Submitted to the Annals of Applied Statistics

MODELLING CORRELATION MATRICES IN MULTIVARIATE DATA, WITH APPLICATION TO RECIPROCITY AND COMPLEMENTARITY OF CHILD-PARENT EXCHANGES OF SUPPORT

By Siliang Zhang[1,a], Jouni Kuha[2,b] and Fiona Steele[2,c]

[1] *Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, School of Statistics, East China Normal University,* [a] *slzhang@fem.ecnu.edu.cn*

[2] *Department of Statistics, London School of Economics and Political Science,* [b] *j.kuha@lse.ac.uk;* [c] *f.a.steele@lse.ac.uk*

We define a model for the joint distribution of multiple continuous latent variables which includes a model for how their correlations depend on explanatory variables. This is motivated by and applied to social scientific research questions in the analysis of intergenerational help and support within families, where the correlations describe reciprocity of help between generations and complementarity of different kinds of help. We propose an MCMC procedure for estimating the model which maintains the positive definiteness of the implied correlation matrices, and describe theoretical results which justify this approach and facilitate efficient implementation of it. The model is applied to data from the UK Household Longitudinal Study to analyse exchanges of practical and financial support between adult individuals and their non-coresident parents.

1. Introduction. Many substantive research questions lead to modelling of multivariate response data. Sometimes the focus of interest is then not just on the means of the response variables, but also on how *associations* between them depend on explanatory variables. In this paper we analyse intergenerational exchanges of family support, where correlations between different types and directions of help correspond to questions about the recriprocity and complementarity of support, and how they may depend on characteristics of the individuals and their families. Other applications where such models for correlations or covariances may be of interest include attitudes of different members of a family, inter-rater agreement in educational and psychological studies, and associations between different measures of health and well-being of an individual.

The methodological literature on such models for associations is much smaller than the one on models for means or variances of individual responses. Specification of a model for correlations or covariances faces a trade-off between two conflicting requirements: ease of interpretation of the model parameters, and ensuring that the association matrices implied by the model are positive definite. In this paper we propose a new modelling framework where a model is specified directly for individual correlations — and is thus easily interpretable — and positive definiteness is monitored and ensured during estimation. We then use it to analyse data on exchanges of help and support between adult individuals and their parents.

In contemporary low-mortality countries, population ageing has led to an increase in the need for help and support for people with age-related functional limitations. At the same time, the need for support may also be increasing among younger people as a result of delayed transitions to adulthood, unstable employment, high cost of living, and rises in divorce and re-partnership rates (Lesthaeghe, 2014; Henretta, Van Voorhis and Soldo, 2018). With limited public resources available to meet these demands, there is a greater reliance on private transfers of support within families, especially between parents and their adult children.

Keywords and phrases: Bayesian estimation, covariance matrix modelling, item response theory models, positive definite matrices, two-step estimation.

The main 'currencies' of such intergenerational exchanges are time (or practical support) and money (Grundy, 2005). Another kind of intergenerational support is coresidence but its overall rate remains low, in spite of a small increase in coresidence between young adults and their parents (e.g. Stone, Berrington and Falkingham, 2011). Transfers of practical and financial support between relatives living in different households are thus the key form of family exchanges. Understanding the nature of these exchanges is important for anticipating which population sub-groups may be at risk of unmet need for support or experience a reduced capacity to provide support due to changes in their circumstances.

Previous research highlights the importance of reciprocity (symmetry) in such exchanges, either contemporaneously or over the life course (Albertini, Kohli and Vogel, 2007; Hogan, Eggebeen and Clogg, 1993; Grundy, 2005; Silverstein et al., 2002), both as a motivating factor for providing support and because of its association with other outcomes. For example, there is evidence that overbenefitting (receiving more than giving) has negative consequences for older parents' well-being (Davey and Eggebeen, 1998) while balanced exchanges are positively associated with parents' mental health (Litwin, 2004). The extent of reciprocity is likely to depend on individual characteristics. In a cross-national European study, Mudrazija (2016) finds that net transfers from parents to adult children follow a similar age pattern across the majority of countries, with declining positive transfers (parents giving more than they receive) for parents aged 50-79, becoming negative from around age 80. There is also evidence from Europe (Mudrazija, 2016) and the U.S. (Hogan, Eggebeen and Clogg, 1993) that reciprocity reflects the geographical proximity of parents and children and gender differences in family roles.

Another question of interest is whether practical and financial support serve as functional substitutes or complements of each other (e.g. Mudrazija, 2016), and how their interdependence depends on individual characteristics. Among the factors that may play a role are income and geographical distance where better-off adult children or children living at a greater distance from their parents may substitute money for time transfers to parents (e.g. Grundy, 2005). Alternatively time and money transfers may be positively associated, with a tendency to give or receive both or neither form of support.

Most previous substantive research has focused on the predictors of giving and receiving support rather than their associations. Many studies have considered only one direction of exchange, combined different types of exchange, or fitted separate models for different types or directions, all of which preclude the study of reciprocity or complementarity. The studies that have considered the associations have employed methods, such as modelling the difference between support given and support received, that are limited or inflexible in some way (we discuss these approaches further in Section 7.2).

A more flexible way to quantify these interrelationships is as residual correlations in a joint model, reciprocity between support given and received, and complementarity between different types of help given or received. Early examples of this are Attias-Donfut, Ogg and Wolff (2005) and Bonsang (2007) who analysed binary indicators of support using multivariate probit models. Later research has extended this joint modelling approach in different ways (Steele and Grundy, 2021; Kuha, Zhang and Steele, 2023; Steele et al., 2024), but no study has allowed the correlations among responses to depend on covariates. This is the development that we focus on. We present a general joint modelling framework that can be used to simultaneously investigate predictors of financial and practical support given and received, as well as predictors of the correlations among these different types of exchange.

We analyse cross-sectional data from the UK Household Longitudinal Study (UKHLS), which contains 16 questions ('items') about exchanges of help on dyads formed of a survey respondent and their non-coresident parent(s). Seven of the items relate to whether or not different kinds of practical help are given to parents (for example, assistance with shopping) and a further seven items indicate forms of practical help received from parents. The remaining two items indicate whether financial help is given and received. The practical help items are treated as multiple binary indicators of two continuous latent variables which are modelled jointly with latent variables taken to underlie the two indicators of financial exchanges. The data are thus 'doubly multivariate' in that we aim to model the joint distribution of four latent variables which are themselves measured by sets of multiple items. We also account for zero inflation, which arises from a high proportion of respondents who report giving or receiving none of these types of support, by including in the model two binary latent variables for the subpopulations with excess zeros.

Two previous papers have used different waves of the UKHLS data to examine questions on intergenerational exchanges of support. Kuha, Zhang and Steele (2023) carry out a cross-sectional analysis of tendencies to give and receive help, treating items on practical and financial help together. Steele et al. (2024) consider them separately, but collapse the seven items on practical help into one binary indicator; their focus is on longitudinal analysis, which requires the specification of appropriate random effects to incorporate the complex multilevel structure of these data. Both of these papers focus on models for mean levels of different types of help given and received, rather than for their correlations. Here we combine and then extend elements of these previous papers. Our analysis is cross-sectional. We start from the model of Kuha, Zhang and Steele (2023), but separating practical and financial help (as in Steele et al. 2024, but without collapsing the practical help items). The residual covariance matrix of the four latent helping tendencies (for giving and receiving practical and financial help) is decomposed into their standard deviations and correlation matrix. We then introduce a model for how the residual correlations depend on predictors (covariates), and develop methods for estimating this model. This is the main focus and contribution of this paper. It allows us to answer questions not only about the predictors of the levels of different forms of support (the mean structure) but also about the predictors of their correlation structure, i.e. the symmetry of exchanges (correlations between giving and receiving help) and complementarity of different forms of help (correlations between giving or receiving financial and practical help) for different population sub-groups.

Methodologically, this paper contributes to the literature on modelling correlation or covariance matrices given covariates. A key technical challenge here is that the estimated matrices should be positive definite. Broadly, two approaches may be taken to achieve this (Pinheiro and Bates, 1996). 'Unconstrained' methods specify a model for some transformation which ensures that the fitted matrix will be positive definite, while 'constrained' methods enforce it during estimation. A disadvantage of the unconstrained approach is that the parameters of the transformation are not easily interpretable. Constrained estimation, in contrast, can use interpretable models for the covariances or correlations themselves, but it faces the challenge of how to actually implement the constraint.

We employ a two-step approach of estimation where the parameters of the measurement model of the latent variables are estimated first, followed by the model for the means and correlations of the latent variables which is the focus of substantive interest. The second step is carried out in the Bayesian framework, using a tailored MCMC algorithm. This uses a constrained approach for estimating the correlation model where the parameters sampled at each MCMC step can only be retained if they imply a positive definite correlation matrix at all relevant values of the covariates. This builds on previously proposed methods (Barnard, McCulloch and Meng, 2000; Wong, Carter and Kohn, 2003), which we extend to models that include covariates for the correlations.

