IRTG 1792 Discussion Paper 2018-064

Semiparametric Estimation and Variable Selection for Single-index Copula Models

Bingduo Yang [*]
Christian M. Hafner [*2]
Guannan Liu [*3]
Wei Long [*4]

[*] Sun Yat-sen University, China
[*2] Université catholique de Louvain, Belgium
[*3] Xiamen University, China
[*4] Tulane University, United States

This research was supported by the Deutsche Forschungsgesellschaft through the International Research Training Group 1792 "High Dimensional Nonstationary Time Series".

http://irtg1792.hu-berlin.de
ISSN 2568-5619

International Research Training Group 1792

Semiparametric Estimation and Variable Selection for Single-index Copula Models

Bingduo Yang* Christian M. Hafner† Guannan Liu‡ Wei Long§

September 17, 2018

Abstract

A copula model with flexibly specified dependence structure can be useful to capture the complexity and heterogeneity in economic and financial time series. However, there exists little methodological guidance for the specification process using copulas. This paper contributes to fill this gap by considering the recently proposed single-index copulas, for which we propose a simultaneous estimation and variable selection procedure. The proposed method allows to choose the most relevant state variables from a comprehensive set using a penalized estimation, and we derive its large sample properties. Simulation results demonstrate the good performance of the proposed method in selecting the appropriate state variables and estimating the unknown index coefficients and dependence parameters. An application of the new procedure identifies six macroeconomic driving factors for the dependence among U.S. housing markets.

Keywords: Semiparametric Copula, Single-Index Copula, Variable Selection, SCAD

JEL classification: **C14, C22**

*Lingnan (University) College, Sun Yat-sen University, Guangzhou, Guangdong 510275, China.

†Institut de statistique and CORE, Université catholique de Louvain, Voie du Roman Pays, 20, 1348 Louvain-la-Neuve, Belgium.

‡Corresponding author (email: gliuecon@gmail.com). School of Economics and Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Xiamen, Fujian 361005, China.

§Department of Economics, Tulane University, New Orleans, LA 70118, United States.

1 Introduction

There is a growing literature on copula-based models for characterizing the dynamics of economic and financial time series data. Copula-based multivariate models, mainly driven by the seminal contribution of Sklar (1959), allow to model multivariate distributions in a simple and flexible way. Every joint distribution with continuous margins can be decomposed into marginal distributions and a copula that explains the dependence structure. In financial econometrics, the copula has become one of the most popular methods to analyze the dependence in financial time series and has been widely applied in different areas such as exchange rates (Patton, 2006), Value-at-Risk (Giacomini, Hardle and Spokoiny, 2009), equity and bond markets (Garcia and Tsafack, 2011), and international financial market (Chollete, Heinen and Valdesogo, 2009), among others.

To account for changes in the dependence structure over time, a copula model should allow for time-variation of the parameters that determine the dependence. Time-varying copulas have been studied via different approaches. One stream of studies assumes that the dependence parameters are parametric functions of either observed variables or latent processes. For example, Patton (2004) specifies the conditional correlation as a function of the lagged risk-free rate, default spread, dividend yield, and the forecasts of the conditional means of two marginal variables. Furthermore, he models exchange rates with an autoregressive moving average (ARMA)-type process in the dependence structure of time-varying copulas (Patton, 2006). Giacomini et al. (2009) assume that there exists an interval of time homogeneity in which the copula parameter is well approximated by a constant. Autoregressive-type dynamic copula models have been proposed by Hafner and Manner (2012) and Bartram, Taylor and Wang (2007). Patton (2012) provides a comprehensive survey of time-varying copulas and their applications in financial time series analysis.

Another line of research investigates the dependence in a nonparametric way to avoid inconsistency due to possible mis-specification. For example, Hafner and Reznikova (2010) specify the copula parameter as a deterministic function of time, while Acar, Craiu and

Yao (2011) assume that it varies as a function of a measured covariate. Both of them employ kernel smoothing methods for estimation.

In this paper, we contribute to the literature by developing a variable selection procedure for the single-index copula, an innovative time-varying copula model introduced by Fermanian and Lopez (2018). The single-index copula model is a semi-parametric conditional copula whose parameter is an unknown link function of an univariate index. This index form avoids the curse of dimensionality because it essentially reduces the multivariate problem to a univariate one. From an economic perspective, this index offers a convenient univariate summary statistic that describes the current state of the various time-varying economic/financial indicators related to the tail dependence in risk management.

Fermanian and Lopez (2018) propose an estimation method for the single-index copula. However, the large sample properties of their estimators are derived for the independently identically distributed (i.i.d.) case, which is potentially restrictive and may not be suitable for financial time series with weakly dependent data. Moreover, the univariate single-index could be a linear combination of many state variables with only a few being statistically and economically important for the dependence structure. We address these issues in this paper by doing the estimation and variable selection simultaneously for the single-index copula model under α-mixing conditions. The variable selection procedure for this index form has been extensively studied in semiparametric models such as partially linear single-index models (Liang, Liu, Li and Tsay, 2010) and functional index coefficient models (Cai, Juhl and Yang, 2015). It has also been applied to semiparametric asset pricing models: Cai, Ren and Yang (2015) propose a semiparametric conditional capital asset pricing model (CAPM) by combining the predictors into an index, while Guo, Wu and Yu (2017) model the conditional market alpha and beta as flexible functions of state variables.

Our contribution is threefold. First, we do the estimation and variable selection simultaneously for single-index copulas. Using a sophisticated estimating function coupled with the smoothly clipped absolute deviation (SCAD) penalty term (Fan and Li, 2001), we can identify the most relevant state variables from a comprehensive set. Based on

α-mixing conditions and other regularity conditions, we provide the large sample properties for the unpenalized estimator and derive the consistency and oracle properties for the penalized estimator.

Second, we propose a three-step procedure to estimate the parametric and nonparametric components of the copula model, where the marginal parameters are estimated by maximizing the marginal log-likelihoods, the index coefficients are estimated by maximizing a global log-likelihood via a profile likelihood approach, and the dependence parameters are obtained by a local log-likelihood function with kernel smoothing. A fixed-point iterative algorithm is provided to facilitate the implementation. We discuss other practical issues such as the selection of the bandwidth and tuning parameters, the choice of the copula family, and the estimation of the asymptotic covariance matrix.

Third, we proceed comprehensive simulations and investigate an empirical application. The simulation results demonstrate that the proposed method works well in selecting the appropriate state variables and estimating the unknown index coefficients and dependence parameters. When applied to housing prices in the United States, the index copula model encompasses eight macroeconomic variables to account for the heterogeneity in dependence. The variable selection results suggest that several macroeconomic factors such as the unemployment rate and GDP growth rate substantially contribute to the dependence among housing markets in Arizona, California, Florida and Nevada, which were devastatingly hit during the 2007-2008 economic recession.

The remaining structure of the paper is as follows. Section 2 introduces the proposed index copula model and the three-step estimation method. In the same section, we establish large sample properties for both unpenalized estimators and penalized estimators. Section 3 reports the finite sample simulation results. Section 4 applies the model to the dependence of housing prices in the United States. The final section draws conclusions. The technical proofs are provided in the Appendix.

2 The single-index copula model

In this paper, we consider the single-index copula model in which its bivariate case can be written as

$$C(F_1(X_1|Z;\psi_1), F_2(X_2|Z;\psi_2)|Z;\theta(\gamma^\intercal Z)), \tag{1}$$

where $Z = (Z_1, Z_2, \cdots, Z_d)^\intercal$ is a vector of state variables, γ is a d-dimensional vector of corresponding loadings, and the dependence parameter θ that controls the strength of the dependence is assumed to be a flexible function of the so called single-index $\gamma^\intercal Z$, i.e., a linear combination of d-dimensional state variables Z. The marginal distributions are of the form $F_i(X_i|Z;\psi_i)$ with q_i-dimensional parameter vectors ψ_i for $i = 1, 2$. We assume that $\|\gamma\| = 1$ and the first element of γ is positive for identification, where $\|\cdot\|$ is the Euclidean norm (L_2-norm). Clearly, the index copula model covers several other existing nonparametric copula models as a special case. For example, when $d = 1$, the index copula model reduces to a nonparametric copula with single exogenous variable as in Acar et al. (2011). When the covariate is time, it becomes to the nonparametric copula model proposed by Hafner and Reznikova (2010).

The single-index copula model in (1) alleviates the "curse of dimensionality" by projecting the d-dimensional state space onto the one-dimensional single index via the dependence parameter θ. This is crucial for empirical studies since there is often a large number of candidate state variables, but only short samples in the case of monthly or quarterly data. Moreover, we can consider the single-index as a composite measure of economic/financial conditions, which are the main drivers of the dynamics of dependence.

By taking first derivatives of the distribution function with respect to X_1 and X_2 sequentially, we obtain its density function as

$$f_1(X_1|Z;\psi_1)f_2(X_2|Z;\psi_2)c(F_1(X_1|Z;\psi_1), F_2(X_2|Z;\psi_2)|Z;\theta(\gamma^\intercal Z)),$$

where $f_i(X_i|Z;\psi_i) = \partial F_i(X_i|Z;\psi_i)/\partial X_i$ are conditional marginal densities and $c(U_1, U_2|Z;\theta(\gamma^\intercal Z)) = \partial^2 C(U_1, U_2|Z;\theta(\gamma^\intercal Z))/\partial U_1 \partial U_2$ is the conditional copula density. The corresponding log-likelihood function for the observed series $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$ is defined as:

$$L(\psi, \theta(\gamma^\intercal Z)) = L_1(\psi_1) + L_2(\psi_2) + L_c(\psi, \theta(\gamma^\intercal Z))$$

where $\psi = (\psi_1^\intercal, \psi_2^\intercal)^\intercal$, $L_i(\psi_i) = \sum_{t=1}^T \ell_{it}(\psi_i)$, $\ell_{it}(\psi_i) := \log f_i(X_{it}|Z_t; \psi_i)$, $i = 1, 2$, and $L_c(\psi, \theta(\gamma^\intercal Z)) = \sum_{t=1}^T \ell_{ct}(\psi, \gamma, \theta)$, $\ell_{ct}(\psi, \gamma, \theta) := \log c(F_1(X_{1t}|Z_t; \psi_1), F_2(X_{2t}|Z_t; \psi_2)|Z_t; \theta(\gamma^\intercal Z_t))$. The joint log-likelihood for the single-index copula model is equal to the sum of the two marginal log-likelihoods and the copula log-likelihood with a single-index form in the dependence parameter θ. To reduce the computational burden for the estimation of the parametric and nonparametric components of the model, we propose a three-step procedure as described in the following.

2.1 Three-step estimator

First, we let $f(\Lambda)$ be the density function of $\Lambda = \gamma^\intercal Z$ and ϵ be a small positive constant. Define the domain of Λ as $A_\Lambda = \{\Lambda : f(\Lambda) \geq \epsilon$; there exist a and b such that $\Lambda \in [a, b]\}$, i.e., A_Λ is the set of bounded Λ whose density is bounded away from 0. Assume that the copula parameter $\theta(\gamma^\intercal Z_t)$ is linked to the conditional Spearman rank correlation $\rho(X_{1t}, \rho_{2t}|Z_t)$ as

$$\theta(\gamma^\intercal Z_t) = g(\rho(X_{1t}, X_{2t}|Z_t))$$

where g is a nonlinear, strictly monotone function. For example, in the bivariate Gaussian copula case it is given by $g(x) = 2\sin(\frac{\pi}{6}x)$, $x \in [-1, 1]$. Many other popular copulas satisfy this assumption. An estimator of the conditional Spearman rank correlation will allow us to obtain an initial estimator of γ.

We propose the following three-step procedure to estimate the parametric and nonparametric components of the model:

Step 1 For the marginal distributions, their parameters $\hat{\psi}_1$ and $\hat{\psi}_2$ are respectively estimated by maximizing the marginal log-likelihoods $L_1(\psi_1)$ and $L_2(\psi_2)$, i.e.,

$$\hat{\psi}_i = \arg\max_{\psi_i} L_i(\psi_i) \quad \text{for} \quad i = 1, 2.$$

This step uses classical parametric estimation and is referred to as the inference function for margins (IFM) approach by Joe (2000). Under weak regularity conditions, the standard convergence rate of the estimators is $\|\hat{\psi}_i - \psi_{0i}\| = O_p(1/\sqrt{T})$ where ψ_{0i} is the true value of ψ_i. Denote the true value vector $\psi_0 = (\psi_{01}^\intercal, \psi_{02}^\intercal)^\intercal$ and its corresponding estimator vector $\hat{\psi} = (\hat{\psi}_1^\intercal, \hat{\psi}_2^\intercal)^\intercal$ hereafter.

Step 2 Let $\hat{U}_{it} := F_i(X_{it}|Z_t; \hat{\psi}_i)$, $i = 1, 2$. An initial estimator of the index coefficient $\tilde{\gamma}$ is obtained by maximizing the global log-likelihood given as

$$\tilde{\gamma} = \arg\max_{\gamma} \frac{1}{T} \sum_{t=1}^{T} \log c(\hat{U}_{1t}, \hat{U}_{2t}|Z_t; \tilde{\theta}_{-t}(\gamma^\intercal Z_t)). \tag{2}$$

where $\tilde{\theta}_{-t}(\Lambda_t)$ is a nonlinear transformation of a nonparametric initial estimator of Spearman's rho, i.e.

$$\tilde{\theta}_{-t}(\Lambda_t) = g(\hat{\rho}_{-t}(\Lambda_t))$$

$$\hat{\rho}_{-t}(\Lambda_t) = \frac{\hat{\Sigma}_{12,-t}(\Lambda_t)}{\sqrt{\hat{\Sigma}_{11,-t}(\Lambda_t)\hat{\Sigma}_{22,-t}(\Lambda_t)}}$$

where $\hat{\Sigma}_{ij,-t}(\Lambda_t)$ is a Nadaraya-Watson leave-one-out estimator of the conditional covariance of $(\hat{U}_{i\tau}, \hat{U}_{j\tau})$ evaluated at Λ_t. That is,

$$\hat{\Sigma}_{ij,-t}(\Lambda_t) = \frac{\sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_b(\gamma^\intercal(Z_\tau - Z_t))\hat{U}_{i\tau}\hat{U}_{j\tau}}{\sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_b(\gamma^\intercal(Z_\tau - Z_t))} - \hat{\mu}_{i,-t}(\Lambda_t)\hat{\mu}_{j,-t}(\Lambda_t), \quad i, j \in \{1, 2\}$$

where $\hat{\mu}_{i,-t}(\Lambda_t)$ is a corresponding estimator of the conditional mean evaluated at Λ_t, i.e.

$$\hat{\mu}_{i,-t}(\Lambda_t) = \frac{\sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_b(\gamma^\intercal(Z_\tau - Z_t))\hat{U}_{i\tau}}{\sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_b(\gamma^\intercal(Z_\tau - Z_t))}, \quad i \in \{1, 2\}$$

and where $k_b(\cdot) = k(\cdot/b)/b$ with $k(\cdot)$ being a kernel function and b a bandwidth.

Step 3 The nonparametric dependence estimator $\hat{\theta}(\Lambda)$ at a given index point Λ is obtained by maximizing the local log-likelihood copula function,

$$\hat{\theta}(\Lambda) = \arg\max_{\theta} \frac{1}{T} \sum_{t=1}^{T} \log c(\hat{U}_{1t}, \hat{U}_{2t} | Z_t; \theta(\Lambda)) k_h(\tilde{\gamma}^T Z_t - \Lambda)$$

and the index coefficient γ is estimated by

$$\hat{\gamma} = \arg\max_{\gamma} \frac{1}{T} \sum_{t=1}^{T} \log c(\hat{U}_{1t}, \hat{U}_{2t} | Z_t; \hat{\theta}(\gamma^\intercal Z_t)). \tag{3}$$

In fact, it is challenging to find a solution for $\hat{\gamma}$ in (3) when its dimension is large. To reduce the computational burden, we obtain the estimator by a fixed-point iterative algorithm which will be described with details in Section 2.3.

To study the asymptotic properties of the unpenalized estimators under α-mixing conditions, we introduce the following assumptions:

(A1) The copula is Lipschitz continuous in its dependence parameter θ. And the function $\theta(\cdot)$ is continuous, bounded, not constant everywhere and has second order continuous derivatives on A_Λ.

(A2) There exists no perfect multicollinearity within the components of Z, and none of the components of Z is constant.

(A3) The parameter γ is element of Γ, a compact subset of $\mathbb{R}^d$, and γ_0 lies in the interior of Γ. The first element of γ is positive and $\|\gamma\| = 1$, where $\|\cdot\|$ is the Euclidean norm (L_2 norm).

