Online Appendix for

Gaussian rank correlation and regression

Dante Amengual

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain

<amengual@cemfi.es>

Enrique Sentana

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain

<sentana@cemfi.es>

Zhanyuan Tian

Boston University

<zhanyuan@bu.edu>

July 2020

B Proofs

For the sake of brevity, Assumption 1 is maintained throughout.

Proposition 1

Consider the differential of (2),

$$\begin{aligned}
\mathrm{d}\phi_K(\mathbf{y}_i,\boldsymbol{\rho}) &= -\frac{1}{2}tr\{\mathbf{P}^{-1}(\boldsymbol{\rho})\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} - \frac{1}{2}\mathbf{y}_i'\mathrm{d}[\mathbf{P}^{-1}(\boldsymbol{\rho})]\mathbf{y}_i \\
&= -\frac{1}{2}vec[\mathbf{P}^{-1}(\boldsymbol{\rho})]'vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} + \frac{1}{2}\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i \\
&= -\frac{1}{2}vec[\mathbf{P}^{-1}(\boldsymbol{\rho})]'vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} + \frac{1}{2}[\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\otimes\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})]vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} \\
&= -\frac{1}{2}\{vec[\mathbf{P}^{-1}(\boldsymbol{\rho})]' - \mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\otimes\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\}vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} \\
&= -\frac{1}{2}\{vec[\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{P}(\boldsymbol{\rho})\mathbf{P}^{-1}(\boldsymbol{\rho})]' - vec[\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})]'\}vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} \\
&= \frac{1}{2}\{[\mathbf{P}^{-1}(\boldsymbol{\rho})\otimes\mathbf{P}^{-1}(\boldsymbol{\rho})]vec[\mathbf{y}_i\mathbf{y}_i' - \mathbf{P}(\boldsymbol{\rho})]\}'vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\}
\end{aligned}$$

Transposing and simplifying terms, we can get the fist order condition:

$$\frac{1}{2}\frac{\partial vec'[\mathbf{P}(\boldsymbol{\rho})]}{\partial\boldsymbol{\rho}}[\mathbf{P}^{-1}(\boldsymbol{\rho})\otimes\mathbf{P}^{-1}(\boldsymbol{\rho})]vec[\mathbf{y}_i\mathbf{y}_i' - \mathbf{P}(\boldsymbol{\rho})] = \mathbf{0}.$$

Then, using the fact that

$$vec[\mathbf{P}(\boldsymbol{\rho})] = vec(\mathbf{I}_K) + \tilde{\mathbf{L}}'\boldsymbol{\rho} + \mathbf{K}\tilde{\mathbf{L}}'\boldsymbol{\rho},$$

so that

$$\frac{\partial\mathbf{P}(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}'} = \tilde{\mathbf{L}}' + \mathbf{K}\tilde{\mathbf{L}}',$$

the result follows. □

Proposition 2

Starting from the expression for the score in (3), we can derive Hessian matrix by differencing once again, namely

$$\begin{aligned}
\mathrm{d}\mathbf{s}_{Ki}(\boldsymbol{\rho}) &= \frac{1}{2}(\tilde{\mathbf{L}} + \tilde{\mathbf{L}}\mathbf{K})[\mathrm{d}\{vec[\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})]\} - \mathrm{d}\{vec[\mathbf{P}(\boldsymbol{\rho})]\}] \\
&= \frac{1}{2}(\tilde{\mathbf{L}} + \tilde{\mathbf{L}}\mathbf{K})[vec\{\mathrm{d}[\mathbf{P}^{-1}(\boldsymbol{\rho})]\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\} \\
&\quad + vec\{\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathrm{d}[\mathbf{P}^{-1}(\boldsymbol{\rho})]\} - \mathrm{d}\{vec[\mathbf{P}^{-1}(\boldsymbol{\rho})]\}] \\
&= \frac{1}{2}(\tilde{\mathbf{L}} + \tilde{\mathbf{L}}\mathbf{K})[-vec\{\mathbf{P}^{-1}(\boldsymbol{\rho})\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\} \\
&\quad - vec\{\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\mathbf{P}^{-1}(\boldsymbol{\rho})\} + vec\{\mathbf{P}^{-1}(\boldsymbol{\rho})\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\mathbf{P}^{-1}(\boldsymbol{\rho})\}] \\
&= \frac{1}{2}(\tilde{\mathbf{L}} + \tilde{\mathbf{L}}\mathbf{K})\{-[\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})]\otimes\mathbf{P}^{-1}(\boldsymbol{\rho}) \\
&\quad - \mathbf{P}^{-1}(\boldsymbol{\rho})\otimes[\mathbf{P}^{-1}(\boldsymbol{\rho})\mathbf{y}_i\mathbf{y}_i'\mathbf{P}^{-1}(\boldsymbol{\rho})] + \mathbf{P}^{-1}(\boldsymbol{\rho})\otimes\mathbf{P}^{-1}(\boldsymbol{\rho})\}vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\}.
\end{aligned}$$

Hence, $\mathrm{d}\mathbf{s}_{Ki}(\boldsymbol{\rho})$ can be written as in (6) after noticing that $vec\{\mathrm{d}[\mathbf{P}(\boldsymbol{\rho})]\} = (\tilde{\mathbf{L}} + \tilde{\mathbf{L}}\mathbf{K})$. □