The UKHLS data are introduced in Section 2, and the specification of the joint model is described in Section 3. Section 4 reviews previous literature on modelling covariance and correlation matrices. Section 5 gives theoretical results that provide the basis of our estimation of the model for the correlations, and estimation of the joint model is then described in Section 6. Results of the analysis of intergenerational exchanges of family support are discussed in Section 7, and a concluding discussion is given in Section 8. Some additional results are given in supplementary materials, as explained in relevant places in the main text.

2. Data. We use data from the Understanding Society survey, also known as the UK Household Longitudinal Study (UKHLS; University of Essex, 2019). This is a long-standing household panel survey. We conduct a cross-sectional analysis of data from wave 9 of UKHLS, collected in 2017–19. This included the 'family network' module which collected information on exchanges of help with relatives living outside a respondent's household.

We consider exchanges from an adult child perspective. Respondents who had at least one non-coresident parent were asked whether they 'nowadays' 'regularly or frequently' gave each of eight types of help to their parent(s): lifts in a car; help with shopping; providing or cooking meals; help with basic personal needs; washing, ironing or cleaning; personal affairs such as paying bills or writing letters; decorating, gardening or house repairs; or financial help. These items are dichotomous, with the response options Yes and No. The same questions were asked about receipt of support from parents, but with personal needs replaced by help with childcare. We will distinguish between financial help (measured by a single item in each direction) and practical help (measured by the remaining seven items). Where a respondent had both biological and step/adoptive parents alive, the respondents were asked to report on the ones that they had most contact with. Although respondents were asked about giving parents a lift in their car 'if they have one', the recorded variable had no missing values for this item. We used other survey information to set this item to missing for respondents who did not have access to a car. Similarly, the childcare item was coded as missing for respondents who did not have coresident dependent children aged 16 or under. For the item on receiving lifts from parents, we do not have information on whether the parents have access to a car, so responses of 'No' to this item will include also cases where they do not.

A notable finding for these data is that less than half of the respondents report that they give (44.4% of our analysis sample) or receive (38.2%) even one of these types of support. This is a feature that we will want to allow for in the modelling of the data.

We consider as covariates a range of individual and household demographic and socioeconomic characteristics that aim to capture an adult child's and their parents' capacities to give support and their potential need for support. Most variables in the survey refer to the respondent (the child in our analysis), as less information was collected on non-coresident relatives, but we also include a small set of characteristics of the parents. The following respondent characteristics were included: age, gender, whether they have a coresident partner, indicators of the presence and age of their youngest biological or adopted coresident children, the number of siblings (as a measure of both alternative sources of support for parents and competition for the receipt of parental support), whether they have a long-term illness that limits their daily activities, employment status (classified as employed or non-employed [unemployed or economically inactive]), education (up to secondary school only, or post-secondary qualifications), household tenure (home-owner or social/private renter), and household income (equivalised, adjusted for inflation using the 2019 Consumer Price Index, and log transformed). The parental characteristics included were the age of the oldest living parent and whether either parent lives alone. We also include the travel time to the nearest parent, dichotomized as 1 hour or less vs. more than 1 hour.

An important limitation of the UKHLS data, shared by other large-scale national studies with information on intergenerational exchanges, is the reliance on reports from one member of each parent-child dyad. This is due to practical obstacles with collecting data from individuals living apart from the sample members. Studies that do collect data on parents and children from the same family include the German *pairfam* study (Huinink et al., 2011), the

Netherlands Kinship Panel Study (Mandermakers and Dykstra, 2008), and the Californian Longitudinal Study of Generations (Bengtson, 2001) and several other US studies (Suitor et al., 2017). Such multi-actor data are not available for the UK. It would be possible to use UKHLS to study exchanges from the perspective of parent respondents (who would be from different families than the child respondents in our study), in effect reversing the focus of our analysis. This would allow analysis of the effects of a richer set of parental characteristics on exchanges with children, but with a correspondingly smaller set of child characteristics (Steele et al., 2024). Apart from the limited information on parents when considering exchanges from a child perspective, previous research suggests that single informant dyadic data are subject to reporting biases, with a tendency to understate help received and overstate help given (Shapiro, 2004; Kim et al., 2011). Although in a traditional dyadic design each member of the pair would thus be interviewed, our data nevertheless have a dyadic structure and can be analysed using methods for dyadic data.

The analysis sample was first restricted to the 15,825 respondents aged 18 or over who had at least one non-coresident parent but no coresident parent. We excluded respondents whose nearest parent lived or worked abroad (1830 of them), because the nature of their exchanges is likely to differ from parents based in the UK, and also omitted 1792 respondents who had missing data on any covariate or on all the help items. The final sample size for analysis is $n = 12,203$. The UKHLS sample can include some respondents who are siblings to each other. However, preliminary analysis indicated that their number was very small for our analysis sample, so we ignore this feature and treat all the respondents as independent of each other. Summary statistics for the helping items and the covariates for this sample are shown in supplementary Appendix A.

3. Latent variable model for multivariate dyadic data. Here we define the joint model for the data. The specification builds on that of Kuha, Zhang and Steele (2023), but with two extensions. First, tendencies to (give as well as receive) financial and practical help are represented by separate latent variables, so that the model includes four rather than two such variables for each respondent. Second, the correlations between these variables are also modelled as functions of covariates.

Let $(\mathbf{X}_i, \mathbf{Y}_{Gi}, \mathbf{Y}_{Ri})$ be observed data for a sample of units $i = 1, \ldots, n$, where $\mathbf{X}_i$ is a $Q \times 1$ vector of covariates (including a constant 1), and $\mathbf{Y}_{Gi} = (\mathbf{Y}_{GPi}^{\mathsf{T}}, Y_{GFi})^{\mathsf{T}}$ and $\mathbf{Y}_{Ri} = (\mathbf{Y}_{RPi}^{\mathsf{T}}, Y_{RFi})^{\mathsf{T}}$ are $(J+1) \times 1$ vectors of binary indicator variables (*items*). In our application, a unit is the dyad of a survey respondent and their non-coresident parent(s), $\mathbf{Y}_{GPi} = (Y_{GPi1}, \ldots, Y_{GPiJ})^{\mathsf{T}}$ are the respondent's answers to $J = 7$ items on different types of *practical help given* to their parents, $\mathbf{Y}_{RPi} = (Y_{RPi1}, \ldots, Y_{RPiJ})^{\mathsf{T}}$ are the items on *practical help received* from the parents, and Y_{GFi} and Y_{RFi} are the single items on *financial help given* and *financial help received* respectively. Each item is coded 1 if that kind of help is given or received, and 0 if not. In other applications Y_{GFi} and Y_{RFi} could also be vectors of multiple indicators, with obvious modifications of the specifications below.

3.1. *Measurement model for the observed items.* The items in $\mathbf{Y}_{GPi}$, $\mathbf{Y}_{RPi}$, Y_{GFi} and Y_{RFi} are regarded as measures of continuous latent variables η_{GPi}, η_{RPi}, η_{GFi} and η_{RFi} respectively. We interpret η_{GPi} and η_{RPi} as an individual's underlying tendencies to give and to receive practical help respectively, and η_{GFi} and η_{RFi} as tendencies to give and receive financial help.

The data that we analyse have a large number of responses where all the items in $\mathbf{Y}_{Gi}$ or $\mathbf{Y}_{Ri}$ are zero (no help given or received). The proportions of these all-zero responses may be higher than can be well accounted for by standard latent variable models given the continuous latent variables alone. To allow for this multivariate zero inflation, the model also includes two binary latent class variables ξ_{Gi} and ξ_{Ri}, for each of which one class represents individuals who are certain not to give (for ξ_{Gi}) or receive (for ξ_{Ri}) any kind of help. For giving help, the *measurement model* for the observed responses $\mathbf{Y}_{Gi}$ given the latent variables $(\eta_{GPi}, \eta_{GFi}, \xi_{Gi})$ is then

$$p(\mathbf{Y}_{Gi} = \mathbf{0}|\xi_{Gi} = 0, \eta_{GPi}, \eta_{GFi}; \boldsymbol{\phi}_G) = p(\mathbf{Y}_{Gi} = \mathbf{0}|\xi_{Gi} = 0) = 1 \quad \text{and} \tag{1}$$

$$\begin{aligned} p(\mathbf{Y}_{Gi}|\xi_{Gi} = 1, \eta_{GPi}, \eta_{GFi}; \boldsymbol{\phi}_G) = \prod_{j=1}^{J} & p(Y_{GPij}|\xi_{Gi} = 1, \eta_{GPi}; \boldsymbol{\phi}_G) \\ & \times p(Y_{GFi}|\xi_{Gi} = 1, \eta_{GFi}) \end{aligned} \tag{2}$$

where $p(\cdot|\cdot)$ denotes a conditional distribution and $\boldsymbol{\phi}_G$ are measurement parameters. When $\xi_{Gi} = 0$, the respondent in dyad i is certain to answer 'No' to all items related to giving help. When $\xi_{Gi} = 1$, the probabilities of responses to Y_{GPij} are determined by the continuous latent variable η_{GPi} and the response to Y_{GFi} is determined by η_{GFi}. Items Y_{GPij} ($j = 1, \ldots, J$) are assumed to be conditionally independent of each other given η_{GPi}. If any items in $\mathbf{Y}_{Gi}$ are missing for respondent i, they are omitted from the product in (2). The measurement models for the individual items are specified as

