On Partial and Conditional Association Measures For Ordinal Contingency Tables

Zheng Wei and Daeyoung Kim

University of Maine and University of Massachusetts-Amherst

Supplementary Material

The supplementary material contains (i) the proofs of Proposition 3.1, Theorem 3.1, Theorem 4.1, Theorem 4.2, and Theorem 4.3; (ii) the detailed information on the parameter values of the simulation log linear models used in Section 5.1.1; (iii) the simulation results of the simulation scenario [I] (conditional independence) and [II] (linear relationship) mentioned in Section 5.1.1; and (iv) the detailed information on the parameter values of the quadratic latent variable regression models and the discrete marginal distributions used in Section 5.2.

S1 Proof of Proposition 3.1

Proof. From Eq. (3.10), we have the following identities,

$$\sigma_{vv\bullet u} = \sigma_{vv} - \frac{\sigma_{vu}^2}{\sigma_{uu}} = \sigma_{vv}\left(1-\rho_{VU}^2\right), \qquad \sigma_{ww\bullet u} = \sigma_{ww} - \frac{\sigma_{wu}^2}{\sigma_{uu}} = \sigma_{ww}\left(1-\rho_{WU}^2\right)$$
$$\sigma_{vw\bullet u} = \sigma_{vw} - \frac{\sigma_{vu}\sigma_{wu}}{\sigma_{uu}}.$$

Therefore, we have

$$\rho_{VW\bullet U} = \frac{\sigma_{vw\bullet u}}{\sqrt{\sigma_{vv\bullet u}}\sqrt{\sigma_{ww\bullet u}}} = \frac{\rho_{VW} - \rho_{VU}\rho_{WU}}{\sqrt{1-\rho_{VU}^2}\sqrt{1-\rho_{WU}^2}}.$$

□

S2 Proof of Theorem 3.1

(I). If Y and Z are conditionally independent given X, then $p_{jk|i} = p_{j|i}p_{k|i}$ for all i, j, and k.

Then the numerator of $\rho^C_{VW\bullet U}$ equals

$$
\begin{aligned}
& \sum_{i,j,k}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(v_j - \sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)\left(w_k - \sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{ijk} \\
= & \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(v_j - \sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)\left(w_k - \sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{jk|i}p_{\bullet\bullet i} \\
= & \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(v_j - \sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)p_{j|i}\left(w_k - \sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{k|i}p_{\bullet\bullet i} \\
= & \sum_{i=1}^{I}\sum_{j=1}^{J}\left(v_j - \sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)p_{j|i}\sum_{k=1}^{K}\left(w_k - \sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{k|i}p_{\bullet\bullet i} \\
= & \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_jp_{j|i} - \sum_{j=1}^{J} p_{j|i}\sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)\left(\sum_{k=1}^{K} w_kp_{k|i} - \sum_{k=1}^{K} p_{k|i}\sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{\bullet\bullet i} \\
= & \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_jp_{j|i} - \sum_{j'=1}^{J} v_{j'}p_{j'|i}\right)\left(\sum_{k=1}^{K} w_kp_{k|i} - \sum_{k'=1}^{K} w_{k'}p_{k'|i}\right)p_{\bullet\bullet i} = 0
\end{aligned}
$$

By same argument, we can show $\rho^C_{VW|U=u_i} = 0$ for each $i = 1, \ldots, I$. □

(II). Because the law of total variance, we have

$$
\begin{aligned}
& Var\left(E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right] - E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\right) - \Sigma_{12}\Sigma_{22}^{-1}(U - E(U))\middle| U\right)\right) \\
= \; & Var\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right] - E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\right) - \Sigma_{12}\Sigma_{22}^{-1}(U - E(U))\right) - \\
& E\left(Var\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right] - E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\right) - \Sigma_{12}\Sigma_{22}^{-1}(U - E(U))\middle| U\right)\right) \\
= \; & \Sigma_{YZ\bullet X} - E\left(\Sigma_{VW|U}\right)
\end{aligned}
$$

Therefore, $\Sigma_{YZ\bullet X} = E\left(\Sigma_{VW|U}\right)$ if and only if

$$
Var\left(E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right] - E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\right) - \Sigma_{12}\Sigma_{22}^{-1}(U - E(U))\middle| U\right)\right) = 0,
$$

which in turn is equivalent to

$$
E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\middle| U\right) - E\left([V, W]^T\right) - \Sigma_{12}\Sigma_{22}^{-1}(U - E(U)) = \boldsymbol{B},
$$

for some 2×1 constant vector $\boldsymbol{B}$ almost surely. Thus, the results applies by setting $\boldsymbol{\beta} = \Sigma_{12}\Sigma_{22}^{-1}$ and

$$
\boldsymbol{\alpha} = E\left(\left[\begin{smallmatrix} V \\ W \end{smallmatrix}\right]\right) + \boldsymbol{B} - \Sigma_{12}\Sigma_{22}^{-1}E(U).
$$

□

(III). Let Y be a binary variable with two categories $\{y_1, y_2\}$ and the corresponding p.m.f., $P(Y = y_1) = p_{\bullet 1 \bullet}$ and $P(Y = y_2) = p_{\bullet 2 \bullet}$, and denote $\tilde{V} = F_{\tilde{Y}}(\tilde{Y})$. Then, the range of the marginal distribution of $\tilde{Y}$ is $\tilde{D}_{\tilde{y}} = \{0, \tilde{v}_1, 1\}$, where $\tilde{v}_1 = p_{2\bullet\bullet} = 1 - v_1$, and we have

$$
\begin{aligned}
E(V) &= p_{1\bullet\bullet}^2 + p_{2\bullet\bullet} = p_{1\bullet\bullet}^2 + (1 - p_{1\bullet\bullet}) = (p_{1\bullet\bullet} - 1)^2 + p_{1\bullet\bullet} = p_{2\bullet\bullet}^2 + p_{1\bullet\bullet} = E(\tilde{V}). \\
Var(V) &= (p_{1\bullet\bullet} - E(V))^2 p_{1\bullet\bullet} + (1 - E(V))^2 p_{2\bullet\bullet} = p_{2\bullet\bullet}^4 p_{1\bullet\bullet} + p_{1\bullet\bullet}^2 p_{2\bullet\bullet}^3 \\
&= p_{1\bullet\bullet} p_{2\bullet\bullet}^3 (p_{2\bullet\bullet} + p_{1\bullet\bullet}) = p_{1\bullet\bullet} p_{2\bullet\bullet}^3.
\end{aligned}
$$

Similarly, we have $Var(\tilde{V}) = p_{1\bullet\bullet}^3 p_{2\bullet\bullet}$. To show this proposition, it will be enough to show that the correlation satisfying $\rho_{\widetilde{V}W} = -\rho_{VW}$, which is equivalent to show

$$\frac{E(VW) - E(V)E(W)}{\sqrt{Var(V)}} = -\frac{E(\tilde{V}W) - E(\tilde{V})E(W)}{\sqrt{Var(\tilde{V})}}. \tag{S2.1}$$

Eq. (S2.1) is equivalent to show

$$\frac{p_{1\bullet\bullet} \sum_{k=1}^{K} w_k p_{1k} + \sum_{k=1}^{K} w_k p_{2k} - E(V)E(W)}{\sqrt{p_{1\bullet\bullet} p_{2\bullet\bullet}^3}} = -\frac{p_{2\bullet\bullet} \sum_{k=1}^{K} w_k p_{2k} + \sum_{k=1}^{K} w_k p_{1k} - E(V)E(W)}{\sqrt{p_{1\bullet\bullet}^3 p_{2\bullet\bullet}}},$$

which in turn is equivalent to

$$\frac{p_{1\bullet\bullet} \sum_{k=1}^{K} w_k p_{1k} + \sum_{k=1}^{K} w_k p_{2k} - E(V)E(W)}{p_{2\bullet\bullet}} = -\frac{p_{2\bullet\bullet} \sum_{k=1}^{K} w_k p_{2k} + \sum_{k=1}^{K} w_k p_{1k} - E(V)E(W)}{p_{1\bullet\bullet}} \tag{S2.2}$$

