Journal of Econometrics 235 (2023) 1500–1521

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Estimation and inference in factor copula models with exogenous covariates[☆]

Alexander Mayer [a,*], Dominik Wied [b]

[a] Department of Economics, Università Ca' Foscari Venezia, Italy
[b] Institute of Econometrics and Statistics, University of Cologne, Germany

ARTICLE INFO

Article history:
Received 21 December 2021
Received in revised form 1 December 2022
Accepted 6 January 2023
Available online 3 February 2023

JEL classification:
C13
C15
C22

Keywords:
Factor analysis
Simulation estimator
Empirical process
Dependence modeling

ABSTRACT

A factor copula model is proposed in which factors are either simulable or estimable from exogenous information. Point estimation and inference are based on a simulated methods of moments (SMM) approach with non-overlapping simulation draws. Consistency and limiting normality of the estimator is established and the validity of bootstrap standard errors is shown. Doing so, previous results from the literature are verified under low-level conditions imposed on the individual components of the factor structure. Monte Carlo evidence confirms the accuracy of the asymptotic theory in finite samples and an empirical application illustrates the usefulness of the model to explain the cross-sectional dependence between stock returns.

1. Introduction

Factor copula models have been successfully introduced as a means to cope with data of high cross-sectional dimensionality; see, e.g., Krupskii and Joe (2013), Creal and Tsay (2015), and Oh and Patton (2017). The use of a latent factor structure offers an economically intuitive yet flexible way to multivariate modeling that parsimoniously handles commonly encountered characteristics of financial time series like, for example, the tail asymmetry and tail dependence described by Hansen (1994). Recently, some research effort has been devoted to incorporate time variation and exogenous information to factor copula models; see, e.g., Creal and Tsay (2015), Oh and Patton (2018), Opschoor et al. (2020), and Krupskii and Joe (2020). For example, Oh and Patton (2018) and Opschoor et al. (2020), by utilizing the generalized autoregressive score (GAS) framework of Creal et al. (2013), consider specifications with latent factors and time-varying loadings that may depend on exogenous information. We contribute to this literature by introducing a class of factor

☆ Financial support by Deutsche Forschungsgemeinschaft, Germany (DFG grant 'Strukturbrüche und Zeitvariation in hochdimensionalen Abhängigkeitsstrukturen') is gratefully acknowledged. We would like to thank the associate editor and two anonymous referees for many helpful comments and suggestions which improved the quality of the paper. Moreover, we are grateful for helpful discussions with Hans Manner, Sven Otto, and Florian Stark as well as for computational support from Sebastian Valet. Suggestions and comments made by participants of research seminars at the universities of Bonn, Cologne, and the Balearic Islands, as well as by participants of the following workshops are highly appreciated: The 7th RCEA Time Series Workshop (virtual), the Symposium of the Hausdorff Center for Mathematics in Bonn 2021, the Workshop on High Dimensional Data Analysis at the University Carlos III de Madrid 2021 (virtual), the European Winter Meeting of the Econometric Society 2021 (virtual), the CFE 2021 (virtual), and the DAGStat 2022.

* Corresponding author.
E-mail address: alexandersimon.mayer@unive.it (A. Mayer).

https://doi.org/10.1016/j.jeconom.2023.01.003

copula models with exogenous, (partly) observable *factors*, an idea reminiscent of Bernanke et al. (2005), Boivin et al. (2009), and Stock and Watson (2005). Contrary to the above cited factor copula models, we take a step back and treat the, possibly group-specific, loadings as time-invariant constants and the SMM estimator employed here is build on the *un*conditional copula—a concession in the name of tractability that frees us from the necessity of specifying parametric marginals (e.g., Oh and Patton (2018)) or a closed form likelihood of the copula (e.g., Opschoor et al. (2020)) and thereby allows for a large variety of 'covariate-augmented' factor copulas, nesting the model (Oh and Patton, 2017) as a special case.

Since the copula likelihood is rarely available in closed form for the model class considered here, an SMM framework for estimation and inference is proposed which uses the general principles outlined by Oh and Patton (2013). Our main contribution is a novel distinction between simulable factors and factors that are estimable from exogenous information. Following the seminal SMM literature of McFadden (1989), Pakes and Pollard (1989), and Lee (1992), we exploit the benefits from non-overlapping simulation draws. The incorporation of exogenous covariates considerably complicates the development of an asymptotic theory as many arguments made by Oh and Patton (2013) do not apply. Nevertheless, we show that all technical hurdles can be overcome by combining recent developments from copula empirical process theory (see, e.g., Bücher and Volgushev (2013), Berghaus et al. (2017), and Neumeyer et al. (2019)) with a seminal result for extremum estimation with nonsmooth objective function due to Newey and McFadden (1994). In consequence, consistency, limiting normality, and validity of bootstrap standard errors are established. In doing so, we derive the stochastic equicontinuity of the objective function from primitive conditions on the distributional characteristics of the factor structure using the functional central limit theorem (FCLT) of Andrews and Pollard (1994) for α-mixing triangular arrays. The theory developed here verifies earlier equicontinuity results from the literature that made use of high-level conditions; see, e.g., Oh and Patton (2013), and Manner et al. (2019, 2021). Since stochastic equicontinuity is an essential ingredient of the asymptotic theory that links pointwise and uniform properties, more primitive conditions are of utmost interest. An application to dependence modeling of a cross-section of stock returns of eleven financial companies illustrates the theoretical results and highlights how the incorporation of estimable factors can help to achieve improvements in model performance.

The remainder of this paper is organized as follows. Section 2 introduces the model. The main results for SMM estimation and inference are contained in Section 3. A small Monte Carlo exercise is conducted in Section 4 and an empirical application can be found in Section 5. Section 6 briefly summarizes and concludes the paper.

2. Model

Our aim is to capture the dependence structure among the cross-sectional entities of the $n \times 1$ vector of financial assets $Y_t := (Y_{1,t}, \dots, Y_{n,t})'$ in time-period $t \in \{1, \dots, T\}$, conditional on the available information $\mathcal{F}_t := \sigma(\{Y_j^*, Y_{j-1} : j \leq t\})$, where Y_t^* represents a vector of exogenous regressors. The number of financial assets n might be large but is assumed finite. If the marginal conditional distributions $Y_{i,t} \mid \mathcal{F}_t \sim \mathsf{H}_{i,t}$ are continuous, we can follow Patton (2006) and uniquely decompose the joint conditional distribution $Y_t \mid \mathcal{F}_t \sim \mathsf{H}_t$ into its n margins and a copula function $\mathsf{C}_t : [0, 1]^n \mapsto [0, 1]$, where $\mathsf{C}_t(\cdot)$ completely describes the dependence conditionally on $\mathcal{F}_t$; i.e., $\mathsf{H}_t(x_1, \dots, x_n) = \mathsf{C}_t\{\mathsf{H}_{1,t}(x_1), \dots, \mathsf{H}_{n,t}(x_n)\}$, $x_i \in \mathbb{R}$, $i \in \{1, \dots, n\}$. Following, among others, Chen and Fan (2006), Oh and Patton (2013), and Fan and Patton (2014, Section 2.1), we assume for the n assets parametric location-scale specifications of the form

$$Y_{i,t} = \mu_{1,i}(\mathcal{F}_t, \lambda_0) + \mu_{2,i}(\mathcal{F}_t, \lambda_0)\eta_{i,t}, \tag{2.1}$$

where, for each $i \in \{1, \dots, n\}$, $\{\eta_{i,t} : t \geq 1\}$ are *i.i.d.* innovations independent of $\mathcal{F}_t$, while $\mu_{1,i}$ and $\mu_{2,i}$ are $\mathcal{F}_t$-measurable parametric specifications of the conditional mean $\mu_{1,i}(\mathcal{F}_t, \lambda) = \mathsf{E}[Y_{i,t} \mid \mathcal{F}_t]$ and the conditional standard deviation $\mu_{2,i}(\mathcal{F}_t, \lambda) = \sqrt{\mathsf{var}[Y_{i,t} \mid \mathcal{F}_t]}$ that are known up to the true $r \times 1$ parameter vector $\lambda = \lambda_0 \in \Lambda_0 \subset \mathbb{R}^r$. In particular, we assume that

$$\mu_{j,i}(\mathcal{F}_t, \lambda) = \mu_{j,i}(R_t(\lambda), \lambda), \quad j \in \{1, 2\},\ i \in \{1, \dots, n\}, \tag{2.2}$$

where, for any $\lambda \in \Lambda$, the $p_{\mathsf{R}} \times 1$ random vector $R_t(\lambda)$ is $\mathcal{F}_t$-measurable with components that may parametrically depend on the entire history $\mathcal{F}_t$ through λ. For example, this model class includes nonlinear autoregressions (AR) with nonlinear autoregressive conditional heteroskedasticity (ARCH), in which case the dependence of $R_t(\lambda)$ on λ is superfluous, as well as nonlinear generalized autoregressive conditional heteroskedasticity (GARCH) as illustrated below for the GJR model of Glosten et al. (1993).

Example. Consider the AR(1) model $Y_{i,t} = \gamma_i Y_{i,t-1} + \mu_{2,i,t}\eta_{i,t}$, with conditional heteroskedasticity $\mu_{2,i,t}^2 = \omega_i + (\beta_i + \eta_i 1\{\varepsilon_{i,t-1}(\gamma_i) < 0\})\varepsilon_{i,t-1}^2(\gamma_i) + \alpha_i \mu_{2,i,t-1}^2$ for $\varepsilon_{i,t}(\gamma_i) := Y_{i,t} - \gamma_i Y_{i,t-1}$. In accordance with our notation, let $\lambda = (\gamma_1, \dots, \gamma_n, \alpha_1, \dots, \alpha_n, \omega_1, \dots, \omega_n, \beta_1, \dots, \beta_n, \eta_1, \dots, \eta_n)'$ be the $r \times 1$ parameter vector collecting all unknowns for $r = 5n$. Following the discussion in Yang (2006), by suitably restricting $\Lambda_0 \subset \mathbb{R}^r$, this specification can be cast in form of (2.1) and (2.2) if we set $\mu_{1,i}(\mathcal{F}_t, \lambda) = \gamma_i Y_{i,t-1}$, $R_t(\lambda) = (Y_{t-1}', \tilde{R}_t(\lambda)')'$, where $\tilde{R}_t(\lambda) = (\tilde{R}_{1,t}(\lambda), \dots, \tilde{R}_{n,t}(\lambda))'$, with $\tilde{R}_{i,t}(\lambda) = \sum_{j=1}^{\infty} \alpha_i^{j-1} g_i\{Y_{i,t-j} - \gamma_i Y_{i,t-1}, \eta_i\}$, $g(x, a) = x^2 + a1\{x < 0\}x^2$, and $\mu_{2,i}(R_t(\lambda), \lambda) = \beta_i \tilde{R}_{i,t}(\lambda) + \omega_i/(1 - \alpha_i)$.

Since we assume that individually $\eta_{i,t}$ are independent of $\mathcal{F}_t$ and, by Eq. (2.1), $\eta_{i,t} = \{Y_{i,t} - \mu_{1,i}(\mathcal{F}_t, \lambda_0)\}/\mu_{2,i}(\mathcal{F}_t, \lambda_0)$, we can—assuming continuous margins $\mathsf{F}_i(\eta_{i,t} \le x_i) := \mathsf{P}(\eta_{i,t} \le x_i)$, $i \in \{1, \dots, n\}$—rephrase the conditional joint distribution of Y_t in terms of the innovation ranks $V_{i,t} := \mathsf{F}_i(\eta_{i,t})$ as

$$\begin{aligned} \mathsf{H}_t(x_1, \dots, x_n) &= \mathsf{C}_t\{\mathsf{F}_1(x_1), \dots, \mathsf{F}_n(x_n)\}, \\ \mathsf{C}_t(u_1, \dots, u_n) &= \mathsf{P}(V_{1,t} \le u_1, \dots, V_{n,t} \le u_n \mid \mathcal{F}_t), \end{aligned} \tag{2.3}$$

$u_i \in [0, 1]$, $i \in \{1, \dots, n\}$. While the effect of $\mathcal{F}_t$ on the margins has been completely removed by the use of the location-scale model (2.1), the *joint* cross-sectional distribution of $\eta_t := (\eta_{1,t}, \dots, \eta_{n,t})'$ is allowed to depend on $\mathcal{F}_t$ through some 'exogenous' vector Z_t. That is, we assume that

$$\mathsf{C}_t(u_1, \dots, u_n) = \mathsf{P}(V_{1,t} \le u_1, \dots, V_{n,t} \le u_n \mid \mathcal{F}_t) = \mathsf{P}(V_{1,t} \le u_1, \dots, V_{n,t} \le u_n \mid Z_t). \tag{2.4}$$

We do not model the conditional copula directly but we assume that the unconditional copula

$$\mathsf{C}(u_1, \dots, u_n) = \mathsf{P}(V_{1,t} \le u_1, \dots, V_{n,t} \le u_n), \tag{2.5}$$

can be generated from an auxiliary factor model *via*

$$\mathsf{G}(x_1, \dots, x_n) = \mathsf{C}\{G_1(x_1), \dots, G_n(x_n)\}, \ \ x_i \in \mathbb{R}, \tag{2.6}$$

where $G_i(x_i) := \mathsf{P}(X_{i,t} \le x_i)$ represents the ith margin of

$$X_{i,t} = a_{0,i}' F_t + b_{0,i}' Z_t + \varepsilon_{i,t} \tag{2.7}$$

and $\mathsf{G}(x) := \mathsf{P}(X_{1,t} \le x_1, \dots, X_{n,t} \le x_n)$, $x := (x_1, \dots, x_n)' \in \mathbb{R}^n$, is the corresponding joint distribution. As defined below, F_t and $\varepsilon_{i,t}$ denote latent factors and the idiosyncratic component, respectively. It is crucial to stress that the margins G_i, $i \in \{1, \dots, n\}$, can differ from the univariate distributions of the observed data and are not of interest here. Rather, Eq. (2.7) serves as a means to generate the copula C that determines the joint distribution. Since any effect of $\mathcal{F}_t$ on the margins has been filtered out while the conditional distribution is only affected by $\mathcal{F}_t$ through the regressor, the unconditional copula obtains directly by taking the expectation with respect to Z_t, i.e. $\mathsf{C}(\cdot) = \mathsf{E}[\mathsf{C}_t(\cdot)]$.

Oh and Patton (2018) or Opschoor et al. (2020) consider a related specification. Contrary to our approach, however, they model the conditional copula directly, i.e. they consider a conditional copula $\mathsf{C}_t = \mathsf{C}(\theta_{0,t})$ indexed by a time-varying copula parameter $\theta_{0,t} := \theta_0(Z_t)$ that is driven by GAS-dynamics so that $Z_t = (\eta_1', \dots, \eta_{t-1}')'$. We, on the other hand, assume that all components of the factor model including Z_t are *i.i.d.*. To make these notions precise, Assumption A formalizes the characteristics of the factor model. It constitutes a naturally extension of Oh and Patton (2017).

Assumption A.

(A1) $\{Z_t : t \ge 1\}$, $\{F_t : t \ge 1\}$ and $\{\varepsilon_t : t \ge 1\}$ are mutually independent *i.i.d.* sequences with $\varepsilon_{i,t} \perp \varepsilon_{j,t}$, $i, j \in \{1, \dots, n\}$ and $F_{t,i} \perp F_{t,j}$, $i, j \in \{1, \dots, p_\alpha\}$;
(A2) $\{X_t, \eta_t : t \ge 1\}$ is an *i.i.d.* sequence.

The peculiar feature of Eq. (2.7), and the main contribution of this paper, is the distinction between *simulable* and *observable* factors: while $F_t := (F_{t,1}, \dots, F_{t,p_\alpha})'$ is a $p_\alpha \times 1$ vector of latent random variables with known parametric distribution, the $p_\beta \times 1$ vector $Z_t := (Z_{t,1}, \dots, Z_{t,p_\beta})'$ can be recovered from observed data based on econometric tools. Therefore, Z_t is also referred to as the estimable factor. More specifically, both F_t and $\varepsilon_{i,t}$ are *i.i.d.* with parametric distributions

$$\mathsf{D}_\varepsilon(x;\ \delta_0) := \mathsf{P}(\varepsilon_{i,t} \le x), \quad \mathsf{D}_{F,j}(x;\ \gamma_{0,j}) := \mathsf{P}(F_{t,j} \le x),\ j \in \{1, \dots, p_\alpha\},$$

which are partially known up to the $p_\delta \times 1$ vector $\delta_0 := (\delta_{0,1}, \dots, \delta_{0,p_\delta})'$ and the $p_\alpha p_\gamma \times 1$ vector $\gamma_0 := (\gamma_{0,1}', \dots, \gamma_{0,p_\alpha}')'$, with $\gamma_{0,j} := (\gamma_{0,j,1}, \dots, \gamma_{0,j,p_\gamma})'$, respectively. On the other hand, the distribution of the vector Z_t is unknown but we assume that its components $Z_{t,j}$ can be represented as *i.i.d.* innovations of an observable $\mathcal{F}_t$-measurable processes $W_{t,j}$ given by

$$Z_{t,j} = W_{t,j} - \sigma_j(\mathcal{W}_t, \nu_0), \ \ j \in \{1, \dots, p_\beta\}, \tag{2.8}$$

where the measurable functions $\sigma_j(\cdot, \nu)$ are known up to the $m \times 1$ vector $\nu = \nu_0 \in \mathcal{V}_0 \subset \mathbb{R}^m$; the *i.i.d.* innovations $Z_{t,j}$ are independent of the history $\mathcal{W}_t \subset \mathcal{F}_{t-1}$. Similar to the specification of Eq. (2.2), we assume that $\sigma_j(\mathcal{W}_t, \nu) = \sigma_j(M_t(\nu), \nu)$, where the $p_{\mathsf{M}} \times 1$ vector $M_t(\nu)$ comprises short-range dependent covariates, possibly including lagged dependent variables, whose components may be parametrically generated from the complete history $\mathcal{W}_t$. We cannot allow for both time-varying conditional means *and* variances to ensure that the limiting distribution of the SMM estimator is unaffected by the first step estimation of ν_0. Note, however, that several models with time-varying conditional variance that obey $\tilde{W}_t = \tilde{\sigma}(M_t(\nu), \nu)\tilde{Z}_t$, $\sup_{x,y} \tilde{\sigma}(x, y) > 0$, can be cast in form of (2.8) by setting $W_t := \log|\tilde{W}_t|$, $\sigma_t(M_t(\nu), \nu) := \log \tilde{\sigma}_t(M_t(\nu), \nu)$, and $Z_t := \log|\tilde{Z}_t|$.[1]

[1] The argument is inspired by Genest et al. (2007), who propose this transformation in their Example 1 to ensure a nuisance-free distribution of the BDS-type test studied there; see also Caporale et al. (2005) for a study of the BDS test based on the logarithm of absolute GARCH(1,1)-residuals.

As in Oh and Patton (2017, Section 4.2) and Opschoor et al. (2020, Section 2.1.1), it is assumed that the factor loadings $a_{0,i} := (a_{0,i,1}, \dots, a_{0,i,p_\alpha})'$ and $b_{0,i} := (b_{0,i,1}, \dots, b_{0,i,p_\beta})'$ can be grouped into a small number of Q group-specific coefficients $\alpha_{0,q} := (\alpha_{0,q,1}, \dots, \alpha_{0,q,p_\alpha})'$ and $\beta_{0,q} := (\beta_{0,q,1}, \dots, \beta_{0,q,p_\beta})'$, $q \in \{1, \dots, Q\}$. Put differently, there exists a finite collection of disjoint sets $\{\mathcal{G}_1, \dots, \mathcal{G}_Q\}$ partitioning the cross-sectional index set $\{1, \dots, n\}$ such that $a_{0,i} = a_{0,j} = \alpha_{0,q}$ and $b_{0,i} = b_{0,j} = \beta_{0,q}$ for any $i, j \in \mathcal{G}_q$, $q \in \{1, \dots, Q\}$. Importantly, this 'block-equidependent' factor structure implies that the number of latent marginals needed to be specified reduces from n to Q distinct distributions $\mathsf{G}_1, \dots, \mathsf{G}_Q$, say, so that $G_i(x) = G_j(x) =: \mathsf{G}_q(x)$ for any $i, j \in \mathcal{G}_q$, $q \in \{1, \dots, Q\}$. Throughout, the group assignment is assumed to be known.