(A4) For any $\gamma \in \mathbb{R}^d$ and $\Lambda \in A_\Lambda$, the density function $f(\Lambda)$ is continuous and there exists $\epsilon > 0$ such that $f(\Lambda) \geq \epsilon$.

(A5) The joint likelihood function L_c is three times differentiable with respect to θ and twice differentiable with respect to ψ_1 and ψ_2. The marginal likelihood function $L_i(\psi_i)$ is twice differentiable with respect to ψ_i.

(A6) The kernel function $k(z)$ is twice continuously differentiable on its support, and its second derivative satisfies a Lipschitz condition. Define the kernel constants $\mu_2 = \int z^2 k(z)dz < \infty$ and $\nu_0 = \int k^2(z)dz < \infty$.

(A7) The bandwidth $h = h_T$ satisfies $h \to 0$ and $Th \to \infty$, as $T \to \infty$. And the bandwidth $b = b_T$ satisfies $b \to 0$ and $Tb/(-\log b) \to \infty$, as $T \to \infty$.

(A8) Assume that $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^{T}$ is a strictly stationary α-mixing sequence. Further, assume that there exists some constant $c > 0$ such that $E|X_{1t}|^{2(2+c)} < \infty$, $E|X_{2t}|^{2(2+c)} < \infty$, $E|Z_t|^{2(2+c)} < \infty$, and the mixing coefficient $\alpha(\ell)$ satisfying $\alpha(\ell) = O(\ell^{-\vartheta})$ with $\vartheta = (2+c)(1+c)/c$.

Conditions (A1)-(A3) are mild conditions for identification. It is obvious that γ cannot be identified if θ is a constant. The no perfect multicollinearity condition in Assumption (A2) is similar to that in classical linear models. A constant is excluded from Z as it can be absorbed by the nonparametric function θ. As γ is identified up to sign and scale, Condition (A3) imposes sign and scale restrictions for identification. Condition (A4) implies that the distance between two ranked values $\gamma^\intercal Z_{(t)}$ is at most of order $O_p(\log T/T)$, see Janson (1987). For any value $\Lambda \in A_\Lambda$, we can find a closest value $\gamma^\intercal Z_t$ to $\Lambda = \gamma^\intercal Z$ such that $|\gamma^\intercal Z_t - \Lambda| = O_p(\log T/T)$. Under Condition (A1), $\|\theta(\Lambda_t) - \theta(\Lambda)\| = O_p(\log T/T)$, which is of smaller order than the nonparametric convergence rate $T^{-2/5}$. This implies that we only need to estimate $\hat{\theta}(\Lambda_t)$ for $t = 1, 2, \cdots, T$, rather than $\hat{\theta}(\Lambda)$ for all values in the domain A_Λ. The detailed arguments can be found in Wang and Xia (2009). Condition (A5) is for deriving the asymptotic distribution and Conditions (A6) and (A7) are common assumptions in nonparametric estimation. In our simulation and empirical study, the commonly adopted Epanechnikov kernel function $k(u) = 3/4(1-u^2)I(|u| \leq 1)$ is used, where $I(|u| \leq 1)$ takes the value 1 if $|u| \leq 1$ and 0 otherwise. The bandwidth b in (A7) is used for obtaining an initial estimator of the index coefficient γ. The conditions in

(A8) are the common conditions with weakly dependent data. Most financial time series models, such as ARMA and GARCH, satisfy these conditions, see e.g. Cai (2002).

To find the asymptotic properties of the estimators, we introduce some notations. Let $\xi = (\psi^\intercal, \gamma^\intercal)^\intercal$, $\xi_0 = (\psi_0^\intercal, \gamma_0^\intercal)^\intercal$, $\Lambda = \gamma^\intercal Z$, $m_z(\Lambda) = E(Z|\Lambda)$ and $\ell'_{it}(\psi_i) = \partial \ell_{it}(\psi_i)/\partial \psi_i$. The efficient score function related to the copula part of the model is given by $\pi(\theta, m_z, \xi) = \sum_{t=1}^T \pi_t(\theta, m_z, \xi)$, where $\pi_t(\theta, m_z, \xi) = \ell'_{ct}(\psi, \gamma, \theta)\theta'(\Lambda_t)(Z_t - m_z(\Lambda_t))$, the derivative $\ell'_{ct}(\psi, \gamma, \theta) = \partial \ell_{ct}(\psi, \gamma, \theta)/\partial\theta$, and $\theta'(\cdot)$ is the first derivative of the function $\theta(\cdot)$ which exists by Assumption (A1). Similarly, we can define $\hat{\xi} = (\hat{\psi}^\intercal, \hat{\gamma}^\intercal)^\intercal$, $\hat{m}_z(\Lambda) = \hat{E}(Z|\Lambda)$ with $\hat{E}(Z|\Lambda)$ being the local constant (Nadaraya-Watson) estimate of Z at Λ, and $\pi(\hat{\theta}, \hat{m}_z, \xi) = \sum_{t=1}^T \ell'_{ct}(\psi, \gamma, \hat{\theta})\hat{\theta}'(\Lambda_t)(Z_t - \hat{m}_z(\Lambda_t))$. Further, we define the efficient score vector for the complete set of parameters as

$$\Pi(\theta, m_z, \xi) = \begin{pmatrix} \partial L_1(\psi_1)/\partial\psi_1 \\ \partial L_2(\psi_2)/\partial\psi_2 \\ \pi(\theta, m_z, \xi) \end{pmatrix} \quad \text{and} \quad \Pi(\hat{\theta}, \hat{m}_z, \xi) = \begin{pmatrix} \partial L_1(\psi_1)/\partial\psi_1 \\ \partial L_2(\psi_2)/\partial\psi_2 \\ \pi(\hat{\theta}, \hat{m}_z, \xi) \end{pmatrix}.$$

Theorem 1. *Let $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$ be a strictly stationary α-mixing sequence following the index copula model in (1). Under Conditions (A1)-(A8), as $T \to \infty$, $\|\hat{\xi} - \xi_0\| = O_p(T^{-1/2})$ and*

$$\sqrt{T}(\hat{\xi} - \xi_0) \xrightarrow{d} N(0, V),$$

where $V = M^{-1}\Omega(M^{-1})^\intercal$ with $M = -E\ \partial\Pi(\theta, m_z, \xi_0)/\partial\xi$ and $\Omega = \sum_{j=-\infty}^\infty \Gamma_j$ with $\Gamma_j = \operatorname{Cov}(\zeta_t, \zeta_{t-j})$ and $\zeta_t := (\ell'_{1t}(\psi_{01})^\intercal, \ell'_{2t}(\psi_{02})^\intercal, \pi_t(\theta, m_z, \xi_0)^\intercal)^\intercal$.

If the random vector sequence $\{\zeta_t\}_{t=1}^\infty$ is either i.i.d. or a martingale difference sequence, then the long-run variance Ω simplifies to $\Omega = \Gamma(0) = \operatorname{Var}(\zeta_t)$. Otherwise, the autocovariance function $\Gamma(j)$ may not be zero at least for some lag order $j \neq 0$ due to the serial correlation of ζ_t.

Remark 1. *To find the asymptotic distribution of the estimator of γ in the index copula model, we define $\iota = (\boldsymbol{0}, I_d)$, where $\boldsymbol{0}$ is a $d \times q$ matrix of zeros with $q = q_1 + q_2$, and*

I_d is the d-dimensional identity matrix, where q and d are the dimensions of ψ and γ, respectively. Then, as $T \to \infty$, $\sqrt{T}(\hat{\gamma} - \gamma_0) \xrightarrow{d} N(0, V_\gamma)$ where $V_\gamma = \iota V \iota^\intercal$.

Theorem 2. *Let $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$ be a strictly stationary α-mixing sequence following the index copula model in (1). Assume that the bandwidth $h = O_p(T^{-1/5})$ and the estimators $\hat{\psi}$ and $\hat{\gamma}$ satisfy $\|\hat{\psi} - \psi_0\| = O_p(1/\sqrt{T})$ and $\|\hat{\gamma} - \gamma_0\| = O_p(1/\sqrt{T})$, respectively. For a fixed point Λ lying in the interior of the support A_Λ, under Conditions (A1)-(A8), as $h \to 0$ and $Th \to \infty$, we have*

$$\sqrt{Th}\{\hat{\theta}(\Lambda) - \theta(\Lambda) - h^2 B(\Lambda)\} \xrightarrow{d} N\left(0, \frac{\nu_0}{f(\Lambda)} \Sigma(\Lambda)^{-1} \Phi(\Lambda) \Sigma(\Lambda)^{-1}\right),$$

where $B(\Lambda) = \frac{1}{f(\Lambda)} \theta'(\Lambda) f'(\Lambda) \mu_2 + \frac{1}{2} \theta''(\Lambda) \mu_2$, $\Sigma(\Lambda) = -E\{\ell''_{ct}(\psi, \gamma, \theta(\Lambda)) | \gamma^\intercal Z = \Lambda\}$ and $\Phi(\Lambda) = \sum_{j=-\infty}^{\infty} R_j(\Lambda)$ with $R_j(\Lambda) = E\{\ell'_{ct}(\psi, \gamma, \theta(\Lambda)) \ell'_{c(t+j)}(\psi, \gamma, \theta(\Lambda))^\intercal | \gamma^\intercal Z = \Lambda\}$.

Remark 2. *The condition $\|\hat{\psi} - \psi_0\| = O_p(1/\sqrt{T})$ can be derived from the marginal log-likelihood estimation and the condition $\|\hat{\gamma} - \gamma_0\| = O_p(1/\sqrt{T})$ is obtained from Theorem 1. From Theorem 2, as expected, the initial estimators of both $\hat{\psi}$ and $\hat{\gamma}$ have little effect on the final estimation of $\hat{\theta}(\cdot)$ in large samples, due to the fact that the parametric parts of the model $\hat{\psi}$ and $\hat{\gamma}$ are estimated at a faster convergence rate than the nonparametric component $\hat{\theta}(\cdot)$. The convergence rate of the nonparametric estimator is $\sqrt{Th}$, and the bias term is $h^2 B(\Lambda)$.*

2.2 Penalized estimating function estimator

In the case where many state variables are included in the dependence function, there is a risk of overfitting and efficiency loss. As pointed out in the discussion part of Acar et al. (2011), "*we recommend a careful selection of the variable prior to estimation of the calibration function if more covariates are of potential interest for the conditional copula model.*" This motivates us to do the estimation and variable selection simultaneously. In the literature, variable selection is usually implemented by maximizing a penalized likelihood function or minimizing a penalized least square criterion. In our case, the likelihood function for the index copula model is composed of several parts pertaining to the marginal distributions and the dependence parameter, and we have to make a choice for the penalization of these parts. In a general framework, penalized estimating functions have been proposed by Johnson, Lin and Zeng (2008), where the functions do not necessarily correspond to the likelihood-based criteria.

In this section, we propose the following penalized estimating function,

$$\Pi^P(\hat{\theta}, \hat{m}_z, \xi) = \Pi(\hat{\theta}, \hat{m}_z, \xi) - T\mathbf{p}'_\lambda(|\xi|)\text{sgn}(\xi) \tag{4}$$

where $\mathbf{p}'_\lambda(|\xi|) = (\mathbf{0}^\intercal, p'_{\lambda_1}(|\gamma_1|), \cdots, p'_{\lambda_d}(|\gamma_d|))^\intercal$, with $\mathbf{0}$ being a q-vector of zeros, and $p'_\lambda(\cdot)$ the derivative of a penalty function with tuning parameter λ. We do not penalize the marginal parameters ψ_1 and ψ_2 since our main focus is the dependence structure. An estimator of ξ will be defined as the solution of the equation

$$\Pi^P(\hat{\theta}, \hat{m}_z, \xi) = 0. \tag{5}$$

Note that the estimating equation (5) can be obtained from classical moment conditions and is referred to as an estimating function method (EFM) in Johnson et al. (2008). By the definition of $\pi_t(\theta, m_z, \xi)$, note that $E[\pi_t(\theta, m_z, \xi)] = 0$. The solution of $\Pi^P(\hat{\theta}, \hat{m}_z, \xi) = 0$ yields a consistent estimator of γ as long as one of $\ell'_{ct}(\psi, \gamma, \theta)$ and $\theta'(\Lambda_t)$ is correctly specified, see Zhu, Dong and Li (2013) for details.

Various penalty functions have been proposed over the last decades. As pointed out by Fan and Li (2001), a good penalty function should have three properties: unbiasedness for large parameters, sparsity, and continuity to avoid instability in model prediction. Here, we will use the smoothly clipped absolute deviation (SCAD) penalty function (Fan and Li, 2001) that enjoys all three properties, although many other penalty functions such as LASSO (Tibshirani,1996) and adaptive LASSO (Zou, 2006) are also applicable. The first-order derivative $p'_{\lambda_k}(|\gamma_k|)$ of the continuous differentiable SCAD penalty function

$p_{\lambda_k}(|\gamma_k|)$ is given by

$$p'_{\lambda_k}(|\gamma_k|) = \lambda_k I(|\gamma_k| \le \lambda_k) + \frac{(a\lambda_k - |\gamma_k|)_+}{(a-1)} I(|\gamma_k| > \lambda_k)$$

for some $a > 2$, where $I(\cdot)$ is the indicator function and $(a\lambda_k - |\gamma_k|)_+$ takes its positive value if $a\lambda_k - |\gamma_k|$ is positive, and zero otherwise. For simplicity, we assume that the tuning parameters λ_k are the same for $k = 1, \ldots, d$ by taking $\lambda_k = \lambda_T$. We select $a = 3.7$ from a Bayesian risk point of view as suggested by Fan and Li (2001). They note that this choice provides a good practical performance for various variable selection problems.

To find the asymptotic properties of the proposed penalized estimating function, without loss of generality, we assume the first d_1 coefficients of γ to be nonzero and all $(d - d_1)$ remaining components of γ to be zero. That is, the true parameter vector is decomposed as $\gamma_0 = (\gamma_{10}^\intercal, \gamma_{20}^\intercal)^\intercal$, where all elements of the d_1-dimensional vector γ_{10} are nonzero and all elements of the $(d - d_1)$-dimensional vector γ_{20} are equal to zero. Moreover, we define $\xi_{10} = (\psi_0^\intercal, \gamma_{10}^\intercal)^\intercal$, and let $Z = (Z_1^\intercal, Z_2^\intercal)^\intercal$ with Z_1 being the d_1 relevant state variables, Z_2 the $(d - d_1)$ irrelevant state variables, $\Lambda_1 = \gamma_1^\intercal Z_1$ and $m_{z_1}(\Lambda_1) = E(Z_1|\Lambda_1)$.

Furthermore, we include the following additional technical conditions:

(B1) $\lim_{T\to\infty} \{\sqrt{T} p'_{\lambda_T}(|\gamma_k|)\} = 0$ and $p''_{\lambda_T}(|\gamma_k|) \to 0$ for $k = 1, \cdots, d_1$. For any $C > 0, \lim_{T\to\infty} \{\sqrt{T} \inf_{\|\gamma\| \le C/\sqrt{T}} p'_{\lambda_T}(|\gamma_k|)\} \to \infty$, for $k = d_1 + 1, \cdots, d$.

Condition (B1) holds when the tuning parameter $\lambda_T \to 0$ and $\sqrt{T}\lambda_T \to \infty$ as $T \to \infty$, which are commonly employed in the SCAD penalty based variable selection. See Fan and Li (2001) for details.

The next two theorems respectively postulate the existence of the $\sqrt{T}$-consistent estimator and its oracle property from the proposed penalized estimating functions.

Theorem 3. *Let $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$ be a strictly stationary α-mixing sequence following the index copula model in (1). Under conditions (A1)-(A8) and (B1), there exists a $\sqrt{T}$-consistent estimator $\hat{\xi}$ satisfying both $\|\hat{\xi} - \xi_0\| = O_p(T^{-1/2})$ and $\Pi^P(\hat{\theta}, \hat{m}_z, \hat{\xi}) = 0$.*

The next theorem states that the estimator of ξ_{10} is as efficient as the maximum likelihood estimator that uses the information $\gamma_{20} = 0$, which is commonly called the oracle property.