To show Eq. (S2.2), it will be enough to show

$$(p_{1\bullet\bullet}^2 + p_{2\bullet\bullet}) \sum_{k=1}^{K} w_k p_{1k} + (p_{1\bullet\bullet} + p_{2\bullet\bullet}^2) \sum_{k=1}^{K} w_k p_{2k} - E(V)E(W) = 0, \tag{S2.3}$$

which is true because

$$\begin{aligned}
&(p_{1\bullet\bullet}^2 + p_{2\bullet\bullet}) \sum_{k=1}^{K} w_k p_{1k} + (p_{1\bullet\bullet} + p_{2\bullet\bullet}^2) \sum_{k=1}^{K} w_k p_{2k} - E(V)E(W) \\
&= E(V) \sum_{k=1}^{K} w_k p_{1k} + E(V) \sum_{k=1}^{K} w_k p_{2k} - E(V)E(W) = E(V)E(W) - E(V)E(W) = 0.
\end{aligned}$$

This implies the proposition holds. □

(IV). From the definition of $\rho_{YZ\bullet X}$, it is straightforward to see that if the order of X is permuted randomly, all the sums in the numerator and denominator of $\rho_{YZ\bullet X}$ will not change. $\square$

S3 Proof of Theorem 4.1

Proof. We denote the relative frequencies of the observed marginal two-way $J \times K$ contingency table by

$$\hat{P}_{YZ} = \left\{\hat{p}_{jk\bullet} \mid j = 1, \ldots, J, k = 1, \ldots, K\right\}.$$

Define $\hat{\boldsymbol{p}}_{YZ}$ to be the $(JK) \times 1$-dimensional estimator for $\boldsymbol{p}_{YZ}$ where

$$\hat{\boldsymbol{p}}_{YZ} = (\hat{p}_{11\bullet}, \hat{p}_{21\bullet}, \ldots, \hat{p}_{J1\bullet}, \ldots, \hat{p}_{JK\bullet})^T .$$

Then the expectation and the covariance matrix of $\hat{\boldsymbol{p}}_{YZ}$ are $E[\hat{\boldsymbol{p}}_{YZ}] = \boldsymbol{p}_{YZ}$ and $\mathrm{Cov}\left(\hat{\boldsymbol{p}}_{YZ}\right) = \left(diag(\boldsymbol{p}_{YZ}) - \boldsymbol{p}_{YZ}\boldsymbol{p}_{YZ}^T\right)/n$. Note that $n\hat{p}_{jk\bullet}$ and $n\hat{p}_{j'k'\bullet}$ where $(j,k) \neq (j',k')$ follows a trinomial distribution with parameters $(p_{jk\bullet}, p_{j'k'\bullet}, 1 - p_{jk\bullet} - p_{j'k'\bullet})$ and $\mathrm{Cov}(n\hat{p}_{jk\bullet}, n\hat{p}_{j'k'\bullet}) = -np_{jk\bullet}p_{j'k'\bullet}$. Therefore, we have

$$\sqrt{n}\left(\hat{\boldsymbol{p}}_{YZ} - \boldsymbol{p}_{YZ}\right) \xrightarrow{D} N_{JK}\left(\boldsymbol{0}, \left(diag(\boldsymbol{p}_{YZ}) - \boldsymbol{p}_{YZ}\boldsymbol{p}_{YZ}^T\right)\right),$$

where $N_d(\boldsymbol{\mu}, \Sigma)$ denotes the d-dimensional multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ.

As the proposed pairwise correlation of (Y, Z), ρ_{YZ}, is a function of $\boldsymbol{p}_{YZ}$, denote it as $\rho_{YZ} = h_{YZ}(\boldsymbol{p}_{YZ})$. We now calculate the gradient for the function h_{YZ}. Let $T(\boldsymbol{p}_{YZ}) = \sum_{j=1}^{J} \sum_{k=1}^{K} v_j w_k p_{jk\bullet} - \mu_v \mu_w$ be the numerator of ρ_{YZ}, and denote

$$B(\boldsymbol{p}_{YZ}) = \sqrt{\sum_{j=1}^{J} (v_j - \mu_v)^2 p_{j\bullet\bullet}} \sqrt{\sum_{k=1}^{K} (w_k - \mu_w)^2 p_{\bullet k\bullet}}$$

to be the denominator of ρ_{YZ}.

Given $s = 1, \ldots, J$ and $t = 1, \ldots, K$, we have (i) $\partial v_j / \partial p_{st\bullet} = 1$ for $j \geq s$ and $\partial v_j / \partial p_{st\bullet} = 0$ for $j < s$; (ii) $\partial w_k / \partial p_{st\bullet} = 1$ for $k \geq t$ and $\partial w_k / \partial p_{st\bullet} = 0$ for $k < t$; (iii) $\partial \mu_v / \partial p_{st\bullet} = v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}$ and $\partial \mu_w / \partial p_{st\bullet} = w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}$. By (i)-(iii), we have

$$
\begin{aligned}
\frac{\partial T(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}} &= \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{st\bullet}} - \frac{\partial \mu_v \mu_w}{\partial p_{st\bullet}} \\
= \quad & \sum_{j=1}^{s-1}\sum_{k=1}^{t-1}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{st\bullet}} + \sum_{j=1}^{s-1}\sum_{k=t}^{K}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{st\bullet}} + \sum_{k=1}^{t-1}\frac{\partial v_s w_k p_{sk\bullet}}{\partial p_{st\bullet}} + \sum_{k=t}^{K}\frac{\partial v_s w_k p_{sk\bullet}}{\partial p_{st\bullet}} \qquad \text{(S3.4)} \\
& + \sum_{j=s+1}^{J}\sum_{k=1}^{t-1}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{st\bullet}} + \sum_{j=s+1}^{J}\sum_{k=t}^{K}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{st\bullet}} - \mu_v\left(w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}\right) - \left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right)\mu_w \\
= \quad & 0 + \sum_{j=1}^{s-1}\sum_{k=t}^{K} v_j p_{jk\bullet} + \sum_{k=1}^{t-1} w_k p_{sk\bullet} + \left(v_s w_t + \sum_{k=t}^{K}(v_s + w_k) p_{sk\bullet}\right) \\
& + \sum_{j=s+1}^{J}\sum_{k=1}^{t-1} w_k p_{jk\bullet} + \sum_{j=s+1}^{J}\sum_{k=t}^{K}\Big((v_j + w_k) p_{jk\bullet}\Big) - \mu_v\left(w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}\right) - \left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right)\mu_w \\
= \quad & \sum_{j=1}^{s-1}\sum_{k=t}^{K} v_j p_{jk\bullet} + v_s w_t + \sum_{j=s}^{J}\sum_{k=1}^{t-1} w_k p_{jk\bullet} + \sum_{j=s}^{J}\sum_{k=t}^{K}(v_j + w_k) p_{jk\bullet} \\
& -\mu_v\left(w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}\right) - \left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right)\mu_w
\end{aligned}
$$

To calculate $\frac{\partial B(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}}$, the following equalities will be useful :

$$
\begin{aligned}
\frac{\partial \sigma_{vv}}{\partial p_{st\bullet}} &= \frac{\partial \sum_{j=1}^{J}(v_j - \mu_v)^2 p_{j\bullet\bullet}}{\partial p_{st\bullet}} = \frac{\partial\left(\sum_{j=1}^{J} v_j^2 p_{j\bullet\bullet} - \mu_v^2\right)}{\partial p_{st\bullet}} \\
&= \sum_{j=1}^{J} 2 v_j \frac{\partial v_j}{\partial p_{st\bullet}} p_{j\bullet\bullet} + \sum_{j=1}^{J} v_j^2 \frac{\partial p_{j\bullet\bullet}}{\partial p_{st\bullet}} - 2\mu_v\left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right) \\
&= \sum_{j=s}^{J} 2 v_j p_{j\bullet\bullet} + v_s^2 - 2\mu_v\left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right), \\
\frac{\partial \sigma_{ww}}{\partial p_{st\bullet}} &= \frac{\partial \sum_{k=1}^{K}(w_k - \mu_w)^2 p_{\bullet k\bullet}}{\partial p_{st\bullet}} = \sum_{k=t}^{K} 2 w_k p_{\bullet k\bullet} + w_t^2 - 2\mu_w\left(w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}\right).
\end{aligned}
$$

By the above two identities, we have

$$\frac{\partial B(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}} = \frac{\sigma_{ww}}{2\sqrt{\sigma_{vv}}}\left[\sum_{j=s}^{J} 2v_j p_{j\bullet\bullet} + v_s^2 - 2\mu_v\left(v_s + \sum_{j=s}^{J} p_{j\bullet\bullet}\right)\right] + \frac{\sigma_{vv}}{2\sqrt{\sigma_{ww}}}\left[\sum_{k=t}^{K} 2w_k p_{\bullet k\bullet} + w_t^2 - 2\mu_w\left(w_t + \sum_{k=t}^{K} p_{\bullet k\bullet}\right)\right] \tag{S3.5}$$