3. SMM-based estimation

The object of interest is the $p \times 1$ vector $\theta_0 := (\alpha'_{0,1}, \beta'_{0,1}, \dots, \alpha'_{0,Q}, \beta'_{0,Q}, \gamma'_0, \delta'_0)' \in \Theta \subseteq \mathbb{R}^p$, which, in view of the latent factor structure (2.6), collects all $p := Q(p_\alpha + p_\beta) + p_\alpha p_\gamma + p_\delta$ unknown copula parameters. A different parameter vector $\theta := (\alpha'_1, \beta'_1, \dots, \alpha'_Q, \beta'_Q, \gamma', \delta')' \in \Theta$ gives rise to an alternative factor structure

$$X_{i,t}(d_i) := a_i' F_t(\gamma) + b_i' Z_t + \varepsilon_{i,t}(\delta), \tag{3.1}$$

say, where the notational conventions $X_{i,t}(d_i)$, $F_t(\gamma) := (F_{t,1}(\gamma_1), \dots, F_{t,p_\alpha}(\gamma_{p_\alpha}))'$, and $\varepsilon_{i,t}(\delta)$ are used to make the dependence of the various quantities on the $(p_\alpha + p_\beta + p_\alpha p_\gamma + p_\delta) \times 1$ vector $d_i := (a_i', b_i', \gamma', \delta')'$ explicit; e.g., $\varepsilon_{i,t}(\delta) \sim \mathsf{D}_\varepsilon(\delta)$ and $F_{t,j}(\gamma_j) \sim \mathsf{D}_{F,j}(\gamma_j)$, $j \in \{1, \dots, p_\alpha\}$. The block-equidependent design ensures that

$$d_i = d_j = \theta_q, \quad \forall\, i, j \in \mathcal{G}_q,\ q \in \{1, \dots, Q\}, \tag{3.2}$$

where the $(p_\alpha + p_\beta + p_\alpha p_\gamma + p_\delta) \times 1$ vector $\theta_q := (\alpha'_q, \beta'_q, \delta', \gamma')'$ contains the parameters specific to the qth group. Thus, with a slight abuse of notation, $\theta = \cup_{q=1}^{Q} \theta_q$. For each $\theta \in \Theta$, Eq. (3.1) generates a differently parametrized copula

$$\begin{aligned} \mathsf{C}(u_1, \dots, u_n; \theta) &:= \mathsf{P}\{U_{1,t}(d_1) \le u_1, \dots, U_{n,t}(d_n) \le u_n\}, \\ U_{i,t}(d_i) &:= G_i(X_{i,t}(d_i); d_i), \\ G_i(x_i; d_i) &:= \mathsf{P}\{X_{i,t}(d_i) \le x_i\},\ u_i \in [0, 1],\ x_i \in \mathbb{R}. \end{aligned} \tag{3.3}$$

Importantly, due to the block-equidependent design and Eq. (3.2), we have for any two cross-sectional indices belonging to the same group

$$G_i(x; d_i) = G_j(x; d_j) =: \mathsf{G}_q(x; \theta_q), \quad i, j \in \mathcal{G}_q,\ q \in \{1, \dots, Q\},$$

say, so that $G_i(X_{i,t}(d_i); d_i) = U_{i,t}(d_i) = U_{i,t}(\theta_q) = \mathsf{G}_q(X_{i,t}(\theta_q); \theta_q)$.

The simulation-based estimation uses Assumption B below to estimate the true value of $\theta \in \Theta$ given by $\theta_0 = \cup_{q=1}^{Q} \theta_{0,q}$, $\theta_{0,q} = (\alpha'_{0,q}, \beta'_{0,q}, \gamma'_0, \delta'_0)'$.

Assumption B.

(B1) $\mathsf{C}(u_1, \dots, u_n) = \mathsf{C}(u_1, \dots, u_n; \theta)$ uniformly in $(u_1, \dots, u_n)' \in [0, 1]^n$ if and only if $\theta = \theta_0$, where $\theta_0 \in \Theta \subset \mathbb{R}^p$. The parameter space Θ is compact.

(B2) (*i*) The joint distribution $\mathsf{F}(x_1, \dots, x_n) = \mathsf{P}(\eta_{1,t} \le x_1, \dots, \eta_{n,t} \le x_n)$ is continuous with continuous marginal distributions $\mathsf{F}_i(x_i)$, $x_i \in \mathbb{R}$, $i \in \{1, \dots, n\}$.

(*ii*) The joint distribution $\mathsf{G}(x_1, \dots, x_n; \theta) = \mathsf{P}(X_{1,t}(d_1) \le x_1, \dots, X_{n,t}(d_n) \le x_n)$ is continuous with continuous marginal distributions $G_i(x_i; d_i)$, $x_i \in \mathbb{R}$, $i \in \{1, \dots, n\}$, uniformly in $\theta \in \Theta$.

Assumption B formalizes the introductory notion of a factor copula; i.e., the unknown copula $\mathsf{C}(u_1, \dots, u_n)$ can be generated by the latent factor structure for a suitable choice of $\theta = \theta_0 \in \Theta$.

3.1. Independent simulations

Akin to the 'independent simulation' scheme known from classical SMM estimation (see, e.g. McFadden (1989), Pakes and Pollard (1989), and Lee (1992)), we generate for a given candidate value $\theta \in \Theta \subset \mathbb{R}^p$ a random sample $\{(F_{t,s}(\gamma)', \varepsilon_{i,t,s}(\delta))' : i = 1, \dots, n,\ t = 1, \dots, T,\ s = 1, \dots, S\}$ to obtain a version of the auxiliary factor model

$$X_{i,t,s}(d_i) := a_i' F_{t,s}(\gamma) + b_i' Z_t + \varepsilon_{i,t,s}(\delta), \tag{3.4}$$

where $F_{t,s}(\gamma) := (F_{t,s,1}(\gamma_1), \dots, F_{t,s,p_\alpha}(\gamma_{p_\alpha}))'$. Hence, we sample for each time period t a new batch of S random variables from D_ε and $\mathsf{D}_{F,j}$, $j \in \{1, \dots, p_\alpha\}$. More specifically, let $\mathsf{D}_\varepsilon^{-1}$ and $\mathsf{D}_{F,j}^{-1}$ denote the inverse distribution functions of ε_t and $F_{t,j}$, $j \in \{1, \dots, p_\alpha\}$, respectively. We may then write $\varepsilon_{i,t,s}(\delta) := \mathsf{D}_\varepsilon^{-1}(\varepsilon^*_{i,t,s}; \delta)$ and $F_{t,s}(\gamma) := \mathsf{D}_F^{-1}(F^*_{t,s}; \gamma)$, with $\mathsf{D}_F^{-1}(F^*_{t,s}; \gamma) := (\mathsf{D}_{F,1}^{-1}(F^*_{t,s,1}; \gamma_1), \dots, \mathsf{D}_{F,p_\alpha}^{-1}(F^*_{t,s,p_\alpha}; \gamma_{p_\alpha}))'$, where $F^*_{t,s} := (F^*_{t,s,1}, \dots, F^*_{t,s,p_\alpha})'$ and $\varepsilon^*_{i,t,s}$ denote independent draws of *i.i.d.* standard uniform random variates which are drawn once. Note that for our estimation procedure we hold the underlying random draws $F^*_{t,s}$ and $\varepsilon^*_{i,t,s}$ fixed while θ is allowed to vary over the compact set Θ. This is important to ensure uniform convergence of simulated moments and to facilitate convergence of the numerical optimization routine used to find θ_0; see Gouriéroux and Monfort (1997, p. 29) and Pakes and Pollard (1989, p. 1048) for further remarks. Since Z_t might be unobservable, we will replace the unknown innovation with $\hat{Z}_t(\hat{\nu}_T) := (\hat{Z}_{t,1}(\hat{\nu}_T), \dots, \hat{Z}_{t,p_\beta}(\hat{\nu}_T))'$, where $\hat{Z}_{t,j}(\nu) := W_{t,j} - \sigma_j(M_t(\nu), \nu)$, $j \in \{1, \dots, p_\beta\}$, represents the generalized residual. The estimator $\hat{\nu}_T$ is assumed to be $\sqrt{T}$-consistent for the $m \times 1$ vector ν_0 satisfying certain mild regularity conditions outlined below; for example, in the empirical application maximum likelihood estimation is used. Therefore, a feasible counterpart of Eq. (3.4) is obtained from

$$\hat{X}_{i,t,s}(d_i, \hat{\nu}_T) := a_i' F_{t,s}(\gamma) + b_i' \hat{Z}_t(\hat{\nu}_T) + \varepsilon_{i,t,s}(\delta). \tag{3.5}$$

3.2. The estimator

Throughout, the cross-sectional dimension n might be large but is considered fixed, while asymptotics are carried out as $T \to \infty$; the number of simulation draws S can either be fixed or a function of T such that $S := S(T) \to \infty$ as $T \to \infty$. For the sake of brevity we report only results for the latter case.

Similar to Oh and Patton (2013), estimation aims at minimizing the difference between empirical and simulated rank dependence measures that only depend on the unknown bivariate marginal copulae. Importantly, Assumption B implies also the equivalence at $\theta = \theta_0$ between each of the $n(n-1)/2$ bivariate marginals copulae of the joint copula from Eq. (2.5), given by $\mathsf{C}_{i,j}(u_i, u_j) := \mathsf{P}(V_{i,t} \le u_i, V_{j,t} \le u_j)$, and the bivariate marginals of the joint factor copula from Eq. (3.3), given by $\mathsf{C}_{i,j}(u_i, u_j; d_i, d_j) := \mathsf{P}(U_{i,t}(d_i) \le u_i, U_{j,t}(d_j) \le u_j)$, $1 \le i < j \le n$. By block-equidependence,

$$\mathsf{C}_{i,j}(u_i, u_j; d_i, d_j) =: \mathsf{C}_q(u_i, u_j; \theta_q), \quad i, j \in \mathcal{G}_q,\ q \in \{1, \dots, Q\}.$$

Put differently, the number of distinct marginal copulae reduces from $n(n-1)/2$ to Q block-specific copulae $\mathsf{C}_1(\theta_1), \dots, \mathsf{C}_Q(\theta_Q)$ for which $\mathsf{C}_{i,j}(u_i, u_j) = \mathsf{C}_q(u_i, u_j; \theta_q)$ if $i, j \in \mathcal{G}_q$, $q \in \{1, \dots, Q\}$, and $\theta_q = \theta_{0,q}$. To illustrate the main idea behind the following SMM estimator, suppose $i, j \in \mathcal{G}_q$ for some $q \in \{1, \dots, Q\}$ and introduce, for some estimator $\sqrt{T}$-consistent estimator $\hat{\lambda}_T$ of λ_0 satisfying some regularity conditions outlined below, the following two $\ell \times 1$ vectors

$$\begin{aligned} \hat{\psi}_{T,i,j}(\hat{\lambda}_T) &:= (\hat{\psi}_{T,i,j,1}(\hat{\lambda}_T), \dots, \hat{\psi}_{T,i,j,\ell}(\hat{\lambda}_T))' \\ \hat{\psi}_{T,S,i,j}(\theta_q, \hat{\nu}_T) &:= (\hat{\psi}_{T,S,i,j,1}(\theta_q, \hat{\nu}_T), \dots, \hat{\psi}_{T,S,i,j,\ell}(\theta_q, \hat{\nu}_T))', \end{aligned} \tag{3.6}$$

which collect bivariate dependence measures like, for example, Spearman's ρ, Blomqvist's β, Gini's γ, or the measures of quantile dependence used by Oh and Patton (2013, p. 691). Formally, these statistics can be expressed with the help of a suitable collection of bivariate functions $\{\varphi_k : [0,1]^2 \mapsto \mathbb{R}, 1 \le k \le \ell\}$ as follows

$$\begin{aligned} \hat{\psi}_{T,i,j,k}(\hat{\lambda}_T) &:= \frac{1}{T} \sum_{t=1}^{T} \varphi_k(\hat{V}_{i,t}(\hat{\lambda}_T), \hat{V}_{j,t}(\hat{\lambda}_T)) \\ \hat{\psi}_{T,S,i,j,k}(\theta_q, \hat{\nu}_T) &:= \frac{1}{TS} \sum_{t=1}^{T} \sum_{s=1}^{S} \varphi_k(\hat{U}_{i,t,s}(\theta_q, \hat{\nu}_T), \hat{U}_{j,t,s}(\theta_q, \hat{\nu}_T)), \end{aligned} \tag{3.7}$$

where $\hat{V}_{i,t}(\hat{\lambda}_T)$ and $\hat{U}_{i,t,s}(\theta_q, \hat{\nu}_T)$ represent the rank of $\hat{\eta}_{i,t}(\hat{\lambda}_T) := \{Y_{i,t} - \mu_{1,i}(\mathcal{F}_t, \hat{\lambda}_T)\}/\mu_{2,i}(\mathcal{F}_t, \hat{\lambda}_T)$ among $\{\hat{\eta}_{i,t}(\hat{\lambda}_T) : t = 1, \dots, T\}$ and the rank of $\hat{X}_{i,t,s}(\theta_q, \hat{\nu}_T)$ among $\{\hat{X}_{i,t,s}(\theta_q, \hat{\nu}_T) : t = 1, \dots, T; s = 1, \dots, S\}$, respectively.

The ℓ different bivariate dependence measures are then aggregated according to the group-specific factor structure. To provide some intuition, note that $\hat{\psi}_{T,i,j,k}(\hat{\lambda}_T)$ and $\hat{\psi}_{T,S,i,j,k}(\theta_q, \hat{\nu}_T)$, $k \in \{1, \dots, \ell\}$, can be viewed as sample estimates of the population statistics $\mathsf{E}[\varphi_k(V_{i,t}, V_{j,t})]$ and $\mathsf{E}[\varphi_k(U_{i,t,s}(\theta_q), U_{j,t,s}(\theta_q))]$. These statistics depend only on the bivariate copulae $\mathsf{C}_q(u_i, u_j)$ and $\mathsf{C}_q(u_i, u_j; \theta_q)$, which, due to the block-equidependence, exhibit within-group homogeneity; i.e., each of the ℓ statistics depends on the cross-sectional index set $\{1, 2, \dots, n\}$ only *via* the group-identifier $q \in \{1, \dots, Q\}$:

$$\begin{aligned} \psi_{q,k}(\theta) &:= \mathsf{E}[\varphi_k(U_{i,t,s}(\theta_q), U_{j,t,s}(\theta_q))] = \int_{[0,1]^2} \varphi_k(u_i, u_j) \mathrm{d}\mathsf{C}_q(u_i, u_j; \theta_q) \\ \psi_{q,k} &:= \mathsf{E}[\varphi_k(V_{i,t}, V_{j,t})] = \int_{[0,1]^2} \varphi_k(u_i, u_j) \mathrm{d}\mathsf{C}_q(u_i, u_j), \quad k \in \{1, \dots, \ell\},\ \forall\, i, j \in \mathcal{G}_q. \end{aligned} \tag{3.8}$$

Therefore, the following aggregation scheme of bivariate dependence measures is justified

$$\begin{aligned} \hat{\psi}_{T,S,q}(\theta, \hat{\nu}_T) &:= \frac{1}{\binom{|\mathcal{G}_q|}{2}} \sum\sum_{\substack{1 \le i < j \le n \\ i,j \in \mathcal{G}_q}} \hat{\psi}_{T,S,i,j}(\theta_q, \hat{\nu}_T) \\ \hat{\psi}_{T,q}(\hat{\lambda}_T) &:= \frac{1}{\binom{|\mathcal{G}_q|}{2}} \sum\sum_{\substack{1 \le i < j \le n \\ i,j \in \mathcal{G}_q}} \hat{\psi}_{T,i,j}(\hat{\lambda}_T), \quad q \in \{1, \dots, Q\}. \end{aligned} \tag{3.9}$$

We thus obtain the $\bar{\ell} \times 1$, $\bar{\ell} := Q\ell$, vector of empirical dependence measures

$$\hat{\psi}_T(\hat{\lambda}_T) := (\hat{\psi}_{T,1}(\hat{\lambda}_T)', \ldots, \hat{\psi}_{T,Q}(\hat{\lambda}_T)')'$$

and the $\bar{\ell} \times 1$ vector of simulated dependence measures

$$\hat{\psi}_{T,S}(\theta, \hat{\nu}_T) := (\hat{\psi}_{T,S,1}(\theta_1, \hat{\nu}_T)', \ldots, \hat{\psi}_{T,S,Q}(\theta_Q, \hat{\nu}_T)')',$$

respectively. Following the literature on extremum estimators (see, e.g., Newey and McFadden (1994, Section 7)), we define, for some stochastically bounded and positive-definite weight matrix $\hat{L}_{T,S}$, an SMM estimator $\hat{\theta}_{T,S}$ as an estimator that minimizes the objective function

$$\hat{A}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T) := \hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T)'\hat{L}_{T,S}\hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T),$$

with $\hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T) := \hat{\psi}_T(\hat{\lambda}_T) - \hat{\psi}_{T,S}(\theta, \hat{\nu}_T)$, in the sense that

$$\hat{A}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) \leq \inf_{\theta \in \Theta} \hat{A}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T) + o_p(1/T). \tag{3.10}$$

3.3. Limiting normality

As pointed out by Oh and Patton (2013), the objective function is non-differentiable and, in general, does not possess a population counterpart in known closed form. Thus, some care is required in deriving the asymptotic distribution of $\hat{\theta}_{T,S}$. Due to the mutual dependence on the covariate, $\hat{\psi}_T(\hat{\lambda}_T)$ and $\hat{\psi}_{T,S}(\theta, \hat{\nu}_T)$ are not independent, which considerably complicates the analysis and precludes a direct application of the arguments developed by the aforementioned authors. To shed some light, let us recall from Newey and McFadden (1994) that two crucial high-level conditions for asymptotic normality of $\sqrt{T}(\hat{\theta}_{T,S} - \theta_0)$ are (1) limiting normality of the normalized sample 'moments' $\sqrt{T}\hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T)$ evaluated at $\theta = \theta_0$ and (2) the stochastic equicontinuity of the map $\theta \mapsto \sqrt{T}\hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T)$. Both conditions are shown to be closely tied to the limiting behavior of the triangular-array empirical process

$$\begin{aligned}
&\hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j; \theta_q, \hat{\lambda}_T, \hat{\nu}_T) \\
&\quad := \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Big[1\{\hat{\eta}_{i,t}(\hat{\lambda}_T) \leq \hat{\mathsf{F}}^-_{T,i}(u_i; \hat{\lambda}_T), \hat{\eta}_{j,t}(\hat{\lambda}_T) \leq \hat{\mathsf{F}}^-_{T,j}(u_j; \hat{\lambda}_T)\} \\
&\qquad - \frac{1}{S}\sum_{s=1}^{S} 1\{\hat{X}_{i,t,s}(\theta_q, \hat{\nu}_T) \leq \hat{\mathsf{G}}^-_{T,S,i}(u_i; \theta_q, \hat{\nu}_T), \hat{X}_{j,t,s}(\theta, \hat{\nu}_T) \leq \hat{\mathsf{G}}^-_{T,S,j}(u_j; \theta_q, \hat{\nu}_T)\} \Big],
\end{aligned} \tag{3.11}$$

where $\mathsf{H}^-(p) := \inf\{x \in \mathbb{R} : \mathsf{H}(x) \geq p,\ p \in (0, 1]\}$ denotes the left-continuous generalized inverse function of a distribution function H, and

$$\hat{\mathsf{F}}_{T,k}(x; \hat{\lambda}_T) := \frac{1}{T}\sum_{t=1}^{T} 1\{\hat{\eta}_{k,t}(\hat{\lambda}_T) \leq x\}, \quad \hat{\mathsf{G}}_{T,S,k}(x; \theta_q, \hat{\nu}_T) := \frac{1}{TS}\sum_{t=1}^{T}\sum_{s=1}^{S} 1\{\hat{X}_{k,t,s}(\theta_q, \hat{\nu}_T) \leq x\},$$

for $x \in \mathbb{R}$, $k \in \{i, j\}$, with $i, j \in \mathcal{G}_q$, $q \in \{1, \ldots, Q\}$. More specifically, taking Fermanian et al. (2004, Theorem 6) and Bücher and Segers (2013, Lemma 7.2) into account, we can express the kth entry of $\hat{\Psi}_{T,S,i,j}(\theta_q, \hat{\lambda}_T, \hat{\nu}_T) := \hat{\psi}_{T,i,j}(\hat{\lambda}_T) - \hat{\psi}_{T,S,i,j}(\theta_q, \hat{\nu}_T)$ as a Lebesgue–Stieltjes integral

$$\sqrt{T}\hat{\Psi}_{T,S,i,j,k}(\theta_q, \hat{\lambda}_T, \hat{\nu}_T) = \int_{[0,1]^2} \hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j; \theta_q, \hat{\lambda}_T, \hat{\nu}_T)\,\mathrm{d}\varphi_k(u_i, u_j) + o_p(1) \tag{3.12}$$

for $k \in \{1, \ldots, \ell\}$. Hence, the limiting distribution of $\sqrt{T}\hat{\Psi}_{T,S}(\theta_0, \hat{\lambda}_T, \hat{\nu}_T)$ can be deduced from the weak convergence of the process $\{\hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j; \theta_{0,q}, \hat{\lambda}_T, \hat{\nu}_T) : u_i, u_j \in [0, 1]\}$, which is readily recognized as the difference between two empirical copula processes. Since filtered data is used, an invariance result with respect to the corresponding statistics based on the unknown counterparts is desirable. If empirical and simulated rank statistics are independent, then it suffices to show that the empirical copula processes based on $\hat{V}_{i,t}(\hat{\lambda}_T)$ and $\hat{U}_{i,t,s}(\theta_{0,q}, \hat{\nu}_T)$ share the same weak limit; a proof strategy employed by Oh and Patton (2013) who argue along the lines of Rémillard (2017). Here, we require the somewhat stronger notion of uniform asymptotic negligibility; i.e., we show, under sufficient regularity of the data, that

$$\sup_{u_i, u_j \in [0,1]} \sup_{\theta_q \in \Theta} |\hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j; \theta_q, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j; \theta_q, \lambda_0, \nu_0)| = o_p(1), \tag{3.13}$$

for each $i, j \in \mathcal{G}_q$, $q \in \{1, \ldots, Q\}$. There exist already similar results in the literature for β-mixing processes (see Neumeyer et al. (2019) and Chen et al. (2020) who rely on Dette et al. (2009) and Akritas and Van Keilegom (2001, Lemma 1)); the underlying stochastic equicontinuity result of Doukhan et al. (1995) is, however, not directly applicable to the triangular-array case considered here. In order to overcome this difficulty, we resort to the FCLT of Andrews and Pollard (1994, Theorem 2.2). The following regularity conditions are assumed to hold:

Assumption C. For each $i \in \{1, \ldots, n\}$, the ith first-order partial derivative $\partial_i \mathsf{C}(u_1, \ldots, u_n; \theta)$ exists and is continuous on the set $\{(u_1, \ldots, u_n)' \in [0, 1]^n : 0 < u_i < 1\}$ uniformly in $\theta \in \Theta$. The same holds for each bivariate copula $\{\mathsf{C}_q(u_i, u_j; \theta_q) : q \in \{1, \ldots, Q\}\}$.

