# Online supplementary material

All equations, lemmas, tables, etc in the main paper are referred to as equation (1), lemma 1, table 1, etc, and in this supplement they are referred to as equation (S1), lemma S1 and table S1, etc.

# S1 Some further technical results for the Gumbel and Clayton copulas

## S1.1 The Gumbel copula

The $J$-dimensional Gumbel copula is another popular example of Archimedean copulas. Its cdf $C(\boldsymbol{u})$ and density $c(\boldsymbol{u})$ are

$$\begin{aligned}
C(\boldsymbol{u}) &:= \exp\left\{-\left[\sum_{j=1}^{J}\left(-\log\left(u_j\right)\right)^{\theta}\right]^{1/\theta}\right\} \\
c(\boldsymbol{u}) &:= \theta^{J}\exp\left\{-\left[\sum_{j=1}^{J}\left(-\log\left(u_j\right)\right)^{\theta}\right]^{\frac{1}{\theta}}\right\} \\
&\times \frac{\prod_{j=1}^{J}\left(-\log\left(u_j\right)\right)^{\theta-1}}{\left(\sum_{j=1}^{J}\left(-\log\left(u_j\right)\right)^{\theta}\right)^{J}\prod_{j=1}^{J}u_j} \times P_{J,\theta}^{G}\left(\left[\sum_{j=1}^{J}\left(-\log\left(u_j\right)\right)^{\theta}\right]^{\frac{1}{\theta}}\right),
\end{aligned}$$

where

$$P_{J,\theta}^{G}(x) = \sum_{k=1}^{J} a_{mk}^{G}(\theta)\, x^{k},$$

and

$$a_{mk}^{G}(\theta) = \frac{J!}{k!}\sum_{j=1}^{k}\binom{k}{j}\binom{j/\theta}{J}(-1)^{J-j}.$$

The dependence parameter $\theta$ is defined on $[1,\infty)$, where a value of 1 represents the independence case. The Gumbel copula is an appropriate choice if the data exhibit weak correlation at lower values and strong correlation at higher values.

If some of the $a_j$ are zero, then directly estimating the integral (3) is computationally inefficient for the same reasons as given in section 2.2 for the Clayton copula. It can be readily checked that

$$\begin{aligned}
D\left(\boldsymbol{u}_{1:K},\boldsymbol{b}_{K+1:J}\right) &:= \theta^{K}\exp\left\{-\left[\sum_{j=1}^{K}\left(-\log\left(u_j\right)\right)^{\theta}+\sum_{j=K+1}^{J}\left(-\log\left(b_j\right)\right)^{\theta}\right]^{\frac{1}{\theta}}\right\} \\
&\times \frac{\prod_{j=1}^{K}\left(-\log\left(u_j\right)\right)^{\theta-1}}{\left(\sum_{j=1}^{K}\left(-\log\left(u_j\right)\right)^{\theta}+\sum_{j=K+1}^{J}\left(-\log\left(b_j\right)\right)^{\theta}\right)^{K}\prod_{j=1}^{K}u_j} \\
&\times P_{K,\theta}^{G}\left(\left[\sum_{j=1}^{K}\left(-\log\left(u_j\right)\right)^{\theta}+\sum_{j=K+1}^{J}\left(-\log\left(b_j\right)\right)^{\theta}\right]^{\frac{1}{\theta}}\right).
\end{aligned}$$

Then, we can rewrite the integral as

$$\int_{a_1}^{b_1} ... \int_{a_K}^{b_K} D\left(\boldsymbol{u}_{1:K}, \boldsymbol{b}_{K+1:J}\right) \mathrm{d}\boldsymbol{u}_{1:K} = \prod_{j=1}^{K} (b_j - a_j)$$
$$\times \int_0^1 ... \int_0^1 D\left(\left(b_1 - a_1\right) v_1 + a_1, ..., \left(b_K - a_K\right) v_K + a_K, \boldsymbol{b}_{K+1:J}\right) d\boldsymbol{v}_{1:K}.$$