Since we have

$$\frac{\partial \rho_{YZ}(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}} = \frac{\frac{\partial T(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}} B(\boldsymbol{p}_{YZ}) - \frac{\partial B(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}} T(\boldsymbol{p}_{YZ})}{B(\boldsymbol{p}_{YZ})^2}, \tag{S3.6}$$

where $\frac{\partial T(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}}$ and $\frac{\partial B(\boldsymbol{p}_{YZ})}{\partial p_{st\bullet}}$ are given in Eq. (S3.4) and (S3.5), respectively, the gradient for the function h_{YZ} is

$$\nabla h_{YZ}(\boldsymbol{p}_{YZ}) = \left[\frac{\partial}{\partial p_{11\bullet}}\rho_{YZ}, \frac{\partial}{\partial p_{21\bullet}}\rho_{YZ}, \ldots, \frac{\partial}{\partial p_{JK\bullet}}\rho_{YZ}\right]^T. \tag{S3.7}$$

Thus, by the delta-method,

$$\sqrt{n}\,(\hat{\rho}_{YZ} - \rho_{YZ}) \xrightarrow{D} N_1\left(0, \nabla h_{YZ}(\boldsymbol{p}_{YZ})^T\left(diag(\boldsymbol{p}_{YZ}) - \boldsymbol{p}_{YZ}\boldsymbol{p}_{YZ}^T\right)\nabla h_{YZ}(\boldsymbol{p}_{YZ})\right).$$

□

S4 Proof of Theorem 4.2

Proof. We denote the relative frequencies of the observed three-way $J \times K \times I$ contingency table by

$$\hat{P} = \left\{\hat{p}_{jki} \mid j = 1, \ldots, J, k = 1, \ldots, K, i = 1, \ldots, I\right\}.$$

Define $\hat{\boldsymbol{p}}$ to be the $(JKI) \times 1$-dimensional estimator for $\boldsymbol{p}$ where

$$\hat{\boldsymbol{p}} = (\hat{p}_{111}, \hat{p}_{211}, \ldots, \hat{p}_{J11}, \hat{p}_{121}, \ldots, \hat{p}_{JKI})^T .$$

Then the expectation and the covariance matrix of $\hat{\boldsymbol{p}}$ are $E[\hat{\boldsymbol{p}}] = \boldsymbol{p}$ and $\mathrm{Cov}\,(\hat{\boldsymbol{p}}) = \left(diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^T\right)/n$, and thus

$$\sqrt{n}\,(\hat{\boldsymbol{p}} - \boldsymbol{p}) \xrightarrow{D} N_{JKI}\left(\mathbf{0}, \left(diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^T\right)\right).$$

The proposed partial correlation of Y and Z given X, $\rho_{YZ\bullet X}$, is a function of $\boldsymbol{p}$ and we denote $\rho_{YZ\bullet X} = h_{YZ\bullet X}(\boldsymbol{p})$. We now calculate the gradient for the function $h_{YZ\bullet X}$. Denote $T_{YZ\bullet X}(\boldsymbol{p})$ to be the numerator of $\rho_{YZ\bullet X}$, and

$$\begin{aligned}
T_{YZ\bullet X}(\boldsymbol{p}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(v_j - \sum_{j'=1}^{J} v_{j'} p_{j'|i}\right)\left(w_k - \sum_{k'=1}^{K} w_{k'} p_{k'|i}\right) p_{jki} \\
&= \sum_{i=1}^{I}\left(\sum_{j=1}^{J}\sum_{k=1}^{K} v_j w_k p_{jk|i} - \left(\sum_{j'=1}^{J} v_{j'} p_{j'|i}\right)\left(\sum_{k'=1}^{K} w_{k'} p_{k'|i}\right)\right) p_{\bullet\bullet i} \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} v_j w_k p_{jk\bullet} - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{k=1}^{K} w_k p_{k|i}\right) p_{\bullet\bullet i}.
\end{aligned}$$

We denote $B_{YZ\bullet X}(\boldsymbol{p})$ to be the denominator of $\rho_{YZ\bullet X}$,

$$\begin{aligned} B_{YZ\bullet X}(\boldsymbol{p}) &= \sqrt{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(v_j - \sum_{j'=1}^{J} v_{j'} p_{j'|i}\right)^2 p_{j\bullet i}}\sqrt{\sum_{i=1}^{I}\sum_{k=1}^{K}\left(w_k - \sum_{k'=1}^{K} w_{k'} p_{k'|i}\right)^2 p_{\bullet ki}} \\ &= \sqrt{\sum_{j=1}^{J} v_j^2 p_{j\bullet\bullet} - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)^2 p_{\bullet\bullet i}}\sqrt{\sum_{k=1}^{K} w_j^2 p_{\bullet k\bullet} - \sum_{i=1}^{I}\left(\sum_{k=1}^{K} w_k p_{k|i}\right)^2 p_{\bullet\bullet i}} \end{aligned}$$

Given $s = 1, \ldots, J$, $t = 1, \ldots, K$, and $r = 1, \ldots, I$, we have (i) $\partial v_j / \partial p_{str} = 1$ for $j \geq s$ and $\partial v_j / \partial p_{str} = 0$ for $j < s$; (ii) $\partial w_k / \partial p_{str} = 1$ for $k \geq t$ and $\partial w_k / \partial p_{str} = 0$ for $k < t$; (iii)

$$\frac{\partial v_j p_{j|i}}{\partial p_{str}} = \begin{cases} p_{j|i}, & \text{for } j \geq s, i \neq r \\ p_{j|i} - v_j \frac{p_{j\bullet i}}{p^2_{\bullet\bullet i}}, & \text{for } j > s, i = r \\ p_{j|i} + v_j \frac{p_{\bullet\bullet i} - p_{j\bullet i}}{p^2_{\bullet\bullet i}}, & \text{for } j = s, i = r \\ -v_j \frac{p_{j\bullet i}}{p^2_{\bullet\bullet i}}, & \text{for } j < s, i = r \\ 0, & \text{for } j < s, i \neq r \end{cases}, \quad \frac{\partial w_k p_{k|i}}{\partial p_{str}} = \begin{cases} p_{k|i}, & \text{for } k \geq t, i \neq r \\ p_{k|i} - w_k \frac{p_{\bullet ki}}{p^2_{\bullet\bullet i}}, & \text{for } k > t, i = r \\ p_{k|i} + w_k \frac{p_{\bullet\bullet i} - p_{\bullet ki}}{p^2_{\bullet\bullet i}}, & \text{for } k = t, i = r \\ -w_k \frac{p_{\bullet ki}}{p^2_{\bullet\bullet i}}, & \text{for } k < t, i = r \\ 0, & \text{for } k < t, i \neq r. \end{cases}$$