Theorem 4 (Oracle Property). *Let $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$ be a strictly stationary α-mixing sequence following the index copula model in (1). Assume that the tuning parameter $\lambda_T \to 0$ and $\sqrt{T}\lambda_T \to \infty$ as $T \to \infty$, under conditions (A1)-(A8) and (B1), the $\sqrt{T}$-consistent estimator for $\Pi_k^P(\hat{\theta}, \hat{m}_z, \hat{\xi}) = 0$, denoted by $\hat{\xi} = (\hat{\xi}_1^\intercal, \hat{\gamma}_2^\intercal)^\intercal$ with $\hat{\xi}_1 = (\hat{\psi}^\intercal, \hat{\gamma}_1^\intercal)^\intercal$, satisfies the following two properties:*

(a) Sparsity: $\hat{\gamma}_2 = 0$,

(b) Asymptotic normality:

$$\sqrt{T}(\hat{\xi}_1 - \xi_{10}) \xrightarrow{d} N(0, V_1),$$

where $V_1 = M_1^{-1}\Omega_1(M_1^{-1})^\intercal$ with $M_1 = -E\,\partial\Pi(\theta, m_{z_1}, \xi_{10})/\partial\xi_1$ and $\Omega_1 = \sum_{j=-\infty}^{\infty}\Gamma_{1j}$ with $\Gamma_{1j} = \mathrm{Cov}\,(\zeta_{1t}, \zeta_{1,t-j})$ and $\zeta_{1t} = (\ell_{1t}'(\psi_{01})^\intercal, \ell_{2t}'(\psi_{02})^\intercal, \pi_t(\theta, m_{z_1}, \xi_{01})^\intercal)^\intercal$.

Sparsity is an important statistical property in high-dimensional statistics. By assuming that only a small subset of state variables is important, the sparsity principle reduces complexity and improves the model's interpretability and predictability. The sparsity property from Theorem 4 demonstrates that the proposed penalized estimating function based copula model shrinks the zero components of the true parameter vector to zero with probability one as the sample size T goes to infinity.

2.3 Practical issues

A. Algorithm In the following, we only present the algorithm for penalized estimators in (5), since we can remove the penalty term in (5) to find the unpenalized estimators in (3).

In Step 3 of the estimation procedure in Section 2.1, it is challenging to find a solution for $\hat{\gamma}$ in (5) since the assumption $\|\gamma\| = 1$ represents a non-standard problem: the true coefficient γ_0 is on the boundary of a unit ball. Moreover, we do not have closed form expressions for $\hat{\gamma}$ and the local likelihood estimator $\hat{\theta}$. Numerical optimization algorithms such as Newton-Raphson might be sensitive to starting values and suffer from heavy computational costs.

We propose to use a fixed-point iterative algorithm (Cui, Härdle and Zhu, 2011), which is less sensitive to initial values, to compute the penalized estimator in (5). First, we standardize all state variables to have mean zero and standard deviation one. Define $\pi^P(\hat{\theta}, \hat{m}_z, (\hat{\psi}^\intercal, \hat{\gamma}^\intercal)^\intercal) = \pi(\hat{\theta}, \hat{m}_z, (\hat{\psi}^\intercal, \hat{\gamma}^\intercal)^\intercal) - T\mathbf{p}'_\lambda(|\gamma|)\mathrm{sgn}(\gamma)$ with $\mathbf{p}'_\lambda(|\gamma|) = (p'_{\lambda_1}(|\gamma_1|), \cdots, p'_{\lambda_d}(|\gamma_d|))^\intercal$ and efficient score function $\pi(\cdot)$ being defined in Section 2.1. Let $\pi^P(\hat{\gamma}^{(0)}) = \pi^P(\hat{\theta}, \hat{m}_z, (\hat{\psi}^\intercal, \hat{\gamma}^{(0)\intercal})^\intercal)$ and $\pi_1^P(\hat{\gamma}^{(0)})$ denotes the first element of $\pi^P(\hat{\gamma}^{(0)})$.

Then, the coefficient γ can be iteratively updated by $\hat{\gamma}^{(j)}/\|\hat{\gamma}^{(j)}\|$ with

$$\hat{\gamma}^{(j)} = \frac{C_1}{\pi_1^P(\hat{\gamma}^{(j-1)})/\|\pi^P(\hat{\gamma}^{(j-1)})\| + C_1}\hat{\gamma}^{(j-1)} + \frac{|\pi_1^P(\hat{\gamma}^{(j-1)})|/\|\pi^P(\hat{\gamma}^{(j-1)})\|^2}{\pi_1^P(\hat{\gamma}^{(j-1)})/\|\pi^P(\hat{\gamma}^{(j-1)})\| + C_1}\pi^P(\hat{\gamma}^{(j-1)}), \tag{6}$$

for $j = 1, 2, \ldots$, where the marginal parameters $\hat{\psi}$ are estimated from Step 1 in Section 2.1, and $\hat{\theta}$ and $\hat{m}_z$ are the nonparametric estimators with given $\hat{\psi}$ and $\hat{\gamma}^{(j-1)}$. The derivative $\hat{\theta}'(\hat{\gamma}^{(0)\intercal}Z_t)$ is numerically approximated by $\hat{\theta}'(\hat{\gamma}^{(0)\intercal}Z_t) = (\hat{\theta}(\hat{\gamma}^{(0)\intercal}Z_{\bar{t}}) - \hat{\theta}(\hat{\gamma}^{(0)\intercal}Z_t))/(\hat{\gamma}^{(0)\intercal}Z_{\bar{t}} - \hat{\gamma}^{(0)\intercal}Z_t)$, where $\hat{\gamma}^{(0)\intercal}Z_{\bar{t}}$ is the closest value to $\hat{\gamma}^{(0)\intercal}Z_t$ in the sequence $\{\hat{\gamma}^{(0)\intercal}Z_t\}_{t=1}^T$ except for $\hat{\gamma}^{(0)\intercal}Z_t$ itself. C_1 is a constant satisfying $\pi_1^P(\hat{\gamma}^{(0)})/\|\pi^P(\hat{\gamma}^{(0)})\| + C_1 \neq 0$ for any $\hat{\gamma}^{(0)}$. We choose $C_1 \in [2/\sqrt{d}, d/2]$ which gives a good practical performance as suggested by Cui et al. (2011). We iterate equation (6) until $\hat{\gamma}$ converges. The final vector $\hat{\gamma}^{(j)}/\|\hat{\gamma}^{(j)}\|$ is the estimate of γ_0.

B. Choosing bandwidth and tuning parameters To do the nonparametric estimation and variable selection simultaneously, we need to choose suitable regularization parameters, i.e., the bandwidth h for the nonparametric estimator and λ_T for the penalty terms. Various methods for the selection of bandwidths and tuning parameters have been discussed in the variable selection literature, such as cross-validation, AIC- and BIC-type criteria, among others. Due to the time series nature of the sequence $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^T$, we propose to use the *forward leave-one-out* cross-validation to select both the bandwidth h and tuning parameter λ_T in the penalty term simultaneously.

Define $\hat{\Theta}(h, \lambda_T) \equiv (\hat{\theta}(h, \lambda_T), \hat{m}_z(h, \lambda_T), \hat{\xi}(h, \lambda_T))$ as the parametric and nonparametric estimators for the penalized index copula models in (5) with known bandwidth h and tuning parameter λ_T. For each data point $t_0 + 1 \leq t^* \leq T$, we use the data $\{X_{1t}, X_{2t}, Z_t, t < t^*\}$ to construct the estimate $\hat{\Theta}_{t^*}(h, \lambda_T)$ at the sample point $\{X_{1t^*}, X_{2t^*}, Z_{t^*}\}$, where t_0 is the minimum window size used to estimate $\hat{\Theta}_{t_0+1}(h, \lambda_T)$. Under this forward recursive scheme, we obtain the sequential estimators $\{\hat{\Theta}_{t^*}(h, \lambda_T)\}_{t^*=t_0+1}^{T}$ and the optimal bandwidth h^* and tuning parameter λ_T^* can be obtained by maximizing the objective function

$$(h^*, \lambda_T^*) = \arg \max_{(h, \lambda_T)} \sum_{t^*=t_0+1}^{T} \{\ell_{ct^*}(\hat{\psi}, \hat{\gamma}, \hat{\theta}) | \hat{\Theta}_{t^*}(h, \lambda_T)\}.$$

It is clear that (h^*, λ_T^*) is the *forward leave-one-out* cross-validation estimator in terms of the log-likehood.

C. Choosing copula families In practice, any copula-based modeling must be accompanied by a strategy to select among a large set of candidate copula families to best approximate the time series data at hand. The selection for copula family in a parametric setting is usually implemented by comparing the likelihood values plus the AIC or BIC penalty term (Patton, 2004 and Patton, 2006). To verify how well the model fits the underlying process, the goodness-of-fit (GoF) test for parametric copula models was developed by using a Rosenblatt probability integral transformation (Dobric and Schmid, 2007). However, the same approach does not apply to the selection of copula family in a semiparametric or nonparametric setting, as the scales of the likelihood vary across families via nonparametric estimation. Following the idea of Acar et al. (2011) who select copula with i.i.d. data, we propose to use a cross-validated prediction error (CVPE) method for time series observations.

Let $\mathcal{C} := \{\mathcal{C}_1, \ldots, \mathcal{C}_N\}$ be a set of N candidate copula families. The aim is to select an optimal copula family $\mathcal{C}^*$ that best fits the time series sequence $\{X_{1t}, X_{2t}, Z_t\}_{t=1}^{T}$. By the same token as in choosing the bandwidth and tuning parameters, we use the data $\{X_{1t}, X_{2t}, Z_t, t < t^*\}$ to construct the estimate $\hat{\Theta}_{t^*}(\mathcal{C}_j)$ for the j-th copula family $\mathcal{C}_j$, where $t_0 + 1 \le t^* \le T$ and t_0 is the minimum window size used to estimate $\hat{\Theta}_{t_0+1}(\mathcal{C}_j)$. Then, the CVPE for the j-th copula family $\mathcal{C}_j$ is defined as

$$\mathrm{CVPE}(\mathcal{C}_j) = \sum_{t^*=t_0+1}^{T} (\hat{U}_{1t^*} - \widetilde{U}_{1t^*}(\mathcal{C}_j))^2 + \sum_{t^*=t_0+1}^{T} (\hat{U}_{2t^*} - \widetilde{U}_{2t^*}(\mathcal{C}_j))^2$$

where $\widetilde{U}_{1t^*}(\mathcal{C}_j) = \int_0^1 U_1 c(U_1, U_2 | \hat{\Theta}_{t^*}(\mathcal{C}_j)) dU_1$ and $\widetilde{U}_{2t^*}(\mathcal{C}_j) = \int_0^1 U_2 c(U_1, U_2 | \hat{\Theta}_{t^*}(\mathcal{C}_j)) dU_2$. The optimal $\mathcal{C}^*$ is then defined as $\mathcal{C}^* = \arg\min_{\mathcal{C}} \mathrm{CVPE}(\mathcal{C}_j)$.

D. Estimation of the variance covariance matrix It is difficult to estimate the variance-covariance matrix $T^{-1}V$ directly since it is not straightforward to calculate the exact form of M. To avoid this difficulty, we adopt the idea of jackknife suggested by Joe (2000) to estimate $T^{-1}V$. Let $\hat{\xi}^{(-t)}$ be the leave-one-out estimator of ξ, $t = 1, 2, \ldots, T$, then the jackknife estimator of $T^{-1}V$ can be calculated by $\sum_{t=1}^{T} (\hat{\xi}^{(-t)} - \hat{\xi})(\hat{\xi}^{(-t)} - \hat{\xi})^T$. If the sample size T is large enough, we can extend the estimator to the leave-one-block-out jackknife estimator. Denote the full data set (X_1, X_2, Z) by $\mathcal{F}$, and the training set and the test set by $\mathcal{F} \backslash \mathcal{F}^b$ and $\mathcal{F}^b$ for $b = 1, 2, \cdots, \kappa$, respectively, where κ is the number of blocks. For each b, we obtain the estimator $\hat{\xi}^{(-b)}$ from the training set $\mathcal{F} \backslash \mathcal{F}^b$. Then the leave-one-block-out jackknife estimator of $T^{-1}V$ is calculated by $\sum_{b=1}^{\kappa} (\hat{\xi}^{(-b)} - \hat{\xi})(\hat{\xi}^{(-b)} - \hat{\xi})^\intercal$. Note that the leave-one-out jackknife estimator is a special case of the leave-one-block-out estimator if there is only one observation in the test set $\mathcal{F}^b$.

3 Numerical studies

In this section, we investigate the finite-sample performance of the proposed estimation and selection method through a series of numerical studies. We consider a bivariate case where the data are generated by an ARMA(1,1)-GARCH(1,1) process:

$$x_{it} = \rho_i x_{i,t-1} + e_{it} + \varphi_i e_{i,t-1}, \qquad i = 1, 2;\ t = 1, ..., T,$$

where $\rho_1 = 0.05$, $\rho_2 = 0.1$, $\varphi_1 = 0.1$, $\varphi_2 = 0.2$, and $e_{it} = \sigma_{it}\epsilon_{it}$. We further assume that

$$\sigma_{it}^2 = \alpha_{i0} + \alpha_{i1} e_{i,t-1}^2 + \beta_{i1}\sigma_{i,t-1}^2,$$

where $\alpha_{10} = 10^{-5}$, $\alpha_{11} = 0.05$, $\beta_{11} = 0.90$ for the first margin, and $\alpha_{20} = 10^{-5}$, $\alpha_{21} = 0.04$, $\beta_{21} = 0.91$ for the second margin. The dependence structure between ϵ_{1t} and ϵ_{2t} is governed by a single-index copula model

$$C(U_1, U_2; \theta(\gamma^\intercal Z)),$$

where U_1 and U_2 respectively denote the marginal distributions of ϵ_{1t} and ϵ_{2t} which are respectively assumed to be student's t-distribution with 3 and 4 degrees of freedoms. The copula dependence parameter θ is a function of $\gamma^\intercal Z$, where $Z = (Z_1, Z_2, \cdots, Z_5)^\intercal$ is a vector of state variables and $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_5)^\intercal$ is a vector of corresponding loadings satisfying the assumptions that γ_1 is positve and $||\gamma|| = 1$. We generate $Z_1, Z_2, \cdots, Z_5$ from a truncated standard normal distribution with the interval of truncation $[-1, 1]$, and then standardize all state variables to have mean zero and standard deviation one. Moreover, three of the five index coefficients $\gamma_1, \gamma_2, \cdots, \gamma_5$ (say $\gamma_3 \sim \gamma_5$) are set to be zeros. The goal of the simulation studies is to check whether the proposed method can accurately estimate the unknown parameters and correctly select the state variables.

We respectively consider three commonly used copulas: the Clayton, Gumbel and Frank copulas as the copula function. These three copulas are suitable for different dependence structures. As is well known, the Clayton copula exhibits strong lower tail dependence, and can well capture cases such as simultaneous crash in two markets. The Gumbel copula shows strong upper tail dependence and can be an appropriate model when two markets tend to boom together. In contrast to the Clayton and Gumbel copulas, the Frank copula shows no tail dependence and strong dependence appears at the center.

For each copula function, we consider two models for the index coefficients and the dependence parameters. Therefore, we have $3 \times 2 = 6$ cases in total:

- Model 1: $\gamma_1 = \frac{\sqrt{2}}{2}$, $\gamma_2 = \frac{\sqrt{2}}{2}$, $\gamma_3 = \gamma_4 = \gamma_5 = 0$; $\theta(\Lambda) = \exp(\frac{5}{12}\Lambda + 1) + 2$ for Clayton and Gumbel, and $\theta(\Lambda) = \exp(\frac{5}{12}\Lambda + 1)$ for Frank;
- Model 2: $\gamma_1 = \frac{1}{2}$, $\gamma_2 = \frac{\sqrt{3}}{2}$, $\gamma_3 = \gamma_4 = \gamma_5 = 0$; $\theta(\Lambda) = 6.6 - 0.7\Lambda^2$ for Clayton and Gumbel, and $\theta(\Lambda) = 4.6 - 0.7\Lambda^2$ for Frank.

In addition, two sample sizes, $T = 500$ and 1000, are considered for each case, and all simulations are repeated 1000 times ($M = 1000$). For each sample, we calculate the estimates for the dependence parameters at 101 equally-spaced grid points $\Lambda_i = -2 + 0.04i$ for $i \in \{0, 1, ..., 100\}$.

