_n(\hat{Y}_{t+h+j-2}) - F(Y_{t+h+j-2}) \right) \right|
\]

\[
\leq \frac{1}{\sqrt{n}} \sum_{h=-M}^{M} \sum_{t=2}^{n} \left| K \left( \frac{h}{M} \right) \right| \sum_{j=1}^{2} \left| \sup_{F \in F_S} \left[ S_t \left( F^*, \hat{\beta} \right) w \left( F^*(Y_{t+h+j-2}) \right) \ell_{\beta_j} \left( F(Y_{t+h-1}), F(Y_{t+h}) \right) \right] \right|
\]

\[
\times \left| \sqrt{n} \left( \hat{F}_n(\hat{Y}_{t+h+j-2}) - F^*(Y_{t+h+j-2}) \right) \right| \left| w \left( F^*(Y_{t+h+j-2}) \right) \right|
\]

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 \( \log c_2(\hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\hat{Y}_t), \beta_2) \)

around \( \hat{\beta}_2 \), and notice that the FOC corresponding to \( \hat{\beta}_2 \) implies

\[
\sum_{t} \frac{\partial \log c_2(\hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\hat{Y}_t), \beta_2)}{\partial \beta} = 0,
\]

67
in the non-nested case,

\[
\text{Pr} \left[ \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} \neq \mathbb{E} \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] \right] > 0
\]

\[
\text{Pr} \left[ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} \neq \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right] > 0
\]

we have

\[
\frac{1}{n} \sum_{t=2}^{n} \log c_2(\hat{F}_n(Y_{t-1}), \hat{F}_n(Y_t), \beta_2)
\]

\[
= \frac{1}{n} \sum_{t=2}^{n} \log c_2(\hat{F}_n(Y_{t-1}), \hat{F}_n(Y_t), \beta_2) - \frac{1}{2n} \sum_{t=2}^{n} \left( \beta_2 - \beta_2 \right) \frac{\partial^2 \log c_2(\hat{F}_n(Y_{t-1}), \hat{F}_n(Y_t), \beta_2)}{\partial \beta \partial \beta'} \left( \beta_2 - \beta_2 \right)
\]

\[
= \frac{1}{n} \sum_{t=2}^{n} \log c_2(U_{t-1}, U_t, \beta_2) + \frac{1}{n} \sum_{j=1}^{2} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} \left[ \hat{F}_n(Y_{t-2+j}) - F(Y_{t-2+j}) \right] + o_p \left( n^{-1/2} \right)
\]

and

\[
\overline{L R}_n = \frac{1}{n} \sum_{t=2}^{n} \log \frac{c_2(\hat{F}_n(Y_{t-1}), \hat{F}_n(Y_t), \beta_2)}{c_1(\hat{F}_n(Y_{t-1}), \hat{F}_n(Y_t), \beta_1)}
\]

\[
= \frac{1}{n} \sum_{t=2}^{n} \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} + \frac{1}{n} \sum_{j=1}^{2} \sum_{t=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right\} \left[ \hat{F}_n(Y_{t-2+j}) - F(Y_{t-2+j}) \right] + o_p \left( n^{-1/2} \right)
\]

Thus

\[
\overline{L R}_n - \mathbb{E} \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right]
\]

\[
= \frac{1}{n} \sum_{t=2}^{n} \left[ \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} - \mathbb{E} \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] \right]
\]

\[
+ \frac{1}{n} \sum_{j=1}^{2} \sum_{t=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right\} \left[ \hat{F}_n(Y_{t-2+j}) - F(Y_{t-2+j}) \right] + o_p \left( n^{-1/2} \right)
\]
\[
\frac{1}{n} \sum_{j=1}^{2} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-2+j}} \left[ \hat{F}_n(\hat{Y}_{t-2+j}) - F(Y_{t-2+j}) \right] \\
= \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-1}} \left[ \hat{F}_n(\hat{Y}_{t-1}) - F_n(Y_{t-1}) \right] \\
+ \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_t} \left[ \hat{F}_n(\hat{Y}_t) - F_n(Y_t) \right] \\
+ \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_{t-1}} \left[ F_n(Y_{t-1}) - F(Y_{t-1}) \right] \\
+ \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_t} \left[ F_n(Y_t) - F(Y_t) \right]
\]