Assumption D.

(D1) (i) For each $i \in \{1, \ldots, n\}$, F_i has density function f_i satisfying $\mathsf{sup}_{x \in \mathbb{R}} \mathsf{f}_i(x) < \infty$, $\mathsf{sup}_{x \in \mathbb{R}} |x \mathsf{f}_i(x)| < \infty$, $\mathsf{f}_i\{\mathsf{F}_i^{-1}(x)\}(1 + \mathsf{F}_i^{-1}(x)) = o(1)$ as $x \to 0$ or $x \to 1$

(ii) For each $1 \leq i < j \leq n$, the bivariate distribution functions $\mathsf{F}_{i,j}(x_i, x_j) := \mathsf{P}(\eta_{i,t} \leq x_i, \eta_{j,t} \leq x_j)$ satisfy $\mathsf{max}_{k,l \in \{i,j\}} \mathsf{sup}_{x_i, x_j \in \mathbb{R}} |\partial_l \partial_k \mathsf{F}_{i,j}(x_i, x_j)(1 + x_k)(1 + x_l)| < \infty$.

(D2) For each $q \in \{1, \ldots, Q\}$, G_q density function g_q satisfying $\mathsf{sup}_{\theta_q \in \Theta} \mathsf{sup}_{x \in \mathbb{R}} |\mathsf{g}_q(x; \theta_q)| < \infty$ and $\mathsf{sup}_{\theta_q \in \Theta} \mathsf{g}_q\{\mathsf{G}_q^{-1}(x; \theta_q); \theta_q\} = o(1)$ as $x \to 0$ or $x \to 1$.

(D3) $\mathsf{D}_\varepsilon(x; \delta)$ and $\mathsf{D}_{\mathsf{F},j}(x; \gamma_j)$, $j \in \{1, \ldots, p_\alpha\}$, are continuous and strictly increasing distribution functions, which are known up to finite dimensional parameter δ and $\gamma := (\gamma_1', \ldots, \gamma_{p_\alpha}')'$, respectively.

(i) $\mathsf{sup}_{\delta \in \Theta} \mathsf{sup}_{x \in \mathbb{R}} \mathsf{d}_\varepsilon(x; \delta) < \infty$, where d_ε denotes the marginal density of D_ε.

(ii) There exists an integrable function $\dot{Q}_F : [0, 1] \mapsto \mathbb{R}_+$ bounding $\mathsf{sup}_{\gamma_j \in \Theta} |\mathsf{D}_{F,j}^{-1}(u, \gamma_j)|$ and $\mathsf{sup}_{\gamma_j \in \Theta} \|\nabla_{\gamma_j} \mathsf{D}_{F,j}^{-1}(u, \gamma_j)\|$ from above for any $j \in \{1, \ldots, p_\alpha\}$ and $u \in [0, 1]$.

Assumption E.

(E1) (i) $\sqrt{T}\|\hat{\lambda}_T - \lambda_0\| = O_p(1)$ and the true parameter λ_0 is element of the compact set $\Lambda_0 \subset \mathbb{R}^r$.

(ii) $\sqrt{T}\|\hat{\nu}_T - \nu_0\| = O_p(1)$ and the true parameter ν_0 is element of the compact set $\mathcal{V}_0 \subset \mathbb{R}^m$.

(E2) (i) Let $\mathcal{U}(\lambda_0) \subset \Lambda_0$ be a neighborhood around λ_0 and set $R_t := R_t(\lambda_0)$. There exists a measurable function $\dot{\mu}(R_t)$ with $\mathsf{E}[\dot{\mu}(R_t)^4] < \infty$ such that *a.s.*

$$\sup_{\lambda \in \mathcal{U}(\lambda_0)} \max\{\|\nabla_\lambda \mu_{j,i}(\mathcal{F}_t, \lambda)/\mu_{2,i}(R_t, \lambda_0)\|, \|\nabla_\lambda^2 \mu_{j,i}(\mathcal{F}_t, \lambda)/\mu_{2,i}(R_t, \lambda_0)\|\} \leq \dot{\mu}(R_t)$$

for each $i \in \{1, \ldots, n\}$, $t \in \{1, \ldots, T\}$, and any $j \in \{1, 2\}$. Moreover, there exists some constant $\underline{b} \in (0, \infty)$ such that $\mathsf{inf}_{\lambda \in \mathcal{U}(\lambda_0)} \mathsf{inf}_{1 \leq t \leq T} \mu_{2,i}(\mathcal{F}_t, \lambda) \geq 1/\underline{b}$ *a.s.* for each $i \in \{1, \ldots, n\}$.

(ii) Let $\mathcal{U}(\nu_0) \subset \nu_0$ be a neighborhood around ν_0 and set $M_t := M_t(\nu_0)$. There exists a measurable function $\dot{\sigma}(M_t)$ with $\mathsf{E}[\dot{\sigma}(M_t)^2] < \infty$ such that *a.s.*

$$\sup_{\nu \in \mathcal{U}(\nu_0)} \max\{\|\nabla_\nu \sigma_j(\mathcal{W}_t, \nu)\|, \|\nabla_\nu^2 \sigma_j(\mathcal{W}_t, \nu)\|\} \leq \dot{\sigma}(M_t),$$

for each $j \in \{1, \ldots, p_\beta\}$, $t \in \{1, \ldots, T\}$.

(E3) (i) The process $\{R_t : t \geq 1\}$ is strictly stationary and α-mixing with mixing number $\alpha_{\mathsf{R}}(j) = O(j^{-c})$, $j \in \mathbb{N}$, where $c > (c_1 - 1)(c_1 + c_2)/c_2$ for $c_1 := \mathsf{min}\{i \in \mathbb{N} : i > 4(1 + r)(2 + c_2)\}$ and some $c_2 > 0$.

(ii) The process $\{M_t : t \geq 1\}$ is strictly stationary and α-mixing with mixing number $\alpha_{\mathsf{M}}(j) = O(j^{-k})$, $j \in \mathbb{N}$, where $k > (k_1 - 1)(k_1 + k_2)/k_2$ for $k_1 := \mathsf{min}\{i \in \mathbb{N} : i > 2(2 + p_\alpha(1 + p_\gamma) + p_\beta(1 + m) + p_\delta)(2 + k_2)\}$ and some $k_2 > 0$.

Assumption F. Define $\mathsf{C}_t(u_1, \ldots, u_n) := \mathsf{P}(V_{1,t} \leq u_1, \ldots, V_{n,t} \leq u_n \mid Z_t)$ and $\mathcal{C}_t(u_1, \ldots, u_n) := \mathsf{P}(U_{1,t,s} \leq u_1, \ldots, U_{n,t,s} \leq u_n \mid Z_t)$. Then, for $w_k = (u_{1,k}, \ldots, u_{n,k})' \in [0, 1]^n$, $k \in \{1, \ldots, m\}$, such that $\mathsf{C}_t(w_1), \ldots, \mathsf{C}_t(w_m)$ are all distinct *a.s.* and $\mathcal{C}_t(w_1), \ldots, \mathcal{C}_t(w_m)$ are all distinct *a.s.* with values in $(0, 1)$ *a.s.* the matrix

$$\Big(\mathsf{E}[\mathsf{C}_t(w_k \wedge w_l) - \mathsf{C}_t(w_k)\mathsf{C}_t(w_l)]\Big)_{1 \leq k,l \leq m}$$

is positive definite and the matrices

$$\Big(\mathsf{E}[\mathcal{C}_t(w_k \wedge w_l) - \mathcal{C}_t(w_k)\mathcal{C}_t(w_l)]\Big)_{1 \leq k,l \leq m}, \quad \Big(\mathsf{E}[(\mathsf{C}_t(w_k) - \mathcal{C}_t(w_k))(\mathsf{C}_t(w_l) - \mathcal{C}_t(w_l))]\Big)_{1 \leq k,l \leq m},$$

are positive semi-definite for any $m \in \mathbb{N}_+$.

Assumption G. $\{\varphi_k : 1 \leq k \leq \ell\}$ are of bounded Hardy–Krause variation; see, e.g., Owen (2005, Definition 2).

The smoothness condition C—due to Segers (2012)—is needed to apply the functional delta method; see also Bücher and Volgushev (2013). Assumptions (D1) and (D2) are similar to regularity conditions imposed by Neumeyer et al. (2019, p. 141); as discussed in Côté et al. (2019) and Omelka et al. (2020), this assumption can be relaxed at the expense of additional technicalities. Assumption (D3) summarizes conditions which, in conjunction with the remaining assumptions, ensure the asymptotic equicontinuity of $\theta_q \mapsto \hat{\mathbb{B}}_{T,S,i,j}(\theta_q)$. When compared to similar conditions used by Manner et al. (2021, Assumption 5), Assumption (D3) is relatively primitive. The assumption is, for example, satisfied if factors and idiosyncratic errors are Gaussian. To give a less trivial example, suppose a scalar factor F_t follows a (standardized) Student's t-distribution with degrees of freedom parameter $2 < \underline{\gamma} \leq \gamma_0 \leq \bar{\gamma} < \infty$. Then, (D3) ($ii$) is satisfied by setting $\dot{Q}_F(u) := \sqrt{\bar{\gamma} \vee 1/(\underline{\gamma}-2)}|\tilde{D}^{-1}(u, \underline{\gamma})|$, where $\tilde{D}^{-1}(u, \underline{\gamma})$ is the inverse of the non-standardized Student's t-distribution so that

$$\frac{\underline{\gamma}-1}{\underline{\gamma}} \int_{[0,1]} \dot{Q}_F(u)\, \mathrm{d}u \leq \sqrt{\frac{2}{\pi}\left(\bar{\gamma} \vee \frac{1}{\underline{\gamma}-2}\right)} < \infty;$$

see the online supplement for details. Assumption E concerns the marginal time-series models: part (E1) is high-level and can be verified for many estimators of AR-GARCH type-models (see Francq and Zakoïan (2004) for a more primitive underpinning); part (E2) is similar to Chen and Fan (2006, Assumption N) and means that the gradient vectors of the location and scale functions are locally dominated. The α-mixing sizes in part (E3) are chosen as to match the conditions of the FCLT in Andrews and Pollard (1994, Theorem 2.2). Assumption E would, for instance, be satisfied by many stationary AR-GARCH processes with geometric mixing rate; see, e.g., Carrasco and Chen (2002), Fryzlewicz and Rao (2011), or Liu and Yang (2016). Assumption F is a regularity condition needed to establish the weak convergence of the 'finite-dimensional distributions' of (3.11) based on an argument borrowed from Boistard et al. (2017). The assumption does not seem overly restrictive as similar results exist for univariate unconditional distributions; see, e.g., Boistard et al. (2017, Lemma 9.5). Bounded variation in the sense of Hardy–Krause, imposed by Assumption G on the functions $\varphi_k : [0, 1]^2 \mapsto \mathbb{R}$, ensures an integration by parts formula for bivariate integrals; see Fermanian et al. (2004) and Radulović et al. (2017). Since the identity map or indicators of axis-parallel boxes are of bounded Hardy–Krause variation (see, e.g., Owen (2005)), dependence measures used here and in Oh and Patton (2013) like Spearman's ρ and quantile dependence can be expressed in terms of Eq. (3.7) using functions that satisfy this assumption. Assumption G implies Riemann-integrability and thus boundedness (see Owen and Rudolf (2020, Lemma 1)), a strong assumption, admittedly, but one which completely suffices here and that could in principle be relaxed as pointed out by Berghaus et al. (2017, Section 3.1).

Proposition 1 distills the main ingredients needed to derive the asymptotic properties of the SMM estimator.

Proposition 1. *Suppose Assumptions A, B, C, D, E, F, and G hold true.*

(a) *For each $\theta \in \Theta$, $\hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T) \xrightarrow{p} \Psi(\theta) := \psi - \psi(\theta)$, where $\psi := (\psi_1', \dots, \psi_Q')'$ and $\psi(\theta) := (\psi_1(\theta_1)', \dots, \psi_Q(\theta_Q)')'$ are $\bar{\ell} \times 1$ vectors; typical elements of the $\ell \times 1$ vectors $\psi_q := (\psi_{q,1}, \dots, \psi_{q,\ell})'$ and $\psi_q(\theta) := (\psi_{q,1}(\theta), \dots, \psi_{q,\ell}(\theta))'$ are given in Eq. (3.8).*

(b) *$\sqrt{T}\hat{\Psi}_{T,S}(\theta_0, \hat{\lambda}_T, \hat{\nu}_T) \xrightarrow{d} \mathcal{N}(0_{\bar{\ell}}, \Sigma_0)$, where Σ_0 is a positive-definite variance–covariance matrix.*

(c) *For any $\epsilon, \eta > 0$, there exists some $\delta > 0$ such that*

$$\overline{\lim_{T\to\infty}} \mathsf{P}\left[\sup_{\theta,\tilde{\theta} \in \Theta: \|\theta-\tilde{\theta}\| \leq \delta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \hat{\nu}_T) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \hat{\nu}_T) + \psi(\tilde{\theta})\| > \eta\right] < \epsilon.$$

Remark 1. It is instructive to take a look at the limiting distribution of part (b). As shown in Appendix A.1, $\Sigma_0 := (\Sigma_0(g, q))_{1\leq g,q\leq Q}$ is a block-symmetric matrix whose (g, q)th block is given by the $\ell \times \ell$ matrix $\Sigma_0(g, q) := (\sigma_0(g, q \mid k, l))_{1\leq k,l\leq \ell}$ with typical element

$$\sigma_0(g, q \mid k, l) := \int_{[0,1]^2}\int_{[0,1]^2} \mathsf{E}[\mathbb{C}_g(u_1, v_1)\mathbb{C}_q(u_2, v_2)]\mathrm{d}\varphi_k(u_1, v_1)\mathrm{d}\varphi_l(u_2, v_2),$$

with

$$\mathbb{C}_q(u, v) := \mathbb{B}_q(u, v) - \partial_u \mathsf{C}_q(u, v)\mathbb{B}_q(u, 1) - \partial_v \mathsf{C}_q(u, v)\mathbb{B}_q(1, v),$$

where $\mathbb{B}_q$ is a mean-zero Gaussian process concentrated on $\mathbb{D}_0 := \{\alpha \in C[0, 1]^2 : \alpha(1, 1) = \alpha(u, 0) = \alpha(0, u) = 0,\ u \in (0, 1)\}$ such that

$$\begin{aligned}\mathrm{cov}[\mathbb{B}_q(u_1, v_1), \mathbb{B}_q(u_2, v_2)] =\ & \mathsf{E}[\mathsf{C}_{q,t}(u_1 \wedge u_2, v_1 \wedge v_2) - \mathsf{C}_{q,t}(u_1, v_1)\mathsf{C}_{q,t}(u_2, v_2)] \\ & + \mathsf{E}[(\mathsf{C}_{q,t}(u_1, v_1) - \mathcal{C}_{q,t}(u_1, v_1))(\mathsf{C}_{q,t}(u_2, v_2) - \mathcal{C}_{q,t}(u_2, v_2))],\end{aligned}$$

where $\mathsf{C}_{q,t}(u, v) = \mathsf{P}(V_{i,t} \leq u, V_{j,t} \leq v \mid Z_t)$ and $\mathcal{C}_{q,t}(u, v) = \mathsf{P}(U_{i,t,s} \leq u, U_{j,t,s} \leq v \mid Z_t)$, $i, j \in \mathcal{G}_q$, are the bivariate counterparts of C_t and $\mathcal{C}_t$ defined in Assumption F. Thus, the limiting distribution of $\sqrt{T}\hat{\Psi}_{T,S}(\theta_0, \hat{\lambda}_T, \hat{\nu}_T)$ is unaffected by the first-step estimation error, a finding which, in view of Eq. (3.13), was to be expected. A closer inspection of the limiting variance–covariance matrix reveals that the limiting distribution depends on the covariate Z_t and the partial derivatives of the copula through the covariance kernel of the Gaussian processes $\mathbb{C}_q$, $q \in \{1, \dots, Q\}$.

The stochastic equicontinuity result of Proposition 1 provides the link between the pointwise properties of $\theta \mapsto \hat{\Psi}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T)$ and the asymptotic behavior of $\hat{\theta}_{T,S}$. To make the argument precise, an additional set of regularity conditions is imposed.

Assumption H.

(H1) $\hat{L}_{T,S} = L_0 + o_p(1)$, L_0 is a deterministic $\bar{\ell} \times \bar{\ell}$ positive definite matrix;
(H2) $\Psi(\theta) \neq 0_{\bar{\ell}}$ for $\theta \neq \theta_0$;
(H3) θ_0 is an interior point of the compact set $\Theta \subset \mathbb{R}^p$, with $p \leq \bar{\ell} < \infty$
(H4) $\psi(\theta)$ is differentiable at θ_0 with $\bar{\ell} \times p$ dimensional Jacobian matrix $\nabla_\theta \psi(\theta)$ such that $\dot{\psi}_0' L_0 \dot{\psi}_0$ is nonsingular for $\dot{\psi}_0 := \nabla_\theta \psi(\theta_0)$
(H5) $\hat{A}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) \leq \inf_{\theta \in \Theta} \hat{A}_{T,S}(\theta, \hat{\lambda}_T, \hat{\nu}_T) + o_p(1/T)$.

Assumption H is common for extremum estimators with non-smooth objective function; see, e.g., Newey and McFadden (1994, Section 7). Analogously to Oh and Patton (2013, Proposition 2), we make use of Newey and McFadden (1994, Theorem 7.2) to derive the asymptotic normality of the SMM estimator.