Proposition 3

Part (a) follows from the *i.i.d.* assumption on $\{\mathbf{y}_i\}$ together with either expression (10) or (11), while part (b) is a direct application of the Delta method to the result in part (a). $\square$

Lemma 1

As in Chen and Fan (2006), we need to compute

$$n_2 = \int_0^1 [1\{u_1 \leq U_1\} - u_1] W^1_{\rho_{12}} dU_1 + \int_0^1 [1\{u_2 \leq U_2\} - u_2] W^2_{\rho_{12}} dU_2,$$

with $W^j_{\rho_{12}} = \int [\partial s_{\rho_{12}}(u_1, u_2; \rho_{12})/\partial u_j] c(u_1, u_2; \rho_{12}) du_j$ for $j = 1, 2$. Then, the result follows from

$$W^j_{\rho_{12}} = \int \left[\frac{1+\rho_{12}^2}{(1-\rho_{12}^2)^2} y_{-j} - \frac{2\rho_{12}}{(1-\rho_{12}^2)^2} y_j \right] \phi(y_j) dy_{-j} = \frac{1+\rho_{12}^2}{(1-\rho_{12}^2)^2} y_j$$

and the fact that

$$\int_{-\infty}^{y} H_1(x)\Phi(x)dx = \frac{H_1(y)}{2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) + \frac{1}{2\sqrt{2}} H_2(y) \left[1 + \text{erf}\left(\frac{y}{\sqrt{2}}\right)\right]$$

and

$$\int_{y}^{\infty} H_1(x)[1-\Phi(x)]dx = \frac{H_1(y)}{2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) - \frac{1}{2\sqrt{2}} H_2(y) \left[\text{erfc}\left(\frac{y}{\sqrt{2}}\right)\right],$$

where $\text{erf}(z) = 2\pi^{-1/2} \int_0^z e^{-t^2} dt$ and $\text{erfc}(z) = 1 - \text{erf}(z)$. Analogous calculations in the trivariate case yield the corresponding correction. $\square$

Proposition 4

Under the maintained assumption of a Gaussian copula, it is straightforward to obtain the variance of the score in (4) using the moments of the bivariate normal, whose reciprocal is $AVar(\hat{\rho})$.

To obtain the asymptotic variance of $\check{\rho} = \sum_i y_{1i}y_{2i} / \sum_i y_{2i}^2$, consider the following vector of influence functions:

$$\check{\mathbf{m}}_{2i}(\boldsymbol{\theta}) = (y_{1i}y_{2i} - \sigma_2^2\rho, y_{2i}^2 - \sigma_2^2)'$$

where $\boldsymbol{\theta} = (\rho, \sigma_2^2)'$. Then, we can easily compute

$$\check{\mathbf{A}}_2 = E\left[\frac{\partial \check{\mathbf{m}}_{2i}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} \sigma_2^2 & \rho \\ 0 & 1 \end{pmatrix} \text{ and } \check{\mathbf{B}}_2 = Var[\check{\mathbf{m}}_{2i}(\boldsymbol{\theta})] = \begin{pmatrix} 1+\rho^2 & 2\rho \\ 2\rho & 2 \end{pmatrix},$$

so that imposing $\sigma^2 = 1$ and applying the sandwich formula yields $Var(\check{\rho}) = (1-\rho^2)^2/(1+\rho^2)$ as the $(1,1)$ element of $\check{\mathbf{A}}_2^{-1}\check{\mathbf{B}}_2\check{\mathbf{A}}_2^{-1\prime}$.

As for $\tilde{\rho} = \sum_i (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2) / \sum_i y_{2i}^2$, where $\bar{y}_j = N^{-1} \sum_i y_{ji}$, we consider the following alternative vector of influence functions:

$$\tilde{\mathbf{m}}_{2i}(\boldsymbol{\theta}) = [y_{1i}y_{2i} - (\mu_1\mu_2 + \sigma_2^2\rho), y_{2i}^2 - (\mu_2^2 + \sigma_2^2), y_{1i} - \mu_1, y_{2i} - \mu_2]'$$

where $\boldsymbol{\theta} = (\rho, \sigma_2^2, \mu_1, \mu_2)'$. Then, we can compute

$$\tilde{\mathbf{A}}_2 = E\left[\frac{\partial \tilde{\mathbf{m}}_{2i}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} \sigma_2^2 & \rho & \mu_2 & \mu_1 \\ 0 & 1 & 0 & 2\mu_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and } \tilde{\mathbf{B}}_2 = Var\left[\tilde{\mathbf{m}}_{2i}(\boldsymbol{\theta})\right] = \begin{pmatrix} \check{\mathbf{B}}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}(\rho) \end{pmatrix},$$

so that imposing the Gaussian copula assumption and applying the sandwich formula yields $Var(\tilde{\rho}) = (1-\rho^2)^2$ as the $(1,1)$ element of $\tilde{\mathbf{A}}_2^{-1}\tilde{\mathbf{B}}_2\tilde{\mathbf{A}}_2^{-1\prime}$. □

Proposition 5

Analogous calculations to the ones used in the proof of Lemma 1 allow us to obtain

$$n_{\mu_j i}(\rho) = H_1(y_{ji}) \quad \text{and} \quad n_{\sigma_j^2 i}(\rho) = \sqrt{2}H_2(y_{ji}) \quad \text{for} \quad j = 1, 2.$$