$$p(Y_{GPij} = 1|\xi_{Gi} = 1, \eta_{GPi}; \boldsymbol{\phi}_G) = \Phi(\tau_{GPj} + \lambda_{GPj}\,\eta_{GPi}) \quad \text{for } j = 1, \ldots, J, \;\text{ and} \tag{3}$$

$$p(Y_{GFi} = 1|\xi_{Gi} = 1, \eta_{GFi}) = \mathbb{1}(\eta_{GFi} > 0), \tag{4}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, $\mathbb{1}(\cdot)$ is the indicator function, τ_{GPj} and λ_{GPj} are parameters, and we fix $\tau_{GP1} = 0$ and $\lambda_{GP1} = 1$ for identification of the scale of η_{GPi}. Here (3) is a standard latent-variable (item response theory) model for binary items, with probit measurement models, and (4), combined with the normal distribution of η_{GFi} defined below, is a latent-variable formulation of a probit model for the single item Y_{GFi}. Thus $\boldsymbol{\phi}_G = (\tau_{GP2}, \ldots, \tau_{GPJ}, \lambda_{GP2}, \ldots, \lambda_{GPJ})^\mathsf{T}$. The measurement model for receiving help $\mathbf{Y}_{Ri}$ given $(\eta_{RPi}, \eta_{RFi}, \xi_{Ri})$ is defined analogously to (3)–(4), with parameters $\boldsymbol{\phi}_R$, and $\mathbf{Y}_{Gi}$ and $\mathbf{Y}_{Ri}$ are assumed to be conditionally independent of each other given the latent variables. Let $\boldsymbol{\phi} = (\boldsymbol{\phi}_G^\mathsf{T}, \boldsymbol{\phi}_R^\mathsf{T})^\mathsf{T}$.

3.2. *Structural model for the latent variables given covariates.* Let $\boldsymbol{\eta}_i = (\eta_{GPi}, \eta_{RPi}, \eta_{GFi}, \eta_{RFi})^\mathsf{T}$ and $\boldsymbol{\xi}_i = (\xi_{Gi}, \xi_{Ri})^\mathsf{T}$. Their conditional distribution $p(\boldsymbol{\eta}_i, \boldsymbol{\xi}_i|\mathbf{X}_i; \boldsymbol{\psi}) = p(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\psi}_\eta) \times p(\boldsymbol{\xi}_i|\mathbf{X}_i; \boldsymbol{\psi}_\xi)$ is the *structural model* for the latent variables given the covariates. Here $\boldsymbol{\eta}_i$ and $\boldsymbol{\xi}_i$ are taken to be conditionally independent given $\mathbf{X}_i$, and $\boldsymbol{\psi} = (\boldsymbol{\psi}_\eta^\mathsf{T}, \boldsymbol{\psi}_\xi^\mathsf{T})^\mathsf{T}$ are parameters. The distribution of the latent class variables $\boldsymbol{\xi}_i$ is specified as multinomial, with probabilities

$$\log\left[\frac{\pi_{k_1k_2}(\mathbf{X}_i)}{\pi_{00}(\mathbf{X}_i)}\right] = \boldsymbol{\gamma}_{k_1k_2}^\mathsf{T}\mathbf{X}_i, \tag{5}$$

where $\pi_{k_1k_2}(\mathbf{X}_i) = p(\xi_{Gi} = k_1, \xi_{Ri} = k_2|\mathbf{X}_i; \boldsymbol{\psi}_\xi)$ for $k_1, k_2 = 0, 1$ and $\boldsymbol{\gamma}_{00} = \mathbf{0}$, so that $\boldsymbol{\psi}_\xi = (\boldsymbol{\gamma}_{01}^\mathsf{T}, \boldsymbol{\gamma}_{10}^\mathsf{T}, \boldsymbol{\gamma}_{11}^\mathsf{T})^\mathsf{T}$. In our application, the coefficients $\boldsymbol{\psi}_\xi$ describe how different covariates are associated with the sizes of latent sub-populations of those are certain not to give and/or receive any kind of help. This could also be interpreted in substantive terms, but in our analysis we use it primarily to allow for the multivariate zero inflation in the observed data.

The main focus of substantive interest is on the structural model for the continuous helping tendencies $\boldsymbol{\eta}_i$ given the covariates $\mathbf{X}_i$. Here $\boldsymbol{\eta}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ is taken to follow a four-variate normal distribution with covariance matrix $\boldsymbol{\Sigma}_i$ and mean vector

$$\boldsymbol{\mu}_i = \mathrm{E}(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\beta}) = \boldsymbol{\beta}^\mathsf{T}\mathbf{X}_i \tag{6}$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}_{GP}, \boldsymbol{\beta}_{RP}, \boldsymbol{\beta}_{GF}, \boldsymbol{\beta}_{RF}]$ is a $Q \times 4$ matrix of coefficients, specifying a separate linear model for each element of $\boldsymbol{\mu}_i$. For the covariance matrix, we first decompose it as

$$\boldsymbol{\Sigma}_i = \text{cov}(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\alpha}, \boldsymbol{\sigma}) = \mathbf{S}_i\, \mathbf{R}_i\, \mathbf{S}_i, \tag{7}$$

where $\boldsymbol{\alpha}$ are parameters of the correlation matrix $\mathbf{R}_i$ and $\boldsymbol{\sigma} = (\sigma_{GP}, \sigma_{RP})^{\mathsf{T}}$ are parameters of $\mathbf{S}_i = \text{diag}(\sigma_{GP}, \sigma_{RP}, 1, 1)$, a diagonal matrix of standard deviations where those of η_{GFi} and η_{RFi} are fixed at 1 to identify the measurement model (4) for η_{GFi} and the corresponding model for η_{RFi}. Here $\boldsymbol{\sigma}$ do not depend on covariates, but they could also be included.

For the correlation matrix, we consider the specification

$$\mathbf{R}_i = \mathbf{R}(\mathbf{X}_i; \boldsymbol{\alpha}) = \begin{bmatrix} 1 & & & \\ \rho_{1i} & 1 & & \\ \rho_{2i} & \rho_{4i} & 1 & \\ \rho_{3i} & \rho_{5i} & \rho_{6i} & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \rho(\mathbf{X}_i; \boldsymbol{\alpha}_1) & 1 & & \\ \rho(\mathbf{X}_i; \boldsymbol{\alpha}_2) & \rho(\mathbf{X}_i; \boldsymbol{\alpha}_4) & 1 & \\ \rho(\mathbf{X}_i; \boldsymbol{\alpha}_3) & \rho(\mathbf{X}_i; \boldsymbol{\alpha}_5) & \rho(\mathbf{X}_i; \boldsymbol{\alpha}_6) & 1 \end{bmatrix}, \tag{8}$$

where only the $L = 6$ distinct correlations in the lower triangular part are shown. We specify separate linear models $\rho_{li} = \rho(\mathbf{X}_i; \boldsymbol{\alpha}_l) = \boldsymbol{\alpha}_l^{\mathsf{T}} \mathbf{X}_i$ for each $l = 1, \ldots, L$, i.e.

$$\boldsymbol{\rho}_i = \boldsymbol{\alpha}^{\mathsf{T}} \mathbf{X}_i \tag{9}$$

where $\boldsymbol{\rho}_i = (\rho_{1i}, \ldots, \rho_{Li})^{\mathsf{T}}$, and $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_L]$ are coefficients. Some variables in $\mathbf{X}_i$ may be included in only one of the models (6) and (9), in which case the corresponding elements of $\boldsymbol{\beta}$ or $\boldsymbol{\alpha}$ are zero. The parameters of the structural model for $\boldsymbol{\eta}_i$ are thus $\boldsymbol{\psi}_\eta = (\text{vec}(\boldsymbol{\beta})^{\mathsf{T}}, \boldsymbol{\sigma}^{\mathsf{T}}, \text{vec}(\boldsymbol{\alpha})^{\mathsf{T}})^{\mathsf{T}}$, where $\text{vec}(\cdot)$ denotes the vectorization of a matrix. We note also that it will be necessary to further constrain the space of $\boldsymbol{\alpha}$ if we want to ensure that correlation matrices defined by (8)–(9) will be positive definite. Our specifications to achieve this are described in the sections below.