By (i)-(iii) above, we have

$$
\begin{aligned}
\frac{\partial T_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}} &= \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial v_j w_k p_{jk\bullet}}{\partial p_{str}} - \sum_{i=1}^{I}\frac{\partial}{\partial p_{str}}\left[\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{k=1}^{K} w_k p_{k|i}\right) p_{\bullet\bullet i}\right] \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial v_j}{\partial p_{str}} w_k p_{jk\bullet} + \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial w_k}{\partial p_{str}} v_j p_{jk\bullet} + \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial p_{jk\bullet}}{\partial p_{str}} v_j w_k \\
&\quad - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{k=1}^{K} w_k p_{k|i}\right)\left[\frac{\partial}{\partial p_{str}} p_{\bullet\bullet i}\right] \\
&\quad - \sum_{i=1}^{I}\left[\frac{\partial}{\partial p_{str}}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\right]\left(\sum_{k=1}^{K} w_k p_{k|i}\right) p_{\bullet\bullet i} - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left[\frac{\partial}{\partial p_{str}}\left(\sum_{k=1}^{K} w_k p_{k|i}\right)\right] p_{\bullet\bullet i} \\
&= \sum_{j=s}^{J}\sum_{k=1}^{K} w_k p_{jk\bullet} + \sum_{j=1}^{J}\sum_{k=t}^{K} v_j p_{jk\bullet} + v_s w_t - \left(\sum_{j=1}^{J} v_j p_{j|r}\right)\left(\sum_{k=1}^{K} w_k p_{k|r}\right) \qquad \text{(S4.8)} \\
&\quad - \left[\sum_{j=1}^{s-1} -v_j \frac{p_{j\bullet r}}{p^2_{\bullet\bullet r}} + p_{s|r} + v_s \frac{p_{\bullet\bullet r} - p_{s\bullet r}}{p^2_{\bullet\bullet r}} + \sum_{j=s+1}^{J}\left(p_{j|r} - v_j \frac{p_{j\bullet r}}{p^2_{\bullet\bullet r}}\right)\right]\left(\sum_{k=1}^{K} w_k p_{k|r}\right) p_{\bullet\bullet r} \\
&\quad - \sum_{i\neq r}^{I}\left[\sum_{j=s}^{J} p_{j|i}\right]\left(\sum_{k=1}^{K} w_k p_{k|i}\right) p_{\bullet\bullet i} \\
&\quad - \left(\sum_{j=1}^{J} v_j p_{j|r}\right)\left[\sum_{k=1}^{t-1} -w_k \frac{p_{\bullet kr}}{p^2_{\bullet\bullet r}} + p_{t|r} + w_t \frac{p_{\bullet\bullet r} - p_{\bullet tr}}{p^2_{\bullet\bullet r}} + \sum_{k=t+1}^{K}\left(p_{k|r} - w_k \frac{p_{\bullet kr}}{p^2_{\bullet\bullet r}}\right)\right] p_{\bullet\bullet r} \\
&\quad - \sum_{i\neq r}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left[\sum_{k=t}^{K} p_{k|i}\right] p_{\bullet\bullet i}
\end{aligned}
$$

To calculate $\frac{\partial B_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}}$, the following equality will be useful :

$$
\sigma'_{vv\bullet u} \equiv \frac{\partial\left(\sum_{j=1}^{J} v_j^2 p_{j\bullet\bullet} - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)^2 p_{\bullet\bullet i}\right)}{\partial p_{str}} = \sum_{j=1}^{J} 2 v_j \frac{\partial v_j}{\partial p_{str}} p_{j\bullet\bullet} + \sum_{j=1}^{J} v_j^2 \frac{\partial p_{j\bullet\bullet}}{\partial p_{str}} \qquad \text{(S4.9)}
$$

$$
\begin{aligned}
& -\sum_{i=1}^{I} 2\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{j=1}^{J} \frac{\partial v_j p_{j|i}}{\partial p_{str}}\right) p_{\bullet\bullet i} - \sum_{i=1}^{I}\left(\sum_{j=1}^{J} v_j p_{j|i}\right)^2 \left(\frac{\partial p_{\bullet\bullet i}}{\partial p_{str}}\right) \\
= \;& \sum_{j=s}^{J} 2 v_j p_{j\bullet\bullet} + v_s^2 - \sum_{i\neq r} 2\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{j=s}^{J} v_j p_{j|i}\right) p_{\bullet\bullet i} \\
& -2\left(\sum_{j=1}^{J} v_j p_{j|r}\right)\left(\sum_{j=1}^{s-1}\left(-v_j \frac{p_{j\bullet r}}{p_{\bullet\bullet r}^2}\right) + p_{s|r} + v_s \frac{p_{\bullet\bullet r} - p_{s\bullet r}}{p_{\bullet\bullet r}^2} + \sum_{j=s+1}^{J}\left(p_{j|r} - v_j \frac{p_{j\bullet r}}{p_{\bullet\bullet r}^2}\right)\right) p_{\bullet\bullet r} \\
& -\left(\sum_{j=1}^{J} v_j p_{j|r}\right)^2
\end{aligned}
$$

$$
\begin{aligned}
\sigma'_{ww\bullet u} \;\equiv\; & \frac{\partial \sum_{k=1}^{K} w_j^2 p_{\bullet k\bullet} - \sum_{i=1}^{I}\left(\sum_{k=1}^{K} w_k p_{k|i}\right)^2 p_{\bullet\bullet i}}{\partial p_{str}} \\
= \;& \sum_{k=t}^{K} 2 w_k p_{\bullet k\bullet} + w_t^2 - \sum_{i\neq r} 2\left(\sum_{k=1}^{K} w_k p_{k|i}\right)\left(\sum_{k=t}^{K} w_k p_{k|i}\right) p_{\bullet\bullet i} \\
& -2\left(\sum_{k=1}^{K} w_k p_{k|r}\right)\left(\sum_{k=1}^{t-1}\left(-w_k \frac{p_{\bullet kr}}{p_{\bullet\bullet r}^2}\right) + p_{t|r} + w_t \frac{p_{\bullet\bullet r} - p_{\bullet tr}}{p_{\bullet\bullet r}^2} + \sum_{k=t+1}^{K}\left(p_{k|r} - w_k \frac{p_{\bullet kr}}{p_{\bullet\bullet r}^2}\right)\right) p_{\bullet\bullet r} \\
& -\left(\sum_{k=1}^{K} w_k p_{k|r}\right)^2 .
\end{aligned}
\tag{S4.10}
$$

By the above two identities, we have

$$
\frac{\partial B_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}} = \frac{\sqrt{\sigma_{ww}}}{2\sqrt{\sigma_{vv}}}\sigma'_{vv\bullet u} + \frac{\sqrt{\sigma_{vv}}}{2\sqrt{\sigma_{ww}}}\sigma'_{ww\bullet u} \tag{S4.11}
$$

where $\sigma'_{vv\bullet u}$ and $\sigma'_{ww\bullet u}$ are given in Eq. (S4.9) and (S4.10), respectively. As we have

$$\frac{\partial \rho_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}} = \frac{\frac{\partial T_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}} B_{YZ\bullet X}(\boldsymbol{p}_{str}) - \frac{\partial B_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}} T_{YZ\bullet X}(\boldsymbol{p})}{B_{YZ\bullet X}(\boldsymbol{p})^2}, \tag{S4.12}$$

where $\frac{\partial T_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}}$ and $\frac{\partial B_{YZ\bullet X}(\boldsymbol{p})}{\partial p_{str}}$ are given in Eq. (S4.8) and (S4.11), respectively, the gradient for the function $h_{YZ\bullet X}(\boldsymbol{p})$ is

$$\nabla h_{YZ\bullet X}(\boldsymbol{p}) = = \left[\frac{\partial}{\partial p_{111}} \rho_{YZ\bullet X}, \frac{\partial}{\partial p_{211}} \rho_{YZ\bullet X}, \ \ldots \ , \frac{\partial}{\partial p_{JKI}} \rho_{YZ\bullet X} \right]^T. \tag{S4.13}$$

Thus, by the delta method, we obtain

$$\sqrt{n}\left(\hat{\rho}_{YZ\bullet X} - \rho_{YZ\bullet X}\right) \xrightarrow{D} N_1\left(0, \nabla h_{YZ\bullet X}(\boldsymbol{p})^T \left(diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^T\right) \nabla h_{YZ\bullet X}(\boldsymbol{p})\right).$$

□

S5 Proof of Theorem 4.3

Proof. We denote the relative frequencies of the observed three-way $J \times K \times I$ contingency table by $\hat{P} = \left\{\hat{p}_{jki} \mid j = 1, \ldots, J, k = 1, \ldots, K, i = 1, \ldots, I\right\}$. Define $\hat{\boldsymbol{p}}$ to be the $(JKI) \times 1$-dimensional estimator for $\boldsymbol{p}$ where $\hat{\boldsymbol{p}} = (\hat{p}_{111}, \hat{p}_{211}, \ldots, \hat{p}_{J11}, \hat{p}_{121}, \ldots, \hat{p}_{JKI})^T$.