Hence, the asymptotic variance of the ML estimator of ρ can be obtained as

$$AVar(\hat{\rho}^{np}) = \frac{Var[s_{\rho i}^{np}(\rho)]}{\{Var[s_{\rho i}(\rho)]\}^2} = (1-\rho^2)^2.$$

As for the other estimators, letting $\mathbf{B}_2^{np} = Var[\mathbf{m}_{2i}(\boldsymbol{\theta}) + \mathbf{n}_{2i}(\boldsymbol{\theta})]$, we can show that $AVar(\hat{\rho}^{np})$ coincides with the common asymptotic variance of $\check{\rho}^{np}$ and $\tilde{\rho}^{np}$, which is given by

$$AVar(\tilde{\rho}^{np}) = AVar(\check{\rho}^{np}) = (1-\rho^2)^2$$

because $\check{\mathbf{B}}_2^{np}$ and $\tilde{\mathbf{B}}_2^{np}$ have all the elements equal to zero except the $(1,1)$ one, which is equal to $(1-\rho^2)$. □

Proposition 6

First, we can obtain the asymptotic variance of $\hat{\boldsymbol{\rho}}$ as $AVar(\hat{\boldsymbol{\rho}}) = \mathcal{A}^{-1}(\boldsymbol{\rho})$, where the expressions for the expected (minus) Hessian $\mathcal{A}(\boldsymbol{\rho})$ are reported in Online Appendix C. Then, regarding the ML estimator $\hat{\boldsymbol{\beta}}^{(1)}$, we can exploit

$$\hat{\boldsymbol{\beta}}^{(1)} = \begin{bmatrix} 1 & \hat{\rho}_{23} \\ \hat{\rho}_{23} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\rho}_{12} \\ \hat{\rho}_{13} \end{bmatrix}$$

to obtain the asymptotic variance of say $\hat{\beta}_2^{(1)}$ by applying the Delta method, namely

$$AVar(\hat{\beta}_2^{(1)}) = \nabla\hat{\beta}_2^{(1)\prime}(\boldsymbol{\rho})AVar(\hat{\boldsymbol{\rho}})\nabla\hat{\beta}_2^{(1)}(\boldsymbol{\rho})$$

where

$$\nabla\hat{\beta}_2^{(1)}(\boldsymbol{\rho}) = \left[\frac{1}{1-\rho_{23}^2}, -\frac{\rho_{23}}{1-\rho_{23}^2}, \frac{2\rho_{12}\rho_{23} - \rho_{13}(1+\rho_{23}^2)}{(1-\rho_{23}^2)^2}\right]'.$$

This yields

$$
\begin{aligned}
AVar(\hat{\beta}_2^{(1)}) \;\; = \;\; & -(((-1+\rho_{23}^2)^3(1+\rho_{23}^2)+\rho_{12}^6(1+3\rho_{23}^2)-2\rho_{12}^5\rho_{13}\rho_{23}(5+7\rho_{23}^2) \\
& +\rho_{13}^6(-1+3\rho_{23}^2+2\rho_{23}^4)+\rho_{13}^2(-1+\rho_{23}^2)^2(1+3\rho_{23}^2+2\rho_{23}^4) \\
& +\rho_{13}^4(1-6\rho_{23}^2+\rho_{23}^4+4\rho_{23}^6)-2\rho_{12}\rho_{13}\rho_{23}(3(-1+\rho_{23}^2)^2(1+\rho_{23}^2) \\
& +4\rho_{13}^2(-1+\rho_{23}^2)(1+\rho_{23}^2)^2+\rho_{13}^4(1+7\rho_{23}^2+4\rho_{23}^4)) \\
& -4\rho_{12}^3\rho_{13}\rho_{23}(4(-1+\rho_{23}^4)+\rho_{13}^2(4+11\rho_{23}^2+5\rho_{23}^4)) \\
& +\rho_{12}^4(-3-2\rho_{23}^2+5\rho_{23}^4+3\rho_{13}^2(1+11\rho_{23}^2+8\rho_{23}^4)) \\
& +\rho_{12}^2(3(-1+\rho_{23}^2)^2(1+\rho_{23}^2)+\rho_{13}^4(1+25\rho_{23}^2+26\rho_{23}^4+8\rho_{23}^6)+ \\
& 2\rho_{13}^2(-2-11\rho_{23}^2+4\rho_{23}^4+9\rho_{23}^6))) \\
& /((-1+\rho_{23}^2)^3(-1+\rho_{12}^4+\rho_{13}^4-2\rho_{12}^2\rho_{13}^2\rho_{23}^2+\rho_{23}^4))).
\end{aligned}
$$