Let $\mathbf{Y} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_n]^{\mathsf{T}}$ denote all the observed data on the items, where $\mathbf{Y}_i = (\mathbf{Y}_{Gi}^{\mathsf{T}}, \mathbf{Y}_{Ri}^{\mathsf{T}})^{\mathsf{T}}$, and $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_n]^{\mathsf{T}}$ the data on the covariates. Define $G_i = \mathbb{1}(\mathbf{Y}_{Gi} \neq \mathbf{0})$ and $R_i = \mathbb{1}(\mathbf{Y}_{Ri} \neq \mathbf{0})$, the indicators for whether responses on giving and on receiving help are not all zero for respondent i. Assuming the observations for different respondents to be independent, the log likelihood function of the model is

$$\begin{aligned}
&\log p(\mathbf{Y}|\mathbf{X}; \boldsymbol{\phi}, \boldsymbol{\psi}) \\
&= \sum_{i=1}^{N} \log \Big\{ \pi_{11}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \Big[\int p(\mathbf{Y}_{Gi}|\xi_{Gi} = 1, \eta_{GPi}, \eta_{GFi}; \boldsymbol{\phi}_G) p(\mathbf{Y}_{Ri}|\xi_{Ri} = 1, \eta_{RPi}, \eta_{RFi}; \boldsymbol{\phi}_R) \\
&\qquad\qquad \times p(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{GPi}\, d\eta_{RPi}\, d\eta_{GFi}\, d\eta_{RFi} \Big] \\
&\quad + (1 - R_i)\, \pi_{10}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \Big[\int p(\mathbf{Y}_{Gi}|\xi_{Gi} = 1, \eta_{GPi}, \eta_{GFi}; \boldsymbol{\phi}_G)\, p(\eta_{GPi}, \eta_{GFi}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{GPi}\, d\eta_{GFi} \Big] \\
&\quad + (1 - G_i)\, \pi_{01}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \Big[\int p(\mathbf{Y}_{Ri}|\xi_{Ri} = 1, \eta_{RPi}, \eta_{RFi}; \boldsymbol{\phi}_R)\, p(\eta_{RPi}, \eta_{RFi}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{RPi}\, d\eta_{RFi} \Big] \\
&\quad + (1 - G_i)(1 - R_i)\, \pi_{00}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \Big\}.
\end{aligned}$$

Estimation of this model is described in Section 6, after some further discussion of questions related to the model for the correlations.

4. Models for correlation and covariance matrices given covariates: Existing approaches. There is a substantial literature on modelling association structures of multivariate distributions. We review here those parts of it that are most relevant to our work, focusing on different ways of specifying models for correlation or covariance matrices given covariates. These approaches can be combined with different specifications for the joint distribution as a whole, and with different methods of estimating its parameters. Our modelling uses a parametric specification of a multivariate normal distribution and Bayesian estimation of the parameters, but the review here is not limited to that case.

We focus on approaches which consider associations directly in terms of pairwise covariances or correlations. This excludes models for conditional associations of some of the variables given the others, such as log-linear models for categorical data or covariance selection models for the inverse covariance matrix of a multivariate normal distribution. We also exclude specifications where the associations are determined indirectly via further latent variables, such as random effects models and common factor models. The multivariate response variable whose covariance or correlation matrix is modelled may itself be a latent variable, as it is in our analysis where we model the correlations of the latent $\boldsymbol{\eta}_i$, but we still focus on models that are specified directly for their associations or transformations of the associations.

Models for associations may have two broad goals. One of them is to define a patterned structure on the associations which is more parsimonious than an unstructured matrix that has separate parameters for each pair of variables. This is the case, for example, when an autocorrelation model is specified for responses that are ordered in time. An extreme version of this occurs in very high-dimensional problems where parsimonious specification is essential for consistent estimation of covariance matrices. We do not consider such regularisation methods here (see Pourahmadi 2011 and Fan, Liao and Liu 2016 for reviews). The second broad type of model specification considers instead an unstructured matrix of associations but allows the correlations or covariances in it to depend on covariates that describe the units of analysis, such as the survey respondents in our application. This is the goal of our modelling.

When the goal is to model covariances or correlations in this way, a key question is how to ensure that the estimated matrices of them will be positive definite. Pinheiro and Bates (1996) pointed out a key distinction between two approaches: *unconstrained* ones where the models are specified for parametrizations (transformations) of the association matrix which are guaranteed to imply a positive definite matrix, and *constrained* ones where positive definiteness is imposed in the estimation process. Our approach is an instance of constrained estimation, but we list first the most important unconstrained methods (see Pourahmadi 2011 and Pan and Pan 2017 for more detailed reviews). They differ in what transformation they use. The most common is the modified Cholesky decomposition of the covariance matrix. It was introduced by Pourahmadi (1999), and general models for it were proposed by Pan and MacKenzie (2006). Other possible transformations include the matrix logarithm (Chiu, Leonard and Tsui, 1996) and the 'alternative Cholesky decomposition' of the covariance matrix (Chen and Dunson, 2003), a variant of the modified Cholesky decomposition proposed by Zhang and Leng (2012), parametrizations of the correlation matrix in terms of partial autocorrelations (Wang and Daniels, 2013) or hyperspherical co-ordinates of its standard Cholesky decomposition (Zhang, Leng and Tang, 2015), and the matrix logarithm of the correlation matrix (Archakov and Hansen, 2021; Hu et al., 2021).

The natural advantage of the unconstrained methods is that they ensure positive definiteness at any values of the covariates. The corresponding disadvantage is that because the models are not specified for the individual association parameters, the model parameters are not easily interpretable. All of the interpretations that are available apply only when the response variables have a natural ordering, most obviously in longitudinal data where they are ordered in time. Then the parameters of the modified Cholesky decomposition can be interpreted in terms of an autoregressive model for each variable given its predecessors, those of the alternative Cholesky decomposition and of Zhang and Leng (2012) in terms of a moving average representation of each variable given residuals of the previous ones (Pourahmadi, 2007; Pan and Pan, 2017), those of Wang and Daniels (2013) as partial autocorrelations of two variables given all the intervening ones, and the hyperspherical co-ordinate parametrization in terms of semi-partial correlations (Ghosh, Mallick and Pourahmadi, 2021).

Turning now to approaches that model individual pairwise association parameters directly, for correlations we could use transformations of them (e.g. Fisher's z) to ensure that the fitted correlations are constrained to $(-1, 1)$. This, however, is not sufficient to ensure that the correlation matrix as a whole is positive definite, except for a bivariate response (for this case, see e.g. Wilding et al. 2011 and references therein). One possible pragmatic approach would be to simply employ such models anyway, ignoring the possibility of some non-positive definite matrices (see e.g. Yan and Fine 2007). This could work well in some applications, in the best case that the fitted correlation matrices end up being positive definite at all relevant values of the covariates. However, it is not a satisfactory general approach. Luo and Pan (2022) suggest post-hoc adjustments to fitted correlation models to make them positive definite; this, however, is unhelpful when the focus is on interpreting coefficients of the model. A different solution is provided by Hoff and Niu (2012) who propose a model where covariances depend on quadratic functions of covariates and the matrix is automatically positive definite.

Most of the literature on constrained estimation considers linear models for covariances or correlations. This is not a limitation even for correlations, because positive definiteness of the matrix also implies that all the correlations in it will be in $(-1, 1)$. The most developed results here are for the linear covariance model for multivariate normal distribution (Anderson, 1973), in which the covariance matrix takes the form $\mathbf{\Sigma} = \sum_k \nu_k \mathbf{G}_k$ where ν_k are parameters and $\mathbf{G}_k$ are known, linearly independent symmetric matrices. Zwiernik, Uhler and Richards (2017) show that although the log-likelihood for this model typically has multiple local maxima, any hill climbing method initiated at the least squares estimator will converge to its global maximum with high probability. Zou et al. (2017) consider the case where the $\mathbf{G}_k$ are similarity matrices between the response variables, and propose maximum likelihood estimates and constrained least squares estimators for this model. In these formulations, $\mathbf{\Sigma}$ is the same for all units i. This is relaxed by Zou et al. (2022), who allow the values of the similarity matrices in Zou et al. (2017) to depend on unit-specific covariates.

We will also consider a linear model, as shown in (9), but for the correlations and given general unit-specific covariates. We estimate it in the Bayesian framework and using Markov chain Monte Carlo (MCMC) methods. Here it is convenient to employ the decomposition (7) and model the standard deviations and correlations separately, as has also been done in most previous literature that has used a Bayesian approach. MCMC also provides an obvious way to implement constrained estimation, at least in principle. This can be done at each sampling step of the estimation, by constraining the prior distribution, the proposal distribution from which the parameters are drawn, or the acceptance probabilities of the sampled values in a way which rules out parameter values that imply non-positive definite matrices. But although the principle is obvious, implementing it is not necessarily easy. Two instances of this approach that we will draw on in particular are those of Barnard, McCulloch and Meng (2000) and Wong, Carter and Kohn (2003), as discussed further below. Other methods of this kind have been proposed by Chib and Greenberg (1998) and Liechty, Liechty and Müller (2004).

What is missing from existing Bayesian implementations is the inclusion of unit-specific covariates in the models for the correlations, which is our focus. In Section 6 we propose an MCMC estimation procedure which accommodates them. This in turn requires some further consideration of the constraint of positive definiteness of the correlation matrix, because it now has to hold at different values of the covariates. This is discussed in the next section.

5. Ensuring a positive definite correlation matrix. The key challenge in our constrained estimation is to ensure that the estimated correlation matrices remain positive definite at all relevant values of the covariates $\mathbf{X}$. To achieve this, we specify the set of values for $\mathbf{X}$ for which this should hold, and define the parameter space for $\boldsymbol{\alpha}$ as the set of values for $\boldsymbol{\alpha}$ for which the correlation matrix will be positive definite for all such $\mathbf{X}$. The prior distribution of $\boldsymbol{\alpha}$ will be zero outside this set, and the Markov chain defined by the MCMC estimation procedure should never transition outside it. In this section we describe some further theoretical results which establish when and how this can be achieved.