Proposition 2. *Suppose Assumptions A, B, C, D, E, F, G, and H hold. Then,*

$$\sqrt{T}(\hat{\theta}_{T,S} - \theta_0) \xrightarrow{d} \mathcal{N}(0_{\bar{\ell}}, \Omega_0), \;\; \Omega_0 := (\dot{\psi}_0' L_0 \dot{\psi}_0)^{-1} \dot{\psi}_0' L_0 \Sigma_0 L_0 \dot{\psi}_0 (\dot{\psi}_0' L_0 \dot{\psi}_0)^{-1},$$

where Σ_0 is the variance–covariance matrix given in Proposition 1.

3.4. Standard errors and inference

We follow Oh and Patton (2013) by resorting to numerical derivatives and the bootstrap to estimate $\dot{\psi}_0$ and the limiting variance–covariance matrix Σ_0, respectively. The former estimator is almost completely analogous to the one used by Oh and Patton (2013, p. 692); i.e., for a step size $\pi_T \to 0^+$ define $\hat{\dot{\psi}}_{T,S}$, whose kth column is given by

$$\hat{\dot{\psi}}_{T,S,k} := \frac{\hat{\psi}_{T,S}(\hat{\theta}_{T,S} + e_k \pi_T, \hat{\nu}_T) - \hat{\psi}_{T,S}(\hat{\theta}_{T,S} - e_k \pi_T, \hat{\nu}_T)}{2\pi_T}, \quad k \in \{1, \dots, p\},$$

where e_k denotes the p-dimensional vector of zeros with one at kth position.

Contrary to the aforementioned authors, however, the estimator of Σ_0 needs to account for the dependence structure induced by the exogenous regressor. Therefore, we propose standard errors based on bootstrap replications of both empirical *and* simulated statistics. More specifically, draw for each $i, j \in \mathcal{G}_q$, $q \in \{1, \dots, Q\}$, with replacement B bootstrap samples $\left\{\mathcal{Z}^{(b)}_{S,t,i,j}(\hat{\theta}_{T,S,q}, \hat{\lambda}_T, \hat{\nu}_T)\right\}_{t=1}^T$, $b \in \{1, \dots, B\}$, from $\left\{\mathcal{Z}_{S,t,i,j}(\hat{\theta}_{T,S,q}, \hat{\lambda}_T, \hat{\nu}_T)\right\}_{t=1}^T$, where

$$\begin{aligned}\mathcal{Z}_{S,t,i,j}(\hat{\theta}_{T,S,q}, \hat{\lambda}_T, \hat{\nu}_T) := (\hat{\eta}_{i,t}(\hat{\lambda}_T), \hat{\eta}_{j,t}(\hat{\lambda}_T), \hat{X}_{i,t,1}(\hat{\theta}_{T,S,q}, \hat{\nu}_T), \hat{X}_{j,t,1}(\hat{\theta}_{T,S,q}, \hat{\nu}_T), \dots \\ \dots, \hat{X}_{i,t,S}(\hat{\theta}_{T,S,q}, \hat{\nu}_T), \hat{X}_{j,t,S}(\hat{\theta}_{T,S,q}, \hat{\nu}_T))',\end{aligned}$$

with $\hat{\theta}_{T,S,q}$ representing the $(p_\alpha + p_\beta + p_\alpha p_\gamma + p_\delta) \times 1$ sub-vector of $\hat{\theta}_{T,S}$ that contains the SMM estimates pertaining to group $\mathcal{G}_q$. Next, denote the corresponding ranks by $\hat{V}^{(b)}_{k,t}(\hat{\lambda}_T)$, $\hat{U}^{(b)}_{k,t,s}(\hat{\theta}_{T,S}, \hat{\nu}_T)$, $k \in \{i, j\}$. In view of Eq. (3.7), introduce the bootstrap rank-based dependence measures

$$\begin{aligned}\hat{\psi}^{(b)}_{T,i,j,k}(\hat{\lambda}_T) &:= \frac{1}{T}\sum_{t=1}^{T} \varphi_k(\hat{V}^{(b)}_{i,t}(\hat{\lambda}_T), \hat{V}^{(b)}_{j,t}(\hat{\lambda}_T)), \\ \hat{\psi}^{(b)}_{T,S,i,j,k}(\hat{\theta}_{T,S,q}, \hat{\nu}_T) &:= \frac{1}{ST}\sum_{s=1}^{S}\sum_{t=1}^{T} \varphi_k(\hat{U}^{(b)}_{i,t,s}(\hat{\theta}_{T,S,q}, \hat{\nu}_T), \hat{U}^{(b)}_{j,t,s}(\hat{\theta}_{T,S,q}, \hat{\nu}_T)).\end{aligned} \tag{3.14}$$

Akin to the discussion surrounding (3.9), let $\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T)$ denote the $\bar{\ell} \times 1$ vector of group-averages of $\hat{\Psi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q}, \lambda, \lambda, \hat{\nu}_T) := \hat{\psi}^{(b)}_{T,i,j}(\hat{\nu}_T) - \hat{\psi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q}, \hat{\lambda}_T)$. Following the arguments made by Fermanian et al. (2004, Theorem 5), we can then show that the conditional distribution of $\sqrt{T}(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\Psi}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T))$ consistently estimates the limiting distribution of $\sqrt{T}\hat{\Psi}_{T,S}(\theta_0, \hat{\lambda}_T, \hat{\nu}_T)$. Importantly, since the first-step estimation of the location-scale parameters does not contaminate the limiting distribution of $\sqrt{T}\hat{\Psi}_{T,S}(\theta_0, \hat{\lambda}_T, \hat{\nu}_T)$, there is no need to adjust for this source of uncertainty as is, for example, done in Gonçalves et al. (forthcoming). Hence, the limiting variance–covariance $\Sigma_{0,S}$ is consistently estimable by the bootstrap second moment[2] $\mathrm{cov}^*[\sqrt{T}\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T)]$, provided $\sqrt{T}(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\Psi}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T))$ is uniformly square integrable; see, e.g., Brown and Wegkamp (2002, Theorem 7) or Cheng (2015, Lemma 1). Since an estimator of $\mathrm{cov}^*[\sqrt{T}\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T)]$ is given by

$$\hat{\Sigma}_{T,S,B} := \frac{T}{B}\sum_{b=1}^{B}(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\Psi}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T))(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\Psi}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T))',$$

2 Throughout, '$*$' indicates that the given probability/moment has been computed under the bootstrap distribution conditional on the original sample.

Table 4.1
Factor structure.

	design 1				design 2			
	Loadings		Estimable factors		Loadings		Estimable factors	
$i \in \mathcal{G}_1$	$\alpha_1 = 1$	$\beta = 1/2$			$\alpha = 2$	$\beta_1 = 1/2$	$Z_{i,t} = \tilde{Z}_t$ $\tilde{W}_t = 0.65\tilde{W}_{t-1} + \tilde{Z}_t$ $\tilde{Z}_t \sim \mathcal{N}(0, 1)$	AR(1)
$i \in \mathcal{G}_2$	$\alpha_2 = 3/2$	$\beta = 1/2$	$\forall i,\ Z_{i,t} = Z_t$ $W_t = 0.65W_{t-1} + Z_t$ $Z_t \sim \mathcal{N}(0, 1)$	AR(1)	$\alpha = 2$	$\beta_2 = 1$	$Z_{i,t} = \log\lvert Z_t^*\rvert$ $W_t^* = \sigma_t Z_t^*$ $\sigma_t^2 = 0.1 + 0.1\sigma_{t-1}^2 + 0.5W_{t-1}^*$ $Z_t^* \sim \mathcal{N}(0, 1)$	GARCH(1,1)
$i \in \mathcal{G}_3$	$\alpha_3 = 2$	$\beta = 1/2$			$\alpha = 2$	$\beta_3 = 3/2$	$Z_{i,t} = Z_t$ $Z_t \sim \mathcal{N}(0, 1)$	WN

we can introduce the following consistent estimator of $\Omega_{0,S}$

$$\hat{\Omega}_{T,S,B} := (\hat{\dot{\psi}}'_{T,S}\hat{L}_{T,S}\hat{\dot{\psi}}_{T,S})^{-1}\hat{\dot{\psi}}'_{T,S}\hat{L}_{T,S}\hat{\Sigma}_{T,S,B}\hat{L}_{T,S}\hat{\dot{\psi}}_{T,S}(\hat{\dot{\psi}}'_{T,S}\hat{L}_{T,S}\hat{\dot{\psi}}_{T,S})^{-1}. \tag{3.15}$$

Corollary 1. *Suppose that* $\mathsf{E}^*[\|\sqrt{T}(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) - \hat{\Psi}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T))\|^{2+\delta}] < \infty$ *a.s. for some* $\delta > 0$ *and* $\sqrt{T}\pi_T \to \infty$. *Then,* $\hat{\Omega}_{T,S,B} \xrightarrow{p} \Omega_0$, *as* $B, T \to \infty$.

Corollary 1 allows to conduct inference about θ_0 and to obtain the two-step SMM estimator with optimal weight matrix $\hat{L}_{T,S} = \hat{\Sigma}^{-1}_{T,S,B}$. More primitive conditions under which the uniform square integrability holds are not readily available. However, as discussed by Hahn and Liao (2021), if this assumption fails, bootstrap standard errors based on $\hat{\Sigma}_{T,S,B}$ are likely to yield conservative tests. Provided $\bar{\ell} > p$, the preceding result can also be used to ascertain overidentifying restrictions based on the Sargan–Hansen type J-statistic

$$J_{T,S} := T\hat{A}_{T,S}(\hat{\theta}_{T,S}, \hat{\lambda}_T, \hat{\nu}_T) \xrightarrow{d} u'\mathcal{A}'_0\mathcal{A}'_0 u, \quad u \sim \mathcal{N}(0_{\bar{\ell}}, I_{\bar{\ell}\times\bar{\ell}}), \tag{3.16}$$

where $\mathcal{A}_0 := L_0^{1/2}\Sigma_0^{1/2}R_0$, with $R_0 := I_{\bar{\ell}\times\bar{\ell}} - \Sigma_0^{-1/2}\dot{\psi}_0(\dot{\psi}'_0 L_0\dot{\psi}_0)^{-1}\dot{\psi}'_0 L_0\Sigma_0^{1/2}$. Critical values for $J_{T,S}$ need to be simulated (using estimators of Σ_0 and $\dot{\psi}_0$) unless the optimal weight matrix is used, in which case the common result $J_{T,S} \xrightarrow{d} \chi^2(\bar{\ell} - p)$ obtains; see also Oh and Patton (2013, Proposition 4).

4. Monte Carlo experiment

The Monte Carlo experiment uses a data generating process similar to that in Oh and Patton (2013, p. 695); that is, we consider an AR(1)-GARCH(1,1) process to describe the evolution of each of n assets over time:

$$Y_{i,t} = 0.01 + 0.05Y_{i,t-1} + \sigma_{i,t}\eta_{i,t}, \quad \sigma^2_{i,t} = 0.05 + 0.85\sigma^2_{i,t-1} + 0.1\sigma^2_{i,t-1}\eta^2_{i,t-1}, \tag{4.1}$$

where $\eta_t := (\eta_{1,t}, \dots, \eta_{n,t}) \sim \mathsf{C}(\Phi_1, \dots, \Phi_n)$, with Φ_i, $i \in \{1, \dots, n\}$, denoting the marginal (Gaussian) distribution function of $\eta_{i,t}$. The copula C is generated by the following 'block-equidependent' factor model

$$X_{i,t} = \alpha_{0,j}F_t + \beta_{0,j}Z_{j,t} + \varepsilon_{i,t}, \quad i \in \mathcal{G}_q,\ q \in \{1, 2, 3\}. \tag{4.2}$$

We consider three groups $\mathcal{G}_1, \mathcal{G}_2$, and $\mathcal{G}_3$, of equal size partitioning the cross-sectional index set; i.e., $\{1, 2, \dots, n\} = \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, with $|\mathcal{G}_q| = n/3$, $q \in \{1, 2, 3\}$ for $n \in \{15, 30\}$. The factor F_t and the idiosyncratic component $\varepsilon_{i,t}$ are latent but simulable. It is assumed that $F_t \overset{\text{IID}}{\sim} t(\zeta_0, \xi_0)$, with $\zeta_0 = 1/4$, $\xi_0 = -1/2$, where $t(\zeta_0, \xi_0)$ denotes Hansen's standardized skewed t-distribution with tail thickness parameter $2 < 1/\zeta < \infty$ and skewness parameter $-1 < \xi < 1$. For the idiosyncratic component, we either consider $\varepsilon_{i,t} \overset{\text{IID}}{\sim} t(\zeta_0, 0)$ (called the skew-t/t specification) or $\varepsilon_{i,t} \overset{\text{IID}}{\sim} \mathcal{N}(0, 1)$ (called the skew-t/normal specification). Turning to the loadings and the estimable factors, we consider two 'block-equidependent' specifications similar to the multilevel model of Bai and Wang (2015, Proposition 3); see the summary of Table 4.1 .

Under design 1, loadings on F_t are group specific while the loading on the single estimable factor is common; the latter factor has to be estimated from an AR(1) model. Under design 2, the loading on F_t is common while three estimable factors have group-specific loadings; the estimable factors have to be estimated from an AR(1) model, from a GARCH(1,1) model, or follow observable white noise. Hence, design 1 and design 2 imply 6 unknown copula parameters $\theta_0 = (\alpha_{0,1}, \alpha_{0,2}, \alpha_{0,3}, \beta_0, \zeta_0, \xi_0)'$ and $\theta_0 = (\alpha_0, \beta_{0,1}, \beta_{0,2}, \beta_{0,3}, \zeta_0, \xi_0)'$, respectively. Estimation of θ_0 is in both cases based on Spearman's rank correlation

$$\hat{\varphi}_{T,i,j,1} := \frac{12}{T}\sum_{t=1}^{T}\hat{V}_{i,t}(\hat{\lambda}_T)\hat{V}_{j,t}(\hat{\lambda}_T) - 3, \ \ \hat{\varphi}_{T,S,i,j,1} := \frac{12}{ST}\sum_{t=1}^{T}\sum_{s=1}^{S}\hat{U}_{i,t,s}(\theta_q, \hat{\nu}_T)\hat{U}_{j,t,s}(\theta_q, \hat{\nu}_T) - 3$$

Table 4.2
Simulation results (design 1: skew-t/normal).

n	T		Feasible						Unfeasible					
			ζ_0	ξ_0	β_0	$\alpha_{0,1}$	$\alpha_{0,2}$	$\alpha_{0,3}$	ζ_0	ξ_0	β_0	$\alpha_{0,1}$	$\alpha_{0,2}$	$\alpha_{0,3}$
			0.25	−0.5	0.5	1.0	1.5	2.0	0.25	−0.5	0.5	1.0	1.5	2.0
15	500	mean	0.275	−0.535	0.574	0.994	1.547	2.083	0.276	−0.531	0.572	0.996	1.548	2.087
		median	0.287	−0.515	0.586	0.971	1.507	2.020	0.282	−0.520	0.582	0.975	1.510	2.021
		var	0.008	0.017	0.028	0.035	0.073	0.133	0.008	0.016	0.030	0.036	0.066	0.133
		rmse	0.092	0.136	0.184	0.187	0.275	0.374	0.092	0.129	0.189	0.190	0.261	0.375
		t	7.60	1.20	14.80	4.20	6.00	7.40	7.60	1.80	16.00	5.80	7.00	8.00
		J			4.20						4.60			
	1000	mean	0.264	−0.524	0.536	1.006	1.533	2.059	0.262	−0.524	0.536	1.006	1.529	2.053
		median	0.257	−0.517	0.550	0.993	1.512	2.010	0.254	−0.518	0.544	0.995	1.504	2.007
		var	0.005	0.007	0.017	0.016	0.033	0.069	0.005	0.007	0.015	0.015	0.031	0.064
		rmse	0.071	0.087	0.135	0.125	0.186	0.269	0.069	0.087	0.129	0.123	0.180	0.258
		t	6.20	1.20	12.00	2.60	4.60	6.00	4.80	1.40	10.20	3.60	5.60	6.00
		J			3.60						4.20			
	2000	mean	0.254	−0.510	0.514	1.000	1.511	2.020	0.253	−0.509	0.510	1.002	1.514	2.023
		median	0.247	−0.508	0.523	0.992	1.498	1.997	0.245	−0.507	0.521	0.994	1.499	1.993
		var	0.002	0.003	0.009	0.006	0.011	0.023	0.002	0.003	0.009	0.006	0.014	0.034
		rmse	0.048	0.056	0.097	0.076	0.108	0.153	0.047	0.054	0.096	0.079	0.120	0.185
		t	6.20	2.40	8.20	3.20	4.80	3.60	5.80	3.40	6.80	3.20	4.80	4.40
		J			4.80						2.60			
30	500	mean	0.240	−0.510	0.505	0.998	1.498	1.999	0.240	−0.509	0.503	1.002	1.506	2.009
		median	0.240	−0.497	0.512	0.982	1.467	1.956	0.240	−0.495	0.518	0.985	1.470	1.959
		var	0.005	0.011	0.017	0.019	0.037	0.073	0.005	0.011	0.018	0.024	0.054	0.109
		rmse	0.068	0.106	0.131	0.136	0.192	0.269	0.069	0.103	0.133	0.154	0.232	0.330
		t	4.20	1.40	3.80	3.40	7.80	9.00	4.20	1.20	4.20	4.40	8.40	9.40
		J			2.00						2.00			
	1000	mean	0.243	−0.508	0.488	1.006	1.505	2.001	0.243	−0.508	0.488	1.006	1.506	2.001
		median	0.240	−0.499	0.498	0.995	1.494	1.982	0.241	−0.498	0.499	0.994	1.493	1.980
		var	0.002	0.005	0.010	0.008	0.016	0.028	0.002	0.005	0.009	0.008	0.016	0.028
		rmse	0.050	0.069	0.100	0.092	0.126	0.167	0.050	0.068	0.097	0.091	0.125	0.168
		t	2.60	3.40	3.00	4.40	5.80	6.60	2.40	3.20	3.80	3.60	6.00	6.80
		J			2.40						1.80			
	2000	mean	0.248	−0.506	0.497	1.003	1.506	2.007	0.248	−0.504	0.496	1.003	1.506	2.006
		median	0.247	−0.501	0.501	1.000	1.497	1.984	0.245	−0.501	0.501	0.998	1.498	1.983
		var	0.002	0.003	0.006	0.005	0.010	0.019	0.001	0.002	0.005	0.004	0.008	0.015
		rmse	0.039	0.051	0.077	0.067	0.099	0.137	0.039	0.046	0.074	0.065	0.091	0.124
		t	4.20	4.60	4.40	3.40	4.80	5.40	4.20	4.60	4.60	2.80	5.00	5.00
		J			3.40						2.80			

and quantile dependence

$$\hat{\varphi}^{(\tau)}_{T,i,j,2} := \begin{cases} \dfrac{1}{T\tau}\displaystyle\sum_{t=1}^{T} 1\{\hat{V}_{i,t}(\hat{\lambda}_T) \le \tau, \hat{V}_{j,t}(\hat{\lambda}_T) \le \tau\} & \text{if } \tau \in (0, 1/2], \\ \dfrac{1}{T(1-\tau)}\displaystyle\sum_{t=1}^{T} 1\{\hat{V}_{i,t}(\hat{\lambda}_T) > \tau, \hat{V}_{j,t}(\hat{\lambda}_T) > \tau\} & \text{if } \tau \in (1/2, 1) \end{cases}$$
$$\hat{\varphi}^{(\tau)}_{T,S,i,j,2} := \begin{cases} \dfrac{1}{TS\tau}\displaystyle\sum_{t=1}^{T}\sum_{s=1}^{S} 1\{\hat{U}_{i,t,s}(\theta_q, \hat{\nu}_T) \le \tau, \hat{U}_{j,t,s}(\theta_q, \hat{\nu}_T) \le \tau\} & \text{if } \tau \in (0, 1/2], \\ \dfrac{1}{TS(1-\tau)}\displaystyle\sum_{t=1}^{T}\sum_{s=1}^{S} 1\{\hat{U}_{i,t,s}(\theta_q, \hat{\nu}_T) > \tau, \hat{U}_{j,t,s}(\theta_q, \hat{\nu}_T) > \tau\} & \text{if } \tau \in (1/2, 1). \end{cases} \tag{4.3}$$

Throughout, we set $T \in \{500, 1000, 2000\}$, $S = 25$, and use the identity weight matrix $L_{T,S} = I_{\bar{\ell}}$. We make use of quantile dependence for $\tau \in \{0.05, 0.10, 0.90, 0.95\}$ alongside Spearman's rank correlation, which yields $\bar{\ell} = 3 \times 5 = 15$ rank-based dependence measures for estimation. Numerical optimization employs a derivative-free simplex search based on MATLAB's `fminsearchbnd` routine; see D'Errico (2021). The starting values are obtained from a first-step surrogate minimization using MATLAB's `surrogateopt` optimization for time-consuming objective functions and the individual time-series models are estimated using maximum likelihood.