In turn, to obtain the asymptotic variance of

$$
\breve{\boldsymbol{\beta}}^{(1)} = \begin{pmatrix} \sum_i y_{2i}^2 & \sum_i y_{2i}y_{3i} \\ \sum_i y_{2i}y_{3i} & \sum_i y_{3i}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_i y_{1i}y_{2i} \\ \sum_i y_{1i}y_{3i} \end{pmatrix}
$$

we consider the following vector of influence functions:

$$
\breve{\mathbf{m}}_{3i}(\boldsymbol{\theta}) = (y_{1i}y_{2i}-\sqrt{\sigma_1^2\sigma_2^2}\rho_{12}, y_{1i}y_{3i}-\sqrt{\sigma_1^2\sigma_3^2}\rho_{13}, y_{2i}y_{3i}-\sqrt{\sigma_2^2\sigma_3^2}\rho_{23}, y_{1i}^2-\sigma_1^2, y_{2i}^2-\sigma_2^2, y_{3i}^2-\sigma_3^2)'
$$

where $\boldsymbol{\theta} = (\rho_{12}, \rho_{13}, \rho_{23}, \sigma_1^2, \sigma_2^2, \sigma_2^3)'$. Then, under the assumption of a Gaussian copula we will have

$$
\breve{\mathbf{A}}_3 = E\left[\frac{\partial \breve{\mathbf{m}}_{3i}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} 1 & 0 & 0 & \rho_{12}/2 & \rho_{12}/2 & 0 \\ 0 & 1 & 0 & \rho_{13}/2 & 0 & \rho_{13}/2 \\ 0 & 0 & 1 & 0 & \rho_{23}/2 & \rho_{23}/2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

and

$$
\breve{\mathbf{B}}_3 = Var[\breve{\mathbf{m}}_{3i}(\boldsymbol{\theta})] = \begin{pmatrix} 1+\rho_{12}^2 & \rho_{12}\rho_{13}+\rho_{23} & \rho_{12}\rho_{23}+\rho_{13} & 2\rho_{12} & 2\rho_{12} & 2\rho_{13}\rho_{23} \\ \rho_{12}\rho_{13}+\rho_{23} & 1+\rho_{13}^2 & \rho_{13}\rho_{23}+\rho_{12} & 2\rho_{13} & 2\rho_{12}\rho_{23} & 2\rho_{13} \\ \rho_{12}\rho_{23}+\rho_{13} & \rho_{13}\rho_{23}+\rho_{12} & 1+\rho_{23}^2 & 2\rho_{12}\rho_{13} & 2\rho_{23} & 2\rho_{23} \\ 2\rho_{12} & 2\rho_{13} & 2\rho_{12}\rho_{13} & 3 & 2\rho_{12}^2 & 2\rho_{13}^2 \\ 2\rho_{12} & 2\rho_{12}\rho_{23} & 2\rho_{23} & 2\rho_{12}^2 & 3 & 2\rho_{23}^2 \\ 2\rho_{13}\rho_{23} & 2\rho_{13} & 2\rho_{23} & 2\rho_{13}^2 & 2\rho_{23}^2 & 3 \end{pmatrix},
$$

which allow us to obtain $AVar(\breve{\boldsymbol{\theta}})$ as $\breve{\mathbf{A}}_3^{-1}\breve{\mathbf{B}}_3\breve{\mathbf{A}}_3^{-1\prime}$. We can then use the Delta method to obtain the asymptotic variance of $\breve{\beta}_2^{(1)}$. Specifically, we have

$$
\nabla\breve{\beta}_2^{(1)}(\boldsymbol{\theta}) = \left[\frac{1}{1-\rho_{23}^2}, \frac{-\rho_{23}}{1-\rho_{23}^2}, \frac{2\rho_{12}\rho_{23}-\rho_{13}-\rho_{13}\rho_{23}^2}{(1-\rho_{23}^2)^2}, \frac{\rho_{12}-\rho_{13}\rho_{23}}{2(1-\rho_{23}^2)}, -\frac{\rho_{12}-\rho_{13}\rho_{23}}{2(1-\rho_{23}^2)}, 0\right]',
$$

and therefore

$$
AVar(\breve{\beta}_2^{(1)}) = \frac{1-\rho_{12}^2-\rho_{13}^2+2\rho_{12}\rho_{13}\rho_{23}-\rho_{23}^2}{(1-\rho_{23}^2)^2}.
$$

Finally, to obtain the asymptotic variance of

$$\tilde{\boldsymbol{\beta}}^{(1)} = \begin{pmatrix} N^{-1}\sum_i y_{2i}^2 - \bar{y}_2^2 & N^{-1}\sum_i y_{2i}y_{3i} - \bar{y}_2\bar{y}_3 \\ N^{-1}\sum_i y_{2i}y_{3i} - \bar{y}_2\bar{y}_3 & N^{-1}\sum_i y_{3i}^2 - \bar{y}_3^2 \end{pmatrix}^{-1} \begin{pmatrix} N^{-1}\sum_i y_{1i}y_{2i} - \bar{y}_1\bar{y}_2 \\ N^{-1}\sum_i y_{1i}y_{3i} - \bar{y}_1\bar{y}_3 \end{pmatrix}$$

we consider the following vector of influence functions:

$$\tilde{\mathbf{m}}_{3i}(\boldsymbol{\theta}) = \begin{pmatrix} y_{1i}y_{2i} - \mu_1\mu_2 - \sqrt{\sigma_1^2\sigma_2^2}\rho_{12} \\ y_{1i}y_{3i} - \mu_1\mu_3 - \sqrt{\sigma_1^2\sigma_3^2}\rho_{13} \\ y_{2i}y_{3i} - \mu_2\mu_3 - \sqrt{\sigma_2^2\sigma_3^2}\rho_{23} \\ y_{1i} - \mu_1 \\ y_{2i} - \mu_2 \\ y_{3i} - \mu_3 \\ y_{1i}^2 - (\mu_1^2 + \sigma_1^2) \\ y_{2i}^2 - (\mu_2^2 + \sigma_2^2) \\ y_{3i}^2 - (\mu_3^2 + \sigma_3^2) \end{pmatrix}$$