Let $\mathbf{R} = \mathbf{R}(\boldsymbol{\rho})$ denote a symmetric matrix where all the diagonal elements equal 1 and the distinct off-diagonal elements $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_L)^\mathsf{T}$ are all in $(-1, 1)$. Let C_ρ denote the set of $\boldsymbol{\rho}$ such that $\mathbf{R}(\boldsymbol{\rho})$ is positive definite, and thus a correlation matrix, for all $\boldsymbol{\rho} \in C_\rho$. It is a convex subset of the hypercube $[-1, 1]^L$ (for the shape of C_ρ in the cases $L = 3$ and $L = 6$, i.e. for 3×3 and 4×4 correlation matrices, see Rousseeuw and Molenberghs, 1994).

We consider model (9) where $\boldsymbol{\rho} = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X}$ (in this section we omit the unit subscript i, and take $\mathbf{X}$ to include only those covariates that are included in the model for $\boldsymbol{\rho}$ rather than only for $\boldsymbol{\mu}$). It is clear that this $\boldsymbol{\rho}$ cannot be in C_ρ for all values of the parameters $\boldsymbol{\alpha}$ and covariates $\mathbf{X}$. What we need to do is to decide first what values of $\mathbf{X}$ are substantively relevant, and then ensure that the estimated models do imply $\boldsymbol{\rho} \in C_\rho$ for all these $\mathbf{X}$. It is useful to introduce here some additional notation. Let $\mathbf{Z}$ be the smallest vector of distinct variables, including a constant term 1, which determines $\mathbf{X} = \mathbf{X}(\mathbf{Z})$. Here $\mathbf{Z}$ may be shorter than $\mathbf{X}$ if some variables in $\mathbf{X}$ are functions of $\mathbf{Z}$, e.g. polynomials or product terms (interactions). Suppose that $\mathbf{Z}$ is a $p \times 1$ vector and $\mathbf{X}$ a $q \times 1$ vector. Below we denote sets $S_Z \subset \mathbb{R}^p$ and $S_X \subset \mathbb{R}^q$ of $\mathbf{Z}$ and $\mathbf{X}$ respectively with appropriate subscripts.

A combination of values $(\mathbf{X}, \boldsymbol{\alpha})$ is said to be feasible if $\boldsymbol{\rho} = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X} \in C_\rho$, and $(\mathbf{Z}, \boldsymbol{\alpha})$ to be feasible if $\boldsymbol{\rho} = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X}(\mathbf{Z}) \in C_\rho$. We aim to identify known sets of $\mathbf{Z}$ and $\boldsymbol{\alpha}$ such that all combinations of values from them are feasible. This will involve the following steps:

(1) Choose a set S_Z for $\mathbf{Z}$ for which the correlation matrices should be positive definite.
(2) Specify a test set $S_{XT} = \{\mathbf{X}_1, \ldots, \mathbf{X}_T\}$ of values for $\mathbf{X}$, which will be used for checking positive definiteness during MCMC estimation. The choice of S_{XT} depends on S_Z.
(3) Carry out MCMC estimation which samples values of $\boldsymbol{\alpha}$ (and the other model parameters). At each iteration, carry out checks to ensure that the value of $\boldsymbol{\alpha}$ that is retained is feasible with all $\mathbf{X} \in S_{XT}$. In the end, this produces an MCMC sample $S_\alpha = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M\}$.
(4) Conclude that $(\mathbf{Z}, \boldsymbol{\alpha})$ is feasible for all combinations of any $\mathbf{Z} \in S_Z$ and any $\boldsymbol{\alpha}$ in the convex hull of S_α.

Step (3) is the computational one where the constraint is enforced. It alone is not enough, however: We can only check feasibility directly for a finite number of values of $\mathbf{X}$ and $\boldsymbol{\alpha}$, but the sets that we want to draw conclusions on are infinite for at least $\boldsymbol{\alpha}$. So some additional results are needed to motivate steps (1) and (2) and to justify the conclusion in step (4).

In step (1), S_Z should include the substantively relevant and interesting values of the covariates for which we want our estimated model to imply valid correlation matrices. For example, this could be a finite set $S_{ZN} = \{\mathbf{Z}_1, \ldots, \mathbf{Z}_N\}$, normally including at least all the distinct values among the $\mathbf{Z}_i$, $i = 1, \ldots, n$, in the observed data. S_Z is always of this form when all the variables in $\mathbf{Z}$ are categorical. If $\mathbf{Z}$ includes continuous variables, we may also expand S_{ZN} to an infinite set, such as its convex hull $S_{Zh} = \{\sum_{j=1}^N \lambda_j \mathbf{Z}_j \mid \sum_{j=1}^N \lambda_j = 1; \lambda_j \geq 0 \text{ for all } j = 1, \ldots, N\}$ or the hyperrectangle $S_{Zr} = \{(Z_1, \ldots, Z_p) \mid Z_s \in [l_s, u_s] \text{ for all } s = 1, \ldots, p\}$, for specified $l_s \leq \min\{Z_{js} \mid j = 1, \ldots, N\}$ and $u_s \geq \max\{Z_{js} \mid j = 1, \ldots, N\}$ for each $s = 1, \ldots, p$. Here $S_{ZN} \subset S_{Zh} \subseteq S_{Zr}$.

For step (2), consider the set $C_{\alpha, S_X} = \{\boldsymbol{\alpha} \in \mathbb{R}^{L \times q} \mid \boldsymbol{\rho} = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X} \in C_\rho \text{ for all } \mathbf{X} \in S_X\}$ of values of $\boldsymbol{\alpha}$ which are feasible when combined with any $\mathbf{X}$ in a set S_X. Basic properties of

C_{α,S_X} are given by Proposition 1, which is included in supplementary Appendix B. Parts (i) and (ii) of it explain how C_{α,S_X} depends on S_X. In particular, they show that if a parameter value $\boldsymbol{\alpha}$ is feasible when combined with any $\mathbf{X}$ in S_X, it is also feasible with any value of $\mathbf{X}$ in the convex hull of S_X. This in turn allows us to choose the test set S_{XT} so that feasibility for it also implies feasibility for all $\mathbf{Z}$ in the set of interest S_Z. For a finite $S_Z = S_{ZN}$ we can choose S_{XT} as all the distinct values of $\mathbf{X}$ implied by $\mathbf{Z} \in S_{ZN}$. If S_Z is infinite, S_{XT} can be chosen so that its convex hull implies coverage of all of S_Z. These choices are discussed in more detail in supplementary Appendix B.

We can thus define the support of $\boldsymbol{\alpha}$ as $C_{\alpha,S_{XT}}$. Step (3) above refers to the implementation of the MCMC algorithm, where we need to ensure that all accepted values of $\boldsymbol{\alpha}$ are in $C_{\alpha,S_{XT}}$. Here the brute-force approach would be to check the whole proposed vector $\boldsymbol{\alpha}$ — i.e. the correlation matrices implied by it — at each MCMC iteration. This, however, would be computationally demanding and likely to lead to high rates of rejection. What we do instead is sample and check one scalar element of $\boldsymbol{\alpha}$ at a time. This is justified by the following proposition, the proof of which is given in supplementary Appendix C:

PROPOSITION 2. *Consider a finite set* $S_{XT} = \{\mathbf{X}_j = (X_{j1},\dots,X_{jq})^\mathsf{T} \mid j = 1,\dots,T\}$ *and any fixed value* $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_L]^\mathsf{T} \in C_{\alpha,S_{XT}}$. *Denote here* $\boldsymbol{\alpha} = (\alpha_{lm}, \boldsymbol{\alpha}_{-lm}^\mathsf{T})^\mathsf{T}$ *where* α_{lm} *is the coefficient of* X_{jm} *in the model for correlation* ρ_l, *for any* $m = 1,\dots,q$ *and* $l = 1,\dots,L$, *and* $\boldsymbol{\alpha}_{-lm}$ *denotes all other elements of* $\boldsymbol{\alpha}$, $\boldsymbol{\rho} = (\rho_l, \boldsymbol{\rho}_{-l}^\mathsf{T})^\mathsf{T}$ *where* $\boldsymbol{\rho}_{-l}$ *denotes all other elements of the distinct correlations* $\boldsymbol{\rho}$ *except* ρ_l, *and* $\mathbf{R}(\rho_l, \boldsymbol{\rho}_{-l})$ *the correlation matrix implied by* $\boldsymbol{\rho}$. *Let* $\boldsymbol{\rho}_{-l}^{(j)}$ *denote the value of* $\boldsymbol{\rho}_{-l}$ *for* $\boldsymbol{\rho}_j = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X}_j$, *for* $j = 1,\dots,T$.