Using similar argument as in the previous Sections, we can show
\[
\frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_t} \left[ \hat{F}_n(\hat{Y}_t) - F_n(Y_t) \right] \\
= \frac{1}{\sqrt{n}} \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_t} f^*(Y_t) \frac{1}{n} \sum_{j=1}^{n} \left[ (X_j' - X_j') D_{n}^{-1} n^{1/2} \right] D_n (\hat{\pi} - \pi^*) + o_p \left( n^{-1/2} \right)
\]
\[
= \frac{1}{\sqrt{n}} \frac{1}{n} \sum_{t=2}^{n} \frac{\partial \log c_2(U_{t-1}, U_t, \overline{\beta}_2)}{\partial U_t} f^*(Y_t) \frac{1}{n} \sum_{j=1}^{n} \left[ X_j' D_{n}^{-1} n^{1/2} \right] D_n (\hat{\pi} - \pi^*)
\]
\ [- \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^*(Y_t) \frac{1}{n} \sum_{j=1}^{n} \left[ X_j' D_{n}^{-1} n^{1/2} \right] D_n (\hat{\pi} - \pi^*) + o_p \left( n^{-1/2} \right)
\]
\[
= o_p \left( n^{-1/2} \right)
\]

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

Let
\[
g_{ij}(U_t, \overline{\beta}_t) = E \left\{ \left[ \frac{\partial \log c_i(U_{t-1}, U_t, \overline{\beta}_t)}{\partial U_{t-2+j}} \right] (1(U_t \leq U_{t-2+j}) - U_{t-2+j}) \right| U_t \right\},
\]
\[
\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \sum_{i=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right\} \left[ F_n(Y_{t-2+j}) - F(Y_{t-2+j}) \right] \\
= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \sum_{i=2}^{n} E \left\{ \left[ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right] \left(1(U_i \leq U_{t-2+j}) - U_{t-2+j} \right) \right\} \\
= \sum_{j=1}^{2} \left[ \frac{1}{\sqrt{n}} \sum_{i=2}^{n} E \left\{ g_{2j}(U_t, \beta_2) - g_{1j}(U_t, \beta_1) \right\} \right],
\]

we have

\[
\sqrt{n} \left( LR_n - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] \right) \\
= \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \left[ \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] \right] \\
+ \frac{1}{\sqrt{n}} \sum_{j=1}^{2} \sum_{i=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right\} \left[ F_n(Y_{t-2+j}) - F(Y_{t-2+j}) \right] + o_p(1) \\
= \frac{1}{\sqrt{n}} \sum_{i=2}^{n} \left[ \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] \right] \\
+ \sum_{j=1}^{2} \left[ \frac{1}{\sqrt{n}} \sum_{i=2}^{n} E \left\{ g_{2j}(U_t, \beta_2) - g_{1j}(U_t, \beta_1) \right\} \right] + o_p(1) \\
\Rightarrow N(0, \omega^2)
\]

In the generalized nested case, denote

\[
H_{jn} = -\frac{1}{n} \sum_{i=2}^{n} \frac{\partial^2 \log c_j(\hat{F}_n(\hat{Y}_{t-1}), \hat{F}_n(\hat{Y}_t), \beta_j)}{\partial \beta \partial \beta'} \to \mathbf{H}_{j, \beta},
\]

Notice that

\[
\Pr \left[ c_2(U_{t-1}, U_t, \beta_2) = c_1(U_{t-1}, U_t, \beta_1) \right] = 1
\]

thus

\[
\Pr \left[ \log \frac{c_2(U_{t-1}, U_t, \beta_2)}{c_1(U_{t-1}, U_t, \beta_1)} = 0 \right] = E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \beta_2)}{c_1(F(Y_{t-1}), F(Y_t), \beta_1)} \right] = 1
\]

\[
\Pr \left[ \frac{\partial \log c_2(U_{t-1}, U_t, \beta_2)}{\partial U_{t-2+j}} = \frac{\partial \log c_1(U_{t-1}, U_t, \beta_1)}{\partial U_{t-2+j}} \right] = 1
\]
thus,
\[
LR_n - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \bar{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \bar{\beta}_1)} \right] \\
= \frac{1}{n} \sum_{t=2}^{n} \left[ \log \frac{c_2(U_{t-1}, U_t, \bar{\beta}_2)}{c_1(U_{t-1}, U_t, \bar{\beta}_1)} - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \bar{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \bar{\beta}_1)} \right] \right] \\
+ \frac{1}{n} \sum_{j=1}^{n} \sum_{t=2}^{n} \left\{ \frac{\partial \log c_2(U_{t-1}, U_t, \bar{\beta}_2)}{\partial U_{t-2+j}} - \frac{\partial \log c_1(U_{t-1}, U_t, \bar{\beta}_1)}{\partial U_{t-2+j}} \right\} \left[ \hat{E}_n(\hat{Y}_{t-2+j}) - F(Y_{t-2+j}) \right] \\
+ \frac{1}{2} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) H_{2n} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) - \frac{1}{2} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) H_{1n} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) + \\
= \frac{1}{2} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) H_{2n} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) - \frac{1}{2} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) H_{1n} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) + o_p \left( \frac{1}{n} \right)
\]