where $\boldsymbol{\theta} = (\rho_{12}, \rho_{13}, \rho_{23}, \sigma_1^2, \sigma_2^2, \sigma_2^3, \mu_1, \mu_2, \mu_3)'$. Then, under the assumption of a Gaussian copula we will have

$$\tilde{\mathbf{A}}_3 = E\left[\frac{\partial \tilde{\mathbf{m}}_{3i}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = \begin{pmatrix} \check{\mathbf{A}}_3 & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_3 \end{pmatrix} \text{ and } \tilde{\mathbf{B}}_3 = Var[\tilde{\mathbf{m}}_{3i}(\boldsymbol{\theta})] = \begin{pmatrix} \check{\mathbf{A}}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}(\boldsymbol{\rho}) \end{pmatrix},$$

which allow us to obtain $AVar(\tilde{\boldsymbol{\theta}})$ as $\tilde{\mathbf{A}}_3^{-1}\tilde{\mathbf{B}}_3\tilde{\mathbf{A}}_3^{-1\prime}$. We can then use the Delta method to obtain the asymptotic variance of $\tilde{\beta}_2^{(1)}$. Specifically, we have

$$\nabla\tilde{\beta}_2^{(1)}(\boldsymbol{\theta}) = \left[\frac{1}{1-\rho_{23}^2}, \frac{-\rho_{23}}{1-\rho_{23}^2}, \frac{2\rho_{12}\rho_{23} - \rho_{13} - \rho_{13}\rho_{23}^2}{(1-\rho_{23}^2)^2}, \frac{\rho_{12} - \rho_{13}\rho_{23}}{2(1-\rho_{23}^2)}, -\frac{\rho_{12} - \rho_{13}\rho_{23}}{2(1-\rho_{23}^2)}, 0, 0, 0, 0\right]',$$

and therefore

$$AVar(\tilde{\beta}_2^{(1)}) = \frac{1 - \rho_{12}^2 - \rho_{13}^2 + 2\rho_{12}\rho_{13}\rho_{23} - \rho_{23}^2}{(1-\rho_{23}^2)^2},$$

as desired. □

Proposition 7

We first compute the variance of the ML correlation estimator by using the correction for the trivariate case given in Lemma 1. Specifically, the resulting diagonal elements for the variance of the corrected scores are

$$V_{11} = \mathcal{V}_i^c(\rho_{12}, \rho_{13}, \rho_{23}),\ V_{22} = \mathcal{V}_i^c(\rho_{13}, \rho_{12}, \rho_{23}) \text{ and } V_{33} = \mathcal{V}_i^c(\rho_{23}, \rho_{12}, \rho_{13}),$$

where

$$\mathcal{V}_i^c(\rho_{12}, \rho_{13}, \rho_{23}) = \frac{1 + 2\rho_{12}^2 + \rho_{12}^4 - \rho_{13}^2 - 4\rho_{12}\rho_{13}\rho_{23} - 2\rho_{12}^3\rho_{13}\rho_{23} - \rho_{23}^2 + 3\rho_{13}^2\rho_{23}^2 + \rho_{12}^2\rho_{13}^2\rho_{23}^2}{(1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12}\rho_{13}\rho_{23})^2}.$$

In turn, the corresponding off-diagonal elements are

$$V_{12} = \mathcal{V}_{ij}^c(\rho_{12}, \rho_{13}, \rho_{23}),\ V_{13} = \mathcal{V}_{ij}^c(\rho_{13}, \rho_{23}, \rho_{12}) \text{ and } V_{23} = \mathcal{V}_{ij}^c(\rho_{23}, \rho_{12}, \rho_{13}),$$

with

$$\begin{aligned} V_{ij}^c(\rho_{12}, \rho_{13}, \rho_{23}) = {} & [5\rho_{12}\rho_{13} + \rho_{12}^3\rho_{13} + \rho_{12}\rho_{13}^3 - 2\rho_{23} - 3\rho_{12}^2\rho_{23} - \rho_{12}^4\rho_{23} - 3\rho_{13}^2\rho_{23} \\ & -2\rho_{12}^2\rho_{13}^2\rho_{23} - \rho_{13}^4\rho_{23} + 2\rho_{12}\rho_{13}\rho_{23}^2 + \rho_{12}^3\rho_{13}\rho_{23}^2 + \rho_{12}\rho_{13}^3\rho_{23}^2 + 2\rho_{23}^3 \\ & -\rho_{12}^2\rho_{23}^3 - \rho_{13}^2\rho_{23}^3 + \rho_{12}\rho_{13}\rho_{23}^4]/[2(1 - \rho_{12}^2 - \rho_{13}^2 + 2\rho_{12}\rho_{13}\rho_{23} - \rho_{23}^2)^2]. \end{aligned}$$

These quantities, together with the expressions for the expected (minus) Hessian in Online Appendix C, allow us to compute the corrected asymptotic variance of the ML estimators via the usual sandwich formula $\mathcal{H}(\boldsymbol{\rho})^{-1}\mathcal{V}^c(\boldsymbol{\rho})\mathcal{H}(\boldsymbol{\rho})^{-1}$.