We use Table 1 to check whether the proposed method can correctly select the state variables. The numbers in Table 1 present the percentages corresponding to the correctly and incorrectly (in parentheses) selected state variables for the 6 cases respectively. The results in Table 1 demonstrate that the proposed method can select the appropriate state variables with high accuracy. For all cases of $T = 500$, the correct state variables are selected with at least 93% probability. When the sample size increases to 1000, the probability of selecting the important state variables is 100% for all cases except for Model 2 of Frank (which is 98.7%). Furthermore, the percentage that the redundant state variables (the variables that are not included in the true index copula models) are selected is small. For example, when $T = 1000$, the probability of selecting the redundant state variables is 7.1% and 6.9% respectively for Models 1 and 2 of Clayton, and 4.9% and 5.1% respectively for Models 1 and 2 of Gumbel. Even for the worst case (Model 2 of Frank), the probability is still not higher than 10%.

To investigate the performance of the proposed method in estimating the unknown index coefficients, we report the means, medians, and mean square errors (MSEs) of the estimators of the nonzero index coefficients (γ_1 and γ_2) for all 6 cases in Table 2. One can easily see that as sample size increases, the means and medians are getting closer to the true values and the MSEs decrease. Moreover, the proposed method can accurately estimate the unknown index coefficients.

The MSEs of the estimated copula dependence parameter θ respectively shown in the 6th, 9th and 12th columns of Table 2 are calculated as

$$MSE(\widehat{\theta}) = \frac{1}{M}\frac{1}{101}\sum_{j=1}^{M}\sum_{i=1}^{101}\left(\widehat{\theta}_j(\Lambda_i) - \theta(\Lambda_i)\right)^2.$$

We observe that as the sample size increases from 500 to 1000, the MSEs of the estimated θ decrease and the estimates become more accurate.

Figure 1 further evaluates the performance of the proposed method in estimating θ by presenting the estimated and the true dependent parameter paths for different models. We depict the true dependence parameter function through a black solid line, and use two curves to respectively represent medians (blue) and means (red) of the 1000 simulation parameter function estimates at the grid points. The two black dashed curves represent the 5% and 95% percentiles of the dependence parameter estimates at the grid points. We only present the results for $T = 1000$ to save space. Figure 1 shows that in all cases, the median and mean curves closely follow the true value paths, which again confirms the good performance of the proposed method in estimating the unknown dependence parameters.

4 Empirical applications

In this section, we use the proposed estimation method to examine the dependence of housing prices in the United States. Asset prices are well known to be closely correlated and housing prices are not an exception. Even though housing markets in different areas are commonly believed to be localized, they are often related across different geographic areas because prices are sensitive to nation-wide macroeconomic factors such as monetary policy and fiscal policy (Del Negro and Otrok, 2007 and Kallberg, Liu and Pasquariello, 2014).

We collect quarterly percentage changes of the Housing Prices Index (HPI) in four states — Arizona, California, Florida and Nevada — from the U.S. Federal Housing Finance Agency (FHFA). These four states were strongly hit during the economic recession of 2008-2010 (Zimmer, 2012). The data span the period from 1975:Q1 to 2018:Q1, a period that witnessed five recessions. Figure 2 shows the paths of the quarterly percentage changes of HPI in the four states and the five economic recessions (indicated by the five shaded areas) identified by the National Bureau of Economic Research (NBER). For the four states, the greatest volatility in housing prices happened during the late 1970s and early 1980s. Since then, housing prices in these states have become less volatile until 2006, the eve of the subprime mortgage crisis which caused substantial decline in housing prices through almost the whole country. For potentially relevant state variables, we collect eight national economic factors from the Federal Reserve Economic Data (FRED): the consumer price index (CPI), the quarterly growth of per capita real GDP (GDP), the quarterly growth of real disposable personal income (INC), the effective real federal funds rate (INT), the quarterly growth of residential investment as a percent of GDP (INV), the quarterly growth of industrial production index (IPI), the quarterly growth of oil price (West Texas Intermediate, OIL), and the civilian unemployment rate (UNE). All these factors can, to some extent, mirror the macroeconomic situation and are tracked closely by investors and policy makers. The descriptive statistics for the percentage changes of housing prices and the eight state variables are documented in Table 3.

Preliminary examinations suggest that autocorrelation and autoregressive conditional heteroscedasticity exist in all four states' housing price series. The Ljung-Box test (not tabulated here) indicates that the percentage changes of housing prices are serially correlated up to the fourth order. To avoid spurious dependence, we follow Chen and Fan (2006) and use AR-GARCH models for filtering. For example, we fit the percentage changes of housing prices to an AR(p)-GARCH(1,1) process specified as

$$y_{it} = \delta_i + \gamma_{i1} y_{i,t-1} + \gamma_{i2} y_{i,t-2} + ... + \gamma_{ip} y_{i,t-p} + e_{it},$$

where y_{it} denotes the growth rate of housing prices at time t for division i and e_{it} is a conditionally heteroskedastic error term. The error term is decomposed as $e_{it} = \sigma_{it}\epsilon_{it}$, where ϵ_{it} is an i.i.d. innovation term with mean zero and variance one. The contemporaneous dependence of the pair $(\epsilon_{1t}, \epsilon_{2t})$ will be characterized by a copula model, to be specified. The conditional variance σ_{it}^2 is specified as a classical GARCH(1,1) model, i.e.,

$$\sigma_{it}^2 = \omega_i + \alpha_{i1}\epsilon_{i,t-1}^2 + \beta_{i1}\sigma_{i,t-1}^2.$$

To capture potentially fat tails in the conditional distribution of the error terms, we assume that the marginal distribution of ϵ_{it} is a standardized Student's t distribution.

Table 4 summarizes the coefficients of the AR-GARCH filtering and most of the results are statistically significant. Ljung-Box tests confirm the zero autocorrelation for the filtered percentage changes. For AZ, CA and NV, we use the AR(1)-GARCH(1,1) model while for FL we use the AR(2)-GARCH(1,1) model. We adopt the Jarque-Bera test to examine whether the four states' housing prices, after the filtering, follow a normal distribution. The test statistics in Figure 3 provide evidence of the non-normality in the four filtered series and support the use of the Student's t distribution for ϵ_{it}.

The set of candidate copulas includes six widely-used copulas in empirical studies: Gaussian, Clayton, Gumbel, Frank, Rotated Clayton and Rotated Gumbel. We first use the proposed CVPE method to choose the optimal copula. Table 5 presents the prediction errors of the six candidate copulas and it shows that Frank is selected for AZ-CA, CA-FL, CA-NV and FL-NV and Gumbel is selected for AZ-FL and AZ-NV. Next, we implement the proposed method to identify state variables which are relevant to the dependence structure of housing prices among the four states. Table 6 documents the estimates of γ associated with the eight state variables. We have two observations on Table 6. Besides several important macroeconomic variables such as the GDP growth and unemployment rate, the share of gross residential investment in GDP also significantly contributes to the dependence structure of housing prices. Shiller (2007) identifies that the residential investment is highly correlated with the business cycle. On average, the magnitude of the γ associated with the share of residential investment in GDP is the largest. As an important channel of housing's contribution to GDP, the residential investment usually includes construction of new single-family and multifamily structures, residential remodeling, production of manufactured homes, and brokers fees. From 1980 to 2005, the residential investment increased from 333 billion dollars to 873 billion dollars, but then sharply dropped to 382 billion dollars in 2010 (U.S. Bureau of Economic Analysis). One can also observe that the industrial production index (IPI) and oil price (OIL) are relatively less important than the other six state variables and are filtered out by SCAD in all six pairs. This finding is consistent with Kallberg, Liu and Pasquariello (2014) who find that the comovements among housing markets can be attributed to the fundamental factors directly influencing real estate prices.

For comparison purposes, we transform the estimated single-index copula parameters into Kendall's τ which ranges between -1 and 1.[1] In Figure 4, we plot the paths of Kendall's τ for the six pairs to examine how the dependence evolves during the sample period. The black solid curve indicates the estimates of Kendall's τ and the red dashed curves represent the 95% confidence interval. We have three observations for Figure 4. First, the dependence of housing prices among the four states are relatively stronger before the 1980s, but the degree of dependence has begun to decline since then. Second, for each of the six pairs, there is an obvious upward trend in the degree of dependence path starting from the end of 1990s and early 2000 when the bubble in the housing market was formed (Shiller, 2007). After the most recent economic recession which led to the burst of housing bubbles, the degrees of dependence (τ) in most pairs have decreased substantially. One exception is the pair of California and Florida: Figure 4(d) shows that the comovement of housing prices in the two states has become strengthened again since 2013.

[1]For Gumbel, Kendall's $\tau = 1 - \frac{1}{\theta}$. For Frank, Kendall's $\tau = 1 + \frac{4(D_1(\theta)-1)}{\theta}$ where $D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{\exp(t)-1}dt$. See Nelson (2006) for details.

5 Conclusion

We propose a simultaneous estimation and variable selection procedure for the single-index copula model, in which the dependence parameter is a deterministic function of a linear combination of multiple state variables. We carry out a three-step procedure to estimate the parametric and nonparametric components of the model, and derive their large sample properties. We then adopt an exhaustive variable selection procedure to select the most relevant state variables from a large candidate set. The simulation results confirm that the proposed method can correctly select the state variables and accurately estimate the unknown parameters. Empirical results demonstrate that, besides several important macroeconomic variables such as the GDP growth rate and unemployment rate, the percentage of residential investment in GDP also significantly contributes to the dependence structure of housing prices. However, the contribution of the industrial production index and oil price is relatively small and these two state variables are filtered out by the SCAD variable selection procedures.

Acknowledgements

We thank professor Liang Peng and professor Jiancheng Jiang for their helpful comments and suggestions. Yang's research was partly supported by NSF of Jiangxi Province (No. 20171BAA208002) and the science and technology project of Jiangxi Province Education Department (No. GJJ170329).

6 Appendix: mathematical proofs

In this appendix we establish the main results of Section 2. We will first show consistency of the initial estimator $\tilde{\gamma}$, for which the following lemmas are useful.

Lemma 1. *Let $\tilde{\Theta}_{-t}(\Lambda)$ denote the probability limit of $\tilde{\theta}_{-t}(\Lambda)$. Under Assumptions (A1)* *to (A8), as* $T \to \infty$,

$$P(\sup_{\gamma} |\tilde{\theta}_{-t}(\Lambda) - \tilde{\Theta}_{-t}(\Lambda)| > \epsilon) \to 0.$$

Proof: Consider first $\hat{\Sigma}_{ij,-t}(\Lambda_t) = \hat{A}_{tT}/B_{tT}$, where

$$\hat{A}_{tT} := \frac{1}{T} \sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_h(\gamma^\intercal (Z_\tau - Z_t)) \hat{U}_{i\tau} \hat{U}_{j\tau}$$

and

$$B_{tT} := \frac{1}{T} \sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_h(\gamma^\intercal (Z_\tau - Z_t)).$$

Note that $\hat{U}_{i\tau} = U_{i\tau} + o_p(1)$, with $U_{it} := F_i(X_{it}|Z_t; \psi_i)$ and the $o_p(1)$ term converging uniformly in Γ. Therefore, uniformly in Γ, $\hat{A}_{tT} = A_{tT} + o_p(1)$ with

$$A_{tT} := \frac{1}{T} \sum_{\tau \neq t} I(\Lambda_\tau \in A_\Lambda) k_h(\gamma^\intercal (Z_\tau - Z_t)) U_{i\tau} U_{j\tau}$$

which converges in probability to A_t. That is to say, we have

$$\sup_{\gamma \in \Gamma} |A_{tT} - A_t| \leq \sup_{\gamma} |A_{tT} - E(A_{tT})| + \sup_{\gamma} |E(A_{tT}) - A_t|.$$

The first term on the right hand size converges to zero in probability under the assumption $Tb_T/(-\log b_T) \to \infty$ using Lemma 8 of Ichimura (1993), and the second term on the right hand size is $O(b_T^2)$ by standard arguments.

Similarly, B_{tT} converges uniformly to its probability limit. This implies that

$$\sup_{\gamma \in \Gamma} |\hat{\rho}_{ij,-t}(\Lambda) - \rho_{ij}(\Lambda)| \to_p 0$$

and

$$\sup_{\gamma \in \Gamma} |\hat{\theta}_{-t}(\Lambda) - \theta(\Lambda)| \to_p 0.$$

Lemma 2. *Under Assumptions (A1) to (A8), as* $T \to \infty$, $\tilde{\gamma} \to_p \gamma$.

Proof: Let

$$\begin{aligned} L_c(\gamma) &= \frac{1}{T}\sum_{t=1}^{T} \log c(U_{1t}, U_{2t}|Z_t; \tilde{\theta}_{-t}(\Lambda)) \\ \tilde{L}_c(\gamma) &= \frac{1}{T}\sum_{t=1}^{T} \log c(U_{1t}, U_{2t}|Z_t; \tilde{\Theta}_{-t}(\Lambda)) \\ \bar{L}_c(\gamma) &= \frac{1}{T}\sum_{t=1}^{T} \mathrm{E} \log c(U_{1t}, U_{2t}|Z_t; \tilde{\Theta}_{-t}(\Lambda)). \end{aligned}$$

Following the same argument as the proof of Theorem 5.1 of Ichimura (1993), it suffices to show that

$$P(\sup_{\gamma\in\Gamma} |L_c(\gamma) - \tilde{L}_c(\gamma)| > \epsilon) \to 0 \tag{7}$$

and

$$P(\sup_{\gamma\in\Gamma} |\tilde{L}_c(\gamma) - \bar{L}_c(\gamma)| > \epsilon) \to 0. \tag{8}$$

First, by Lipschitz continuity in Assumption A1,

$$|L_c(\gamma) - \tilde{L}_c(\gamma)| \leq C|\tilde{\theta}_{-t}(\Lambda) - \tilde{\Theta}_{-t}(\Lambda)|$$

for some constant $C > 0$. Thus, (7) follows by Lemma 1. Second, (8) follows by the uniform convergence theorem of Andrews (1987) as his assumptions A1, A2 and A4 hold under our set of assumptions.

Lemma 3. *Under Assumptions (A1) to (A8), as $T \to \infty$,*

$$\tilde{\gamma} - \gamma_0 = O_p(T^{-1/2}).$$

Proof: By the mean value theorem, for some $\bar{\gamma}$ on the line segment between $\tilde{\gamma}$ and γ_0,

$$\mathcal{L}_T(\tilde{\gamma}) = \mathcal{L}_T(\gamma_0) + \frac{\partial \mathcal{L}_T(\gamma_0)}{\partial \gamma^\intercal}(\tilde{\gamma} - \gamma_0) + \frac{1}{2}(\tilde{\gamma} - \gamma_0)^\intercal \frac{\partial^2 \mathcal{L}_T(\bar{\gamma})}{\partial \gamma \partial \gamma^\intercal}(\tilde{\gamma} - \gamma_0).$$

As $\mathcal{L}_T(\tilde{\gamma}) \geq \mathcal{L}_T(\gamma_0)$ by definition, we have

$$(\tilde{\gamma} - \gamma_0)^\intercal \frac{\partial \mathcal{L}_T(\gamma_0)}{\partial \gamma} - \frac{1}{2}(\tilde{\gamma} - \gamma_0)^\intercal V(\tilde{\gamma} - \gamma_0) + o_p(|\tilde{\gamma} - \gamma_0|^2) \geq 0 \tag{9}$$

where V is the probability limit of $-\frac{\partial^2 \mathcal{L}_T(\bar{\gamma})}{\partial \gamma \partial \gamma^\intercal}$, which is positive definite by assumption. Multiplying both sides of (9) by $T(1 + \sqrt{T}|\tilde{\gamma} - \gamma_0|)^{-2}$, we obtain

$$c_T^\intercal(\tilde{\gamma})\sqrt{T}\frac{\partial \mathcal{L}_T(\gamma_0)}{\partial \gamma^\intercal}(1 + \sqrt{T}|\tilde{\gamma} - \gamma_0|)^{-1} - \frac{1}{2}(\tilde{\gamma} - \gamma_0)^\intercal V(\tilde{\gamma} - \gamma_0) + o_p(1) \geq 0 \tag{10}$$

where $c_T := (1+\sqrt{T}|\tilde{\gamma}-\gamma_0|)^{-1}\sqrt{T}(\tilde{\gamma}-\gamma_0)$. Suppose that $\sqrt{T}|\tilde{\gamma}-\gamma_0|$ diverges in probability to infinity. Then, (10) implies that $c_T^\intercal(\tilde{\gamma})Vc_T(\tilde{\gamma}) \leq 0$. Because V is positive definite, this implies that $|c_T(\tilde{\gamma})| = o_p(1)$, or $\sqrt{T}|\tilde{\gamma} - \gamma_0| = o_p(1)$, which is a contradiction. Therefore, we must have $\sqrt{T}|\tilde{\gamma} - \gamma_0| = O_p(1)$.