## S1.2 The VBIL approximation distribution

For the Clayton copula, the VB approximation to the posterior of $\theta$ is the inverse gamma density

$$q_\lambda(\theta) = \frac{a^b}{\Gamma(a)} (\theta)^{-1-a} \exp\left(-b/\theta\right), \ \theta > 0,$$

and for the Gumbel copula

$$q_\lambda(\theta) = \frac{a^b}{\Gamma(a)} (\theta - 1)^{-1-a} \exp\left(-b/\left(\theta - 1\right)\right), \ \theta > 1,$$

with the natural parameters $a$ and $b$. The Fisher information matrix for the inverse gamma is

$$I_F(a, b) = \begin{pmatrix} \nabla_{aa}\left[\log \Gamma(a)\right] & -1/b \\ -1/b & a/b^2 \end{pmatrix}$$

with gradient

$$\nabla_a\left[\log q_\lambda(\theta)\right] = -\log(\theta) + \log(b) - \nabla_a\left[\log \Gamma(a)\right] \quad \text{and} \quad \nabla_b\left[\log q_\lambda(\theta)\right] = -\frac{1}{\theta} + \frac{a}{b}.$$

# S2 Further description and analysis of the well-being and life-shock events dataset

This section gives further details of the of the well-being and life-shock events dataset (abbreviated to 'well-being dataset') described in section 5.1. The health data used in this paper is obtained from the SF-36 data collected by the HILDA survey. The SF-36 (Medical Outcome Trust, Boston, MA) is a multipurpose and short form health survey with 36 items. Each item provides multiple choice answers for respondents to select from in regard to different aspects of their health. SF-36 is one of the most widely used generic measures of health-related quality of life (HRQoL) in clinical research and general population health. It is a standardised questionnaire used to assess patient health across eight attributes (Ware et al., 1993). These are physical functioning (PF, items 3 to 12), role-physical (RP, items 13 to 16), bodily pain (BP, items 21 and 22), general health (GH, items 1, 2, 33-36), vitality (VT, items 28-31), social functioning (SF, items 20 and 32), role-emotional (RE, items 17 to 19), and mental health (MH, items 23-27). The details of the survey questions can be found in Ware et al. (1993).

# S3 Details of the data augmentation approach

This section gives further details of Algorithm 3. The conditional distribution of $p\left(\boldsymbol{u}_{(j)}|\boldsymbol{\theta}, \boldsymbol{u}_{(k\neq j)}, \boldsymbol{x}\right)$ is given by

$$\begin{aligned} p\left(\boldsymbol{u}_{(j)}|\boldsymbol{\theta}, \boldsymbol{u}_{(k\neq j)}, \boldsymbol{x}\right) &\propto p\left(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{u}\right) p\left(\boldsymbol{u}_{(j)}|\boldsymbol{\theta}, \boldsymbol{u}_{(k\neq j)}\right) \\ &\propto \prod_{i=1}^{n} I\left(a_{i,j} \leq u_{i,j} < b_{i,j}\right) c\left(\boldsymbol{u}_i; \boldsymbol{\theta}\right) \\ &\propto \prod_{i=1}^{n} I\left(a_{i,j} \leq u_{i,j} < b_{i,j}\right) c_{j|k\neq j}\left(u_{i,j}|u_{i,k\neq j}; \boldsymbol{\theta}\right) \end{aligned}$$

The latents $u_{i,j}$ are generated from the conditional densities $c_{j|k\neq j}$ constrained to $[a_{i,j}, b_{i,j})$ and an iterate of $\boldsymbol{u}_{(j)}$ obtained. In this sampling scheme, the copula parameter $\boldsymbol{\theta}$ is generated conditional on $\boldsymbol{u}$ from

$$p\left(\boldsymbol{\theta}|\boldsymbol{u}, \boldsymbol{x}\right) = p\left(\boldsymbol{\theta}|\boldsymbol{u}\right) \propto \prod_{i=1}^{n} c\left(\boldsymbol{u}_i; \boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right)$$

The following algorithm is used to generate the latent variables one margin at a time.

For $j = 1, ..., J$ and for $i = 1, ..., n$

- Compute
$$A_{ij} = C_{j|\{1,...,J\}\setminus j}\left(a_{i,j}|\left\{u_{i1}, ..., u_{iJ}\right\} \setminus u_{ij}, \boldsymbol{\theta}\right)$$
and
$$B_{ij} = C_{j|\{1,...,J\}\setminus j}\left(b_{i,j}|\left\{u_{i1}, ..., u_{iJ}\right\} \setminus u_{ij}, \boldsymbol{\theta}\right)$$
- Generate $w_{i,j} \sim Uniform\left(A_{i,j}, B_{i,j}\right)$
- Compute $u_{i,j} = C^{-1}_{j|\{1,...,J\}\setminus j}\left(w_{i,j}|\left\{u_{i1}, ..., u_{iJ}\right\} \setminus u_{ij}, \boldsymbol{\theta}\right)$