As for the moment-based estimators, we can also correct the corresponding moment conditions using the following terms:

$$n_{\mu_j i}(\boldsymbol{\theta}) = -H_1(y_{ji}) \text{ and } n_{\sigma_j^2 i}(\boldsymbol{\theta}) = -\sqrt{2}H_2(y_{ji}) \text{ for } j = 1,2,3$$

and

$$n_{\sigma_{jh}i}(\boldsymbol{\theta}) = -\frac{1}{2}(y_{ji}^2 + y_{hi}^2 - 2)\rho_{hj}, \text{ for } h = 1,2,3, \text{ and } h \neq j.$$

As in the bivariate case, if we define $\mathbf{B}_3^{np} = Var[\mathbf{m}_{3i}(\boldsymbol{\theta}) + \mathbf{n}_{3i}(\boldsymbol{\theta})]$, then we will have

$$\check{\mathbf{B}}_3^{np} = \begin{pmatrix} r_{12} & r_{123} & r_{132} & \mathbf{0} \\ r_{123} & r_{13} & r_{231} & \mathbf{0} \\ r_{132} & r_{231} & r_{23} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ and } \tilde{\mathbf{B}}_3^{np} = \begin{pmatrix} \check{\mathbf{B}}_3^{np} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where

$$r_{jh} = (1-\rho_{jh}^2)^2 \text{ and } r_{jhk} = \frac{1}{2}[\rho_{jh}^3\rho_{jk} - 2\rho_{jh}^2\rho_{hk} + 2(1-\rho_{jk}^2)\rho_{hk} + \rho_{jh}\rho_{jk}(\rho_{jk}^2 + \rho_{hk}^2 - 1)].$$

Finally, the corrected variance of both moment estimators of the regression coefficients $\boldsymbol{\beta}$ can be obtained by combining the Delta method with the sandwich formula, and it turns out to be the same as the corrected variance of the ML estimators. □

Proposition 8

The combination of *i.i.d.* data with Assumption 1 implies that under standard regularity conditions we can effectively prove consistency by showing that the expected value of the score in (3) is zero. Let us start by considering the case in which $\mathbf{P}(\boldsymbol{\rho})$ is unrestricted, so that $\boldsymbol{\rho}$ contains the $K(K-1)/2$ off-diagonal elements of the correlation matrix. But since

$$E(y_i^2) = 1 \text{ and } \rho_{ij} = E(y_i y_j),$$

then $\mathbf{P}(\boldsymbol{\rho}_\infty) = E(\mathbf{yy})$. More generally, consider $\mathbf{P}(\boldsymbol{\rho})$, where $\boldsymbol{\rho}$ is a $p \times 1$ vector with $p < K(K-1)/2$. In this case,

$$E[\mathbf{s}_{\boldsymbol{\rho}i}(\mathbf{y};\boldsymbol{\rho})] = \frac{\partial vecl'[\mathbf{P}(\boldsymbol{\rho})]}{\partial \boldsymbol{\rho}} E[\mathbf{s}_{Ki}\{vecl[\mathbf{P}(\boldsymbol{\rho})]\}] = \mathbf{0},$$

where the first equality follows from the chain rule and the last one from the fact that $\mathbf{P}(\boldsymbol{\rho})$ is correctly specified. □

C Trivariate copula expressions

C.1 Score

Applying the general formula in (3) to the trivariate case yields

$$
\begin{aligned}
s_{\rho_{12}}(y_1, y_2, y_3, \rho_{12}, \rho_{13}, \rho_{23}) \quad = \quad & \frac{1}{(1-\rho_{12}^2-\rho_{13}^2+2\rho_{12}\rho_{13}\rho_{23}-\rho_{23}^2)^2} \\
& \times \left[y_1^2(\rho_{12}-\rho_{13}\rho_{23})(\rho_{23}^2-1)-\rho_{12}^3+\rho_{12}^2\rho_{13}[(3+y_3^2)\rho_{23}-y_2y_3]\right. \\
& +\rho_{13}[y_2^2(\rho_{23}-\rho_{13}^2\rho_{23})+y_2y_3(\rho_{13}^2-\rho_{23}^2-1)+\rho_{23}(y_3^2+\rho_{13}^2+\rho_{23}^2-1)] \\
& -\rho_{12}[-1-y_2^2(-1+\rho_{13}^2)-2y_2y_3\rho_{23}+\rho_{23}^2+y_3^2\rho_{23}^2+\rho_{13}^2(1+y_3^2+2\rho_{23}^2)] \\
& +y_1\{-y_3(\rho_{23}+\rho_{12}^2\rho_{23}+\rho_{13}^2\rho_{23}-2\rho_{12}\rho_{13}-\rho_{23}^3) \\
& \left.+y_2[1+\rho_{12}^2-2\rho_{12}\rho_{13}\rho_{23}-\rho_{23}^2-\rho_{13}^2(1-2\rho_{23}^2)]\}\right]
\end{aligned}
$$

$$s_{\rho_{13}}(y_1, y_2, y_3, \rho_{12}, \rho_{13}, \rho_{23}) = s_{\rho_{12}}(y_1, y_3, y_2, \rho_{13}, \rho_{12}, \rho_{23}),$$

and

$$s_{\rho_{23}}(y_1, y_2, y_3, \rho_{12}, \rho_{13}, \rho_{23}) = s_{\rho_{12}}(y_2, y_3, y_1, \rho_{23}, \rho_{12}, \rho_{13}).$$