Let $f_{jl}(\rho_l') = |\mathbf{R}(\rho_l', \boldsymbol{\rho}_{-l}^{(j)})|$, *treated as a function of* ρ_l', *where* $|\cdot|$ *denotes the determinant of a matrix. If* $X_{jm} \neq 0$, *let*

$$
\begin{aligned}
a_{lm}^{(j)} &= \frac{g_{jl} - \sum_{k\neq m} \alpha_{lk} X_{jk} - sgn(X_{jm})\, h_{jl}}{X_{jm}}, \\
b_{lm}^{(j)} &= \frac{g_{jl} - \sum_{k\neq m} \alpha_{lk} X_{jk} + sgn(X_{jm})\, h_{jl}}{X_{jm}}
\end{aligned}
\tag{10}
$$

for each $j = 1,\dots,T$, *where* $g_{jl} = -d_{jl}/(2c_{jl})$ *and* $h_{jl} = [(d_{jl}^2 - 4c_{jl}e_{jl})/(4c_{jl}^2)]^{1/2}$ *for* $c_{jl} = [f_{jl}(1) + f_{jl}(-1) - 2f_{jl}(0)]/2$, $d_{jl} = [f_{jl}(1) - f_{jl}(-1)]/2$ *and* $e_{jl} = f_{jl}(0)$. *If* $X_{jm} = 0$, *set* $a_{lm}^{(j)} = -\infty$ *and* $b_{lm}^{(j)} = +\infty$.

Define the interval $(a_{lm}, b_{lm}) = \cap_{j=1}^{T}(a_{lm}^{(j)}, b_{lm}^{(j)})$. *Then* $(a_{lm}, b_{lm}) = \{\alpha_{lm}' \mid (\alpha_{lm}', \boldsymbol{\alpha}_{-lm}^\mathsf{T})^\mathsf{T} \in C_{\alpha,S_{XT}}\}$. *This interval is non-empty because it contains at least the current value* α_{lm}.

In other words, the values of α_{lm} that imply a positive definite correlation matrix, holding other elements of $\boldsymbol{\alpha}$ at their previous values, are a continuous non-empty interval with known end points, which are calculated using all the values of $\mathbf{X}$ in the test set S_{XT}. We thus need to check only that a proposed value of an α_{lm} is in this interval. This builds on results by Barnard, McCulloch and Meng (2000) and Wong, Carter and Kohn (2003) for models without covariates, which we extend here to apply to the individual coefficients in $\boldsymbol{\alpha}$ and multiple values of the covariates $\mathbf{X}$. Even for one scalar α_{lm} at a time this procedure is still computationally non-trivial, and it needs to be carried out efficiently. Our implementation of it is described in Section 6 below and in supplementary Appendix E.

When step (3) is completed, we thus know that all the values of $\boldsymbol{\alpha}$ in the MCMC sample $S_\alpha = \{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_M\}$ are feasible when combined with any value of $\mathbf{Z}$ in the target set S_Z. Finally, we can extend this conclusion to other values of $\boldsymbol{\alpha}$ that were not sampled, specifically to the convex hull of S_α. This is justified by parts (iii)–(v) of Proposition 1 in supplementary

Appendix B, thus completing step (4) above. In particular, the convex hull of the MCMC sample includes the summary statistics that we would use for estimation of the parameters in $\boldsymbol{\alpha}$, such as their (posterior) means and quantile-based interval estimates. These estimates are thus also guaranteed to imply positive definite correlation matrices given any values of the covariates in the pre-specified set of interest S_Z.

6. Estimation of the model. We use a two-step procedure to estimate the latent variable model defined in Section 3. In it, the parameters of the measurement model are estimated first, and they are then fixed at their estimated values in the second step where the structural model is estimated. The same approach was used by Kuha, Zhang and Steele (2023) in their analysis which focused on models for the mean levels of intergenerational support.

The general idea of two-step estimation of latent variable models has been described for different types of models by Bakk and Kuha (2018), Rosseel and Loh (2022), and Kuha and Bakk (2023). As discussed there, the motivation of the approach is twofold. Practically, it can substantially simplify the estimation of complex models. Conceptually, it clarifies the meaning of such models by separating the definition of the latent variables from their use. In our analysis the practical simplification means that sampling of the measurement parameters is not included in the MCMC procedure that is used for the second step. This is convenient, but for us the more important motivation is the conceptual one. In essence, the estimates from the first step provide the operational definition of the latent variables, which then remains fixed even if several different structural models are subsequently considered — for example if we compare models with different predictors for the correlations.

In our analysis the first step was done using maximum likelihood estimation, as explained in Section 6.1, while the second step was carried out in the Bayesian framework, using MCMC estimation as described in Section 6.2. The choice for the first step was determined by the convenient availability of standard software for estimating latent variable measurement models. This hybrid approach does not affect the interpretation of the results of the second step, because it treats the measurement parameters as fixed parameters. In other words, what we obtain from the second step is a sample from the posterior distribution of the parameters of the structural model for the specific latent variables whose scales and meaning are defined and fixed by the values of the measurement parameters estimated from the first step.

6.1. *Estimation of the measurement model.* In the first step, the measurement parameters $\boldsymbol{\phi}_G$ and $\boldsymbol{\phi}_R$ are estimated separately. For $\boldsymbol{\phi}_G$, the data are $\mathbf{Y}_{Gi}$, the measurement model is specified by (1)–(4), and the structural model for ξ_{Gi} and $\boldsymbol{\eta}_{Gi} = (\eta_{GPi}, \eta_{GFi})^{\mathsf{T}}$ is obtained from (5)–(8) by omitting $\boldsymbol{\eta}_{Ri}$ and $\mathbf{X}_i$. The log likelihood for $\boldsymbol{\phi}_G$ is then

$$\begin{aligned}
&\log p(\mathbf{Y}_G|\boldsymbol{\phi}_G, \pi_G, \mu_{\eta_{GP}}, \mu_{\eta_{GF}}, \sigma^2_{\eta_{GP}}, \rho_{\eta_G}) \\
&= \sum_{i=1}^{n} \log \Big[\pi_G \int p(\mathbf{Y}_{Gi}|\xi_{Gi}=1, \eta_{GPi}, \eta_{GFi}; \boldsymbol{\phi}_G)\, p(\boldsymbol{\eta}_{Gi}; \mu_{\eta_{GP}}, \mu_{\eta_{GF}}, \sigma^2_{\eta_{GP}}, \rho_{\eta_G})\, d\eta_{GPi}\, d\eta_{GFi} \\
&\qquad + (1-G_i)(1-\pi_G) \Big]
\end{aligned} \tag{11}$$

where $\pi_G = p(\xi_{Gi} = 1)$ and $p(\boldsymbol{\eta}_{Gi}; \mu_{\eta_{GP}}, \mu_{\eta_{GF}}, \sigma^2_{\eta_{GP}}, \rho_{\eta_G})$ is a bivariate normal density with $\mathrm{E}(\eta_{GPi}) = \mu_{\eta_{GP}}$, $\mathrm{Var}(\eta_{GPi}) = \sigma^2_{\eta_{GP}}$, $\mathrm{E}(\eta_{GFi}) = \mu_{\eta_{GF}}$, $\mathrm{Var}(\eta_{GFi}) = 1$ and $\mathrm{Corr}(\eta_{GPi}, \eta_{GFi}) = \rho_{\eta_G}$. The estimates $\tilde{\boldsymbol{\phi}}_G$ of $\boldsymbol{\phi}_G$ are obtained by maximizing (11), while the estimates of $\mu_{\eta_{GP}}$, $\mu_{\eta_{GF}}$, $\sigma^2_{\eta_{GP}}$ and ρ_{η_G} from this step are discarded. We have used Mplus 6.12 software (Muthén and Muthén, 2010) to carry out this step. The estimates $\tilde{\boldsymbol{\phi}}_R$ of $\boldsymbol{\phi}_R$ are obtained analogously, using the data on $\mathbf{Y}_{Ri}$.

6.2. *Estimation of the structural model.* In the second step, the structural parameters $\boldsymbol{\psi}$ are estimated. Let $\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\eta})$, where $\boldsymbol{\xi}$ denotes all the values of the latent $\boldsymbol{\xi}_i$ for the units i in the sample, and $\boldsymbol{\eta}$ all the values of $\boldsymbol{\eta}_i$. We use a Bayesian approach to estimation, using MCMC methods to draw a sample of $\boldsymbol{\psi}$ and $\boldsymbol{\zeta}$ from their posterior distribution

$$p(\boldsymbol{\psi}, \boldsymbol{\zeta} \mid \mathbf{Y}, \mathbf{X}; \tilde{\boldsymbol{\phi}}) \propto p(\mathbf{Y} \mid \mathbf{X}, \boldsymbol{\zeta}; \tilde{\boldsymbol{\phi}})\, p(\boldsymbol{\zeta} \mid \mathbf{X}; \boldsymbol{\psi})\, p(\boldsymbol{\psi}) \tag{12}$$

given the observed data $\mathbf{Y}$ and $\mathbf{X}$. What we use from this are the values of $\boldsymbol{\psi}$, which are a sample from the posterior $p(\boldsymbol{\psi} \mid \mathbf{Y}, \mathbf{X}; \tilde{\boldsymbol{\phi}})$. This is conditional on the estimated values $\tilde{\boldsymbol{\phi}} = (\tilde{\boldsymbol{\phi}}_G^{\mathsf{T}}, \tilde{\boldsymbol{\phi}}_R^{\mathsf{T}})^{\mathsf{T}}$ of the measurement parameters from the first step. They are treated as known and fixed numbers, as discussed above, and for simplicity we omit $\tilde{\boldsymbol{\phi}}$ from the notation below.