Proof of Theorem 1: The proof of consistency for $\hat{\xi}$ is similar to the proof of Theorem 3 with penalty term, so we omit it here. To show asymptotic normality, we first give regularity conditions for the semiparametric method of moments in our index copula model. Let us rewrite the vector of moment functions ζ_t defined in Theorem 1 as

$$\zeta_t(f_t, g_t) = \begin{pmatrix} \ell'_{1t}(\psi_1) \\ \ell'_{2t}(\psi_2) \\ f_t g_t \end{pmatrix}$$

with $f_t = \ell'_{ct}(\psi, \gamma, \theta)$ and $g_t = \theta'(\Lambda_t)(Z_t - m_z(\Lambda_t))$ so that $\Pi(\theta, m_z, \xi) = \sum_{t=1}^T \zeta_t(f_t, g_t)$. It is clear that the moment condition $\mathrm{E}(\zeta_t(f_0, g_0)) = 0$ is satisfied. In the following we suppress the time index for notational simplicity.

Let

$$D(f - f_0, g - g_0) = \frac{\partial \zeta(f_0, g_0)}{\partial f}(f - f_0) + \frac{\partial \zeta(f_0, g_0)}{\partial g}(g - g_0),$$

where $\partial \zeta(f_0, g_0)/\partial f$ and $\partial \zeta(f_0, g_0)/\partial g$ are the Frechet derivatives with $\partial \zeta(f_0, g_0)/\partial f = (0^\intercal, 0^\intercal, g^\intercal)^\intercal$ and $\partial \zeta(f_0, g_0)/\partial g = \big(0^\intercal, 0^\intercal, f^\intercal)^\intercal$.

By construction, we have

$$\|\zeta(f,g)) - \zeta(f_0,g_0) - D(f-f_0,g-g_0)\|_\infty = O_p(\|f-f_0\|_\infty^2 + \|g-g_0\|_\infty^2) \tag{11}$$

where $\|\cdot\|_\infty$ is the supremum norm of a function and its derivatives (Sobolev norm). Assumption 5.1(i) of Newey (1994) holds using equation (11). Using similar methods as in Mack and Silverman (1982), we can show that

$$\begin{aligned} & \|\hat{\theta}(\Lambda) - \theta(\Lambda)\|_\infty = o_p(T^{-1/4}), \qquad \|\hat{\theta}'(\Lambda) - \theta'(\Lambda)\|_\infty = o_p(T^{-1/4}) \\ \text{and} \qquad & \|\hat{m}_z(\Lambda) - m_z(\Lambda)\|_\infty = o_p(T^{-1/4}). \end{aligned}$$

The first term implies that $\|\hat{f}(\Lambda)-f(\Lambda)\|_\infty = o_p(T^{-1/4})$ and the last two terms imply that $\|\hat{g}(\Lambda) - g(\Lambda)\|_\infty = o_p(T^{-1/4})$. Together with equation (11), they imply that Assumption 5.1 (ii) of Newey (1994) holds. Furthermore, his Assumption 5.2 holds by the expression of $D(f-f_0,g-g_0)$.

The estimating equation yields $\frac{1}{T}\sum_{t=1}^{T} \ell'_{ct}(\psi,\gamma,\theta(\Lambda))k_h(\gamma^\intercal Z_t - \Lambda) = 0$ since $\theta(\Lambda)$ is the local maximizer of the objective function $\frac{1}{T}\sum_{t=1}^{T} \ell_{ct}(\psi,\gamma,\theta(\Lambda))k_h(\gamma^\intercal Z_t - \Lambda)$. It follows that $\mathrm{E}(\ell'_c(\psi,\gamma,\theta(\Lambda))|\Lambda) = 0$ by the law of large numbers and $\mathrm{E}(f) = \mathrm{E}(\ell'_c(\psi,\gamma,\theta)) = 0$ by iterated expectations. Moreover, $\mathrm{E}(g) = \mathrm{E}\{\theta'(Z-m_z)\} = 0$ by iterated expectations. Thus, Assumption 5.3 of Newey (1994) holds because of $\mathrm{E}(D(f-f_0,g-g_0)) = 0$.

After having verified Assumptions 5.1-5.3 of Newey (1994) for the index copula model 1, we conclude that $\hat{\xi}$ has the same asymptotic distribution as the solution of the estimating equation

$$\Pi(\theta_0, m_{0z}, \xi) = \sum_{t=1}^{T} \zeta_t(f_0,g_0) = 0.$$

Asymptotic normality then follows directly by Lemma 5.1 of Newey (1994).

$\square$

For the proof of Theorem 2 we will need the following lemma.

Lemma 4. *Assume that the parametric estimators $\hat{\psi}$ and $\hat{\gamma}$ and the local constant esti-*

mator $\hat{\theta}$ are obtained from the three-step procedure of Section 2.1 and satisfy $\|\hat{\psi}-\psi_0\| = O_p(1/\sqrt{T})$, $\|\hat{\gamma}-\gamma_0\| = O_p(1/\sqrt{T})$ and $\|\hat{\theta}-\theta_0\| = O_p(1/\sqrt{Th})$. Define the local log likelihood function as

$$L_h(\psi,\gamma,\theta) = \frac{1}{T}\sum_{t=1}^{T} \ell_{ct}(\psi,\gamma,\theta)k_h(\gamma^\intercal Z_t - \Lambda)$$

where $\ell_{ct}(\psi,\gamma,\theta) = \log c(F_1(X_{1t}|Z_t;\psi_1), F_2(X_{2t}|Z_t;\psi_2)|Z_t;\theta(\gamma^\intercal Z_t))$. Under conditions A1-A7, we have

$$L_h(\hat{\psi},\hat{\gamma},\hat{\theta}) - L_h(\psi,\gamma,\theta) = L_h(\psi,\gamma,\hat{\theta}) - L_h(\psi,\gamma,\theta) + o_p(T^{-1}).$$

Proof: Let

$$\begin{aligned} & L_h(\hat{\psi},\hat{\gamma},\hat{\theta}) - L_h(\psi,\gamma,\theta) \\ = \quad & \underbrace{L_h(\hat{\psi},\hat{\gamma},\hat{\theta}) - L_h(\hat{\psi},\gamma,\hat{\theta})}_{I_1} + \underbrace{L_h(\hat{\psi},\gamma,\hat{\theta}) - L_h(\psi,\gamma,\hat{\theta})}_{I_2} + \underbrace{L_h(\psi,\gamma,\hat{\theta}) - L_h(\psi,\gamma,\theta)}_{I_3}. \end{aligned}$$

By Taylor expansion and the conditions $\|\hat{\psi}-\psi\| = O_p(1/\sqrt{T})$, $\|\hat{\gamma}-\gamma\| = O_p(1/\sqrt{T})$ and $\|\hat{\theta}-\theta\| = O_p(1/\sqrt{Th})$, the first term I_1 is given by

$$\begin{aligned} I_1 = \quad & L_h(\hat{\psi},\hat{\gamma},\hat{\theta}(\hat{\psi},\hat{\gamma})) - L_h(\hat{\psi},\gamma,\hat{\theta}(\hat{\psi},\gamma)) \\ = \quad & \left[\sqrt{T}\frac{\partial L_h(\hat{\psi},\gamma,\hat{\theta}(\hat{\psi},\gamma))}{\partial\gamma}\right]\frac{1}{\sqrt{T}}(\hat{\gamma}-\gamma)\{1+o_p(1)\} \\ = \quad & \left[\sqrt{T}\frac{\partial L_h(\psi,\gamma,\hat{\theta}(\psi,\gamma))}{\partial\gamma}\{1+o_p(1)\}\right]\frac{1}{\sqrt{T}}(\hat{\gamma}-\gamma)\{1+o_p(1)\} \\ = \quad & \left[\sqrt{T}\frac{\partial L_h(\psi,\gamma,\theta(\psi,\gamma))}{\partial\gamma}\{1+o_p(1)\}\{1+o_p(1)\}\right]\frac{1}{\sqrt{T}}(\hat{\gamma}-\gamma)\{1+o_p(1)\} \end{aligned}$$

which is of order $O_p(1/T)$. In the same vein, we can show that the second term,

$$I_2 = \quad L_h(\hat{\psi},\gamma,\hat{\theta}(\hat{\psi},\gamma)) - L_h(\psi,\gamma,\hat{\theta}(\psi,\gamma))$$

$$
\begin{aligned}
&= \left[\sqrt{T}\frac{\partial L_h(\psi,\gamma,\hat{\theta}(\psi,\gamma))}{\partial\psi}\right]\frac{1}{\sqrt{T}}(\hat{\psi}-\psi)\{1+o_p(1)\} \\
&= \left[\sqrt{T}\frac{\partial L_h(\psi,\gamma,\theta(\psi,\gamma))}{\partial\psi}\{1+o_p(1)\}\right]\frac{1}{\sqrt{T}}(\hat{\psi}-\psi)\{1+o_p(1)\}
\end{aligned}
$$

is of order $O_p(1/T)$. The term on the right hand side $\sqrt{T}\partial L_h(\psi,\gamma,\theta(\psi,\gamma))/\partial\psi$ is of order $O_p(1)$ since the first derivative of the marginal likelihood $\sqrt{T}\partial L_m(\psi)/\partial\psi$ and the first derivative of the full likelihood $\sqrt{T}\partial L_m(\psi)/\partial\psi + \sqrt{T}\partial L_h(\psi,\gamma,\theta(\psi,\gamma))/\partial\psi$ are of order $O_p(1)$. This implies that $L_h(\hat{\psi},\gamma,\hat{\theta}) - L_h(\psi,\gamma,\hat{\theta})$ is of order $O_p(1/T)$.

Furthermore, by a Taylor expansion and the condition $\|\hat{\theta}-\theta\| = O_p(1/\sqrt{Th})$, the last term

$$
I_3 = \left[\sqrt{Th}\frac{\partial L_h(\psi,\gamma,\theta)}{\partial\theta}\right]\frac{1}{\sqrt{Th}}(\hat{\gamma}-\gamma)\{1+o_p(1)\}
$$

which is of order $O_p(1/(Th))$, dominates the other two terms. This completes the proof. □

Lemma 4 suggests that we can derive the asymptotic distribution of $\hat{\theta}$ without considering the errors from parametric estimation. The estimators of $(\hat{\psi},\hat{\gamma})$ have little effect on the estimation of $\hat{\theta}$ if the sample size T is large. This result is in line with the fact that the convergence rate of the parametric part of the model is faster than that of the nonparametric component.

Proof of Theorem 2: Using Lemma 4 we can assume that ψ is known for simplicity. Define the kernel constants $\mu_2 = \int z^2 k(z)dz$, $\nu_0 = \int k^2(z)dz$ and $\nu_2 = \int z^2 k^2(z)dz$. Let $\Lambda_t = \gamma^\intercal Z_t$, $\ell_{ct}(\theta(\Lambda)) = \ell_{ct}(\psi,\gamma,\theta)$, $L(\theta(\Lambda)) = \frac{1}{T}\sum_{t=1}^{T}\ell_{ct}(\theta(\Lambda))k_h(\Lambda_t-\Lambda)$, $L'(\theta(\Lambda)) = \frac{1}{T}\sum_{t=1}^{T}\ell'_{ct}(\theta(\Lambda))k_h(\Lambda_t-\Lambda)$ and $L''(\theta(\Lambda)) = \frac{1}{T}\sum_{t=1}^{T}\ell''_{ct}(\theta(\Lambda))k_h(\Lambda_t-\Lambda)$. For a fixed point Λ lying in the interior of the support A_Λ, the normal equation for the local likelihood-based estimator is given by $L'(\hat{\theta}(\Lambda)) = 0$. By a Taylor expansion, it can be written as

$$
L'(\theta(\Lambda)) + L''(\theta(\Lambda))(\hat{\theta}(\Lambda)-\theta(\Lambda)) + o_p(1/\sqrt{Th}) = 0
$$

which leads to

$$\hat{\theta}(\Lambda) - \theta(\Lambda) = -[L''(\theta(\Lambda))]^{-1} L'(\theta(\Lambda)) + o_p(1/\sqrt{Th}).$$

By the moment condition, we have

$$\begin{aligned} 0 &= E\{\ell'_{ct}(\theta(\Lambda_t))|\Lambda_t = \Lambda\} \\ &= E\{\ell'_{ct}(\theta(\Lambda) + r_t)|\Lambda_t = \Lambda\} \\ &\approx E\{\ell'_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} + r_t E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} \end{aligned}$$

where $r_t = \theta'(\Lambda)(\Lambda_t - \Lambda) + \frac{1}{2}\theta''(\Lambda)(\Lambda_t - \Lambda)^2 + o_p(\Lambda_t - \Lambda)^2$. By construction, we have $E\{\ell'_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} \approx -r_t E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\}$. Thus,

$$\begin{aligned} E\{L'(\theta(\Lambda))|\Lambda_t = \Lambda\} &= -\frac{1}{T}\sum_{t=1}^{T} r_t E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} k_h(\Lambda_t - \Lambda) \\ &= \frac{1}{T}\Sigma(\Lambda)\sum_{t=1}^{T} r_t k_h(\Lambda_t - \Lambda) \end{aligned}$$

where $\Sigma(\Lambda) = -E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\}$. Note that

$$\begin{aligned} E\{L''(\theta(\Lambda))|\Lambda_t = \Lambda\} &= \frac{1}{T}\sum_{t=1}^{T} E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} k_h(\Lambda_t - \Lambda) \\ &\approx f(\Lambda) E\{\ell''_{ct}(\theta(\Lambda))|\Lambda_t = \Lambda\} \\ &= -f(\Lambda)\Sigma(\Lambda). \end{aligned}$$

To find the expression for $\mathrm{Var}\{L'(\theta(\Lambda))|\Lambda_t = \Lambda\}$, using the same argument as in Cai (2007) , we can show that

$$\mathrm{Var}\{L'(\theta(\Lambda))|\Lambda_t = \Lambda\} = \frac{\nu_0}{Th}\{R_0(\Lambda) + 2\sum_{j=1}^{\infty} R_j(\Lambda)\} + o_p(\frac{1}{Th})$$

where $R_j(\Lambda) = E\{\ell'_{ct}(\psi, \gamma, \theta(\Lambda))\ell'_{c(t+j)}(\psi, \gamma, \theta(\Lambda))^\intercal | \gamma^\intercal Z = \Lambda\}$.

It follows by a Taylor expansion and the Riemann sum approximation of an integral that the bias term of $\hat{\theta}(\Lambda)$ can be expressed as

$$\begin{aligned}
&\mathrm{E}\{\hat{\theta}(\Lambda)|\Lambda_t = \Lambda\} - \theta(\Lambda) \\
&= -[\mathrm{E}\{L''(\theta(\Lambda))|\Lambda_t = \Lambda\}]^{-1}\mathrm{E}\{L'(\theta(\Lambda))|\Lambda_t = \Lambda\} \\
&= \frac{1}{f(\Lambda)}\frac{1}{T}\sum_{t=1}^{T}\left[\theta'(\Lambda)(\Lambda_t - \Lambda) + \frac{1}{2}\theta''(\Lambda)(\Lambda_t - \Lambda)^2\right] k_h(\Lambda_t - \Lambda) \\
&\approx \frac{1}{f(\Lambda)}\int \theta'(\Lambda)(\Lambda_t - \Lambda)f(\Lambda_t)k_h(\Lambda_t - \Lambda)d\Lambda_t + \frac{1}{2f(\Lambda)}\int \theta''(\Lambda)(\Lambda_t - \Lambda)^2 f(\Lambda_t)k_h(\Lambda_t - \Lambda)d\Lambda_t \\
&= \frac{h}{f(\Lambda)}\int \theta'(\Lambda)uf(\Lambda + uh)k(u)du + \frac{h^2}{2f(\Lambda)}\int \theta''(\Lambda)u^2 f(\Lambda + uh)k(u)du \\
&= \frac{h^2}{f(\Lambda)}\theta'(\Lambda)f'(\Lambda)\mu_2 + \frac{h^2}{2f(\Lambda)}\theta''(\Lambda)\mu_2 + o_p(h^2) \\
&= h^2 B(\Lambda)
\end{aligned}$$

where $B(\Lambda) = \frac{1}{f(\Lambda)}\theta'(\Lambda)f'(\Lambda)\mu_2 + \frac{1}{2}\theta''(\Lambda)\mu_2$.