C.2 Hessian

The expected value of the (minus) Hessian under correct of specification of the correlation matrix is given by

$$E[-\mathbf{h}_i(\boldsymbol{\rho}_\infty)] = \begin{bmatrix} h_{11}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) & h_{12}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) & h_{13}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) \\ & h_{22}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) & h_{23}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) \\ & & h_{33}(\rho_{12\infty}, \rho_{13\infty}, \rho_{23\infty}) \end{bmatrix}$$

where

$$
\begin{aligned}
h_{11}(\rho_{12}, \rho_{13}, \rho_{23}) &= \frac{1+\rho_{12}^2-2\rho_{12}\rho_{13}\rho_{23}-\rho_{23}^2-\rho_{13}^2(1-2\rho_{23}^2)}{(1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23})^2} \\
h_{22}(\rho_{12}, \rho_{13}, \rho_{23}) &= h_{11}(\rho_{13}, \rho_{12}, \rho_{23}), \\
h_{33}(\rho_{12}, \rho_{13}, \rho_{23}) &= h_{11}(\rho_{23}, \rho_{12}, \rho_{13}), \\
h_{12}(\rho_{12}, \rho_{13}, \rho_{23}) &= \frac{\rho_{23}^3+2\rho_{12}\rho_{13}-\rho_{23}(1+\rho_{12}^2+\rho_{13}^2)}{(1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23})^2}, \\
h_{13}(\rho_{12}, \rho_{13}, \rho_{23}) &= h_{12}(\rho_{12}, \rho_{23}, \rho_{13})
\end{aligned}
$$

and

$$h_{23}(\rho_{12}, \rho_{13}, \rho_{23}) = h_{12}(\rho_{13}, \rho_{23}, \rho_{12}).$$

D Spearman's calculations

D.1 Alternative estimators

Alternative estimators to $\tilde{\rho}_S^I$, which is based on the moment conditions (14), can be obtained as follows.

Given that u_{1i} and u_{2i} are uniform by definition, one could exploit the fact that $E(u_{ji}) = 1/2$ and $Var(u_{ji}) = 1/12$ to estimate ρ based on the single moment condition

$$E\left(u_{1i}u_{2i} - \frac{1}{4} - \frac{1}{12}\rho\right) = 0,$$

whence

$$\tilde{\rho}_S^{II} = 12\left(\frac{1}{N}\sum_{i=1}^{N} u_{1i}u_{2i} - \frac{1}{4}\right). \tag{D1}$$

A third estimator in which the mean of each component is substracted before computing the cross-moment is given by

$$\tilde{\rho}_S^{III} = \frac{1}{1/12}\sum_{i=1}^{N}\left(u_{1i} - \frac{1}{2}\right)\left(u_{2i} - \frac{1}{2}\right). \tag{D2}$$

Finally, the fourth estimator we could consider, which is the closest to the one Matlab implements, is

$$\tilde{\rho}_S^{IV} = 1 - \frac{6(N+1)^2}{N(N^2-1)}\sum_{i=1}^{N}(u_{1i} - u_{2i})^2,$$

which in large samples can be interpreted in terms of the following moment conditions

$$E[\mathbf{m}_i^{IV}(\boldsymbol{\theta})] = E\left[\begin{pmatrix} u_{1i}u_{2i} - \frac{1}{2}(\mu_1^2+\sigma_1^2) - \frac{1}{2}(\mu_2^2+\sigma_2^2) - \frac{1}{12}(\rho-1) \\ u_{1i} - \mu_1 \\ u_{2i} - \mu_2 \\ u_{1i}^2 - (\mu_1^2+\sigma_1^2) \\ u_{2i}^2 - (\mu_2^2+\sigma_2^2) \end{pmatrix}\right] = \mathbf{0}. \tag{D3}$$

D.2 Asymptotic variances

Regarding $\tilde{\rho}_S^{I}$, we can easily compute the expected value of the Jacobian and variance of the moment conditions to obtain the asymptotic variance for $\boldsymbol{\theta}$ in (14). In particular,

$$\mathcal{A}^I(\boldsymbol{\theta}) = E\left[\frac{\partial \mathbf{m}_i^I(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} \sqrt{\sigma_1^2\sigma_2^2} & \mu_2 & \mu_1 & \frac{1}{2}\rho\sqrt{\sigma_2^2/\sigma_1^2} & \frac{1}{2}\rho\sqrt{\sigma_1^2/\sigma_2^2} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2\mu_1 & 0 & 1 & 0 \\ 0 & 0 & 2\mu_2 & 0 & 1 \end{pmatrix},$$

and

$$\begin{aligned} \mathcal{B}^I(\boldsymbol{\theta}) &= Var[\mathbf{m}_i^I(\boldsymbol{\theta})] \\ &= \begin{pmatrix} E_{22} - E_{11}^2 & E_{21} - E_{11}E_{10} & E_{12} - E_{11}E_{01} & E_{31} - E_{11}E_{20} & E_{13} - E_{11}E_{02} \\ & Var(u_{1i}) & cov(u_{1i}, u_{2i}) & E_{30} - E_{20}E_{10} & cov(u_{1i}, u_{2i}^2) \\ & & Var(u_{2i}) & cov(u_{1i}^2, u_{2i}) & E_{03} - E_{02}E_{01} \\ & & & E_{40} - E_{20}^2 & cov(u_{1i}^2, u_{2i}^2) \\ & & & & E_{04} - E_{02}^2 \end{pmatrix}, \end{aligned}$$

where $E_{h,j}$ denotes $E(u_{1i}^h u_{2i}^j)$.