Tables 4.2, 4.3, 4.4, and 4.5 contain Monte Carlo estimates of mean, median, and variance of the SMM estimator using 500 Monte Carlo iterations.[3] Moreover, we report rejection frequencies of two-sided t-tests under the null $\theta = \theta_0$ and rejection frequencies of the test of overidentifying restrictions (3.16). Both hypothesis tests are investigated at a nominal significance level of five percent and the test statistics are equipped with the bootstrap standard error (3.15). We use $B = 500$ bootstrap replications and set the tuning parameter for the numerical derivative to $\pi_T = 0.05$; moreover, we use 1000 random draws to obtain critical values for the test of overidentifying restrictions (3.16). We report results for the

[3] The computations were implemented in Matlab, parallelized and performed using CHEOPS, the DFG funded (Funding number: INST 216/512/1FUGG) High Performance Computing (HPC) system of the Regional Computing Center at the University of Cologne (RRZK).

Table 4.3
Simulation results (design 1: skew-t/t).

n	T		Feasible						Unfeasible					
			ζ_0	ξ_0	β_0	$\alpha_{0,1}$	$\alpha_{0,2}$	$\alpha_{0,3}$	ζ_0	ξ_0	β_0	$\alpha_{0,1}$	$\alpha_{0,2}$	$\alpha_{0,3}$
			0.25	−0.5	0.5	1.0	1.5	2.0	0.25	−0.5	0.5	1.0	1.5	2.0
15	500	mean	0.264	−0.514	0.488	0.969	1.471	1.971	0.267	−0.512	0.486	0.966	1.47	1.967
		median	0.268	−0.500	0.492	0.974	1.466	1.971	0.273	−0.497	0.497	0.963	1.469	1.963
		var	0.009	0.015	0.023	0.021	0.024	0.038	0.009	0.015	0.022	0.021	0.025	0.039
		rmse	0.098	0.124	0.151	0.147	0.158	0.197	0.098	0.123	0.149	0.150	0.161	0.199
		t	6.80	2.40	4.80	5.00	6.60	7.00	5.80	1.80	3.20	5.20	6.60	7.40
		J			1.20						1.00			
	1000	mean	0.27	−0.515	0.486	0.979	1.487	1.994	0.269	−0.513	0.485	0.980	1.487	1.993
		median	0.268	−0.505	0.496	0.982	1.484	1.990	0.266	−0.504	0.493	0.989	1.481	1.990
		var	0.008	0.008	0.013	0.010	0.012	0.021	0.008	0.006	0.013	0.010	0.012	0.020
		rmse	0.090	0.089	0.116	0.101	0.108	0.144	0.089	0.080	0.113	0.102	0.108	0.142
		t	6.80	2.40	3.60	6.40	5.00	5.60	6.80	3.20	3.40	5.40	5.20	5.20
		J			3.20						2.60			
	2000	mean	0.272	−0.504	0.485	0.983	1.493	1.998	0.267	−0.505	0.486	0.986	1.495	1.998
		median	0.262	−0.500	0.496	0.984	1.491	1.994	0.260	−0.502	0.492	0.993	1.495	1.994
		var	0.006	0.003	0.009	0.006	0.007	0.011	0.005	0.003	0.008	0.005	0.006	0.010
		rmse	0.079	0.053	0.094	0.078	0.082	0.105	0.075	0.051	0.092	0.075	0.08	0.102
		t	7.40	2.00	5.40	8.20	6.60	5.20	6.20	2.60	4.40	6.40	5.80	5.00
		J			5.80						3.60			
30	500	mean	0.235	−0.507	0.487	0.993	1.484	1.971	0.235	−0.508	0.485	0.993	1.484	1.971
		median	0.242	−0.491	0.502	0.997	1.480	1.958	0.240	−0.494	0.492	0.991	1.481	1.956
		var	0.008	0.010	0.015	0.012	0.017	0.029	0.008	0.011	0.015	0.012	0.017	0.029
		rmse	0.088	0.102	0.122	0.109	0.131	0.174	0.088	0.105	0.123	0.110	0.132	0.173
		t	4.20	2.00	3.00	4.60	5.40	10.40	4.40	2.20	2.80	4.40	5.20	10.00
		J			1.40						1.40			
	1000	mean	0.241	−0.507	0.481	1.004	1.499	1.991	0.241	−0.508	0.480	1.004	1.500	1.991
		median	0.240	−0.499	0.490	1.005	1.496	1.979	0.241	−0.500	0.485	1.004	1.497	1.978
		var	0.005	0.004	0.008	0.005	0.008	0.014	0.005	0.005	0.008	0.006	0.008	0.015
		rmse	0.069	0.066	0.090	0.074	0.090	0.119	0.070	0.068	0.094	0.076	0.091	0.121
		t	4.20	3.80	2.40	4.40	4.40	4.60	4.40	3.20	1.80	5.00	5.20	5.00
		J			2.00						2.00			
	2000	mean	0.247	−0.504	0.490	1.001	1.501	1.999	0.247	−0.505	0.490	1.001	1.501	1.999
		median	0.246	−0.500	0.492	1.003	1.500	1.991	0.243	−0.500	0.491	1.003	1.498	1.990
		var	0.003	0.002	0.004	0.003	0.004	0.008	0.003	0.003	0.004	0.003	0.004	0.008
		rmse	0.055	0.046	0.064	0.052	0.064	0.087	0.055	0.051	0.066	0.052	0.064	0.087
		t	4.60	3.80	2.60	2.80	5.20	4.60	3.80	4.40	3.20	3.40	6.40	4.60
		J			3.80						3.60			

feasible and the unfeasible SMM estimator that differ with respect to whether the parameters governing the estimable factors reported in Table 4.1 are estimated (feasible) or treated as known constants (unfeasible). The simulation evidence reveals that, in accordance with the theory, the estimation accuracy increases with T. Although some size distortions can be observed for $T = 500$, rejection frequencies are close to the nominal significance level when $T \geq 1000$. Moreover, the feasible estimator performs almost equally well as its unfeasible counterpart.

5. Empirical application

We apply the above to study the cross-sectional dependence between $n = 43$ companies matching the first four largest groups of the S&P 100 found by Oh and Patton (2021, Table 4). Specifically, we consider daily close prices, adjusted for stock splits and dividends from January 2014 to January 2020 resulting in $T = 1461$ trading days.

Since gold often acts as an hedge and/or a safe haven for stock markets, its price may convey information about the inter-dependencies between stock returns; see, e.g., Baur and McDermott (2010). Hence, we examine the extent to which information on gold prices can help us to describe the cross-sectional dependence structure among the 43 companies. The conditional mean of the (percentage) logarithmic return of the ith stock price $Y_{i,t}$, $i \in \{1, \ldots, 43\}$, is modeled as an AR(1) process augmented with the first lag of the (percentage) logarithmic change of the three *p.m.* gold fixing price in London bullion market W_t

$$Y_{i,t} = \lambda_{1,i} + \lambda_{2,i} Y_{i,t-1} + \lambda_{3,i} W_{t-1} + \epsilon_{i,t}, \quad \epsilon_{i,t} := \mu_{i,t}\eta_{i,t}, \tag{5.1}$$

while, similar to Oh and Patton (2013, 2017, 2021), the conditional variance $\mu_{i,t}^2$ is assumed to follow a GJR-GARCH(1,1) model

$$\mu_{i,t}^2 = \lambda_{4,i} + \lambda_{5,i}\mu_{i,t-1}^2 + \lambda_{6,i}\epsilon_{i,t-1}^2 + \lambda_{7,i}\epsilon_{i,t-1}^2 1\{\epsilon_{i,t-1} < 0\}. \tag{5.2}$$

As we clearly fail to reject the null hypothesis[4] of a zero conditional mean based on an AR(1) specification with unrestricted constant, a GJR-GARCH(1,1) model is considered for the gold price

$$W_t = \sigma_t Z_t, \quad \sigma_t^2 = \nu_1 + \nu_2\sigma_{t-1}^2 + \nu_3 W_{t-1}^2 1\{W_{t-1} < 0\}. \tag{5.3}$$

[4] The p-value of a Wald test with Newey–West standard errors is about 0.77.

Table 4.4
Simulation results (design 2: skew-t/normal).

n	T		Feasible						Unfeasible					
			ζ_0	ξ_0	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	α_0	ζ_0	ξ_0	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	α_0
			0.25	−0.5	0.5	1.0	1.5	2.0	0.25	−0.5	0.5	1.0	1.5	2.0
15	500	mean	0.258	−0.544	0.665	1.162	1.567	1.978	0.255	−0.541	0.656	1.156	1.564	1.966
		median	0.263	−0.515	0.696	1.159	1.570	1.906	0.267	−0.514	0.682	1.150	1.555	1.905
		var	0.008	0.019	0.067	0.063	0.048	0.170	0.008	0.018	0.072	0.061	0.047	0.155
		rmse	0.090	0.144	0.306	0.298	0.230	0.413	0.089	0.141	0.311	0.292	0.225	0.395
		t	8.40	1.80	14.60	9.40	4.80	13.00	8.60	1.60	12.80	7.80	3.80	11.00
		J			1.60						1.40			
	1000	mean	0.257	−0.520	0.598	1.102	1.541	1.984	0.258	−0.520	0.602	1.107	1.545	1.977
		median	0.256	−0.509	0.632	1.106	1.547	1.946	0.258	−0.510	0.638	1.113	1.547	1.955
		var	0.005	0.006	0.052	0.038	0.029	0.066	0.005	0.007	0.054	0.037	0.029	0.048
		rmse	0.072	0.083	0.248	0.219	0.174	0.257	0.071	0.085	0.253	0.221	0.176	0.221
		t	7.40	0.80	10.80	8.60	4.60	7.00	6.60	0.80	13.20	8.00	5.40	6.00
		J			2.20						2.80			
	2000	mean	0.256	−0.506	0.547	1.057	1.521	1.996	0.256	−0.508	0.548	1.059	1.522	1.997
		median	0.250	−0.504	0.567	1.047	1.526	1.975	0.249	−0.503	0.563	1.058	1.527	1.972
		var	0.003	0.003	0.042	0.026	0.016	0.029	0.003	0.003	0.043	0.025	0.016	0.035
		rmse	0.054	0.053	0.209	0.172	0.127	0.171	0.054	0.057	0.212	0.169	0.128	0.186
		t	6.00	2.00	13.60	10.20	5.00	6.00	6.40	3.00	12.40	9.80	4.00	6.00
		J			3.20						3.80			
30	500	mean	0.243	−0.523	0.587	1.083	1.530	1.970	0.241	−0.527	0.589	1.085	1.531	1.971
		median	0.240	−0.507	0.622	1.079	1.530	1.906	0.238	−0.508	0.611	1.083	1.534	1.905
		var	0.008	0.014	0.060	0.049	0.036	0.103	0.008	0.014	0.058	0.047	0.036	0.106
		rmse	0.088	0.119	0.260	0.235	0.193	0.322	0.088	0.121	0.258	0.232	0.192	0.326
		t	9.20	1.80	10.60	6.80	3.80	9.80	9.20	1.80	10.20	6.20	4.20	9.20
		J			1.40						1.00			
	1000	mean	0.249	−0.510	0.543	1.048	1.521	1.996	0.248	−0.509	0.532	1.038	1.519	1.997
		median	0.241	−0.501	0.566	1.048	1.519	1.952	0.241	−0.498	0.543	1.041	1.514	1.964
		var	0.005	0.005	0.046	0.032	0.020	0.055	0.004	0.005	0.049	0.035	0.020	0.057
		rmse	0.068	0.073	0.218	0.186	0.142	0.235	0.067	0.071	0.223	0.190	0.144	0.239
		t	6.80	3.00	9.20	7.00	5.00	6.20	8.80	2.40	10.40	8.00	5.40	5.80
		J			2.00						2.60			
	2000	mean	0.247	−0.502	0.512	1.019	1.505	1.990	0.247	−0.501	0.509	1.018	1.504	1.990
		median	0.243	−0.499	0.520	1.020	1.503	1.981	0.238	−0.499	0.516	1.017	1.499	1.981
		var	0.002	0.002	0.028	0.018	0.010	0.017	0.002	0.002	0.028	0.018	0.010	0.016
		rmse	0.045	0.049	0.167	0.137	0.102	0.130	0.044	0.046	0.168	0.135	0.102	0.126
		t	5.60	3.20	7.20	6.80	3.80	5.80	4.60	3.60	7.40	6.60	3.60	6.00
		J			4.80						4.20			

Thus, the use of $\log|\hat{Z}_{t-1}|$ as an estimable factor is justified because, as mentioned earlier, the logarithmic transformation fits into the location specification (2.8). Table 5.2 summarizes descriptive statistics alongside the results from quasi maximum-likelihood estimation with skewed Student's t-distributed innovations. The stock returns are left-skewed and leptokurtic with conditional mean and variance dynamics that are similar to findings from the literature; see, e.g., Bollerslev et al. (1994). Note that the distribution of W_t (gold), while also leptokurtic, is right-skewed.

Turning to the specification of the cross-sectional distribution of the 43 companies, we use various skewed-t factor models, inspired by Oh and Patton (2013, 2017), as our benchmark specifications for the copula. Based on Table 5.1 we consider the following block-equidependent design

$$X_{i,t} = \alpha_{1,j}F_t + \alpha_{2,j}F_{j,t} + \varepsilon_{i,t}, \quad i \in \mathcal{G}_j,\ j \in \{1, 2, 3, 4\}, \tag{5.4}$$

where $\mathcal{G}_1 = \{1, \ldots, 13\}$, $\mathcal{G}_2 = \{14, \ldots, 24\}$, $\mathcal{G}_3 = \{25, \ldots, 35\}$, and $\mathcal{G}_4 = \{36, \ldots, 43\}$. We assume that $F_t \overset{\text{IID}}{\sim} t(\zeta, \xi)$, $F_{j,t} \overset{\text{IID}}{\sim} t(\zeta)$, and $\varepsilon_{i,t} \overset{\text{IID}}{\sim} t(\zeta)$, while factors and idiosyncratic errors are mutually independent. We use the following four versions of Eq. (5.4), labeled A1, A2, A3, and A4, imposing certain restrictions on the loadings: specification A1 is a one-factor equidependent model so that $\alpha_{1,j} = \alpha$ and $\alpha_{2,j} = 0$; specification A2 is a one-factor block-equidependent model with $\alpha_{2,j} = 0$; specification A3 has a common factor with common loading $\alpha_{1,j} = \alpha$ and group-specific factors with group-specific loadings; specification A4 does not impose any restrictions and allows for a common factor and group-specific factors, both with group-specific loadings. The following competitor for the copula is generated from a factor model with estimable gold factor:

$$X_{i,t} = \alpha_j F_{j,t} + \beta_j \log|Z_{t-1}| + \varepsilon_{i,t}, \quad i \in \mathcal{G}_j,\ j \in \{1, 2, 3, 4\}, \tag{5.5}$$

where $\varepsilon_{i,t} \overset{\text{IID}}{\sim} t(\zeta)$. Similar to Eq. (5.4), four versions of Eq. (5.5), labeled B1, B2, B3, and B4, are considered: specification B1 allows for group-specific loadings on a common simulable factor $F_{j,t} = F_t \overset{\text{IID}}{\sim} t(\zeta, \xi)$ and imposes $\beta_j = \beta$; specification B2 imposes $\beta_j = \beta$ but allows for group-specific simulable factors $F_{j,t} \overset{\text{IID}}{\sim} t(\zeta, \xi)$ with group-specific loadings; specifications B3 and B4 allow for group-specific loadings assuming a common simulable factor $F_{j,t} = F_t \overset{\text{IID}}{\sim} t(\zeta, \xi)$, $\xi = 0$ (B3) and group-specific simulable factors $F_{j,t} \overset{\text{IID}}{\sim} t(\zeta, \xi)$, $\xi = 0$ (B4), respectively.

Table 5.3 summarizes the SMM estimation results for the benchmark specification (5.4) as well as for the counterpart with estimable gold factor (5.5) based on the same rank-based dependence measures used in the Monte Carlo study.

Table 4.5
Simulation results (design 2: skew-t/t).

n	T		Feasible						Unfeasible					
			ζ_0	ξ_0	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	α_0	ζ_0	ξ_0	$\beta_{0,1}$	$\beta_{0,2}$	$\beta_{0,3}$	α_0
			0.25	−0.5	0.5	1.0	1.5	2.0	0.25	−0.5	0.5	1.0	1.5	2.0
15	500	mean	0.237	−0.534	0.613	1.106	1.519	1.888	0.238	−0.533	0.613	1.106	1.521	1.883
		median	0.245	−0.517	0.614	1.099	1.521	1.887	0.255	−0.516	0.609	1.100	1.525	1.896
		var	0.008	0.016	0.059	0.053	0.054	0.037	0.008	0.016	0.062	0.056	0.054	0.045
		rmse	0.089	0.130	0.267	0.253	0.234	0.224	0.089	0.130	0.273	0.259	0.234	0.242
		t	5.40	1.00	6.60	3.40	2.60	9.00	5.60	0.80	7.40	3.00	2.00	9.40
		J	1.20						1.00					
	1000	mean	0.241	−0.515	0.563	1.064	1.513	1.937	0.242	−0.516	0.567	1.065	1.515	1.935
		median	0.247	−0.503	0.586	1.075	1.516	1.939	0.247	−0.506	0.588	1.074	1.527	1.934
		var	0.005	0.006	0.040	0.035	0.032	0.020	0.005	0.007	0.039	0.035	0.031	0.021
		rmse	0.072	0.082	0.211	0.197	0.178	0.155	0.071	0.085	0.209	0.199	0.177	0.159
		t	5.80	1.40	8.00	3.40	2.80	5.60	6.00	1.60	8.00	4.40	2.60	6.00
		J	1.60						1.60					
	2000	mean	0.248	−0.506	0.532	1.034	1.500	1.968	0.249	−0.507	0.536	1.035	1.500	1.967
		median	0.248	−0.506	0.547	1.029	1.502	1.966	0.249	−0.505	0.556	1.033	1.505	1.967
		var	0.003	0.003	0.028	0.020	0.015	0.011	0.003	0.003	0.027	0.019	0.015	0.011
		rmse	0.058	0.052	0.170	0.144	0.122	0.108	0.057	0.053	0.170	0.144	0.124	0.109
		t	4.60	3.60	8.00	3.80	2.80	5.60	4.40	3.20	7.60	4.40	2.80	5.40
		J	1.20						1.20					
30	500	mean	0.224	−0.523	0.567	1.053	1.501	1.914	0.225	−0.523	0.564	1.048	1.501	1.914
		median	0.229	−0.510	0.585	1.060	1.503	1.915	0.230	−0.509	0.577	1.052	1.496	1.912
		var	0.008	0.011	0.046	0.039	0.044	0.029	0.008	0.012	0.045	0.039	0.043	0.030
		rmse	0.094	0.109	0.224	0.204	0.210	0.191	0.094	0.111	0.222	0.203	0.207	0.194
		t	7.80	0.80	5.00	2.20	3.00	8.40	7.20	1.00	4.40	2.00	2.00	8.00
		J	1.40						1.60					
	1000	mean	0.236	−0.511	0.534	1.025	1.500	1.955	0.234	−0.510	0.530	1.023	1.503	1.956
		median	0.230	−0.504	0.534	1.034	1.503	1.953	0.232	−0.504	0.526	1.026	1.509	1.957
		var	0.005	0.005	0.031	0.027	0.022	0.017	0.005	0.005	0.033	0.025	0.022	0.016
		rmse	0.074	0.069	0.179	0.166	0.150	0.136	0.071	0.069	0.183	0.160	0.147	0.135
		t	7.40	1.80	5.60	2.40	2.00	5.80	6.80	2.60	6.20	2.20	2.80	5.00
		J	1.60						1.60					
	2000	mean	0.244	−0.504	0.518	1.018	1.494	1.977	0.244	−0.504	0.514	1.014	1.493	1.978
		median	0.237	−0.501	0.538	1.016	1.500	1.973	0.239	−0.502	0.532	1.013	1.497	1.976
		var	0.003	0.002	0.021	0.014	0.011	0.009	0.003	0.002	0.022	0.015	0.012	0.009
		rmse	0.053	0.047	0.147	0.121	0.107	0.097	0.054	0.047	0.150	0.125	0.108	0.097
		t	5.40	3.60	6.20	4.40	3.00	5.20	6.40	3.20	7.20	4.40	3.20	5.20
		J	3.00						3.00					

Table 5.1
Companies.

i	Group 1 (‘*Pharma*’)		i	Group 2 (‘*Finance*’)		i	Group 3 (‘*Oil & Gas*’)		i	Group 4 (‘*Transport*’)	
1	*Abbott Lab.*	ABT	14	*Bank Of Am*	BAC	25	*Apache*	APA	36	*Caterpillar*	CAT
2	*AbbVie*	ABBV	15	*Bank Of NY*	BK	26	*Baker Hughes*	BHI	37	*Emerson Ele*	EMR
3	*Amgen*	AMGN	16	*Citigroup Inc*	C	27	*Conocophillips*	COP	38	*Fedex*	FDX
4	*Baxter*	BAX	17	*Capital One*	COF	28	*Chevron*	CVX	39	*Honeywell Int*	HON
5	*Biogen*	BIIB	18	*Goldman Sachs*	GS	29	*Devon*	DVN	40	*3M*	MMM
6	*Bristol–Myers*	BMY	19	*Jpmorgan*	JPM	30	*Halliburton*	HAL	41	*Norfolk South*	NSC
7	*Gilead*	GILD	20	*Metlife*	MET	31	*Nat. Oilwell*	NOV	42	*Union Pacific*	UNP
8	*Johnson & J*	JNJ	21	*Morgan Stanley*	MS	32	*Occidental*	OXY	43	*United Parcel*	UPS
9	*Lilly Eli*	LLY	22	*Regions Fin*	RF	33	*Schlumberger*	SLB			
10	*Medtronic*	MDT	23	*US Bancorp*	USB	34	*Williams Co*	WMB			
11	*Merck*	MRK	24	*Wells Fargo*	WFC	35	*Exxon Mobil*	XOM			
12	*Pfizer*	PFE									
13	*Unitedhealth*	UNH									

Table 5.2
Summary statistics.