The variance term is given by

$$\begin{aligned}
&\mathrm{Var}\{\hat{\theta}(\Lambda)|\Lambda_t = \Lambda\} \\
&= \mathrm{E}\{L''(\theta(\Lambda))|\Lambda_t = \Lambda\}^{-1}\mathrm{Var}\{L'(\theta(\Lambda))|\Lambda_t = \Lambda\}\mathrm{E}\{L''(\theta(\Lambda))|\Lambda_t = \Lambda\}^{-1} \\
&= \frac{1}{Thf(\Lambda)}\nu_0\Sigma(\Lambda)^{-1}\Omega^*(\Lambda)\Sigma(\Lambda)^{-1}.
\end{aligned}$$

□

Lemma 5. *Assume that the sequence of random variables $b_t(X_{1t}, X_{2t}, Z_t)$, denoted by b_t, is bounded, the sequence $\omega_t(X_{1t}, X_{2t}, Z_t)$, denoted by ω_t, is stationary with mean zero and finite variance, and the sequence $e_t(X_{1t}, X_{2t}, Z_t)$, denoted by e_t, satisfies $\|e_t\|_\infty = o_p(T^{-1/4})$. Then,*

$$\sum_{t=1}^{T} b_t\omega_t e_t = o_p(\sqrt{T}).$$

The proof of Lemma 5 is similar to that of Lemma A.1 in Liang and Li (2009), and therefore omitted here.

$\square$

Lemma 6. *For any given constant C,*

$$\sup_{\|\xi-\xi_0\|\leq CT^{-1/2}} \left\|T^{-1/2}\Pi(\theta,m_z,\xi)-T^{-1/2}\Pi(\theta_0,m_{0z},\xi_0)+T^{1/2}M(\xi-\xi_0)\right\|=o_p(1)$$

where $M=-\frac{1}{T}\partial\Pi(\theta_0,m_{0z},\xi_0)/\partial\xi$. Moreover, $T^{-1/2}\Pi(\theta_0,m_{0z},\xi_0)\xrightarrow{d}N(0,\Omega)$ with $\Omega=\Gamma(0)+2\sum_{k=1}^{\infty}\Gamma(k)$, $\Gamma(k)=\mathrm{Cov}\,(\zeta_t,\zeta_{t-k})$ and $\zeta_t=(\ell'_{1t}(\psi_{10})^\intercal,\ell'_{2t}(\psi_{20})^\intercal,\pi_t(\theta_0,m_{0z},\xi_0)^\intercal)^\intercal$.

Proof: We have

$$\begin{aligned}
& T^{-1/2}\Pi(\theta,m_z,\xi)-T^{-1/2}\Pi(\theta_0,m_{0z},\xi_0)\\
= \quad & T^{-1/2}\Pi(\theta,m_z,\xi)-T^{-1/2}\Pi(\theta,m_{0z},\xi)+T^{-1/2}\Pi(\theta,m_{0z},\xi)-T^{-1/2}\Pi(\theta_0,m_{0z},\xi)\\
& +T^{-1/2}\Pi(\theta_0,m_{0z},\xi)-T^{-1/2}\Pi(\theta_0,m_{0z},\xi_0)\\
\doteq \quad & A_1+A_2+A_3
\end{aligned}$$

where

$$\begin{aligned}
A_1 = \quad & T^{-1/2}\Pi(\theta,m_z,\xi)-T^{-1/2}\Pi(\theta,m_{0z},\xi)\\
= \quad & \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix}\ell'_{1t}\\ \ell'_{2t}\\ \ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t)(Z_t-m_z(\Lambda_t))\end{pmatrix}-\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix}\ell'_{1t}\\ \ell'_{2t}\\ \ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t)(Z_t-m_{0z}(\Lambda_t))\end{pmatrix}\\
= \quad & \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix}0\\ 0\\ \ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t)(m_{0z}(\Lambda_t)-m_z(\Lambda_t))\end{pmatrix}.
\end{aligned}$$

By Lemma 5, A_1 is of order $o_p(1)$. Next,

$$A_2 = \quad T^{-1/2}\Pi(\theta,m_{0z},\xi)-T^{-1/2}\Pi(\theta_0,m_{0z},\xi)$$

$$
= \quad \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} 0 \\ 0 \\ \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta_0)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) \end{pmatrix}
$$

and

$$
\begin{aligned}
& T^{-1/2} \sum_{t=1}^{T} \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta_0)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) \\
= \quad & T^{-1/2} \sum_{t=1}^{T} \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) \\
& + T^{-1/2} \sum_{t=1}^{T} \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'_0(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta_0)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) \\
\doteq \quad & A_{21} + A_{22}
\end{aligned}
$$

where $A_{21} = T^{-1/2} \sum_{t=1}^{T} \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) = o_p(1)$ by Lemma 5, and

$$
\begin{aligned}
A_{22} = \quad & T^{-1/2} \sum_{t=1}^{T} \left[\ell'_{ct}(\psi,\gamma,\theta)\theta'_0(\Lambda_t) - \ell'_{ct}(\psi,\gamma,\theta_0)\theta'_0(\Lambda_t)\right](Z_t - m_{0z}(\Lambda_t)) \\
= \quad & T^{-1/2} \sum_{t=1}^{T} \left[\{\ell''_{ct}(\psi,\gamma,\theta_0)(\theta - \theta_0)\}\{1 + o_p(1)\}\right]\theta'_0(\Lambda_t)(Z_t - m_{0z}(\Lambda_t)).
\end{aligned}
$$

is again $o_p(1)$ by Lemma 5. This implies that A_2 is $o_p(1)$. Finally,

$$
\begin{aligned}
A_3 = \quad & T^{-1/2}\Pi(\theta_0, m_{0z}, \xi) - T^{-1/2}\Pi(\theta_0, m_{0z}, \xi_0) \\
= \quad & T^{-1/2}\{(\partial\Pi(\theta_0, m_{0z}, \xi_0)/\partial\xi)(\xi - \xi_0)\}\{1 + o_p(1)\} \\
= \quad & \sqrt{T}(-M)(\xi - \xi_0) + o_p(1).
\end{aligned}
$$

Therefore, $T^{-1/2}\Pi(\theta, m_z, \xi) - T^{-1/2}\Pi(\theta_0, m_{0z}, \xi_0) + T^{1/2}M(\xi - \xi_0) = A_1 + A_2 + A_3 + T^{1/2}M(\xi - \xi_0) = o_p(1)$, which implies the stated result.

□

Proof of Theorem 3:

To show that there exists a $\sqrt{T}$-consistent estimator $\hat{\xi}$ satisfying both $||\hat{\xi}-\xi_0|| = O_p(1/\sqrt{T})$ and $\Pi^P(\hat{\theta}, \hat{m}_z, \hat{\xi}) = 0$, it is sufficient to show that $(\hat{\xi}-\xi_0)^T M^T \Pi^P(\hat{\theta}, \hat{m}_z, \hat{\xi}) < 0$. By Lemma 6, we have

$$
\begin{aligned}
& \sqrt{T}(\hat{\xi}-\xi_0)^T M^T \frac{1}{\sqrt{T}} \Pi^P(\hat{\theta}, \hat{m}_z, \hat{\xi}) \\
= \quad & \sqrt{T}(\hat{\xi}-\xi_0)^T M^T \left[\frac{1}{\sqrt{T}} \Pi(\theta_0, m_{0z}, \xi_0) - \sqrt{T} M(\hat{\xi}-\xi_0) - \sqrt{T}\mathbf{p}'_\lambda(|\hat{\xi}|)\mathrm{sgn}(\hat{\xi}) + o_p(1) \right] \\
\leq \quad & \sqrt{T}(\hat{\xi}-\xi_0)^T M^T \frac{1}{\sqrt{T}} \Pi(\theta_0, m_{0z}, \xi_0) - \sqrt{T}(\hat{\xi}-\xi_0)^T M^T M \sqrt{T}(\hat{\xi}-\xi_0) \\
& -\sqrt{T}(\hat{\xi}-\xi_0)^T M_2^T \sqrt{T}\mathbf{p}'_\lambda(|\hat{\gamma}_1|)\mathrm{sgn}(\hat{\gamma}_1) \\
\leq \quad & \sqrt{T}(\hat{\xi}-\xi_0)^T M^T \frac{1}{\sqrt{T}} \Pi(\theta_0, m_{0z}, \xi_0) - \sqrt{T}(\hat{\xi}-\xi_0)^T M^T M \sqrt{T}(\hat{\xi}-\xi_0) \\
& -\sqrt{T}(\hat{\xi}-\xi_0)^T M_2^T \{\sqrt{T}\mathbf{p}'_\lambda(|\gamma_{10}|)\mathrm{sgn}(\gamma_{10}) + \sqrt{T}\mathbf{p}''_\lambda(|\gamma_{10}|)(\hat{\gamma}_1-\gamma_{10})\}\{1+o_p(1)\} \quad (12)
\end{aligned}
$$

where M_2 is a submatrix of the partition $M^T = (M_1, M_2, M_3)$ with M_1, M_2 and M_3 being $(q+d) \times q$,$(q+d) \times d_1$ and $(q+d) \times (d-d_1)$ matrices, respectively.

The first term on the right hand side of the last inequality in (12) is of order $C * O_p(1)$ and the second term is of order $C^2 * O_p(1)$. Using Condition (B1), $\sqrt{T}\mathbf{p}'_\lambda(|\gamma_{10}|) \to 0$ and $\mathbf{p}''_\lambda(|\gamma_{10}|) \to 0$. By choosing the constant C sufficiently large, the second term will dominate the other two terms. This completes the proof.

□

Proof of Theorem 4: First, we show the sparsity with $\hat{\gamma}_2 = 0$. Let us assume that there exists a $\sqrt{T}$-consistent estimator $\hat{\xi}^* = (\hat{\psi}^\intercal, \hat{\gamma}_1^\intercal, \hat{\gamma}_2^\intercal)^\intercal$ with $\hat{\gamma}_2 \neq 0$ such that $\Pi^P(\hat{\theta}, \hat{m}_z, \hat{\xi}^*) = 0$. By Lemma 6, we have

$$
\frac{1}{\sqrt{T}} \Pi(\theta_0, m_{0z}, \xi_0) - \sqrt{T} M(\hat{\xi}^* - \xi_0) + o_p(1) = \sqrt{T}\mathbf{p}'(\hat{\xi}^*)\mathrm{sgn}(\hat{\xi}^*). \quad (13)
$$

The first two components on the left hand side of (13) are of order $O_p(1)$. However, the last $d - d_1$ elements of $\sqrt{T}\mathbf{p}'(\hat{\xi}^*)$ on the right hand side diverge to ∞ by the conditions in (B1). Therefore, by contradiction, we conclude that $\hat{\gamma}_2 = 0$ must hold.

Second, we show asymptotic normality. By Lemma 6, we have

$$\frac{1}{\sqrt{T}}\Pi(\theta_0, m_{0z_1}, \xi_{10}) - \sqrt{T}M(\hat{\xi}_1 - \xi_{10}) + o_p(1) = \sqrt{T}\mathbf{p}'(\hat{\xi}_1)\text{sgn}(\hat{\xi}_1)$$

where $\hat{\xi}_1 = (\hat{\psi}^\intercal, \hat{\gamma}_1^\intercal)^\intercal$. The term $\sqrt{T}\mathbf{p}'(\hat{\xi}_1) = 0$ as $T \to \infty$ according to the conditions in (B1). It follows that

$$\frac{1}{\sqrt{T}}\Pi(\theta_0, m_{0z_1}, \xi_{10}) - \sqrt{T}M(\hat{\xi}_1 - \xi_{10}) + o_p(1) = 0.$$

Asymptotic normality is obtained by the central limit theorrem, which proves the theorem.

$\square$

References

[1] Acar, E.F., Craiu, R.V. and Yao, F. 2011. Dependence calibration in conditional copulas: A nonparametric approach. *Biometrics,* **67**, 445-453.

[2] Andrews, D.W.K. 1987. Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica*, **55**, 1465-1471.

[3] Bartram, S.M., Taylor, S.J. and Wang, Y.H. 2007. The Euro and European financial market dependence. *Journal of Banking and Finance,* **31**, 1461-1481.

[4] Cai, Z. 2002. Regression quantiles for time series. *Econometric Theory,* **18**, 169-192.

[5] Cai, Z. 2007. Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics,* **136**, 163-188.

[6] Cai, Z., Juhl, T. and Yang, B. 2015. Functional index coefficient models with variable selection. *Journal of Econometrics,* **189**, 272-284.

[7] Cai, Z., Ren, Y. and Yang, B. 2015. A semiparametric conditional capital asset pricing model. *Journal of Banking and Finance,* **61**, 117-126.

[8] Chen, X. and Fan, Y. 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, **135**, 125-154.

[9] Chollete, L., Heinen, A. and Valdesogo, A. 2009. Modeling international financial returns with a multivariate regime switching copula. *Journal of Financial Econometrics,* **7**, 437-480.

[10] Cui, X., Härdle, W.K. and Zhu, L. 2011. The EFM approach for single-index models. *The Annals of Statistics,* **39**, 1658-1688.

[11] Del Negro, M. and Otrok, C. 2007. Luftballons: Monetary policy and the house price boom cross U.S. states. *Journal of Monetary Economics,* **54**, 1962-1985.

[12] Dobric, J. and Schmid, F. 2007. A goodness of fit test for copulas based on rosenblatt's transformation. *Computational Statistics and Data Analysis,* **51**, 4633-4642.

[13] Fan, J. and Li, R. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association,* **96**, 1348-1360.

[14] Fermanian, J.D. and Lopez, O. 2018. Single-index copulas. *Journal of Multivariate Analysis,* **165**, 27-55.

[15] Garcia, R. and Tsafack, G. 2011. Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking and Finance,* **35**, 1954-1970.

[16] Giacomini, E., Hardle, W. and Spokoiny, V. 2009. Inhomogeneous dependency modeling with time varying copulae. *Journal of Business and Economic Statistics,* **27**, 224-234.

[17] Guo, H., Wu, C. and Yu, Y. 2017. Time-varying beta and the value premium. *Journal of Financial and Quantitative Analysis,* **52**, 1551-1576.

[18] Hafner, C.M. and Manner, H. 2012. Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics,* **27**, 269-295.

[19] Hafner, C.M. and Reznikova, O. 2010. Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics and Data Analysis,* **54**, 2609-2627.

[20] Ichimura, H. 1993. Semiparametric least squares (SLS) and weighted SLS estimation of single index model. *Journal of Econometrics,* **58**, 71-120.

[21] Janson, S. 1987. Maximal spacing in several dimensions. *Annals of Probability,* **15**, 274-280.

[22] Joe, H. 2000. Multivariate models and dependence concepts, Series: Monographs on Statistics and Applied Probability Series, Publisher: Taylor and Francis, Inc.

[23] Johnson, B.A., Lin, D.Y. and Zeng, D. 2008. Penalized estimating functions and variable selection in semiparametric regression model. *Journal of the American Statistical Association,* **103**, 672-680.

[24] Kallberg, J, Liu, C. and Pasquariello, P. 2014. On the price comovement of U.S. residential real estate markets. *Real Estate Economics*, **42**, 71-108.

[25] Liang, H. and Li, R. 2009. Variable selection for partially linear models with measurement errors. *Journal of the American Statistical Association,* **104**, 234-248.

[26] Liang, H., Liu, X., Li, R. and Tsay, C. 2010. Estimation and testing for partially linear single-index models. *The Annals of Statistics,* **38**, 3811-3836.

[27] Mack, Y.P. and Silverman, B.W. 1982. Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrscheinlichkeitstheorie View. Geb., 61*, 405-415.

[28] Nelson, R. 2006. An Introduction to Copulas, 2nd ed. (New York: Springer, 2006).

[29] Newey, W.K. 1994. The asymptotic variance of semiparametric estimators. *Econometrica,* **62**, 1349-1382.

[30] Patton, A.J. 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics,* **2**, 130-168.

[31] Patton, A.J. 2006. Modeling asymmetric exchange rate dependence. *International Economic Review,* **47**, 527-556.

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

[33] Shiller, R. 2007. Understanding recent trends in house prices and home ownership. *NBER Working Paper No. 13553*.

[34] Sklar, A. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris,* **8**, 229-231.

[35] Tibshirani, R.J. 1996. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Ser. B,* **58**, 267-288.

[36] Wang, H. and Xia, Y. 2009. Shrinkage estimation of the varying coefficient model. *Journal of the American Statistical Association,* **104**, 747-757.

[37] Zhu, L., Dong, Y. and Li, R. 2013. Semiparametric estimaion of conditional heteroscedasticity via single-index modeling. *Statistica Sinica,* **23**, 1235-1255.

[38] Zimmer, D. 2012. The role of copulas in the housing crisis. *Review of Economics and Statistics*, **94**, 1607-620.

[39] Zou, H. 2006. The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, **101**, 1418-1429.