As for $\tilde{\rho}_S^{II}$, it is straightforward to prove that (D1) implies $AVar(\hat{\rho}) = 144 \times Var(u_{1i}u_{2i})$.

To obtain the asymptotic variance of $\tilde{\rho}_S^{III}$ from (D2), it is convenient to use the following moment conditions

$$E\begin{bmatrix} u_{1i}u_{2i} - \frac{1}{2}\mu_1 - \frac{1}{2}\mu_2 - \frac{1}{12}\rho + \frac{1}{4} \\ u_{1i} - \mu_1 \\ u_{2i} - \mu_2 \\ u_{1i}^2 - (\mu_1^2 + \sigma_1^2) \\ u_{2i}^2 - (\mu_2^2 + \sigma_2^2) \end{bmatrix} = E[\mathbf{m}_i^{III}(\boldsymbol{\theta})] = \mathbf{0},$$

whence

$$\mathcal{A}^{III}(\boldsymbol{\theta}) = E\left[\frac{\partial \mathbf{m}_i^{III}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} 1/12 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2\mu_1 & 0 & 1 & 0 \\ 0 & 0 & 2\mu_2 & 0 & 1 \end{pmatrix}.$$

In addition, it is easy to see that $\mathcal{B}^{III}(\boldsymbol{\theta}) = V[\mathbf{m}_i^{III}(\boldsymbol{\theta})]$ coincides with $\mathcal{B}^{I}(\boldsymbol{\theta})$.

Finally, we can use (D3) to show that $\mathcal{B}^{IV}(\boldsymbol{\theta}) = V[\mathbf{m}_i^{IV}(\boldsymbol{\theta})]$ is equal to $\mathcal{B}^{I}(\boldsymbol{\theta})$ and

$$\mathcal{A}^{IV}(\boldsymbol{\theta}) = E\left[\frac{\partial \mathbf{m}_i^{IV}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right] = -\begin{pmatrix} 1/12 & \mu_1 & \mu_2 & 1/2 & 1/2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2\mu_1 & 0 & 1 & 0 \\ 0 & 0 & 2\mu_2 & 0 & 1 \end{pmatrix},$$

whence we can obtain the asymptotic variance of $\tilde{\rho}_S^{IV}$.

E Description of the marginal distributions used in Section 5

E.1 Tukey distribution

The Tukey lambda distribution is a continuous, symmetric probability distribution defined in terms of its quantile function

$$Q(p,\lambda) = \begin{cases} \frac{1}{\lambda}[p^\lambda - (1-p)^\lambda], & \text{if } \lambda \neq 0 \\ \ln[p/(1-p)], & \text{if } \lambda = 0, \end{cases}$$

where λ is its single shape parameter. It nests the logistic distribution for $\lambda = 0$ and the uniform distribution for both $\lambda = 1$ and $\lambda = 2$. In Figure 6a, we plot the density of a Tukey random variable with parameter $\lambda = 1.5$.

E.2 Asymmetric Laplace distribution

The Asymmetric Laplace distribution is a continuous probability distribution consisting of two exponential distributions of unequal scale, adjusted to ensure continuity and normalization. Its density is

$$f(x; m, \kappa, \lambda) = \frac{\lambda}{\kappa + 1/\kappa}\begin{cases} \exp[(\lambda/k)(x-m)], & x \leq m \\ \exp[-\lambda\kappa(x-m)], & x > m \end{cases}$$

The quantiles for this distribution can be easily obtained from those of the two underlying exponential distributions. In Figure 6b, we plot the density of an Asymmetric Laplace random variable with parameters $m = 0$, $k = 2$ and $\lambda = 1$.

E.3 Weibull distribution

The probability density function of the Weibull distribution is

$$f(x;k,\lambda) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}\exp[-(x/\lambda)^k], & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{E1}$$

where $k > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter. Its quantile function is $F^{-1}(p;k,\lambda) = \lambda[-\ln(1-p)]^k$. When $k = 0$, it particularizes to the exponential distribution with parameter λ^{-1}. We plot the density of a Weibull random variable with parameters $k = 0.75$ and $\lambda = 1$ in Figure 6c.

E.4 Mixture of Weibull distributions

This distribution is generated by mixing a regular Weibull distribution and a mirror image of another Weibull distribution whose support is the negative real line. Suppose that x_1 follows a Weibull distribution with shape and scale parameters k_1 and λ_1, and that $-x_2$ follows a Weibull distribution with shape and scale parameters k_2 and λ_2. Further, let α denote the mixing probability associated to the first component. Then, the nonstandardized mixture x has density given by

$$f(x;k_1,k_2,\lambda_1,\lambda_2,\alpha) = \alpha f(x;k_1,\lambda_1) + (1-\alpha)f(x;k_2,\lambda_2),$$

where $f(x;k,\lambda)$ is given in (E1). We standardize x to achieve zero mean and unit variance. The quantiles for this distribution can be easily obtained from those of the two underlying Weibull distributions. In Figure 6d we plot the density of a mixture of Weibull random variables with parameters $k_1 = 5$, $\lambda_1 = 10$, $k_2 = 5$, $\lambda_1 = 2$ and mixing probability $\alpha = .98$.