Mean	10%	25%	Median	75%	90%	Gold
Mean	−0.0399	−0.0012	0.0277	0.0520	0.0132	
Standard Dev	1.1518	1.2614	1.4962	1.6701	2.2298	0.8119
Skewness	−1.1644	−0.8172	−0.3015	−0.0717	0.0902	0.2363
Kurtosis	5.2379	5.7804	6.4064	10.2722	16.7054	5.1414
Constant	−0.0387	−0.0005	0.0243	0.0503	0.0617	
AR(1)	−0.0172	0.0003	0.0189	0.0370	0.0453	
Gold	−0.0138	0.0057	0.0315	0.0427	0.0645	
Constant	0.0217	0.0349	0.0916	0.1656	0.2758	0.0043
ARCH	0.0000	0.0033	0.0171	0.0414	0.0546	0.0265
Leverage	0.0555	0.0699	0.1141	0.1465	0.2021	−0.0021
GARCH	0.7383	0.8050	0.8766	0.9335	0.9575	0.9680
ξ	3.9098	4.2967	4.8267	5.8873	6.5596	5.6674
$1/\zeta$	−0.1088	−0.0825	−0.0634	−0.0358	−0.0187	0.0433

Table 5.3
Copula parameter estimates.

Panel A: *w/o gold factor*	A1		A2		A3		A4	
	Identity	Optimal	Identity	Optimal	Identity	Optimal	Identity	Optimal
Restrictions	$\alpha_{1,j} = \alpha$ $\alpha_{2,j} = 0$		– $\alpha_{2,j} = 0$		$\alpha_{1,j} = \alpha$ –		– –	
ξ	−0.2207	−0.2172	−0.1782	−0.1563	−0.3652	−0.2687	−0.2780	−0.2730
	0.0320	*0.0334*	*0.0300*	*0.0283*	*0.1726*	*0.1168*	*0.0976*	*0.0926*
ζ	0.1740	0.1596	0.0749	0.0689	0.1200	0.1205	0.1119	0.1133
	0.0368	*0.0384*	*0.0263*	*0.0250*	*0.0372*	*0.0340*	*0.0418*	*0.0393*
$\alpha_{1,1}$			0.9655	0.9217			0.9121	0.8632
			0.0246	*0.0206*			*0.1308*	*0.1155*
$\alpha_{1,2}$	0.8497	0.8282	1.8174	1.7402	0.8751	0.8449	1.3347	1.3255
	0.0216	*0.0200*	*0.0581*	*0.0395*	*0.1258*	*0.1335*	*0.2905*	*0.2409*
$\alpha_{1,3}$			1.2697	1.2995			0.9557	0.9657
			0.0339	*0.0279*			*0.2003*	*0.1860*
$\alpha_{1,4}$			1.2467	1.1851			1.0764	1.0130
			0.0366	*0.0287*			*0.2046*	*0.1559*
$\alpha_{2,1}$					0.4007	0.3744	0.2747	0.3414
					0.2775	*0.2956*	*0.4178*	*0.2795*
$\alpha_{2,2}$					1.5668	1.5200	1.2292	1.1235
					0.0994	*0.0891*	*0.3208*	*0.2813*
$\alpha_{2,3}$					0.9156	0.9470	0.8182	0.8397
					0.1302	*0.1210*	*0.2173*	*0.2027*
$\alpha_{2,4}$					0.8729	0.8151	0.6090	0.6121
					0.1355	*0.1421*	*0.3404*	*0.2515*
J	0.4961	15.7855	9.8333	34.4458	6.2357	25.6570	4.6589	25.7659
p-value	0.0030	0.0030	0.0340	0.0025	0.1940	0.0189	0.0100	0.0041

Panel B: *w/ gold factor*	B1		B2		B3		B4	
	Identity	Optimal	Identity	Optimal	Identity	Optimal	Identity	Optimal
Restrictions	$\beta_j = \beta$ $F_{j,t} = F_t$ –		$\beta_j = \beta$ – –		– $F_{j,t} = F_t$ $\xi = 0$		– – $\xi = 0$	
ξ	−0.0976	−0.1175	−0.0788	−0.0560				
	0.0517	*0.0460*	*0.0467*	*0.0428*				
ζ	0.1155	0.0949	0.1378	0.1212	0.1561	0.1548	0.1556	0.1498
	0.0515	*0.0428*	*0.0512*	*0.0424*	*0.0694*	*0.0643*	*0.0662*	*0.0597*
α_1	0.6043	0.6081	0.6028	0.5824	0.5570	0.5389	0.5613	0.5298
	0.0724	*0.0695*	*0.0680*	*0.0632*	*0.0638*	*0.0612*	*0.0634*	*0.0608*
α_2	1.6260	1.5912	1.6364	1.6022	1.4083	1.3771	1.4243	1.3509
	0.0581	*0.0497*	*0.0588*	*0.0478*	*0.1208*	*0.1065*	*0.1350*	*0.1226*
α_3	1.0245	1.0750	1.0197	1.0591	0.9353	0.9544	0.9404	0.9792
	0.0594	*0.0501*	*0.0572*	*0.0468*	*0.0813*	*0.0725*	*0.0796*	*0.0783*
α_4	0.9903	0.9550	0.9901	0.9387	0.8504	0.8275	0.8521	0.8324
	0.0611	*0.0548*	*0.0571*	*0.0495*	*0.0852*	*0.0846*	*0.0857*	*0.0797*
β_1					0.6666	0.6512	0.6619	0.6613
					0.0487	*0.0441*	*0.0486*	*0.0457*
β_2	0.6538	0.6227	0.6419	0.6253	0.9378	0.9432	0.9598	0.9839
	0.0557	*0.0556*	*0.0510*	*0.0489*	*0.1384*	*0.1169*	*0.1516*	*0.1259*
β_3					0.7151	0.7442	0.7084	0.7246
					0.0814	*0.0699*	*0.0851*	*0.0822*
β_4					0.7675	0.7412	0.7779	0.7345
					0.0769	*0.0710*	*0.0746*	*0.0732*
J	4.3537	22.2472	4.7058	23.5429	3.9102	21.8012	4.1245	22.1737
p-value	0.3350	0.0516	0.2680	0.0356	0.0925	0.0260	0.0805	0.0225

Note: SMM point estimates with standard errors below in italics using $S = 25$. Standard errors and the p-values for the overidentifying restrictions test are based on $B = 2000$ and $\pi_T = 0.05$. The value of the latter test statistic is labeled $J := J_{T,S} = T\hat{A}_{T,S}$; see Eq. Eq. (3.16).

We set $S = 25$, $B = 2000$, and $\pi_T = 0.05$ and report estimation results using identity weighting $L_{T,S} = I_{\bar{p}}$ and 'optimal' weighting $L_{T,S} = \hat{\Sigma}_{T,S}^{-1}$. The point estimates for the benchmark specifications in the upper panel are in line with the values reported in Oh and Patton (2017, Table 3) and suggest significant (negative) asymmetric dependence and significant tail dependence. As can be deduced from the results of the overidentifying restrictions test (3.16), all specifications but A3 are clearly rejected by the data. Moving to the lower panel, we find results for our competitors with estimable factor. The specifications improve the performance of the benchmark models and cannot (at least for the identity weight matrix) be rejected by the data. Interestingly, conditionally on the estimable gold factor, the asymmetry parameter is smaller in size and no longer statistically significant different from zero. This can be explained by the fact that the estimable gold factor already accounts for (some) asymmetry. To illustrate, Fig. 5.1 depicts the distribution of the logarithmic absolute

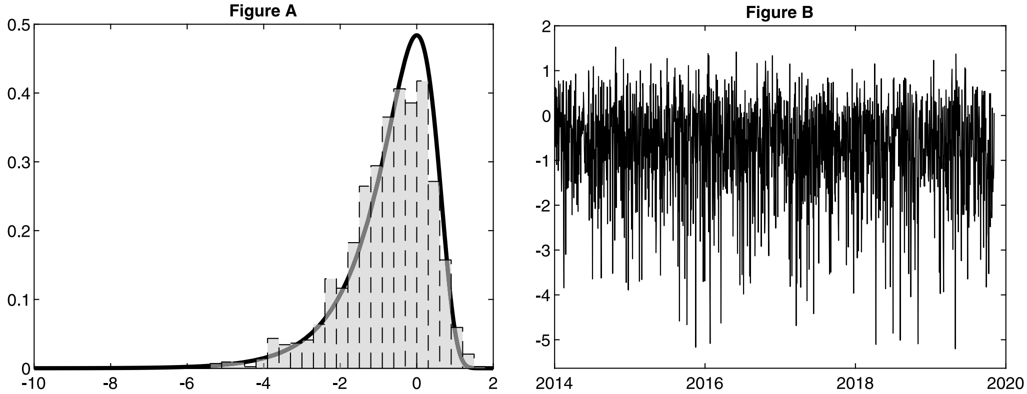

Fig. 5.1. Histogram and time-series plot of $\log|\hat{Z}_t|$.

residuals, which is seen to be left-skewed; for comparison, the density of $\log|Z_t^*|$, $Z_t^* \sim \mathcal{N}(0, 1)$, is depicted as the solid line in panel A.[5]

6. Conclusion

We derive the asymptotic properties of an SMM estimator of the unknown parameter vector governing a factor copula model with estimable factors and show how to estimate its limiting variance–covariance matrix consistently. The asymptotic theory is derived from primitive conditions, thereby complementing also the earlier work of Oh and Patton (2013), whose model is nested in our framework. One avenue for future research that can be pursued is to consider quasi-Bayesian estimation to alleviate the difficulties of having to deal with a non-smooth objective function. For example, the sample criterion function considered may be shown to fulfill the regularity conditions needed for Laplace-type estimation in Chernozhukov and Hong (2003, section 4.1); see also Hong et al. (2021) for a recent application of this idea to an SMM objective function with overlapping simulation draws.

Appendix. Technical appendix

Preliminaries: To begin with, suppose $i, j \in \mathcal{G}_q$ for $1 \leq i < j \leq n$ and some $q \in \{1, \ldots, Q\}$. Next, define the empirical copulae

$$\begin{aligned}\hat{\mathsf{C}}_{T,S,i,j}(u_i, u_j;\ \theta_q, \nu)\\ := \frac{1}{TS}\sum_{t=1}^{T}\sum_{s=1}^{S} 1\{\hat{X}_{i,t,s}(\theta_q, \nu) \leq \hat{\mathsf{G}}^{-}_{T,S,i}(u_i; \theta_q, \nu), \hat{X}_{j,t,s}(\theta_q, \nu) \leq \hat{\mathsf{G}}^{-}_{T,S,j}(u_j; \theta_q, \nu)\}\end{aligned} \tag{A.1}$$

and

$$\hat{\mathsf{C}}_{T,i,j}(u_i, u_j; \lambda) := \frac{1}{T}\sum_{t=1}^{T} 1\{\hat{\eta}_{i,t}(\lambda) \leq \hat{\mathsf{F}}^{-}_{T,i}(u_i; \lambda), \hat{\eta}_{j,t}(\lambda) \leq \hat{\mathsf{F}}^{-}_{T,j}(u_j; \lambda)\}, \tag{A.2}$$

where $\theta_q = (\alpha_q', \beta_q', \gamma', \delta')' \in \Theta$, $u_i, u_j \in [0, 1]$, while $\nu \in \mathcal{U}(\nu_0)$ and $\lambda \in \mathcal{U}(\lambda_0)$, with $\mathcal{U}(\cdot)$ denoting neighborhoods around a given parameter as defined by Assumption E. Below, we make frequently use of the identities [see, e.g., Tsukahara (2005) and Segers (2012)]:

$$\begin{aligned}\hat{\mathsf{C}}_{T,S,i,j}(u_i, u_j;\ \theta_q, \nu) &= \tilde{\mathsf{C}}_{T,S,i,j}\{\tilde{\mathsf{C}}^{-}_{T,S,i}(u_i; \theta_q, \nu), \tilde{\mathsf{C}}^{-}_{T,S,j}(u_j; \theta_q, \nu); \theta_q, \nu\}\\ \hat{\mathsf{C}}_{T,i,j}(u_i, u_j;\ \lambda) &= \tilde{\mathsf{C}}_{T,i,j}\{\tilde{\mathsf{C}}^{-}_{T,i}(u_i; \lambda), \tilde{\mathsf{C}}^{-}_{T,j}(u_j; \lambda); \lambda\},\end{aligned} \tag{A.3}$$

5 If $Z^* \sim \mathcal{N}(0, 1)$, then $Z := \log|Z^*|$ has absolutely continuous density f

$$\mathsf{f}(z) := \sqrt{\frac{2\exp\{2z - \exp\{2z\}\}}{\pi}}, \quad z \in \mathbb{R},$$

with mean $\mathsf{E}[Z] = (-\gamma - \log 2)/2$ and variance $\mathsf{var}[Z] = \pi^2/8$, where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. The density given by the preceding display is depicted as the solid line in Fig. 5.1.

where

$$\begin{aligned}\tilde{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu) &:= \frac{1}{TS}\sum_{t=1}^{T}\sum_{s=1}^{S} 1\{\hat{X}_{i,t,s}(\theta_q,\nu) \le \mathsf{G}_q^{-1}(u_i;\theta_q), \hat{X}_{j,t,s}(\theta_q,\nu) \le \mathsf{G}_q^{-1}(u_j;\theta_q)\} \\ \tilde{\mathsf{C}}_{T,i,j}(u_i,u_j;\lambda) &:= \frac{1}{T}\sum_{t=1}^{T} 1\{\hat{\eta}_{i,t}(\lambda) \le \mathsf{F}_i^{-1}(u_i), \hat{\eta}_{j,t}(\lambda) \le \mathsf{F}_j^{-1}(u_j)\},\end{aligned} \tag{A.4}$$

and $\tilde{\mathsf{C}}_{T,S,k}(u_i;\theta_q,\nu) = \tilde{\mathsf{C}}_{T,S,i,j}(\vec{u}_k;\theta_q,\nu)$, $\tilde{\mathsf{C}}_{T,k}(u_k;\lambda) = \tilde{\mathsf{C}}_{T,i,j}(\vec{u}_k;\lambda)$, with $\vec{u}_i = (u_i,1)'$, $\vec{u}_j = (1,u_j)'$. Moreover, define the non-centered processes

$$\begin{aligned}\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu) &:= \sqrt{T}(\hat{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu) - \mathsf{C}_q(u_i,u_j;\theta_q)) \\ \hat{\mathbb{C}}_{T,i,j}(u_i,u_j;\lambda) &:= \sqrt{T}(\hat{\mathsf{C}}_{T,i,j}(u_i,u_j;\lambda) - \mathsf{C}_q(u_i,u_j)),\end{aligned} \tag{A.5}$$

and the centered processes

$$\begin{aligned}\tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu) &:= \sqrt{T}(\tilde{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu) - \mathsf{E}[\tilde{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu)]) \\ \tilde{\mathbb{C}}_{T,i,j}(u_i,u_j;\lambda) &:= \sqrt{T}(\tilde{\mathsf{C}}_{T,i,j}(u_i,u_j;\lambda) - \mathsf{E}[\tilde{\mathsf{C}}_{T,i,j}(u_i,u_j;\lambda)]).\end{aligned} \tag{A.6}$$

A.1. Proof of Proposition 1

Proof of Proposition 1(*a*). Recall the definition of the population rank statistics from Eq. (3.8) and note that

$$\psi_{q,k} = \int_{[0,1]^2} \varphi_k(u_i,u_j)\mathrm{d}\mathsf{C}_q(u_i,u_j), \quad \psi_{q,k}(\theta) = \int_{[0,1]^2} \varphi_k(u_i,u_j)\mathrm{d}\mathsf{C}_q(u_i,u_j;\theta_q)$$

for any $i,j \in \mathcal{G}_q$, $q \in \{1,\ldots,Q\}$. Hence, one gets, in view of Lemma S.3 and Lemma S.4 supplementary material (SM), the following representation

$$\sqrt{T}(\hat{\psi}_{T,i,j,k}(\hat{\lambda}_T) - \psi_{q,k}) = \int_{[0,1]^2} \hat{\mathbb{C}}_{T,i,j}(u_i,u_j)\mathrm{d}\varphi_k(u_i,u_j) + o_p(1), \tag{A.7}$$

with $\hat{\mathbb{C}}_{T,i,j}(u_i,u_j) := \sqrt{T}(\hat{\mathsf{C}}_{T,i,j}(u_i,u_j) - \mathsf{C}_q(u_i,u_j))$, and, uniformly in $\theta \in \Theta$,

$$\sqrt{T}(\hat{\psi}_{T,S,i,j,k}(\theta_q,\hat{\nu}_T) - \psi_{q,k}(\theta)) = \int_{[0,1]^2} \hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q)\mathrm{d}\varphi_k(u_i,u_j) + o_p(1), \tag{A.8}$$

with $\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q) := \sqrt{T}(\hat{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q) - \mathsf{C}_q(u_i,u_j;\theta_q))$, $\hat{\mathsf{C}}_{T,i,j}(u_i,u_j) := \hat{\mathsf{C}}_{T,i,j}(u_i,u_j;\lambda_0)$, $\hat{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q) := \hat{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta,\nu_0)$; see Eqs. (A.1) and (A.2). Thus, the claim is due to the weak convergence of $\hat{\mathbb{C}}_{T,i,j}$ and $\hat{\mathbb{C}}_{T,S,i,j}$. To see this, note that the functional delta method (cf. van der Vaart and Wellner (1996, Theorem 3.9.4)) in conjunction with Assumption C and Bücher and Volgushev (2013, Theorem 2.4.) yields

$$\hat{\mathbb{C}}_{T,i,j}(u_i,u_j) = \tilde{\mathbb{C}}_{T,i,j}(u_i,u_j) - \sum_{k\in\{i,j\}} \partial_k\mathsf{C}_q(u_i,u_j)\tilde{\mathbb{C}}_{T,i,j}(\vec{u}_k) + R(u_i,u_j), \tag{A.9}$$

and

$$\begin{aligned}\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q) &= \tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q) \\ &\quad - \sum_{k\in\{i,j\}} \partial_k\mathsf{C}_q(u_i,u_j;\theta_q)\tilde{\mathbb{C}}_{T,S,i,j}(\vec{u}_k;\theta_q) + R_1(u_i,u_j;\theta_q),\end{aligned} \tag{A.10}$$

where $\sup_{u_i,u_j\in[0,1]}|R(u_i,u_j)| = o_p(1)$, $\sup_{\theta_q\in\Theta}\sup_{u_i,u_j\in[0,1]}|R_1(u_i,u_j;\theta_q)| = o_p(1)$, and $\tilde{\mathbb{C}}_{T,i,j}(u_i,u_j) := \tilde{\mathbb{C}}_{T,i,j}(u_i,u_j;\lambda_0)$, $\tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q) := \tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q,\nu_0)$; see Eq. (A.6). The claim follows from Lemma S.1 and Lemma S.2 of the SM.