Table 1: Percentages that the corresponding state variables were chosen correctly (incorrectly) for all 6 cases in simulations

Model	T	Z_1	Z_2	$Z_3 - Z_5$
Model 1 (Clayton)	500	0.996	0.998	(0.179)
	1000	1.000	1.000	(0.071)
Model 2 (Clayton)	500	0.954	1.000	(0.100)
	1000	0.990	1.000	(0.069)
Model 1 (Gumbel)	500	1.000	1.000	(0.104)
	1000	1.000	1.000	(0.049)
Model 2 (Gumbel)	500	0.990	1.000	(0.091)
	1000	1.000	1.000	(0.051)
Model 1 (Frank)	500	0.993	1.000	(0.192)
	1000	1.000	1.000	(0.093)
Model 2 (Frank)	500	0.930	1.000	(0.189)
	1000	0.987	1.000	(0.099)

Note: Values without parentheses are the percentages that state variables in the true index copula models were chosen correctly. Values with parentheses are the percentages that state variables not in the true index copula models were chosen incorrectly.

Table 2: Means, medians, and mean square errors (MSEs) of the estimated (nonzero) index coefficients (γ_1 and γ_2), and MSEs of the estimated copula dependence parameters (θ) for all 6 cases in simulations

			Clayton			Gumbel			Frank		
Model	T		γ_1	γ_2	θ	γ_1	γ_2	θ	γ_1	γ_2	θ
Model 1	500	Mean	0.671	0.679	—	0.692	0.686	—	0.657	0.692	—
		Median	0.708	0.705	—	0.707	0.707	—	0.704	0.708	—
		MSE	0.025	0.024	0.550	0.015	0.015	0.361	0.029	0.019	0.804
	1000	Mean	0.692	0.696	—	0.701	0.694	—	0.684	0.700	—
		Median	0.707	0.707	—	0.706	0.708	—	0.707	0.707	—
		MSE	0.011	0.011	0.382	0.007	0.008	0.247	0.013	0.011	0.567
Model 2	500	Mean	0.498	0.811	—	0.508	0.827	—	0.477	0.824	—
		Median	0.500	0.865	—	0.500	0.865	—	0.500	0.866	—
		MSE	0.033	0.026	0.517	0.022	0.019	0.327	0.028	0.037	0.730
	1000	Mean	0.502	0.836	—	0.495	0.854	—	0.497	0.840	—
		Median	0.500	0.866	—	0.500	0.866	—	0.501	0.865	—
		MSE	0.016	0.008	0.358	0.011	0.005	0.233	0.019	0.010	0.525

Table 3: The summary statistics of the quarterly percentage changes of housing prices in Arizona, California, Florida and Nevada and the eight state variables. CPI = Quarterly consumer price index. GDP = Quarterly growth rate of GDP. INC = Quarterly growth rate of per capita disposable income. INT = The effective federal funds rate. INV = Quarterly growth rate of gross private domestic investment in GDP. IPI = Quarterly growth rate of industrial production index. OIL = Quarterly growth rate of oil price (West Texas intermediate). UNE = Quarterly growth rate of unemployment rate. ADF indicates the p-value of the Augmented Dickey-Fuller test for null of non-stationary series. The sample spans from 1975:Q1 - 2018:Q1.

		Mean	Median	Std. Dev	Kurtosis	Skewness	Min.	Max.	ADF
% Change of Housing Price	AZ	1.155	1.216	2.894	1.379	-0.122	-7.256	10.368	0.013
	CA	1.634	1.893	2.629	1.771	-0.398	-8.182	10.180	0.052
	FL	1.179	1.245	3.770	16.903	0.630	-19.790	26.705	0.039
	NV	1.223	1.130	3.716	7.425	0.881	-10.328	22.997	0.019
State Variables	CPI	0.912	0.785	0.773	3.552	0.775	-2.290	3.946	0.086
	GDP	0.451	0.498	0.751	3.386	-0.362	-2.342	3.624	0.010
	INC	2.890	2.900	3.655	5.887	-0.391	-15.700	19.900	0.010
	INT	1.125	0.186	19.710	10.039	1.280	-73.883	125.000	0.010
	INV	0.104	0.000	4.220	2.573	-0.172	-16.000	17.647	0.018
	IPI	0.551	0.669	1.366	3.377	-1.093	-5.622	4.026	0.010
	OIL	1.967	1.339	13.599	2.446	-0.088	-50.527	48.070	0.010
	UNE	0.007	-1.382	9.520	0.006	0.680	-15.789	34.010	0.010

Table 4: Results of AR(p)-GARCH(1,1) filtering. LB represents the Ljung-Box test statistic for examining the null hypothesis of independence in the series filtered by AR(p)-GARCH(1,1).

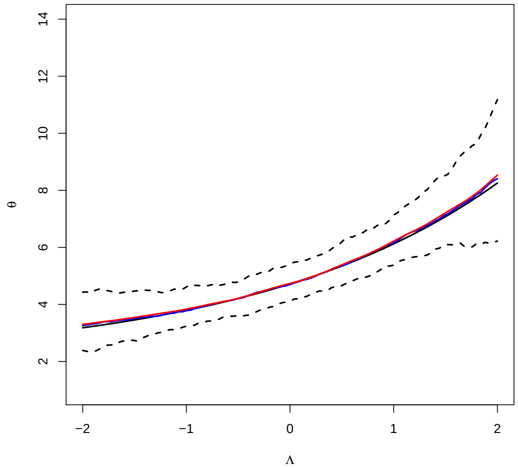
	AR(p) part (up to order 2)	GARCH(1,1) part			
State	γ	ω	α_1	β_1	LB
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
AZ	0.757	0.083	0.530	0.558	95.187
	(0.000)	(0.014)	(0.001)	(0.000)	(0.617)
CA	0.933	0.322	0.438	0.407	79.403
	(0.000)	(0.015)	(0.000)	(0.001)	(0.936)
FL	0.404, 0.507	0.120	0.376	0.623	83.756
	(0.000), (0.000)	(0.029)	(0.000)	(0.000)	(0.879)
NV	0.832	0.079	0.448	0.639	111.170
	(0.000)	(0.276)	(0.001)	(0.000)	(0.209)

Table 5: This table reports the cross-validated prediction errors (CVPEs) of Gaussian, Clayton, Gumbel, Frank, Rotated Clayton and Rotated Gumbel copulas for all six pairs. The best two models are in bold.

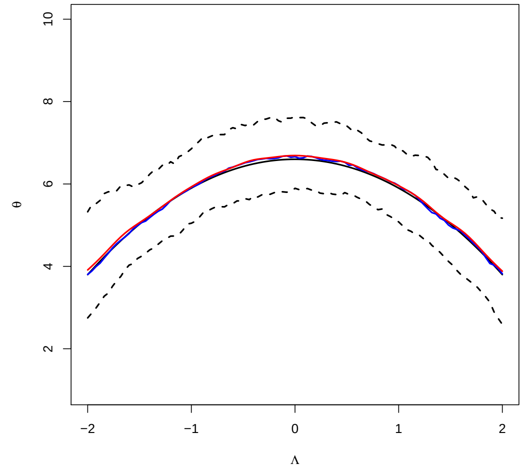
	Gaussian	Clayton	Gumbel	Frank	Rotated Clayton	Rotated Gumbel
AZ - CA	11.087	13.328	12.020	**10.704**	11.749	12.604
AZ - FL	10.734	11.416	**10.485**	11.931	11.948	10.574
AZ - NV	10.752	11.394	**8.309**	11.919	11.987	11.274
CA - FL	13.600	13.469	12.493	**12.039**	14.433	14.508
CA - NV	11.918	14.548	10.951	**9.041**	15.247	14.417
FL - NV	5.963	10.017	9.892	**5.872**	7.223	8.959

Table 6: This table reports the estimated γs associated with the eight state variables for each of the six pairs. Values in parentheses indicate standard errors. CPI = Quarterly consumer price index. GDP = Quarterly growth rate of GDP. INC = Quarterly growth rate of per capita disposable income. INT = The effective federal funds rate. INV = Quarterly growth rate of gross private domestic investment in GDP. IPI = Quarterly growth rate of industrial production index. OIL = Quarterly growth rate of oil price (West Texas intermediate). UNE = Quarterly growth rate of unemployment rate. All state variables span between 1975:Q1 and 2018:Q1.

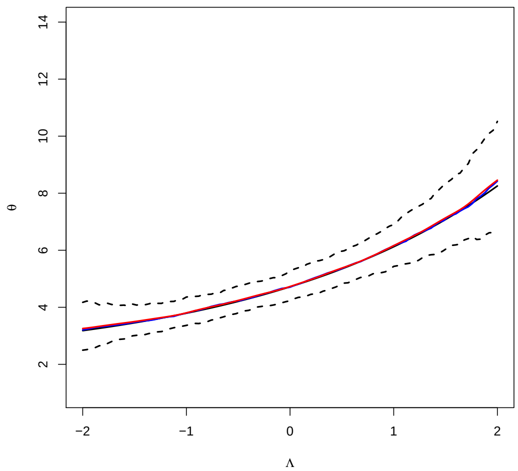
	CPI (S.E.)	GDP (S.E.)	INC (S.E.)	INT (S.E.)	INV (S.E.)	IPI (S.E.)	OIL (S.E.)	UNE (S.E.)
AZ - CA	0.482	0.441	0.361	0.201	0.522	0	0	0.361
	(0.200)	(0.234)	(0.111)	(0.130)	(0.212)	—	—	(0.088)
AZ - FL	0.485	0.253	0.414	0.220	0.590	0	0	0.364
	(0.155)	(0.118)	(0.103)	(0.170)	(0.062)	—	—	(0.095)
AZ - NV	0.496	0.438	0.355	0.196	0.522	0	0	0.354
	(0.071)	(0.022)	(0.042)	(0.016)	(0.011)	—	—	(0.052)
CA - FL	0.482	0.441	0.362	0.202	0.521	0	0	0.361
	(0.243)	(0.032)	(0.013)	(0.013)	(0.010)	—	—	(0.029)
CA - NV	0.373	0.225	0.485	0.325	0.600	0	0	0.331
	(0.200)	(0.112)	(0.243)	(0.100)	(0.088)	—	—	(0.080)
FL - NV	0.449	0.407	0.432	0.643	0.142	0	0	0.117
	(0.286)	(0.184)	(0.209)	(0.170)	(0.071)	—	—	(0.038)



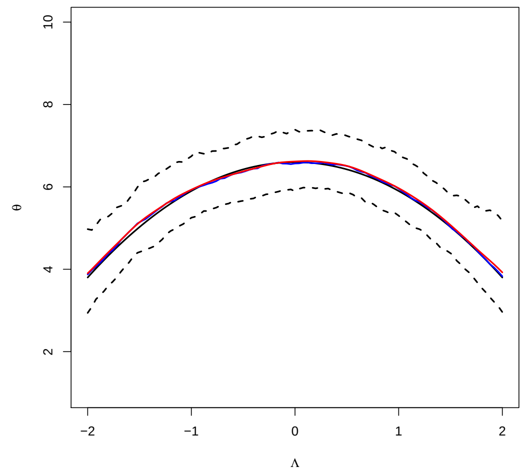
(a) Model 1 (Clayton)



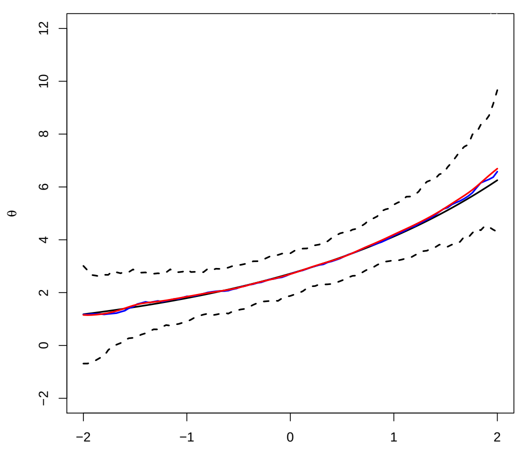
(b) Model 2 (Clayton)



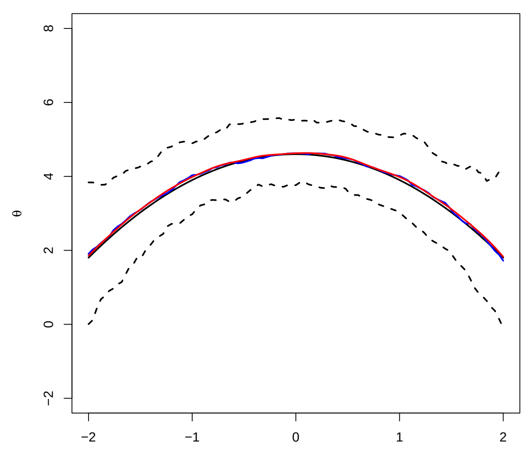
(c) Model 1 (Gumbel)



(d) Model 2 (Gumbel)



(e) Model 1 (Frank)



(f) Model 2 (Frank)

Figure 1: Simulation results of the copula dependence parameters for all 6 cases: true values (black solid lines), mean and median estimates (red and blue lines), and 5% and 95% percentile curves (black dashed lines). The sample size is 1000.

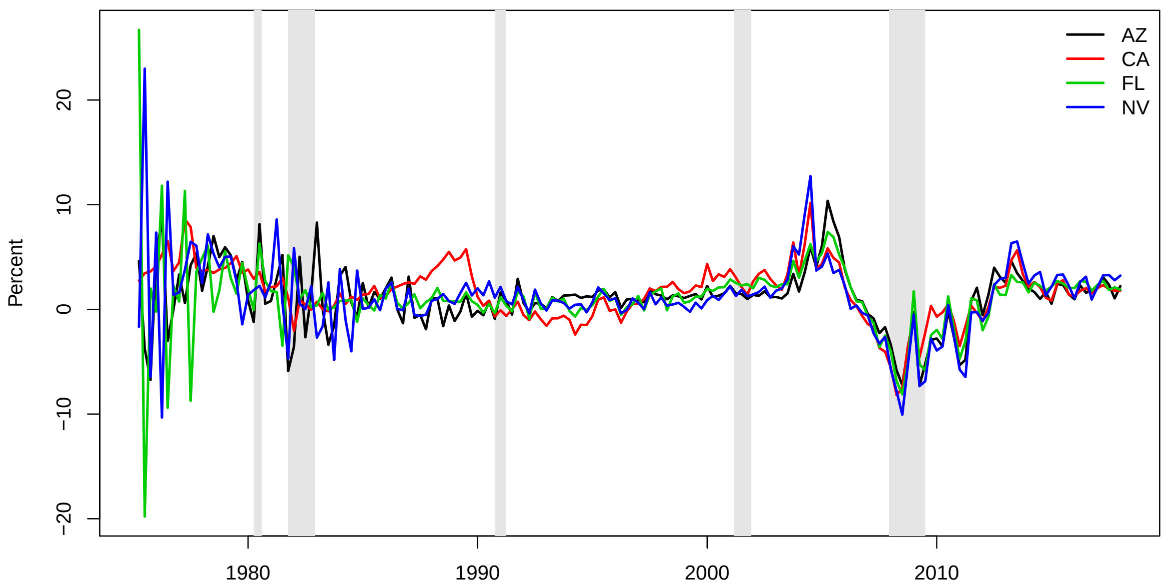
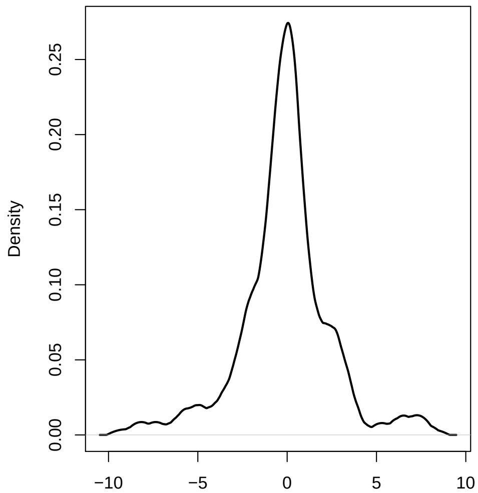
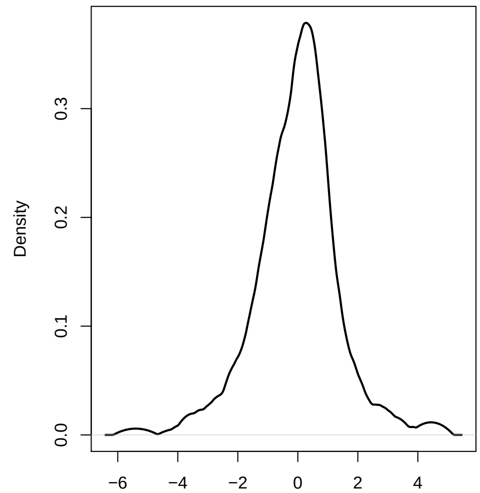


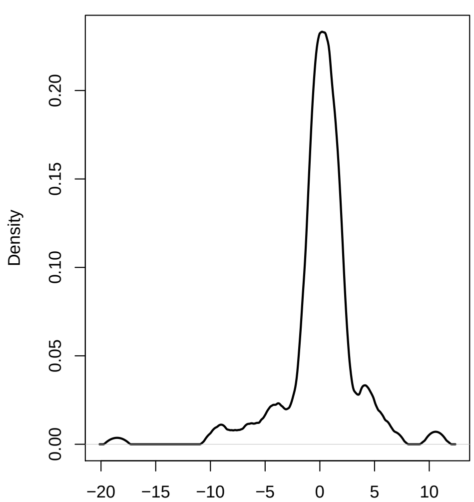
Figure 2: Quarterly percentage changes of HPI in AZ, CA, FL and NV. The shaded areas indicate the economic recessions by NBER.