As shown in the proof of Theorem 4.2, we have

$$\sqrt{n}\left(\hat{\boldsymbol{p}} - \boldsymbol{p}\right) \xrightarrow{D} N_{JKI}\left(\boldsymbol{0}, \left(diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^T\right)\right).$$

After simple calculations, we obtain an alternative formula for the estimator given in Eq. (4.23):

$$\hat{\rho}_{YZ|X=x_i} = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K}\hat{v}_j\hat{w}_k\hat{p}_{jk|i} - \left(\sum_{j=1}^{J}\hat{v}_j\hat{p}_{j|i}\right)\left(\sum_{k=1}^{K}\hat{w}_k\hat{p}_{k|i}\right)}{\sqrt{\sum_{j=1}^{J}\hat{v}_j^2\hat{p}_{j|i} - \left(\sum_{j=1}^{J}\hat{v}_j\hat{p}_{j|i}\right)^2}\sqrt{\sum_{k=1}^{K}\hat{w}_k\hat{p}_{k|i} - \left(\sum_{k=1}^{K}\hat{w}_k\hat{p}_{k|i}\right)^2}}.$$

The proposed conditional correlation of Y and Z given X, $\rho_{YZ|X=x_i}$, is a function of $\boldsymbol{p}$ and we denote $\rho_{YZ|X=x_i} = h_{YZ|X=x_i}(\boldsymbol{p})$. We now calculate the gradient for the function $h_{YZ|X=x_i}$. Denote $T_{YZ|X=x_i}(\boldsymbol{p})$ to be the numerator of $\rho_{YZ|X=x_i}$, and

$$T_{YZ|X=x_i}(\boldsymbol{p}) = \sum_{j=1}^{J}\sum_{k=1}^{K} v_j w_k p_{jk|i} - \left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{k=1}^{K} w_k p_{k|i}\right).$$

We denote $B_{YZ|X=x_i}(\boldsymbol{p})$ to be the denominator of $\rho_{YZ|X=x_i}$,

$$B_{YZ|X=x_i}(\boldsymbol{p}) = \sqrt{\sum_{j=1}^{J} v_j^2 p_{j|i} - \left(\sum_{j=1}^{J} v_j p_{j|i}\right)^2}\sqrt{\sum_{k=1}^{K} w_k p_{k|i} - \left(\sum_{k=1}^{K} w_k p_{k|i}\right)^2}$$

Given $s = 1, \ldots, J$, $t = 1, \ldots, K$, and $r = 1, \ldots, I$, and by (i)-(iii) in the proof of Theorem 4.2, we have

$$\frac{\partial T_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}} = \sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\partial v_j w_k p_{jk|i}}{\partial p_{str}} - \frac{\partial}{\partial p_{str}}\left[\left(\sum_{j=1}^{J} v_j p_{j|i}\right)\left(\sum_{k=1}^{K} w_k p_{k|i}\right)\right] \quad \text{(S5.14)}$$

$$
\begin{aligned}
= &\sum_{j=s}^{J}\sum_{k=1}^{K} w_k p_{jk|i} + \sum_{j=1}^{J}\sum_{k=t}^{K} v_j p_{jk|i} + \left[v_s w_t \left(\frac{p_{\bullet\bullet r} - p_{str}}{p_{\bullet\bullet r}^2} \right) - \sum_{j\neq s}^{J}\sum_{k\neq t}^{K} \left(\frac{v_j w_k p_{jkr}}{p_{\bullet\bullet r}^2} \right) \right] I(r=i) \\
&- \left[\left(\sum_{j=1}^{s-1} \left(-v_j \frac{p_{j\bullet i}}{p_{\bullet\bullet i}^2} \right) + p_{s|i} + v_s \frac{p_{\bullet\bullet i} - p_{s\bullet i}}{p_{\bullet\bullet i}^2} + \sum_{j=s+1}^{J} \left(p_{j|i} - v_j \frac{p_{j\bullet i}}{p_{\bullet\bullet i}^2} \right) \right) I(r=i) \right. \\
&\left. + \left(\sum_{j=s}^{J} p_{j|i} \right) I(r\neq i) \right] \left(\sum_{k=1}^{K} w_k p_{k|i} \right) - \left(\sum_{j=1}^{J} v_j p_{j|i} \right) \\
&\times \left[\left(\sum_{k=1}^{t-1} \left(-w_k \frac{p_{\bullet ki}}{p_{\bullet\bullet i}^2} \right) + p_{t|i} + w_t \frac{p_{\bullet\bullet i} - p_{\bullet ti}}{p_{\bullet\bullet i}^2} + \sum_{k=t+1}^{K} \left(p_{k|i} - w_k \frac{p_{\bullet ki}}{p_{\bullet\bullet i}^2} \right) \right) I(r=i) \right. \\
&\left. + \left(\sum_{k=t}^{K} p_{k|i} \right) I(r\neq i) \right],
\end{aligned}
$$

where $I(\cdot)$ is the indicator function.

To calculate $\frac{\partial B_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}}$, the following two equalities will be useful :

$$
\begin{aligned}
\sigma'_{vv|u_i} &\equiv \frac{\partial \left(\sum_{j=1}^{J} v_j^2 p_{j|i} - \left(\sum_{j=1}^{J} v_j p_{j|i} \right)^2 \right)}{\partial p_{str}} = \frac{\partial \left(\sum_{j=1}^{J} v_j^2 p_{j|i} \right)}{\partial p_{str}} - 2 \left(\sum_{j=1}^{J} v_j p_{j|i} \right) \frac{\partial \left(\sum_{j=1}^{J} v_j p_{j|i} \right)}{\partial p_{str}} \\
&= \sum_{j=s}^{J} 2 v_j p_{j|i} + \sum_{j=s}^{J} v_j^2 \frac{\partial p_{j|i}}{\partial p_{str}} - 2 \left(\sum_{j=1}^{J} v_j p_{j|i} \right) \left[\frac{\partial \left(\sum_{j=1}^{J} v_j p_{j|i} \right)}{\partial p_{str}} I(r=i) + \frac{\partial \left(\sum_{j=1}^{J} v_j p_{j|i} \right)}{\partial p_{str}} I(r\neq i) \right] \\
&= \sum_{j=s}^{J} 2 v_j p_{j|i} + I(r=i) \left(\sum_{j=1}^{J} v_j^2 \frac{p_{\bullet\bullet i} I(j=s) - p_{j\bullet i}}{p_{\bullet\bullet i}^2} \right) \qquad \text{(S5.15)} \\
&-2 \left(\sum_{j=1}^{J} v_j p_{j|i} \right) \left[\left(\sum_{j=1}^{s-1} \left(-v_j \frac{p_{j\bullet i}}{p_{\bullet\bullet i}^2} \right) + p_{s|i} + v_s \frac{p_{\bullet\bullet i} - p_{s\bullet i}}{p_{\bullet\bullet i}^2} + \sum_{j=s+1}^{J} \left(p_{j|i} - v_j \frac{p_{j\bullet i}}{p_{\bullet\bullet i}^2} \right) \right) I(r=i) \right.
\end{aligned}
$$

$$+\left(\sum_{j=s}^{J} p_{j|i}\right) I(r \neq i)\Bigg]$$

$$\begin{aligned}
\sigma'_{ww|u_i} &\equiv \frac{\partial\left(\sum_{k=1}^{K} w_k p_{k|i} - \left(\sum_{k=1}^{K} w_k p_{k|i}\right)^2\right)}{\partial p_{str}} \\
&= \sum_{k=t}^{K} 2w_k p_{k|i} + I(r=i)\left(\sum_{k=1}^{K} w_k^2 \frac{p_{\bullet\bullet i} I(k=t) - p_{\bullet ki}}{p_{\bullet\bullet i}^2}\right) \qquad \text{(S5.16)} \\
&\quad -2\left(\sum_{k=1}^{K} w_k p_{k|i}\right)\left[\left(\sum_{k=1}^{t-1}\left(-w_k \frac{p_{\bullet ki}}{p_{\bullet\bullet i}^2}\right) + p_{t|i} + w_t \frac{p_{\bullet\bullet i} - p_{\bullet ti}}{p_{\bullet\bullet i}^2} + \sum_{k=t+1}^{K}\left(p_{k|i} - w_k \frac{p_{\bullet ki}}{p_{\bullet\bullet i}^2}\right)\right) I(r=i)\right. \\
&\quad \left. + \left(\sum_{k=t}^{K} p_{k|i}\right) I(r \neq i)\right]
\end{aligned}$$