Proof of Proposition 1 (*b*). Taking Eq. (3.12), Lemma S.3, and Lemma S.4 into account, one gets

$$\sqrt{T}\hat{\Psi}_{T,S,i,j,k}(\theta_q,\hat{\lambda}_T,\hat{\nu}_T) = \int_{[0,1]^2} \hat{\mathbb{B}}_{T,S,i,j}(u_i,u_j)\,\mathrm{d}\varphi_k(u_i,u_j) + o_p(1), \tag{A.11}$$

where $\hat{\mathbb{B}}_{T,S,i,j}(u_i,u_j) := \hat{\mathbb{B}}_{T,S,i,j}(u_i,u_j;\theta_{0,q},\lambda_0,\nu_0)$. We show next that $\hat{\mathbb{B}}_{T,S,i,j}$ converges weakly to the tight Gaussian process $\mathbb{B}_S$ concentrated on $\mathbb{D}_0 := \{\alpha \in C[0,1]^n : \alpha(1,\ldots,1) = 0,\ \ \alpha(x) = 0$ if some of the components of x are equal to zero$\}$ by establishing (1) asymptotic tightness and (2) finite dimensional ('*fidi*', henceforth) convergence. *(1) Stochastic equicontinuity:* The functional delta method yields

$$\hat{\mathbb{B}}_{T,S,i,j}(u_i,u_j) = \tilde{\mathbb{B}}_{T,S,i,j}(u_i,u_j) - \sum_{k\in\{i,j\}} \partial_k\mathsf{C}_q(u_i,u_j)\tilde{\mathbb{B}}_{T,S,i,j}(\vec{u}_k) + o_p(1), \tag{A.12}$$

where

$$\tilde{\mathbb{B}}_{T,S,i,j}(u_i,u_j)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{S}\sum_{s=1}^{S}(1\{V_{i,t}\le u_i,V_{j,t}\le u_j\}-1\{U_{i,t,s}\le u_i,U_{j,t,s}\le u_j\}),\tag{A.13}$$

with $V_{k,t}:=\mathsf{F}_k(\eta_{k,t})$, $U_{k,t,s}:=\mathsf{G}_q(X_{k,t,s}(\theta_{0,q});\theta_{0,q})$ for $k\in\{i,j\}$. Let $\xi_{i,j,t,S}=(V_{i,t},V_{j,t},\mathsf{U}'_{S:i,j,t})'$, with $\mathsf{U}_{S:i,j,t}:=(U_{i,t,1},U_{j,t,1},\dots,U_{i,t,S},U_{j,t,S})'$. We can view $\tilde{\mathbb{B}}_{T,S,i,j}$ as an empirical process indexed by $\bar f\in\bar{\mathcal{F}}$:

$$\bar{\mathcal{F}}:=\{\xi_{S:i,j,t}\mapsto\bar f(\xi_{S:i,j,t}):=\frac{1}{S}\sum_{s=1}^{S}(f(V_{i,t},V_{j,t})-f(U_{i,t,s},U_{j,t,s})):f\in\mathcal{F}\},\tag{A.14}$$

where $\mathcal{F}:=\big\{(x_1,x_2)\mapsto 1\{x_1\le u_1,x_2\le u_2\},\ u_1,u_2\in[0,1]\big\}$. Clearly, $\bar{\mathcal{F}}$ has envelope 1. We use Theorem S1 of the SM to establish asymptotic equicontinuity. Specifically, the bracketing number $\mathcal{N}_{[\,]}(\epsilon,\bar{\mathcal{F}},\rho(\cdot))$, with $\rho(\bar f):=\sup_{t,T}\|\bar f(\xi_{i,j,t,S})\|_2$, shall be determined; see the discussion surrounding Theorem S1 for details. It is well-known (see, e.g. van der Vaart (1994, Example 19.6)) that $\mathcal{N}_{[\,]}(\epsilon,\mathcal{F}_k,\|\cdot\|_2)=O(\varepsilon^{-2})$, with $\mathcal{F}_k:=\{x_k\mapsto 1\{x_k\le u_k\},u_k\in[0,1]\}$. Since $\mathcal{F}\subseteq\mathcal{F}_1\cdot\mathcal{F}_2$, one gets $\mathcal{N}_{[\,]}(\epsilon,\mathcal{F},\|\cdot\|_2)=O(\varepsilon^{-4})$; see, e.g., Kosorok (2008, Lemma 9.25). Suppose $[l_k,u_k]$, $k=1,\dots,m:=\mathcal{N}_{[\,]}(\epsilon,\mathcal{F},\|\cdot\|_2)$, represent the brackets needed to cover $\mathcal{F}$. We can then cover $\bar{\mathcal{F}}$ with $[\bar l_k,\bar u_k]$, $k=1,\dots,m$, where

$$\begin{aligned}\bar l_k(\xi_{S:i,j,t})&:=\frac{1}{S}\sum_{s=1}^{S}(l_k(V_{i,t},V_{j,t})-u_k(U_{i,t,s},U_{j,t,s}))\\ \bar u_k(\xi_{S:i,j,t})&:=\frac{1}{S}\sum_{s=1}^{S}(u_k(V_{i,t},V_{j,t})-l_k(U_{i,t,s},U_{j,t,s})).\end{aligned}\tag{A.15}$$

Note, that $\rho(\bar u_k-\bar l_k)\le 2\epsilon$. Thus, $\mathcal{N}_{[\,]}(\epsilon,\bar{\mathcal{F}},\rho(\cdot))=O(\epsilon^{-4})$. Now, since $\{\xi_{i,j,t,S}:t\ge 1\}$ is *i.i.d.* and $\bar{\mathcal{F}}$ is uniformly bounded, the conditions of Theorem S1 are satisfied. *(2) 'Fidi'-convergence:* By the Cramér–Wold device (see, e.g., White (2001, Proposition 5.1)), it suffices to fix some $c:=(c_1,\dots,c_m)'\in\mathbb{R}^m$, with $\|c\|=1$, $(\{u_1,v_1\}\dots,\{u_m,v_m\})'\in[0,1]^{2m}$, and to consider

$$Z_{T,S}(m):=\sum_{l=1}^{m}c_l\tilde{\mathbb{B}}_{T,S,i,j}(u_l,v_l)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}B_{t,S}(m),\tag{A.16}$$

where $B_{t,S}(m):=\sum_{l=1}^{m}c_l\zeta_{t,S}(u_l,v_l)$, with

$$\zeta_{t,S}(u_l,v_l):=1\{V_{i,t}\le u_l,V_{j,t}\le v_l\}-\frac{1}{S}\sum_{s=1}^{S}1\{U_{i,t,s}\le u_l,U_{j,t,s}\le v_l\}.\tag{A.17}$$

The sequence $\{B_{t,S}(m):t\ge 1\}$ is *i.i.d*, bounded, and, by Assumption B, centered. It thus follows from White (2001, Theorem 5.11) that $Z_{T,S}(m)\xrightarrow{d}\mathcal{N}(0,\sigma_S^2(m))$, with $\sigma_S^2(m):=\lim_{T\to\infty}\sigma_{T,S}^2(m)$, provided $\inf_{T,S\ge 1}\sigma_{T,S}^2(m)>0$ for $\sigma_{T,S}^2(m):=\mathrm{var}[Z_{T,S}(m)]$. Now,

$$\sigma_{T,S}^2(m)=\sum_{k,l=1}^{m}c_kc_l\hat\gamma_{T,S}(k,l),\quad \hat\gamma_{T,S}(k,l):=\frac{1}{T}\sum_{t,h=1}^{T}\mathrm{cov}[\zeta_{t,S}(w_k),\zeta_{t,S}(w_l)],\tag{A.18}$$

where $w_j=:(u_j,v_j)'\in[0,1]^2$, $j\in\{1,\dots,m\}$. Since, $\{\zeta_{t,S}(w_k):t\ge 1\}$ is *i.i.d.*, one gets $\hat\gamma_{1,S}(k,l)=\mathrm{cov}[\zeta_{1,S}(w_k),\zeta_{1,S}(w_l)]$. Next, for any $i,j\in\mathcal{G}_q$, $q\in\{1,\dots,Q\}$, one gets

$$\begin{aligned}\mathrm{cov}[\zeta_{1,S}(w_k),\zeta_{1,S}(w_l)\mid Z_1]=&\,\mathsf{C}_{q,1}(u_k\wedge u_l,v_k\wedge v_l)-\mathsf{C}_{q,1}(u_k,v_k)\mathsf{C}_{q,1}(u_l,v_l)\\&+[\mathcal{C}_{q,1}(u_k\wedge u_l,v_k\wedge v_l)-\mathcal{C}_{q,1}(u_k,v_k)\mathcal{C}_{q,1}(u_l,v_l)]/S,\end{aligned}\tag{A.19}$$

where the penultimate equality uses that $\{V_{i,t},V_{j,t}\}\perp\{U_{i,t,s},U_{j,t,s}\}\mid Z_t$ and $\{U_{i,t,r},U_{j,t,r}\}\perp\{U_{i,t,s},U_{j,t,s}\}\mid Z_t$, $r\ne s$. Since $\mathsf{E}[\zeta_{1,S}(w_k)\mid Z_1]=\mathsf{C}_{q,1}(w_k)-\mathcal{C}_{q,1}(w_k)$, we get, by the law of total covariance,

$$\begin{aligned}\mathrm{cov}[\zeta_{1,S}(w_k),\zeta_{1,S}(w_l)]=&\,\mathsf{E}[\mathsf{C}_{q,t}(w_k\wedge w_l)-\mathsf{C}_{q,t}(w_k)\mathsf{C}_{q,t}(w_l)]\\&+\mathsf{E}[(\mathsf{C}_{q,t}(w_k)-\mathcal{C}_{q,t}(w_k))(\mathsf{C}_{q,t}(w_l)-\mathcal{C}_{q,t}(w_l))]\\&+\mathsf{E}[\mathcal{C}_{q,t}(w_k\wedge w_l)-\mathcal{C}_{q,t}(w_k)\mathcal{C}_{q,t}(w_l)]/S\\=:&\,\gamma_q^{(1)}(k,l)+\gamma_q^{(2)}(k,l)+\gamma_q^{(3)}(k,l)/S,\end{aligned}$$

say. We are left with showing

$$\sigma_{T,S}^2(m):=\sigma_S^2(m)=c'(\gamma_q^{(1)}(k,l))_{1\le k,l\le m}c+c'(\gamma_q^{(2)}(k,l))_{1\le k,l\le m}c/S+c'(\gamma_q^{(3)}(k,l))_{1\le k,l\le m}c>0.$$

Clearly, if $\mathsf{C}_{q,t}(w_1), \dots, \mathsf{C}_{q,t}(w_m)$ are all distinct *a.s.* and $\mathcal{C}_{q,t}(w_1), \dots, \mathcal{C}_{q,t}(w_m)$ are all distinct *a.s.* with values in $(0, 1)$ *a.s.*, then the claim follows by Assumption F. For the cases where not all are distinct or take values in $\{0, 1\}$ the claim follows from the argument used in the proof of Boistard et al. (2017, Lemma 9.6). Therefore, combining (1) and (2) yields $\tilde{\mathbb{B}}_{T,S,i,j}(u, v) \rightsquigarrow \mathbb{C}_{q,S}(u, v)$, so that

$$\sqrt{T}\hat{\Psi}_{T,S,i,j,k}(\theta, \hat{\lambda}_T, \hat{\nu}_T) \xrightarrow{d} \mathcal{N}(0, \sigma_{0,S}(q, q \mid k, k)),$$

with

$$\sigma_{0,S}(q, q \mid k, k) := \int_{[0,1]^2} \int_{[0,1]^2} \mathsf{E}[\mathbb{C}_{q,S}(u_1, v_1)\mathbb{C}_{q,S}(u_2, v_2)]\mathrm{d}\varphi_k(u_1, v_1)\mathrm{d}\varphi_k(u_2, v_2);$$

see also Berghaus et al. (2017, Theorem 3.3). To conclude from here, use that $\{\hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j) : 1 \le i < j \le n\}$ are jointly normal as $T \to \infty$. To see this, note that $\hat{\mathbb{B}}_{T,S,i,j}(u_i, u_j) = \hat{\mathbb{B}}_{T,S}(\vec{u}_{i,j})$, where $\vec{u}_{i,j}$ denotes the n-dimensional vector of ones with u_i (u_j) at ith (jth) position and $\tilde{\mathbb{B}}_{T,S}(u_1, \dots, u_n) = \hat{\mathbb{B}}_{T,S}(u_1, \dots, u_n; \theta_0, \lambda_0, \nu_0)$, with

$$\begin{aligned}
&\hat{\mathbb{B}}_{T,S}(u_1, \dots, u_n; \theta, \lambda, \nu) \\
&\quad := \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Big[1\{\hat{\eta}_{1,t}(\lambda_1) \le \mathsf{F}^{-}_{1,T}(u_1; \lambda_1), \dots, \hat{\eta}_{n,t}(\lambda_n) \le \mathsf{F}^{-}_{n,T}(u_n; \lambda_n)\} \\
&\qquad\qquad - \frac{1}{S} \sum_{s=1}^{S} 1\{\hat{X}_{1,t,s}(d_1, \hat{\nu}_T) \le G_1^{-1}(u_1; d_1, \nu), \dots, \hat{X}_{n,t,s}(d_n, \nu) \le G_n^{-1}(u_n; d_n, \nu)\} \Big].
\end{aligned}$$

Weak convergence of $\tilde{\mathbb{B}}_{T,S}(u_1, \dots, u_n)$ follows by the same arguments used to establish weak convergence of $\tilde{\mathbb{B}}_{T,S,i,j}(u_i, u_j)$. Since n is finite, the claim follows.

Proof of Proposition 1 (*c*). Note that we can restrict the event inside the probability to the case where $\hat{\nu}_T$ lies in a $\sqrt{T}$ neighborhood of ν_0. To see this, recall from Assumption E that $\sqrt{T}\|\hat{\nu}_T - \nu_0\| = O_p(1)$; i.e., there exists some constant $K := K(\varepsilon) < \infty$, independent of T, such that $\overline{\lim}_{T\to\infty}\mathsf{P}(\hat{\nu}_T \notin \mathcal{V}_{K,T}) < \epsilon$ for any $\epsilon > 0$, where

$$\mathcal{V}_{K,T} := \{\nu \in \mathcal{V}_0 : \|\nu - \nu_0\| \le K/\sqrt{T}\}. \tag{A.20}$$

Clearly, for sufficiently large T, $\mathcal{V}_{K,T} \subseteq \mathcal{U}(\lambda_0)$, where the neighborhood has been defined in Assumption E. Hence,

$$\begin{aligned}
&\overline{\lim_{T\to\infty}} \mathsf{P}\Bigg[\sup_{\theta,\tilde{\theta} \in \Theta: \|\theta-\tilde{\theta}\| \le \delta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \hat{\nu}_T) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \hat{\nu}_T) + \psi(\tilde{\theta})\| > \eta \Bigg] \\
&\quad \le \overline{\lim_{T\to\infty}} \mathsf{P}(\hat{\nu}_T \notin \mathcal{V}_{K,T}) \\
&\qquad + \overline{\lim_{T\to\infty}} \mathsf{P}\Bigg[\Bigg\{ \sup_{\theta_1,\theta_2 \in \Theta: \|\theta-\tilde{\theta}\| \le \delta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \nu_T) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \nu_T) + \psi(\tilde{\theta})\| > \eta \Bigg\} \cap \Big\{ \hat{\nu}_T \in \mathcal{V}_{K,T} \Big\} \Bigg] \\
&\quad \le \epsilon + \overline{\lim_{T\to\infty}} \mathsf{P}\Bigg[\sup_{\nu \in \mathcal{V}_{K,T},\, \theta,\tilde{\theta} \in \Theta: \|\theta-\tilde{\theta}\| \le \delta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \nu) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \nu) + \psi(\tilde{\theta})\| > \eta \Bigg].
\end{aligned}$$

Since, by Lemma S.3 and Lemma S.4, one has

$$\overline{\lim_{T\to\infty}} \mathsf{P}\Bigg[\sup_{\nu \in \mathcal{V}_{K,T}} \sup_{\theta \in \Theta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \nu) - \hat{\psi}_{T,S}(\theta, \nu_0)\| > \eta \Bigg] < \epsilon,$$

it suffices to show that

$$\overline{\lim_{T\to\infty}} \mathsf{P}\Bigg[\sup_{\theta,\tilde{\theta} \in \Theta: \|\theta-\tilde{\theta}\| \le \delta} \sqrt{T}\|\hat{\psi}_{T,S}(\theta, \nu_0) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \nu_0) + \psi(\tilde{\theta})\| > \eta \Bigg] < \epsilon. \tag{A.21}$$

To begin with, recall that $\theta_q = (\alpha_q', \beta_q', \gamma', \delta')'$, $q = 1, \dots, Q$. Thus, by a slight abuse of notation, $\theta = (\alpha_1', \beta_1', \dots, \alpha_Q', \beta_Q', \gamma', \delta')' = \cup_{q=1}^{Q}\theta_q$. Therefore, by the triangle inequality

$$\begin{aligned}
&\|\hat{\psi}_{T,S}(\theta, \nu_0) - \psi(\theta) - \hat{\psi}_{T,S}(\tilde{\theta}, \nu_0) + \psi(\tilde{\theta})\| \\
&\quad \le \sum_{q=1}^{Q} \frac{1}{\binom{|\mathcal{G}_q|}{2}} \sum\sum_{\substack{1 \le i<j \le n \\ i,j \in \mathcal{G}_q}} \sum_{k=1}^{\ell} |\hat{\psi}_{T,S,i,j,k}(\theta_q, \nu_0) - \psi_{q,k}(\theta_q) - \hat{\psi}_{T,S,i,j,k}(\tilde{\theta}_q, \nu_0) + \psi_{q,k}(\tilde{\theta}_q)|.
\end{aligned}$$

Hence,

$$\begin{aligned} &\mathsf{P}\left[\sup_{\theta,\tilde{\theta}\in\Theta:\|\theta-\tilde{\theta}\|\leq\delta}\sqrt{T}\|\hat{\psi}_{T,S}(\theta,\nu_0)-\psi(\theta)-\hat{\psi}_{T,S}(\tilde{\theta},\nu_0)+\psi(\tilde{\theta})\|>\eta\right] \\ &\quad\leq\sum_{q=1}^{Q}\sum\sum_{\substack{1\leq i<j\leq n\\ i,j\in\mathcal{G}_q}}\sum_{k=1}^{\ell}\mathsf{P}\left[\sup_{\theta_q,\tilde{\theta}_q\in\Theta:\|\theta-\tilde{\theta}\|\leq\delta}|\hat{\psi}_{T,S,i,j,k}(\theta_q,\nu_0)-\psi_{q,k}(\theta_q)\right. \\ &\qquad\qquad\left.-\hat{\psi}_{T,S,i,j,k}(\tilde{\theta}_q,\nu_0)+\psi_{q,k}(\tilde{\theta}_q)|>\frac{\eta}{Q\ell}\right]. \end{aligned}$$