The $p(\boldsymbol{\psi})$ in (12) denotes the prior distribution of the structural parameters. We take it to be of the form $p(\boldsymbol{\psi}) = p(\boldsymbol{\psi}_\eta)\, p(\boldsymbol{\psi}_\xi) = p(\boldsymbol{\beta})\, p(\boldsymbol{\sigma})\, p(\boldsymbol{\alpha})\, p(\boldsymbol{\psi}_\xi)$ where the different blocks of parameters are a priori independent of each other (and individual parameters within the bloks are also independent, as explained in the further details below and in supplementary Appendix D). The prior distribution of the parameters $\boldsymbol{\alpha}$ of the correlation model is a joint uniform distribution $p(\boldsymbol{\alpha}) \propto \mathbb{1}(\boldsymbol{\alpha} \in C_{\alpha, S_{XT}})$ over the set $C_{\alpha, S_{XT}}$ which defines the support of $\boldsymbol{\alpha}$ as explained in Section 5.

The estimation algorithm has a data augmentation structure which alternates between imputing the latent variables given the observed variables and the parameters, and sampling the parameters from their posterior distributions given the observed and latent variables:

- **Sampling of the latent variables**: At MCMC iteration t, sample $\boldsymbol{\zeta}^{(t)}$ from the distribution $p(\boldsymbol{\zeta}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\psi}^{(t-1)}) \propto p(\mathbf{Y}|\boldsymbol{\zeta})\, p(\boldsymbol{\zeta}|\mathbf{X}, \boldsymbol{\psi}^{(t-1)})$ given the observed data $(\mathbf{Y}, \mathbf{X})$ and the values of $\boldsymbol{\psi}^{(t-1)}$ of the parameters from the previous iteration.
- **Sampling of the parameters**: Sample $\boldsymbol{\psi}^{(t)}$ from the posterior distribution $p(\boldsymbol{\psi}|\boldsymbol{\zeta}^{(t)}, \mathbf{X}) \propto p(\boldsymbol{\zeta}^{(t)}|\mathbf{X}, \boldsymbol{\psi})\, p(\boldsymbol{\psi})$, given $\mathbf{X}$ and $\boldsymbol{\zeta}^{(t)}$. This divides into

  $$p(\boldsymbol{\psi}|\boldsymbol{\zeta}^{(t)}, \mathbf{X}) = p(\boldsymbol{\psi}_\eta|\boldsymbol{\eta}^{(t)}, \mathbf{X})\, p(\boldsymbol{\psi}_\xi|\boldsymbol{\xi}^{(t)}, \mathbf{X}) \propto [p(\boldsymbol{\eta}^{(t)}|\mathbf{X}; \boldsymbol{\psi}_\eta) p(\boldsymbol{\psi}_\eta)]\, [p(\boldsymbol{\xi}^{(t)}|\mathbf{X}; \boldsymbol{\psi}_\xi) p(\boldsymbol{\psi}_\xi)],$$

  i.e. into separate posteriors for $\boldsymbol{\psi}_\eta$ and $\boldsymbol{\psi}_\xi$.

These steps split further into separate steps for different components of $\boldsymbol{\zeta}$ and $\boldsymbol{\psi}$. For $\boldsymbol{\zeta}$ and all the parameters except for $\boldsymbol{\alpha}$, the algorithm is similar to the one in Kuha, Zhang and Steele (2023), with adjustments to allow for the differences that here $\boldsymbol{\eta}_i$ has four variables and that their correlations vary by unit i. These steps are described in supplementary Appendix D.

What is new here is sampling the coefficients $\boldsymbol{\alpha}$ of the model for the conditional correlations of $\boldsymbol{\eta}_i$, in such a way that they imply positive definite correlation matrices at all relevant values of $\mathbf{X}$. Here all the other parameters in $\boldsymbol{\psi}$ and all the latent variables $\boldsymbol{\eta}_i$ are taken as known and fixed at their most recently sampled values. The latent variables are thus also treated as observed response variables in this model for their correlations. The other parameters $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}$ in $\boldsymbol{\psi}_\eta$ are omitted from the notation here, so that the posterior distribution that we need is written as $p(\boldsymbol{\alpha}|\mathbf{X}, \boldsymbol{\eta}) \propto p(\boldsymbol{\eta}|\mathbf{X}; \boldsymbol{\alpha})\, p(\boldsymbol{\alpha})$. As explained in Section 5, we need to ensure that all the sampled values are in a convex and bounded set $C_{\alpha, S_{XT}}$, where S_{XT} is a finite test set of values for $\boldsymbol{X}$. The prior $p(\boldsymbol{\alpha})$ and thus also the posterior are non-zero only in $C_{\alpha, S_{XT}}$.

We propose a tailored Metropolis-Hastings (MH) sampling procedure to implement this step efficienctly. This samples one element of $\boldsymbol{\alpha}$ at a time, relying on the result in Proposition 2 in Section 5 that the feasible values for any such parameter given the rest are a known interval. Let α_{lm} denote a single element of $\boldsymbol{\alpha}$, for $l = 1, \ldots, L$, $m = 1, \ldots, q$. The sampling algorithm updates α_{lm}, taking all the other elements $\boldsymbol{\alpha}_{-lm}$ fixed at their most recently sampled values. Denote $\mathbf{R}_i(\alpha_{lm}) = \mathbf{R}(\mathbf{X}_i; \alpha_{lm}, \boldsymbol{\alpha}_{-lm})$ and define the standardized residuals

$\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n]^\mathsf{T} = \mathbf{S}^{-1}(\boldsymbol{\eta} - \mathbf{X}\boldsymbol{\beta})$ where $\boldsymbol{\eta} = [\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n]^\mathsf{T}$ and $\mathbf{S} = \mathrm{diag}(\sigma_{GP}, \sigma_{RP}, 1, 1)$. The conditional posterior distribution from which α_{lm} should be drawn is then

$$p(\alpha_{lm}|\boldsymbol{\alpha}_{-lm}, \boldsymbol{\epsilon}, \mathbf{X}) \propto \prod_{i=1}^{n} p(\boldsymbol{\epsilon}_i|\boldsymbol{\alpha}, \mathbf{X})\, p(\alpha_{lm}|\boldsymbol{\alpha}_{-lm})$$

$$\propto \prod_{i=1}^{n} |\mathbf{R}_i(\alpha_{lm})|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_i^\mathsf{T}\mathbf{R}_i(\alpha_{lm})^{-1}\boldsymbol{\epsilon}_i\right) \mathbb{1}(a_{lm} < \alpha_{lm} < b_{lm}), \tag{13}$$

where (a_{lm}, b_{lm}) is the range of α_{lm} in the subset of $C_{\alpha,S_{XT}}$ given $\boldsymbol{\alpha}_{-lm}$. This involves n matrix determinants and inverses, plus further determinants to obtain the interval (a_{lm}, b_{lm}) as described in Proposition 2. This would be computationally demanding. However, these demands are reduced because the sampling updates only one parameter α_{lm} at a time. The determinant and inverse of $\mathbf{R}_i(\alpha_{lm})$ can then be updated using numerically cheap rules rather than calculated from scratch, reducing the computational complexity from $O(K^3)$ to $O(K^2)$ where K is the dimension of $\mathbf{R}$. These features are included in the general elementwise MH procedure that we propose. It is given in Algorithm 1, together with Remarks 1–4 in supplementary Appendix E.

Based on results of Tierney (1994, 1996), certain regularity conditions — irreducibility, aperiodicity, and positive Harris recurrence — ensure a unique stationary distribution for Markov chains. The Markov chain constructed by our estimation procedure adheres to these conditions as a special case of random walk Metropolis algorithm (Gelman, Gilks and Roberts, 1997). Moreover, through the design of the acceptance probability, the detailed balance condition holds for the chain, thereby ensuring that the desired posterior distribution serves as its unique stationary distribution. This convergence extends to all structural model parameters, as they are sampled from desired conditional distributions within a blockwise Gibbs sampling framework.

7. Analysis of child-parent exchanges of support.

7.1. *Introduction and research questions.* The model defined in Section 3 was fitted to the UKHLS data on exchanges of support between respondents and their non-coresident parents that were introduced in Section 2, using the method of estimation described in Section 6. Receiving and giving help are modelled jointly, treating practical and financial support as distinct but correlated outcomes. We investigate three broad research questions: (a) What individual characteristics are associated with higher or lower levels of giving help to the parents, and receiving help from them? (b) To what extent are exchanges reciprocated and how does reciprocity vary by individual characteristics? (c) Are practical and financial support substitutes for one another or are they complementary, and how does this depend on individual characteristics? Questions (b) and (c) refer to within-person correlations between the helping tendencies. For (b), higher levels of reciprocity would correspond to positive correlations between giving and receiving help. For (c), positive correlations between the tendencies to give (or to receive) practical and financial help would suggest that the two types of support are complementary (i.e. given together), and negative correlations that they are substitutes.

7.2. *Alternative approaches to the analysis.* Before we describe the results of our models below, in this section we briefly discuss other possible methods that could be employed to try to answer the research questions. These methods are ostensibly simpler than the joint modelling approach that we use, but they are ultimately limited and inflexible in ways which make them inadequate for our goals. We discuss them with reference to previous studies that have applied such methods to questions on intergenerational support, but the same ideas could

Algorithm 1: Elementwise Metropolis-Hastings procedure for sampling $\boldsymbol{\alpha}$. Further information is given in Remarks 1–4 in supplementary Appendix E.