By the above two identities, we have

$$\frac{\partial B_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}} = \frac{\sqrt{\sigma_{ww|u_i}}}{2\sqrt{\sigma_{vv|u_i}}}\sigma'_{vv|u_i} + \frac{\sqrt{\sigma_{vv|u_i}}}{2\sqrt{\sigma_{ww|u_i}}}\sigma'_{ww|u_i} \qquad \text{(S5.17)}$$

where $\sigma'_{vv|u_i}$ and $\sigma'_{ww|u_i}$ are given in Eq. (S5.15) and (S5.16), respectively. As we have

$$\frac{\partial \rho_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}} = \frac{\frac{\partial T_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}} B_{YZ|X=x_i}(\boldsymbol{p}) - \frac{\partial B_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}} T_{YZ|X=x_i}(\boldsymbol{p})}{B_{YZ|X=x_i}(\boldsymbol{p})^2}, \qquad \text{(S5.18)}$$

where $\frac{\partial T_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}}$ and $\frac{\partial B_{YZ|X=x_i}(\boldsymbol{p})}{\partial p_{str}}$ are given in Eq. (S5.14) and (S5.17), respectively, the gradient for the function $h_{YZ|X=x_i}(\boldsymbol{p})$ is

$$\nabla h_{YZ|X=x_i}(\boldsymbol{p}) = \left[\frac{\partial}{\partial p_{111}}\rho_{YZ|X=x_i}, \frac{\partial}{\partial p_{211}}\rho_{YZ|X=x_i}, \ldots, \frac{\partial}{\partial p_{JKI}}\rho_{YZ|X=x_i}\right]^T. \qquad \text{(S5.19)}$$

Thus, by the delta method, we obtain

$$\sqrt{n}\left(\hat{\rho}_{YZ|X=x_i} - \rho_{YZ|X=x_i}\right) \xrightarrow{D} N_1\left(0, \nabla h_{YZ|X=x_i}(\boldsymbol{p})^T \left(diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^T\right) \nabla h_{YZ|X=x_i}(\boldsymbol{p})\right).$$

□

S6 Simulations

S6.1 Simulation Design

In this section, we provide detailed information on how the simulation study in Section 5.1.1, of the paper was designed.

We utilized the homogeneous linear-by-linear association loglinear model given in Eq. (5.25) to simulate the three-way contingency tables for conditional independence, linear pattern with different magnitude of association levels, and monotone nonlinear pattern, respectively. To simulate tables with nonmonotone nonlinear pattern, we applied the homogeneous column effect loglinear models with quadratic terms in Eq. (5.26).

(I) For conditional independence scenario, we set $\beta = 0$ and used the values given in Table S6.1 for the parameters in the homogeneous linear-by-linear association loglinear model of Eq. (5.25).

Table S6.1: Parameters in the homogeneous linear-by-linear association loglinear model with conditional independence in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	7.0644	8.4956
λ_2^X	−1.9457	−2.5559
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [-0.3719, -0.9729]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [-0.2321, -0.4474, -0.7219, -1.2779]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [-0.3719, -0.9729]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [-0.2321, -0.4474, -0.7219, -1.2779]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [0.9729, 1.9457]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [0.7881, 1.2779, 1.7678, 2.5559]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [0.9729, 1.9457]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [0.7881, 1.2779, 1.7678, 2.5559]$

(II) For strong, moderate, and weak linear pattern scenarios, we employed the values in Tables (S6.2), (S6.3), (S6.4) for the parameter values in the homogeneous linear-by-linear association loglinear model of Eq. (5.25), respectively.

Table S6.2: Parameters in the homogeneous linear-by-linear association loglinear model for strong linear pattern in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	2.19503	2.2225
λ_2^X	0.0594	−0.2213
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [-6.7331, -18.0214]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [-3.6486, -9.4986, -17.8784, -28.7676]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [-6.3978, -17.2230]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [-4.1540, -9.8150, -18.3336, -29.2190]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [0.2819, 0.1408]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [0.1081, -0.5004, -0.8620, -0.9104]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [-0.3879, -0.4658]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [0.5959, 0.5089, 0.9146, 0.3767]$
β	4.3937	2.4069

Table S6.3: Parameters in the homogeneous linear-by-linear association loglinear model for moderate linear pattern in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	3.2154	2.0376
λ_2^X	−0.1002	−0.1244
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [-2.0816, -5.4615]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [-1.0063, -2.6788, -4.2281, -6.6602]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [-2.3021, -6.0437]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [-0.7338, -1.9922, -3.4758, -5.7510]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [-0.6283, -0.4706]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [1.0111, 1.0965, 0.4547, 0.3917]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [0.4654, 0.2857]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [-0.7310, -0.4222, -1.2901, -0.8818]$
β	1.4254	0.5107

Table S6.4: Parameters in the homogeneous linear-by-linear association loglinear model for weak linear pattern in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	4.2398	2.1281
λ_2^X	−0.4028	−0.2002
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [-0.2655, -0.9196]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [-0.0020, 0.5666, -0.6436, -1.5774]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [-0.9059, -0.9856]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [-0.3078, -0.3312, -0.8833, -1.1794]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [0.2618, 0.1693]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [-1.0055, 0.0015, -0.5140, -0.8016]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [0.4507, -0.2031]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [0.3379, 0.0744, 0.5459, 0.6586]$
β	0.2581	0.0956

(III) For monotone nonlinear pattern scenario, we chose the unit-spaced scores $v_j = 1, \ldots, J$ for Y and the quadratic spaced scores $w_k = 1^2, \ldots, K^2$ for Z, and used the parameter values in Tables (S6.5) for the homogeneous linear-by-linear association loglinear model in Eq. (5.25).

Table S6.5: Parameters in the homogeneous linear-by-linear association loglinear model for monotone nonlinear pattern in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	4.0401	3.8439
λ_2^X	0.1904	0.0899
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [-2.9248, -7.6477]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [-0.8954, -2.6147, -5.1746, -8.3871]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [0.6519, -3.9253]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [1.4865, 1.6237, 0.2866, -4.0266]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [0.2357, 0.2763]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [0.0923, 0.1322, 0.1722, 0.2299]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [-0.5134, -0.4899]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [-0.4287, -0.4606, -0.4035, -0.3777]$
β	0.2961	0.0548

(IV) For nonmonotone nonlinear pattern scenario, we utilized the homogeneous column effect loglinear models with quadratic terms in Eq. (5.26) where the unit-spaced scores $v_j = 1, \ldots, J$ were used for Y, and ω_k and η_k are the parameters with constraints $\omega_1 = \eta_1 = 0$. We used the parameter values given in Tables (S6.6) to simulate the contingency table with the clear nonmonotone nonlinear pattern.

Table S6.6: Parameters in the homogeneous column effect loglinear models with quadratic terms for nonmonotone nonlinear pattern in each table size.

Table size $J \times K \times I$	$3 \times 3 \times 2$	$5 \times 5 \times 2$
λ	3.391	−11.2290
λ_2^X	0.3537	0.5700
λ_j^Y	$[\lambda_2^Y, \lambda_3^Y] = [4.9931, 0.0327]$	$[\lambda_2^Y, \lambda_3^Y, \lambda_4^Y, \lambda_5^Y] = [15.3898, 18.7921, 15.3223, 0.2368]$
λ_k^Z	$[\lambda_2^Z, \lambda_3^Z] = [84.4988, 16.1853]$	$[\lambda_2^Z, \lambda_3^Z, \lambda_4^Z, \lambda_5^Z] = [44.2209, 41.3418, 40.9474, 35.4598]$
λ_{j2}^{YX}	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}] = [-0.3514, 0.1605]$	$[\lambda_{22}^{YX}, \lambda_{32}^{YX}, \lambda_{42}^{YX}, \lambda_{52}^{YX}] = [-0.7441, -0.5798, -0.6370, -0.2152]$
λ_{k2}^{ZX}	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}] = [-0.6089, -0.2275]$	$[\lambda_{22}^{ZX}, \lambda_{32}^{ZX}, \lambda_{42}^{ZX}, \lambda_{52}^{ZX}] = [-0.3351, -0.2711, -0.6193, 0.1345]$
ω_k	$[\omega_2, \omega_3] = [-112.7854, -19.3217]$	$[\omega_2, \omega_3, \omega_4, \omega_5] = [-36.5816, -33.0860, -32.6211, -26.0243]$
η_k	$[\eta_2, \eta_3] = [28.1720, 4.8130]$	$[\eta_2, \eta_3, \eta_4, \eta_5] = [6.0854, 5.4973, 5.4389, 4.3396]$

S6.2 Simulation results for scenario [I] and [II]

Figures 1 and 2 show the boxplots of the conditional/partial association measures for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables with $n = (1000, 5000)$ and conditional independence. Note that Figures 1 and 2 include the four conditional measures and the three partial measures, respectively. First, we observe that in both $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables, each of all conditional/partial association measures symmetrically scatters around zero, and as the sample size n increases, the biases and variabilities of all four measures decrease. Second, the subcopula regression-based conditional/partial correlation perform similar to the conditional/partial Spearman correlation, regardless of the sizes of the table and the sample size. Third, for $3 \times 3 \times 2$ table the conditional Gamma measure shows larger variabilities than the other three conditional measures, and for $5 \times 5 \times 2$ table, the conditional/partial tau measures have smaller variabilities than the other conditional/partial measures. These phenomena become pronounced as n decreases.