Now, suppose $i,j\in\mathcal{G}_q$ for some $q\in\{1,\dots,Q\}$. Moreover, let us recall from Eq. (A.8) that uniformly in $\theta_q\in\Theta$

$$\begin{aligned} \sqrt{T}|\hat{\psi}_{T,S,i,j,k}(\theta_q,\nu_0)-&\psi_{q,k}(\theta_q)| \\ &\leq\left|\int_{[0,1]^2}\mathrm{d}\varphi_k(u_i,u_j)\right|\sup_{u_i,u_j\in[0,1]}|\hat{\mathbb{C}}_{i,j,T,S}(u_i,u_j;\theta_q)|+|R(\theta_q)| \end{aligned}\tag{A.22}$$

for $1\leq i<j\leq n$ and $k\in\{1,\dots,\ell\}$, where $\sup_{\theta\in\Theta}|R(\theta_q)|=o_p(1)$. Since n and ℓ are fixed, we are left with showing that for any $\epsilon,\eta>0$, there exists some $\delta>0$ such that

$$\overline{\lim_{T\to\infty}}\,\mathsf{P}\left[\sup_{\theta_q,\tilde{\theta}_q\in\Theta:\|\theta_q-\tilde{\theta}_q\|\leq\delta}\sup_{u_i,u_j\in[0,1]}|\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q)-\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\tilde{\theta}_q)|>\eta\right]<\epsilon.\tag{A.23}$$

In view of Eqs. (A.3) and (A.4), we see that

$$\begin{aligned} \hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q)=\sqrt{T}(&\tilde{\mathsf{C}}_{T,S,i,j}\{\tilde{\mathsf{C}}^{-}_{T,S,i}(u_i;\theta_q),\tilde{\mathsf{C}}^{-}_{T,S,j}(u_j;\theta_q);\theta_q\}-\mathsf{C}_{i,j}\{\tilde{\mathsf{C}}^{-}_{T,S,i}(u_i;\theta_q),\tilde{\mathsf{C}}^{-}_{T,S,j}(u_j;\theta_q);\theta_q\}) \\ &+\sqrt{T}(\mathsf{C}_{i,j}\{\tilde{\mathsf{C}}^{-}_{T,S,i}(u_i;\theta_q),\tilde{\mathsf{C}}^{-}_{T,S,j}(u_j;\theta_q);\theta_q\}-\mathsf{C}_q(u_i,u_j)). \end{aligned}$$

with $\tilde{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta_q):=\tilde{\mathsf{C}}_{T,S,i,j}(u_i,u_j;\theta,\nu_0)$ and $\tilde{\mathsf{C}}^{-}_{T,S,k}(u_k;\theta_q):=\tilde{\mathsf{C}}^{-}_{T,S,k}(u_k;\theta_q,\nu_0)$ for $k\in\{i,j\}$. Therefore, using an argument similar that in Tsukahara (2005, Appendix B), one obtains

$$\begin{aligned} &\mathsf{P}\left[\sup_{\theta_q,\tilde{\theta}_q\in\Theta:\|\theta_q-\tilde{\theta}_q\|\leq\delta}\sup_{u_i,u_j\in[0,1]}|\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q)-\hat{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\tilde{\theta}_q)|>\eta\right] \\ &\quad\leq\mathsf{P}\left[\sup_{\theta_q,\tilde{\theta}_q\in\Theta:\|\theta_q-\tilde{\theta}_q\|\leq\delta}\sup_{u_i,u_j\in[0,1]}|\tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\theta_q)-\tilde{\mathbb{C}}_{T,S,i,j}(u_i,u_j;\tilde{\theta}_q)|>\frac{\eta}{2}\right] \\ &\qquad+\sum_{k\in\{i,j\}}\mathsf{P}\left[\sup_{\theta_q,\tilde{\theta}_q\in\Theta:\|\theta_q-\tilde{\theta}_q\|\leq\delta}\sup_{u_k\in[0,1]}|\tilde{\mathbb{C}}_{T,S,i,j}(\vec{u}_k;\theta_q)-\tilde{\mathbb{C}}_{T,S,i,j}(\vec{u}_k;\tilde{\theta}_q)|>\frac{\eta}{4}\right]. \end{aligned}\tag{A.24}$$

The claim thus follows from part (c) of Lemma S.2.

A.2. Proof of Corollary 1

Proof of Corollary 1. As argued in Cheng (2015, Lemma 1), it suffices to show that the conditional distribution of $\sqrt{T}(\hat{\Psi}^{(b)}_{T,S}(\hat{\theta}_{T,S},\hat{\lambda}_T,\hat{\nu}_T)-\hat{\Psi}_{T,S}(\hat{\theta}_{T,S},\hat{\lambda}_T,\hat{\nu}_T))$ converges in probability to the limiting distribution of $\sqrt{T}\hat{\Psi}_{T,S}(\theta_0,\hat{\lambda}_T,\hat{\nu}_T)$ given in Proposition 1. Specifically, define for any $1\leq i<j\leq n$, with $i,j\in\mathcal{G}_q$, $q\in\{1,\dots,Q\}$:

$$\begin{aligned} \sqrt{T}(\hat{\Psi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)-&\hat{\Psi}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)) \\ &=\hat{\xi}^{(b)}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T) \\ &\quad+\hat{\xi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)-\hat{\xi}^{(b)}_{i,j,T,S}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T), \end{aligned}$$

where

$$\hat{\xi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T):=\sqrt{T}(\hat{\Psi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)-\hat{\Psi}_{i,j,T,S}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T))\tag{A.25}$$

represents the bootstrap analogue of

$$\hat{\xi}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)=\sqrt{T}(\hat{\Psi}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)-\Psi_{T,S,i,j}(\hat{\theta}_{T,S,q})).\tag{A.26}$$

We proceed in two steps: (1) the conditional distribution of $\hat{\xi}^{(b)}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T)$ converges in probability to that of $\hat{\xi}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T)=\sqrt{T}\hat{\Psi}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T)$; (2) $\hat{\xi}^{(b)}_{T,S,i,j}(\hat{\theta}_{T,S,q},\hat{\lambda}_T,\hat{\nu}_T)-\hat{\xi}^{(b)}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T)=o_p(1)$. *Step (1):* Mimicking the derivation of Eq. (3.12), note that the kth element of $\hat{\xi}^{(b)}_{T,S,i,j}(\theta_{0,q},\hat{\lambda}_T,\hat{\nu}_T)$ can be written as:

$$\int_{[0,1]^2}(\hat{\mathbb{B}}^{(b)}_{T,S,i,j}-\hat{\mathbb{B}}_{T,S,i,j})(u_i,u_j;\theta_{0,q},\lambda_0,\nu_0)\,\mathrm{d}\varphi_k(u_i,u_j)+o_p(1),$$

where $\hat{\mathbb{B}}^{(b)}_{T,S,i,j}$ is the bootstrap analogue of $\hat{\mathbb{B}}_{T,S,i,j}$ defined in Eq. (3.11). We are thus left with showing that the conditional distribution of $\hat{\mathbb{B}}^{(b)}_{T,S,i,j} - \hat{\mathbb{B}}_{T,S,i,j}$ converges weakly in probability to the same limiting process that governs the weak limit of $\hat{\mathbb{B}}_{T,S,i,j}$. But, by the functional delta method, this is due the weak convergence of $\hat{\mathbb{B}}_{T,S,i,j}$; see the proof of Proposition 1 in conjunction with Fermanian et al. (2004, Theorem 5). *Step (2):* We can deduce the bootstrap stochastic equicontinuity of (A.25) from that of (A.26) established in Proposition 1 (c); see also Giné and Zinn (1990), Brown and Wegkamp (2002), or Chen et al. (2003) for a similar argument. The claim then follows because, by Proposition 2, $\sqrt{T}\|\hat{\theta}_{T,S} - \theta_0\| = O_p(1)$.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2023.01.003.

References

Akritas, M.G., Van Keilegom, I., 2001. Non-parametric estimation of the residual distribution. Scand. J. Statist. 28, 549–567. http://dx.doi.org/10.1111/1467-9469.00254.

Andrews, D.W.K., Pollard, D., 1994. An introduction to functional central limit theorems for dependent stochastic processes. Internat. Statist. Rev. 62, 119–132. http://dx.doi.org/10.2307/1403549.

Bai, J., Wang, P., 2015. Identification and bayesian estimation of dynamic factor models. J. Bus. Econom. Statist. 33, 221–240. http://dx.doi.org/10.1080/07350015.2014.941467.

Baur, D.G., McDermott, T.K., 2010. Is gold a safe haven? International evidence. J. Bank. Financ. 34, 1886–1898. http://dx.doi.org/10.1016/j.jbankfin.2009.12.008.

Berghaus, B., Bücher, A., Volgushev, S., 2017. Weak convergence of the empirical copula process with respect to weighted metrics. Bernoulli 23, 743–772. http://dx.doi.org/10.3150/15-BEJ751.

Bernanke, B.S., Boivin, J., Eliasz, P., 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. Q. J. Econ. 120, 387–422. http://dx.doi.org/10.1162/0033553053327452.

Boistard, H., Lopuhaä, H.P., Ruiz-Gazen, A., 2017. Functional central limit theorems for single-stage sampling designs. Ann. Statist. 45, 1728–1758. http://dx.doi.org/10.1214/16-AOS1507.

Boivin, J., Giannoni, M.P., Mihov, I., 2009. Sticky prices and monetary policy: evidence from disaggregated US data. Amer. Econ. Rev. 99, 350–384. http://dx.doi.org/10.1257/aer.99.1.350.

Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH model. In: Engle, R.F., McFadden, D. (Eds.), Handbook of Econometrics, Vol. 4. North-Holland, Amsterdam.

Brown, D.J., Wegkamp, M.H., 2002. Weighted minimum mean-square distance from independence estimation. Econometrica 70, 2035–2051. http://dx.doi.org/10.1111/1468-0262.00362.

Bücher, A., Segers, J., 2013. Extreme value copula estimation based on block maxima of a multivariate stationary time series. Extremes 17, 495–528. http://dx.doi.org/10.1007/s10687-014-0195-8.

Bücher, A., Volgushev, S., 2013. Empirical and sequential empirical copula processes under serial dependence. J. Multivariate Anal. 119, 61–70. http://dx.doi.org/10.1016/j.jmva.2013.04.003.

Caporale, G.M., Ntantamis, C., Pantelidis, T., Pittis, N., 2005. The BDS test as a test for the adequacy of a GARCH(1, 1) specification: A Monte Carlo study. J. Financ. Econom. 3, 282–309. http://dx.doi.org/10.1093/jjfinec/nbi010.

Carrasco, M., Chen, X., 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econom. Theory 18, 17–39. http://dx.doi.org/10.1017/S0266466602181023.

Chen, X., Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. J. Econometrics 135, 125–154. http://dx.doi.org/10.1016/j.jeconom.2005.07.027.

Chen, X., Huang, Z., Yi, Y., 2020. Efficient estimation of multivariate semi-nonparametric GARCH filtered copula models. J. Econometrics http://dx.doi.org/10.1016/j.jeconom.2020.07.012.

Chen, X., Linton, O., Van Keilegom, I., 2003. Estimation of semiparametric models when the criterion function is not smooth. Econometrica 71, 1591–1608. http://dx.doi.org/10.1111/1468-0262.00461.

Cheng, G., 2015. Moment consistency of the exchangeably weighted bootstrap for semiparametric M-estimation. Scand. J. Statist. 42, 665–684. http://dx.doi.org/10.1111/sjos.12128.

Chernozhukov, V., Hong, H., 2003. An MCMC approach to classical estimation. J. Econometrics 115, 293–346. http://dx.doi.org/10.1016/S0304-4076(03)00100-3.

Côté, M.-P., Genest, C., Omelka, M., 2019. Rank-based inference tools for copula regression, with property and casualty insurance applications. Insurance Math. Econom. 89, 1–15. http://dx.doi.org/10.1016/j.insmatheco.2019.08.001.

Creal, D.D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. J. Appl. Econometrics 28, 777–795. http://dx.doi.org/10.1002/jae.1279.

Creal, D.D., Tsay, R.S., 2015. High dimensional dynamic stochastic copula models. J. Econometrics 189, 335–345. http://dx.doi.org/10.1016/j.jeconom.2015.03.027.

D'Errico, J., 2021. Fminsearchbnd, fminsearchcon. MATLAB Central File Exchange, Retrieved April 17, 2021.

Dette, H., Pardo-Fernández, J.C., Keilegom, V.I., 2009. Goodness-of-fit tests for multiplicative models with dependent data. Scand. J. Statist. 36, 782–799. http://dx.doi.org/10.1111/j.1467-9469.2009.00648.x.

Doukhan, P., Massart, P., Rio, E., 1995. Invariance principles for absolutely regular empirical processes. Ann. l'Inst. H. P–Probab. Statist. 31, 393–427.

Fan, Y., Patton, A.J., 2014. Copulas in econometrics. Annu. Rev. Econ. 6, 179–200. http://dx.doi.org/10.1146/annurev-economics-080213-041221.

Fermanian, J.-D., Radulović, D., Wegkamp, M.H., 2004. Weak convergence of empirical copula processes. Bernoulli 10, 847–860. http://dx.doi.org/10.3150/bj/1099579158.

Francq, C., Zakoïan, J.-M., 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. Bernoulli 4, 605–637. http://dx.doi.org/10.3150/bj/1093265632.

Fryzlewicz, P., Rao, S.S., 2011. Mixing properties of ARCH and time-varying ARCH processes. Bernoulli 17, 320–346. http://dx.doi.org/10.3150/10-BEJ270.

Genest, C., Ghoudi, K., Rémillard, B., 2007. Rank-based extensions of the Brock, Dechert, and Scheinkman test. J. Amer. Statist. Assoc. 102, 1363–1376. http://dx.doi.org/10.1198/016214507000001076.

Giné, E., Zinn, J., 1990. Bootstrapping general empirical measures. Ann. Probab. 18, 851–869. http://dx.doi.org/10.1214/aop/1176990862.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48, 1779–1801. http://dx.doi.org/10.1111/j.1540-6261.1993.tb05128.x.
Gonçalves, S., Hounyo, U., Patton, A.J., Sheppard, K., 2023. Bootstrapping two-stage quasi-maximum likelihood estimators of time series models. J. Bus. Econom. Statist. http://dx.doi.org/10.1080/07350015.2022.2058949, (forthcoming).
Gouriéroux, C., Monfort, A., 1997. Simulation-Based Econometric Methods. In: CORE Lectures, Oxford University Press, Oxford.
Hahn, J., Liao, Z., 2021. Bootstrap standard error estimates and inference. Econometrica 89, 1963–1977. http://dx.doi.org/10.3982/ECTA17912.
Hansen, B., 1994. Autoregressive conditional densityestimation. Internat. Econom. Rev. 35, 705–730. http://dx.doi.org/10.2307/2527081.
Hong, H., Li, H., Li, J., 2021. BLP estimation using Laplace transformation and overlapping simulation draws. J. Econometrics 222, 56–72. http://dx.doi.org/10.1016/j.jeconom.2020.07.026.
Kosorok, M.R., 2008. Introduction to Empirical Processes and Semiparametric Inference. In: Springer Series in Statistics, Springer, Berlin.
Krupskii, P., Joe, H., 2013. Factor copula models for multivariate data. J. Multivariate Anal. 120, 85–101. http://dx.doi.org/10.1016/j.jmva.2013.05.001.
Krupskii, P., Joe, H., 2020. Flexible copula models with dynamic dependence and application to financial data. Econ. Statist. 16, 148–167. http://dx.doi.org/10.1016/j.ecosta.2020.01.00.
Lee, J.F., 1992. On efficiency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models. Econom. Theory 8, 518–552. http://dx.doi.org/10.1017/S0266466600013207.
Liu, R., Yang, L., 2016. Spline estimation of a semiparametric garch model. Econom. Theory 32, 1023–1054. http://dx.doi.org/10.1017/S0266466615000055.
Manner, H., Stark, F., Wied, D., 2019. Testing for structural breaks in factor copula models. J. Econometrics 208, 324–345. http://dx.doi.org/10.1016/j.jeconom.2018.10.001.
Manner, H., Stark, F., Wied, D., 2021. A monitoring procedure for detecting structural breaks in factor copula models. Stud. Nonlinear Dyn. Econom. http://dx.doi.org/10.1515/snde-2019-0081.
McFadden, D., 1989. A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica 57, 995–1026. http://dx.doi.org/10.2307/1913621.
Neumeyer, N., Omelka, M., Hudecová, Š., 2019. A copula approach for dependence modeling in multivariate nonparametric time series. J. Multivariate Anal. 171, 139–162. http://dx.doi.org/10.1016/j.jmva.2018.11.016.
Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R., McFadden, D. (Eds.), Handbook of Econometrics, Vol. 4. North Holland, Amsterdam, pp. 2113–2247. http://dx.doi.org/10.1016/S1573-4412(05)80005-4.
Oh, D.H., Patton, A.J., 2013. Simulated method of moments estimation for copula-based multivariate models. J. Amer. Statist. Assoc. 108, 689–700. http://dx.doi.org/10.1080/01621459.2013.785952.
Oh, D.H., Patton, A.J., 2017. Modeling dependence in high dimensions with factor copulas. J. Bus. Econom. Statist. 35, 139–154. http://dx.doi.org/10.1080/07350015.2015.1062384.
Oh, D.H., Patton, A.J., 2018. Time-varying systemic risk: evidence from a dynamic copula model of CDS spreads. J. Bus. Econom. Statist. 36, 181–195. http://dx.doi.org/10.1080/07350015.2016.1177535.
Oh, D.H., Patton, A.J., 2021. Factor Copula Models with Estimated Cluster Assignments. In: Finance and Economics Discussion Series, 2021-029, Board of Governors of the Federal Reserve System, Washington, http://dx.doi.org/10.17016/FEDS.2021.029.
Omelka, M., Neumeyer, N., Hudecová, Š., 2020. Maximum pseudo-likelihood estimation based on estimated residuals in copula semiparametric models. Scand. J. Statist. 1–41. http://dx.doi.org/10.1111/sjos.12498.
Opschoor, A., Lucas, A., Barra, I., van Dijk, D., 2020. Closed-form multi-factor copula models with observation-driven dynamic factor loadings. J. Bus. Econom. Statist. http://dx.doi.org/10.1080/07350015.2020.1763806.
Owen, A.B., 2005. Multidimensional variation for quasi-Monte Carlo. In: Fan, J., Li, G. (Eds.), Int. Conf. Statistics in Honour of Professor Kai-Tai Fang's 65th Birthday, Vol. 4. Hong Kong Baptist University, Hong Kong.
Owen, A.B., Rudolf, D., 2020. A strong law of large numbers for scrambled net integration. SIAM Rev. 63, 360–372. http://dx.doi.org/10.1137/20M1320535.
Pakes, A., Pollard, D., 1989. Simulation and the asymptotics of optimization estimators. Econometrica 57, 1027–1057. http://dx.doi.org/10.2307/1913622.
Patton, A., 2006. Modelling asymmetric exchange rate dependence. Internat. Econom. Rev. 47, 527–556. http://dx.doi.org/10.1111/j.1468-2354.2006.00387.x.
Radulović, D., Wegkamp, M., Zhao, Y., 2017. Weak convergence of empirical copula processes indexed by functions. Bernoulli 23, 3346–3384. http://dx.doi.org/10.3150/16-BEJ849.
Rémillard, B., 2017. Goodness-of-fit tests for copulas of multivariate time series. Econometrics 5, 1–23. http://dx.doi.org/10.3390/econometrics5010013.
Segers, J., 2012. Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. Bernoulli 18, 764–782. http://dx.doi.org/10.3150/11-BEJ387.
Stock, J.H., Watson, M.W., 2005. Implications of Dynamic Factor Models for VAR Analysis. NBER Working Paper, 11467, http://dx.doi.org/10.3386/w11467.
Tsukahara, H., 2005. Semiparametric estimation in copula models. Canad. J. Statist. 33, 357–375. http://dx.doi.org/10.1002/cjs.5540330304.
van der Vaart, A.W., 1994. Asymptotic Statistics. In: Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.
van der Vaart, A.W., Wellner, J.A., 1996. Weak Convergence and Empirical Processes. Springer, New York.
White, H., 2001. Asymptotic Theory for Econometricians, second ed. Emerald, Bingley.
Yang, L., 2006. A semiparametric GARCH model for foreign exchange volatility. J. Econometrics 130, 365–384. http://dx.doi.org/10.1016/j.jeconom.2005.03.006.