Let \( U_t = F^*(Y_t) \), and \( \bar{\mathbf{G}}_{j,n} = n^{-1/2} \sum_{j=2}^{n} \bar{\mathbf{g}}_{j,\beta} (U_{j-1}, U_j, \bar{\beta}_j) \), \( j = 1, 2 \), where
\[
\bar{\mathbf{g}}_{j,\beta} (U_{j-1}, U_j, \bar{\beta}_j) = \frac{\partial \log c_j (U_{j-1}, U_j, \bar{\beta})}{\partial \beta} + \sum_{i=0}^{1} E \left[ \frac{\partial^2 \log c_j (U_{t-1}, U_t, \bar{\beta})}{\partial \beta \partial U_{t-i}} \right] \left[ 1 \left( U_j \leq U_{t-i} \right) - U_{t-i} \right] \left| U_j \right|
\]

\( \bar{\mathbf{G}}_{j,\beta} = \lim_{n \to \infty} \text{Var} (\bar{\mathbf{G}}_{j,n}) \), \( \bar{\mathbf{H}}_{j,\beta} = -E \ell_{j,\beta} \left( F^*(Y_{t-1}), F^*(Y_t), \bar{\beta}_j \right) \)

Using the results of Section 4,
\[
\sqrt{n} \left( \bar{\beta}_j - \bar{\beta}_j \right) \Rightarrow N \left( 0, \bar{\mathbf{H}}_{j,\beta}^{-1} \bar{\mathbf{G}}_{j,\beta}^{-1} \right)
\]

and
\[
n \left[ LR_n - E \left[ \log \frac{c_2(F(Y_{t-1}), F(Y_t), \bar{\beta}_2)}{c_1(F(Y_{t-1}), F(Y_t), \bar{\beta}_1)} \right] \right] \\
= \frac{1}{2} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) H_{2n} \left( \bar{\beta}_2 - \bar{\beta}_2 \right) - \frac{1}{2} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) H_{1n} \left( \bar{\beta}_1 - \bar{\beta}_1 \right) + o_p \left( 1 \right)
\]

\[
= \frac{1}{2} \bar{\mathbf{G}}_{2,n} \bar{\mathbf{H}}_{2,\beta}^{-1} (H_{2n}) \bar{\mathbf{H}}_{2,\beta}^{-1} \bar{\mathbf{G}}_{2,n} - \frac{1}{2} \bar{\mathbf{G}}_{1,n} \bar{\mathbf{H}}_{1,\beta}^{-1} (H_{1n}) \bar{\mathbf{H}}_{1,\beta}^{-1} \bar{\mathbf{G}}_{1,n} + o_p \left( 1 \right)
\]

\[
= \frac{1}{2} \begin{bmatrix} \bar{\mathbf{G}}_{2,n} & \bar{\mathbf{G}}_{1,n} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_{2,\beta}^{-1} & 0 \\ 0 & -\bar{\mathbf{H}}_{1,\beta}^{-1} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{G}}_{2,n} \\ \bar{\mathbf{G}}_{1,n} \end{bmatrix} + o_p \left( 1 \right)
\]

where
\[
\begin{bmatrix} \bar{\mathbf{G}}_{2,n} \\ \bar{\mathbf{G}}_{1,n} \end{bmatrix} \Rightarrow N \left( 0, \begin{bmatrix} \bar{\mathbf{G}}_{2,\beta} & \bar{\mathbf{G}}_{2,\beta} \\ \bar{\mathbf{G}}_{2,\beta} & \bar{\mathbf{G}}_{1,\beta} \end{bmatrix} \right)
\]

Thus, under the null, \( 2nLR_n \) converges to a weighted sum of independent \( \chi^2 \) random variables in which the weights \( (\lambda_1, \ldots, \lambda_{k_1+k_2}) \) is the vector of eigenvalues of the following matrix
\[
\begin{bmatrix} \bar{\mathbf{G}}_{2,\beta} \bar{\mathbf{H}}_{2,\beta}^{-1} & -\bar{\mathbf{G}}_{2,\beta} \bar{\mathbf{H}}_{1,\beta}^{-1} \\ \bar{\mathbf{G}}_{1,\beta} \bar{\mathbf{H}}_{2,\beta}^{-1} & -\bar{\mathbf{G}}_{1,\beta} \bar{\mathbf{H}}_{1,\beta}^{-1} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{G}}_{2,\beta} & \bar{\mathbf{G}}_{2,\beta} \\ \bar{\mathbf{G}}_{2,\beta} & \bar{\mathbf{G}}_{1,\beta} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_{2,\beta}^{-1} & 0 \\ 0 & -\bar{\mathbf{H}}_{1,\beta}^{-1} \end{bmatrix}
\]

71