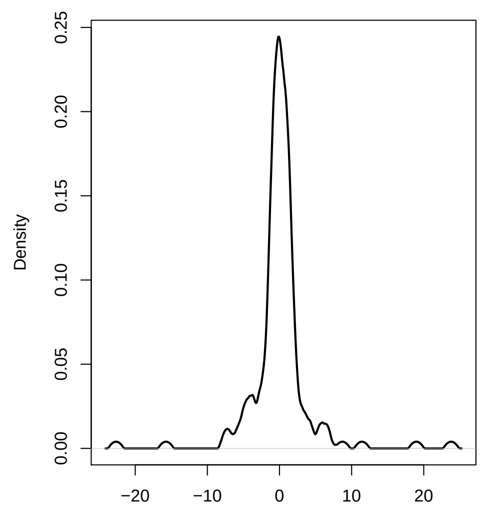
(a) Arizona, $JB = 39.23$ (0.00)



(b) California, $JB = 35.6$ (0.00)

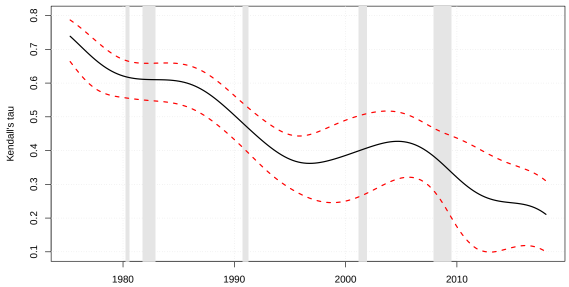


(c) Florida, $JB = 1101.90$ (0.00)

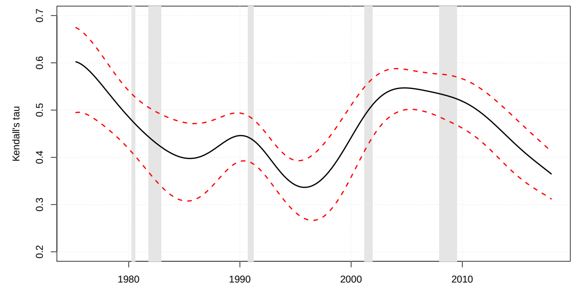


(d) Nevada, $JB = 1682$ (0.00)

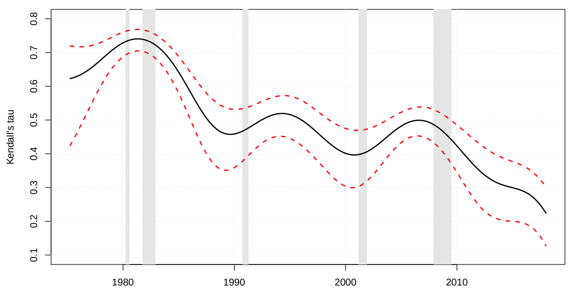
Figure 3: The kernel density curves of the filtered percentage changes of housing prices in Arizona, California, Florida and Nevada during 1975:Q1 - 2018:Q1. The smoothing kernel function is Epanechnikov. JB denotes the Jarque-Bera test statistic for null of normality. Values in parentheses are p-values.



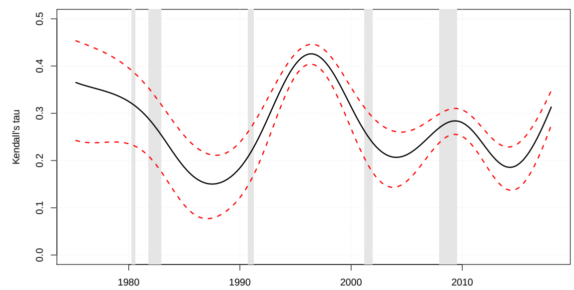
(a) Arizona - California



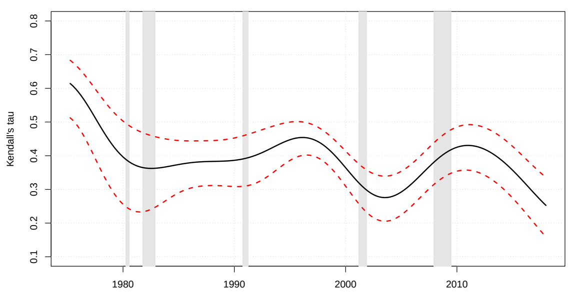
(b) Arizona - Florida



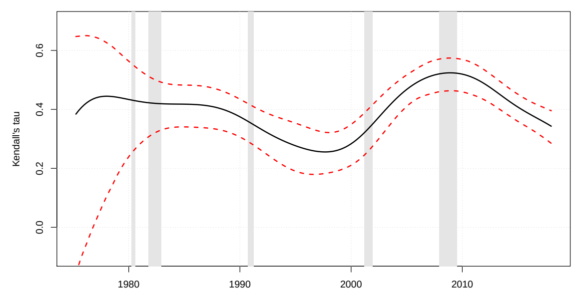
(c) Arizona - Nevada



(d) California - Florida



(e) California - Nevada



(f) Florida - Nevada

Figure 4: Quarterly estimates of Kendall's τ of the selected copula for six pairs of housing price indices (black curve) with the 95% confidence intervals (red dashed lines). For the pairs of AZ-FL and AZ-NV, the selected copula is Gumbel. For the pairs of AZ-CA, CA-FL, CA-NV and FL-NV, the selected copula is Frank. The data span between 1975:Q1 and 2018:Q1. The shaded areas represent the recession periods by NBER.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

001 "Data Driven Value -at -Risk Forecasting using a SVR -GARCH -KDE Hybrid" by Marius Lux, Wolfgang Karl Härdle, Stefan Lessmann, January 2018
002 "Nonparametric Variab le Selection and Its Application to Additive Models" by Zheng-Hui Feng, Lu Lin, Ruo -Qing Zhu as nd Li -Xing Zhu, January 2018
003 "Systemic Risk in Global Volatility Spillover Networks: Evidence from Option - implied Volatility Indices" by Zihui Yang, Yinggang Zhou, January 2018
004 "Pricing Cryptocurrency options: the case of CRIX and Bitcoin" by Cathy YH Chen, Wolfgang Karl Härdle, Ai Jun Hou, Weining Wang, January 2018
005 "Testing for b ubbles in cryptocurrencies with time -varying volatility" by Christian M. Hafner, January 2018
006 "A Note on Cryptocurrencies and Currency Competition" by Anna Almosova, January 2018
007 "Knowing me, knowing you: inventor mobility and the formation of technology -oriented alliances" by Stefan Wagner, Martin C. Go ossen, February 2018
008 "A Monetary Model of Blockchain" by Anna Almosova, February 2018
009 "Deregulated day -ahead electricity markets in Southeast Europe: Price forecasting and comparative structural analysis" by Antanina Hryshchuk, Stefan Lessmann, February 2018
010 "How Sensitive are Tail -related Risk Measures in a Contamination Neighbourh ood?" by Wolfgang Karl Härdle, Chengxiu Ling, February 2018
011 "How to Measure a Performance of a Collaborative Research Centre" by Alona Zharova, Janine Tellinger-Rice, Wolfgang Karl Härdle, February 2018
012 "Targeting customers for profit: An ensemble learning framework to support marketing decision making" by Stefan Lessmann, Kristof Coussement, Koen W. De Bock, Johannes Haupt, February 2018
013 "Improving C rime Count Forecasts Using Twitter and Taxi Data" by Lara Vomfell, Wolfgang Karl Härdle, Stefan Lessmann, February 2018
014 "Price Discovery on Bitcoin Markets" by Paolo Pagnottoni, Dirk G. Baur, Thomas Dimpfl, March 2018
015 "Bitcoin is not the New Gold - A Comparison of Volatility, Correlation, and Portfolio Performance" by Tony Klein, Hien Pham Thu, Thomas Walther, March 2018
016 "Time -varying Limit Order Book Networks" by Wolfgang Karl Härdle, Shi Chen, Chong Liang, Melanie Schienle, April 2018
017 "Regularization Approach for NetworkModeling of German EnergyMarket" by Shi Chen, Wolfgang Karl Härdle, Brenda López Cabrera, May 2018
018 "Adaptive Nonparametric Clustering" by Kirill Efimov, Larisa Adamyan, Vladimir Spokoiny, May 2018
019 "Lasso, knock off and Ga ussian covariates: a comparison" by Laurie Davies, May 2018
020 "A Regime Shift Model with Nonparametric Switching Mechanism" by Haiqiang Chen, Yingxing Li, Ming Lin, Yanli Zhu, May 2018
021 "LASSO -Driven Inference in Time and Space" by Victor Chernozhukov, Wolfgang K. Härdle, Chen Huang, Weining Wang, June 2018

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

022 "Learning from Errors: The case of monetary and fiscal policy regimes" by Andreas Tryphonides, June 2018
023 "Textual Sentiment , Option Characteristics, and Stock Return Predictability" by Cathy Yi-Hsuan Chen, Matthias R. Fengler, Wolfgang Karl Härdle, Yanchu Liu, June 2018
024 "Bootstrap Confidence Sets For Spectral Projectors Of Sample Covariance" by A. Naumov, V. Spokoiny, V. Ulyanov, June 2018
025 "Construction of Non -asymptotic Confidence Sets in 2 -Wasserstein Space" by Johannes Ebert, Vladimir Spokoiny, Alexandra Suvorikova, June 2018
026 "Large ball pro babilities, Gaussian comparison and anti -concentration" by Friedrich Götze, Alexey Naumov, Vladimir Spokoiny, Vladimir Ulyanov, June 2018
027 "Bayesian inference for spectral projectors of covariance matrix" by Igor Silin, Vladimir Spokoiny, June 2018
028 "Toolbox: Gaussian comparison on Eucledian balls" by Andzhey Koziuk, Vladimir Spokoiny, June 2018
029 "Pointwise adaptation via stagewise ag gregation of local estimates for multiclass classification" by Nikita Puchkin, Vladimir Spokoiny, June 2018
030 "Gaussian Process Forecast with multidimensional distributional entri es" by Francois Bachoc, Alexandra Suvorikova, Jean-Michel Loubes, Vladimir Spokoiny, June 2018
031 "Instrumental variables regression" by Andzhey Koziuk, Vladimir Spokoiny, June 2018
032 "Understanding Latent Group Structure of Cryptocurrencies Market: A Dynamic Network Perspective" by Li Guo, Yubo Tao, Wolfgang Karl Härdle, July 2018
033 "Optimal contracts under competition when uncertainty from adverse selection and moral hazard are present" by Natalie Packham, August 2018
034 "A factor -model appro ach for correlation scenarios and correlation stress -testing" by Natalie Packham, Fabian Woebbeking, August 2018
035 "Correlation Under Stress In Normal Variance Mixture Models" by Michael Kalkbrener, Natalie Packham, August 2018
036 "Model risk of contingent claims" by Nils Detering, Natalie Packham, August 2018
037 "Default probabilities and default correlations under stress" by Natalie Packham, Michael Kalkbrener, Ludger Overbeck, August 2018
038 "Tail -Risk Protection Trading Strategies" by Natalie Packham, Jochen Papenbrock, Peter Schwendner, Fabian Woebbeking, August 2018
039 "Penalized Adaptive Forecasting with Large Information Sets and Structural Changes" by Lenka Zbonakova, Xinjue Li, Wolfgang Karl Härdle, August 2018
040 "Complete Convergenc e and Complete Moment Convergence for Maximal Weighted Sums of Extended Negatively Dependent Random Variables" by Ji Gao YAN, August 2018
041 "On complete convergence in Marcinkiewicz -Zygmund type SLLN for random variables" by Anna Kuczmaszewska, Ji Gao YAN, August 2018

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit
http://irtg1792.hu-berlin.de.

042 "On Complete Convergence in Marcinkiewicz -Zygmund Type SLLN for END Random Variables and its Applications" by Ji Gao YAN, August 2018

043 "Textual Sentiment and Sector specific reaction" by Elisabeth Bommes, Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle, September 2018

044 "Understanding Cryptocurrencies" by Wolfgang Karl Härdle, Campbell R. Harvey, Raphael C. G. Reule, September 2018

045 "Predicative Ability of Similarity -based Futures Trading Strategies" by Hsin-Yu Chiu, Mi-Hsiu Chiang, Wei-Yu Kuo, September 2018

046 "Forecasting the Term Structure of Option Implied Volatility: The Power of an Adaptive Method" by Ying Chen, Qian Han, Linlin Niu, September 2018

047 "Inferences for a Partially Varying Coefficient Model With Endoge nous Regressors" by Zongwu Cai, Ying Fang, Ming Lin, Jia Su, October 2018

048 "A Regime Shift Model with Nonparametric Switching Mechanism" by Haiqiang Chen, Yingxing Li, Ming Lin, Yanli Zhu, October 2018

049 "Strict Stationarity Testing and GLAD Estimat ion of Double Autoregressive Models" by Shaojun Guo, Dong Li, Muyi Li, October 2018

050 "Variable selection and direction estimation for single -index models via DC -TGDR method" by Wei Zhong, Xi Liu, Shuangge Ma, October 2018

051 "Property Investment and Rental Rate under Housing Price Uncertainty: A Real Options Approach" by Honglin Wang, Fan Yu, Yinggang Zhou, October 2018

052 "Nonparametric Additive Instrumental Variable Estimator: A Group Shrinkage Estimation Perspective" by Qingliang Fan, Wei Zhong, October 2018

053 "The impact of temperature on gaming productivity: evidence from online games" by Xiaojia Bao, Qingliang Fan, October 2018

054 "Topic Modeling for Analyzing Open -Ended Survey Responses" by Andra-Selina Pietsch, Stefan Lessmann, October 2018

055 "Estimation of the discontinuous leverage effect: Evidence from the NASDAQ order book" by Markus Bibinger, Christopher Neely, Lars Winkelmann, October 2018

056 "Cryptocurrencies, Metcalfe's law and LPPL models" by Daniel Traian Pele, Miruna Mazurencu-Marinescu-Pele, October 2018

057 "Trending Mixture Copula Models with Copula Selection" by Bingduo Yang, Zongwu Cai, Christian M. Hafner, Guannan Liu, October 2018

058 "Investing with cryptocurrencies Œ evaluating the potential of portfolio allocation strategies" by Alla Petukhina, Simon Trimborn, Wolfgang Karl Härdle, Hermann Elendner, October 2018

059 "Towards the interpretation of time -varying regularization parameters in streaming penalized regression models" by Lenka Zbonakova, Ricardo Pio Monti, Wolfgang Karl Härdle, October 2018

060 "Residual's Influence Index (Rinfin), Bad Leverage And Unmaskin g In High Dimensional L2 -Regression" by Yannis G. Yatracos, October 2018

061 "Plug -In L2 -Upper Error Bounds In Deconvolution, For A Mixing Density Estimate In R d And For Its Derivatives" by Yannis G. Yatracos, October 2018

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche
Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

062 "Conversion uplift in e -com merce: A systematic benchmark of modeling strategies" by Robin Gubela, Artem Bequé, Fabian Gebert, Stefan Lessmann, November 2018

063 "Causal I nference using Machine Learning. An Evaluation of recent Methods through Simulations" by Daniel Jacob, Stefan Lessmann, Wolfgang Karl Härdle, November 2018

064 "Semiparametric Estimation and Variable Selection for Single-index Copula Models" by Bingduo Yang, Christian M. Hafner, Guannan Liu, Wei Long, December 2018

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.