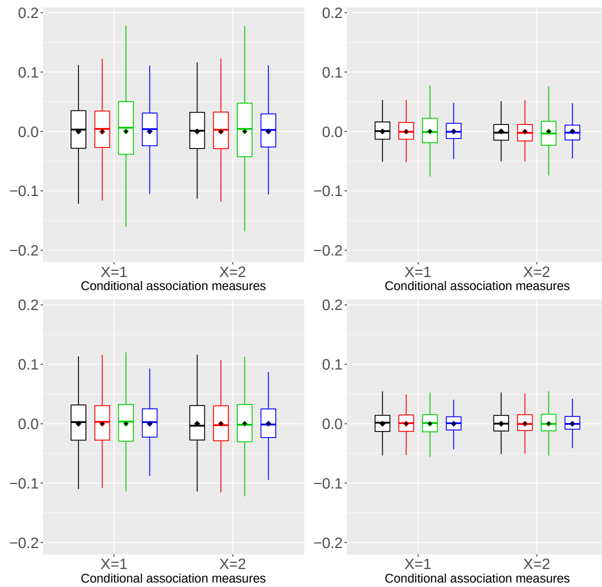

Figure 1: Conditional independence - Boxplots of the conditional association measures (ordered as subcopula regression based correlation, Spearman correlation, Gamma and Tau) for 3×3×2 and 5×5×2 tables (from top to bottom panel) with n=(1000,5000) (from left column to right column).

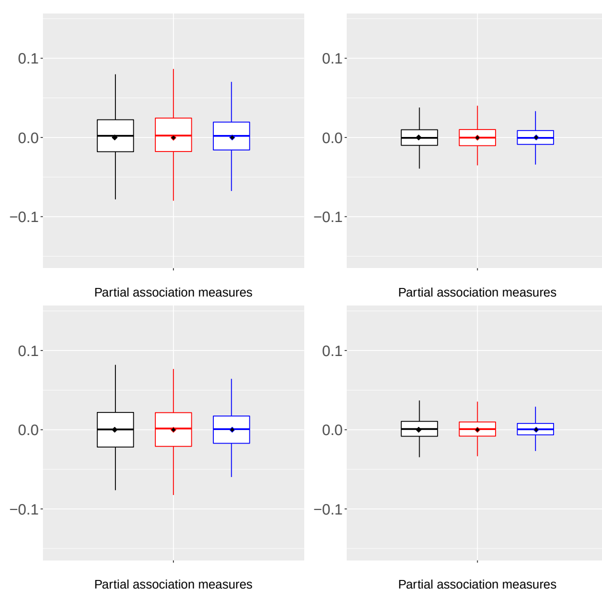

Figure 2: Conditional independence - Boxplots of the partial association measures (ordered as subcopula regression based correlation, Spearman correlation and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $n = (1000, 5000)$ (from left column to right column).

Figures 3 - 8 give the boxplots of the association measures (subcopula regression based correlation, Spearman correlation, Gamma and Tau measures) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables with n=(1000,5000) and linear patterns. Note that Figures 3 and 4 show the results for a weak linear association, Figures 5 and 6 for the results for a moderate linear association, and Figures 7 and 8 for the results for a strong linear pattern. Figures 3, 5, 7 include the boxplots for the four conditional measures at each level of X, and Figures 4, 6, 8 for the three partial measures.

First, we observe that the biases and variabilities of all association measures decrease as n increases and/or the degree of linear association increases, regardless of the size of contingency table (the values of (J, K)). Second, the subcopula regression-based correlation and the Spearman correlation perform similarly irrespective of the magnitude of linear association, the sample size and the size of contingency table. Third, for $3 \times 3 \times 2$ table with weak/moderate/strong linear pattern, the conditional Gamma measure tends to be much larger than the other three conditional measures and the conditional/partial tau measure tends to be smaller than the conditional/partial subcopula regression-based correlation and the conditional/partial Spearman correlation. For $5 \times 5 \times 2$ table, the performances of the association measures are different, depending on the magnitude of association. For the weak linear pattern, the conditional/partial tau measure is smaller than the other conditional/partial measures. For the moderate/strong linear patterns, the conditional/partial tau measure appears to be smaller than the other conditional/partial measures, and the conditional Gamma measure tends to be much larger than the subcopula regression-based conditional correlation and the conditional Spearman correlation.

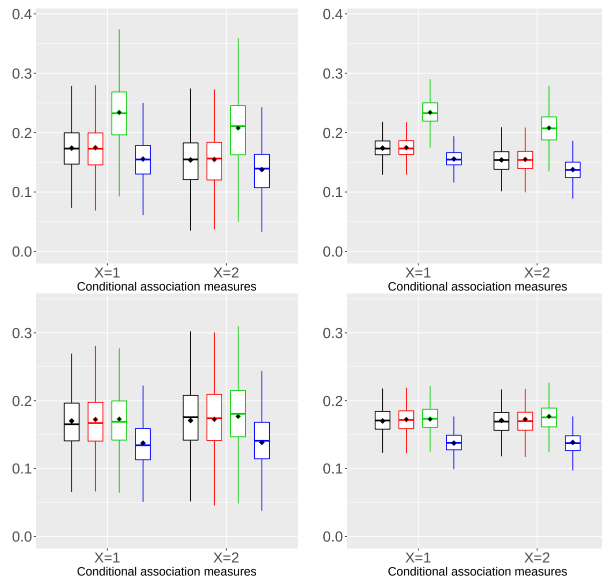

Figure 3: Weak linear pattern - Boxplots of the conditional association measures (subcopula regression based correlation, Spearman correlation, Gamma and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

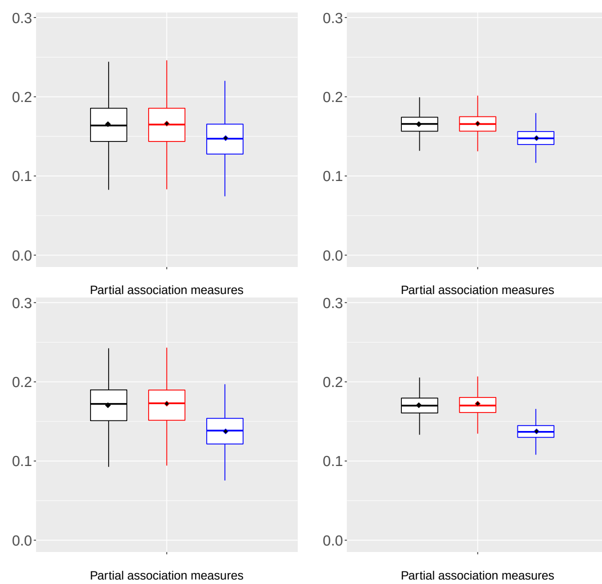

Figure 4: Weak linear pattern - Boxplots of the partial association measures (subcopula regression based correlation, Spearman correlation and Tau) for $3\times3\times2$ and $5\times5\times2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

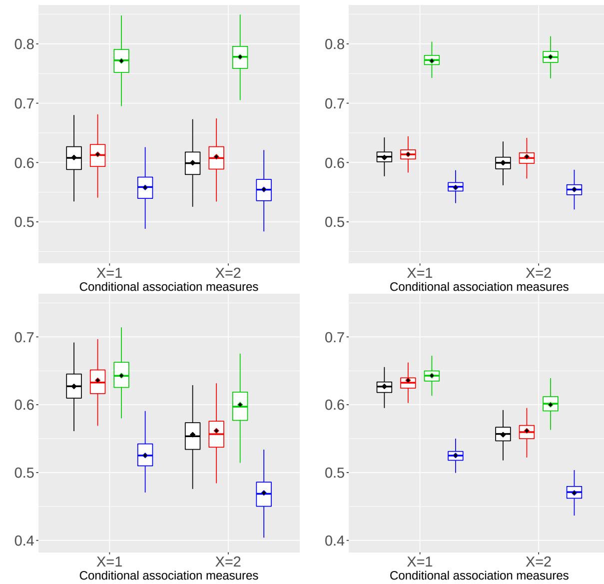

Figure 5: Moderate linear pattern - Boxplots of the conditional association measures (subcopula regression based correlation, Spearman correlation, Gamma and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

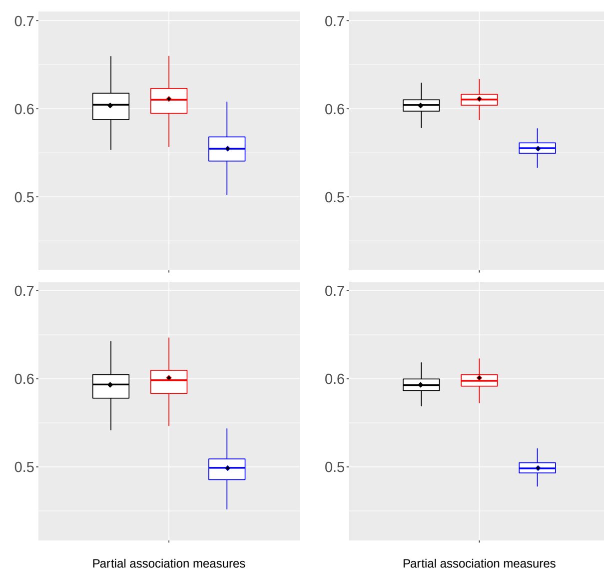

Figure 6: Moderate linear pattern - Boxplots of the partial association measures (subcopula regression based correlation, Spearman correlation and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

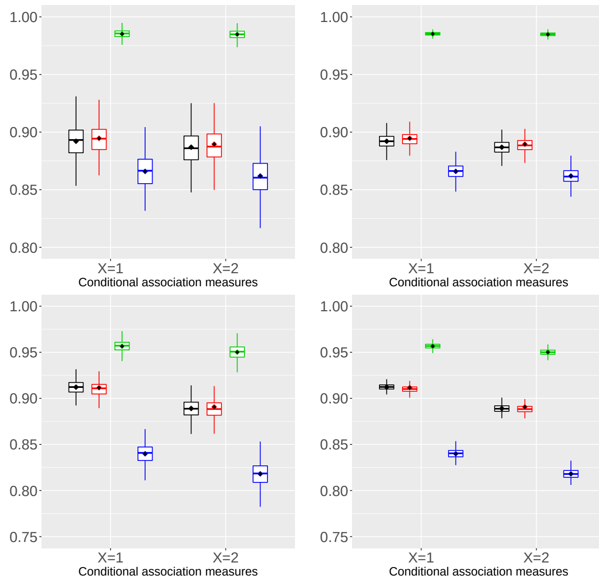

Figure 7: Strong linear pattern - Boxplots of the conditional association measures (subcopula regression based correlation, Spearman correlation, Gamma and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

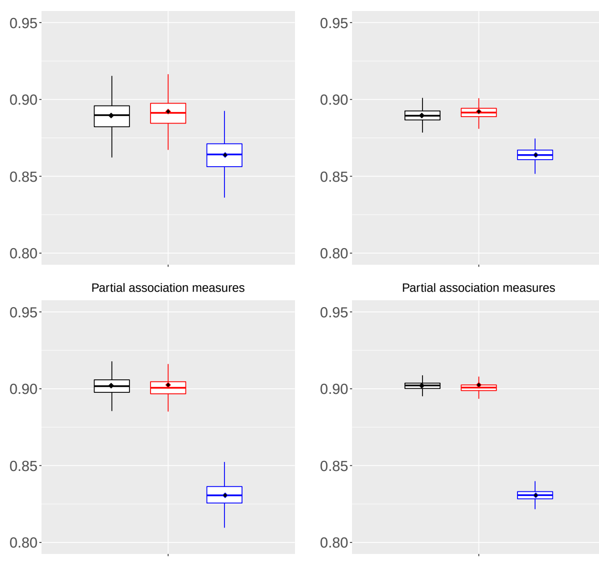

Figure 8: Strong linear pattern - Boxplots of the partial association measures (subcopula regression based correlation, Spearman correlation and Tau) for $3 \times 3 \times 2$ and $5 \times 5 \times 2$ tables (from top panel to bottom panel) with $N = (1000, 5000)$ (from left column to right column).

S6.3 Simulation design for further investigation on the U-shaped associations

In this section, we provide detailed information on how the contingency tables with different types of U-shaped associations were simulated in Section 5.2 of the paper.

To simulate the contingency tables for two ordinal variables Y and Z with different types of U-shaped associations at each level of binary covariate X, we consider the quadratic latent variable regression model based on the underlying continuous variables Y^* and Z^* for Y and Z. That is, we generate $n(=10^6)$ values of (Y^*, Z^*, X) from the quadratic latent variable regression with a normal error distribution and discretize them to construct a $5 \times 5 \times 2$ contingency table for (Y, Z, X) of sample size n.

Specifically, we first employ the four quadratic latent variable regression models to consider the relationships between Y^* and Z^* conditional on a value of X :

$$\mathrm{M}_1 : Z_i^* = -X_i^2[(Y_i^* - \beta_1)^2 - \beta_0] + \epsilon_i \text{ and } Y_i^* \sim N(\beta_1, 1),$$

$$\mathrm{M}_2 : Z_i^* = X_i^2[(Y_i^* - \beta_1)^2 - \beta_0] + \epsilon_i \text{ and } Y_i^* \sim N(\beta_1, 1),$$

$$\mathrm{M}_3 : Y_i^* = -X_i^2[(Z_i^* - \beta_1)^2 - \beta_0] + \epsilon_i \text{ and } Z_i^* \sim N(\beta_1, 1),$$

$$\mathrm{M}_4 : Y_i^* = X_i^2[(Z_i^* - \beta_1)^2 - \beta_0] + \epsilon_i \text{ and } Z_i^* \sim N(\beta_1, 1),$$

where i=$1, \ldots, n$, β_1=1, β_0= 10, X_i (taking 1 or -1) $\sim$ Bernoulli(0.5) and $\epsilon_i \sim N(0, 0.001)$. For the model M_1(M_2), Z^* increases (decreases) and then decreases (increases) as Y^* increases, and under the model M_3(M_4), Y^* increases (decreases) and then decreases (increases) as Z^* increases.

Given $n(= 10^6)$ observations of (Y^*, Z^*, X) generated from a quadratic latent variable regression model and conditional on a value of X, we discretize Y^* and Z^* using the quantiles of the normal distribution based on suitable discrete marginal distributions for Y and Z. We then construct three $5 \times 5 \times 2$ contingency tables, each with one of three U-shaped patterns: (a) smooth, (b) pointy and (c) flat patterns. Note that the discrete marginal distributions for Y and Z, given in Table S6.7, are chosen to obtain the targeted U-shaped pattern.

Table S6.7: List of discrete marginal distributions for Y and Z employed to simulate the contingency tables with targeted relationships and U-shaped patterns.

Relationship \ U-shaped pattern	Smooth	Pointy	Flat
M_1	(0.008, 0.011, 0.962, 0.011, 0.008) (0.003, 0.003, 0.013, 0.024, 0.957)	(0.076, 0.063, 0.722, 0.063, 0.076) (0.152, 0.127, 0.218, 0.252, 0.252)	(0.096, 0.003, 0.801, 0.003, 0.097) (0.028, 0.028, 0.055, 0.083, 0.807)
M_2	(0.008, 0.011, 0.962, 0.011, 0.008) (0.957, 0.024, 0.013, 0.003, 0.003)	(0.076, 0.063, 0.722, 0.063, 0.076) (0.252, 0.252, 0.219, 0.126, 0.152)	(0.097, 0.003, 0.801, 0.003, 0.097) (0.807, 0.083, 0.055, 0.028, 0.028)
M_3	(0.003, 0.003, 0.013, 0.024, 0.957) (0.008, 0.011, 0.962, 0.011, 0.008)	(0.152, 0.127, 0.218, 0.252, 0.252) (0.076, 0.063, 0.722 ,0.063, 0.076)	(0.028, 0.028, 0.055, 0.083, 0.807) (0.097, 0.003, 0.801, 0.003, 0.097)
M_4	(0.957, 0.024, 0.013, 0.003, 0.003) (0.008, 0.011, 0.962, 0.011, 0.008)	(0.252, 0.252, 0.218, 0.127, 0.152) (0.076, 0.063, 0.722, 0.063, 0.076)	(0.807, 0.083, 0.055, 0.028, 0.028) (0.096, 0.003, 0.801, 0.002, 0.097)