1. **Input:** Current parameters $\boldsymbol{\alpha} = (\alpha_{lm})$ for $l = 1, \ldots, L$, $m = 1, \ldots, q$.
   For units $i = 1, \ldots, n$: Standardized residuals $\boldsymbol{\epsilon}_i = \mathbf{S}^{-1}(\boldsymbol{\eta}_i - \boldsymbol{\beta}^{\mathsf{T}}\mathbf{X}_i)$;
   $\mathbf{R}_i^{-1}$ and $|\mathbf{R}_i|$ for correlation matrices $\mathbf{R}_i = \mathbf{R}(\mathbf{X}_i; \boldsymbol{\alpha})$.
   For a test set $S_{XT} = \{\mathbf{X}_j \,|\, j = 1, \ldots, T\}$: Upper triangular matrices $\boldsymbol{\Gamma}_j$ from the Cholesky decompositions $\mathbf{R}_j = \boldsymbol{\Gamma}_j^{\mathsf{T}}\boldsymbol{\Gamma}_j$ of $\mathbf{R}_j = \mathbf{R}(\mathbf{X}_j; \boldsymbol{\alpha})$.
2. **Metropolis-Hastings sampling:**

for $l = 1, \ldots, L$ **do**
  for $m = 1, \ldots, q$ **do**
    Proposal generation:
    Calculate (a_{lm}, b_{lm}) based on $\boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_T$. See Remark 1 for more on this.
    Generate α'_{lm} from a proposal distribution $g(\alpha'_{lm}|\alpha_{lm})$. See Remark 2 for more on how the proposal can be created.
    Rejection:
    Calculate $\mathbf{R}_i(\alpha'_{lm})^{-1}$ by updating $\mathbf{R}_i(\alpha_{lm})^{-1}$ and $|\mathbf{R}_i(\alpha'_{lm})|$ by updating $|\mathbf{R}_i(\alpha_{lm})|$, for $i = 1, ..., n$; see Remark 3.
    Calculate the acceptance probability

$$\pi(\alpha_{lm} \to \alpha'_{lm}) = \min\left\{1, \frac{p(\alpha'_{lm}|\boldsymbol{\alpha}_{-lm}, \boldsymbol{\epsilon}, \mathbf{X})\, g(\alpha_{lm}|\alpha'_{lm})}{p(\alpha_{lm}|\boldsymbol{\alpha}_{-lm}, \boldsymbol{\epsilon}, \mathbf{X})\, g(\alpha'_{lm}|\alpha_{lm})}\right\}$$

    where $p(\alpha_{lm}|\boldsymbol{\alpha}_{-lm}, \boldsymbol{\epsilon}, \mathbf{X})$ is given by equation (13).
    Sample $u \sim U(0,1)$.
    if $u > \pi(\alpha_{lm} \to \alpha'_{lm})$ **then**
      Reject α'_{lm};
      continue
    end
    Accept α'_{lm} and update
      $\alpha_{lm} \to \alpha'_{lm}$, $\mathbf{R}_i(\alpha_{lm})^{-1} \to \mathbf{R}_i(\alpha'_{lm})^{-1}$, $|\mathbf{R}_i(\alpha_{lm})| \to |\mathbf{R}_i(\alpha'_{lm})|$.
    Update $\boldsymbol{\Gamma}_j(\alpha_{lm}) \to \boldsymbol{\Gamma}_j(\alpha'_{lm})$ for $j = 1, \ldots, T$; see Remark 4.
  end
end

3. **Output:** Updated $\boldsymbol{\alpha}$, $\mathbf{R}_i^{-1}$, $|\mathbf{R}_i|$ and $\boldsymbol{\Gamma}_j$.

be used to examine associations in any context. We are not aware of research that has used these methods to explicitly model the complementarity of financial and practical support, so the studies that are mentioned here concern reciprocity of exchanges. We discuss two simple approaches: (i) reducing two variables on giving and receiving into one, and (ii) using an indicator of one type of exchange as a predictor of another type of exchange.

The most common version of approach (i) operationalises reciprocity as the net balance of transfers between parents and children, with the direction of the difference between giving and receiving indicating whether exchanges are from the older to younger generation, or the reverse. Previous research has calculated such difference scores from overall indices of giving and receiving help that combine different types of support after monetarising non-financial transfers (Litwin et al., 2008; Mudrazija, 2016). This approach cannot really be used when, as in our analysis, helping tendencies are treated as continuous latent variables. A different version of the same idea uses latent class analysis of indicators of support given and received to derive a joint categorical outcome for whether exchanges are mutual or one-way (e.g. Hogan, Eggebeen and Clogg, 1993; Silverstein and Bengtson, 1997). Both of these methods allow the modelling of the association between giving and receiving to some extent, but without clear quantification of its strength. Another limitation of this approach is that because it first combines the two variables, it rules out a separate analysis of the predictors of giving and receiving themselves.

A basic version of approach (ii) has been the most widely used in previous research (e.g. Silverstein et al., 2002; Grundy, 2005; Albertini, Kohli and Vogel, 2007; Deindl and Brandt, 2011; Cheng et al., 2015; Evandrou et al., 2018). Here a measure of, say, giving support is included in a model for a measure of receiving support, together with other predictors. The coefficient of giving in this model, possibly suitable standardised, can then be interpreted as a measure of conditional association between giving and receiving (and the coefficients of other predictors of receiving support are also conditional on the level of giving).

This basic version of approach (ii) does not yet provide a model for how the associations depend on predictors. To get that, we would need to include interactions between a helping variable and other covariates. We have found no examples of this approach in previous research on intergenerational support, but it is easy to see how it could be done. For example, an interaction between giving support and a respondent's age would provide a measure of how the association between giving and receiving depends on age. Doing this, however, is not appealing for our purposes, because it would reduce the interpretability of the results without any compensating simplification of the modelling. The model would need to include multiple interaction terms, one for each covariate that was a predictor of an association. In our application the implementation would be further complicated because the helping variables in these interactions would be latent variables and because we would need to do this for six different associations. A further, conceptual problem with this approach is that it is asymmetric: a model for giving conditional on receiving and covariates estimates a different conditional association than one for receiving conditional on giving and covariates. For these reasons it is preferable to model the correlations and means directly and separately, rather than mix them up in an interaction specification.

We note, finally, that another way to explore variation in correlations would be to simply split the data into subsets by levels of covariates (by age group, for example) and estimate the correlations separately for each of them. This, however, would only allow us to consider small numbers of categorical variables, but not to examine multiple categorical and continuous explanatory variables for the correlations together.

7.3. *Estimation of the models.* Estimates $\tilde{\boldsymbol{\phi}}$ of the parameters of the measurement model were obtained first, as explained in Section 6.1. They are shown in supplementary Appendix F. The loading parameters are positive, meaning that the latent variables η_{GP} and η_{RP} are defined in such a way that larger values of them imply higher tendencies to give and receive practical help (and the same is true by construction for the financial help variables η_{GF} and η_{RF}). The measurement parameters were then fixed at $\tilde{\boldsymbol{\phi}}$ in the estimation of the rest of the model below.

The structural model for the joint distribution of the latent variables was estimated using the MCMC algorithm described in Section 6.2 and supplementary Appendices D and E. Estimated parameters and some predicted values for these models are shown in Tables 1-4 and in supplementary Appendix G. They are based on 380,000 draws of the parameters $\boldsymbol{\psi}$, obtained by pooling two MCMC chains of 200,000 iterations, with a burn-in sample of 10,000 omitted from each chain. Convergence was assessed by visual inspection of trace plots of the two chains which suggested adequate mixing. In the role of the target set S_Z for the covariates, we used the simple choice of all the n observed values of $\mathbf{Z}_i$ in the data, and as the test set S_{XT} all the distinct values of $\mathbf{X}_i = \mathbf{X}(\mathbf{Z}_i)$ implied by them.

Estimated parameters of model (5) for the binary latent class variables (ξ_G, ξ_R), and fitted class probabilities $p(\xi_G = 1)$ and $p(\xi_R = 1)$ from it, are shown in supplementary Appendix G. This model component is included primarily to allow for zero-inflation in the observed item responses, so it is not our main focus. We could, however, also interpret the classes defined by $\xi_G = 1$ and $\xi_R = 1$ as latent sub-populations of potential 'givers' and 'receivers' of help respectively. The estimated overall proportions of these classes, averaged over the sample distribution of the covariates, are 0.67 for 'givers' and 0.62 for 'receivers'.

The focus of interest is the model for the joint distribution of $\boldsymbol{\eta} = (\eta_{GP}, \eta_{GF}, \eta_{RP}, \eta_{RF})^{\mathsf{T}}$, which we interpret as continuous latent tendencies for the adult respondents to give and to receive practical and financial help, after accounting for the zero-inflation. We consider first results for the model (6) for the means of $\boldsymbol{\eta}$, which is used to answer research question (a) stated in Section 7.1, and then model (8)–(9) for their correlations, corresponding to questions (b) and (c).

7.4. *Predictors of levels of giving and receiving help.* Table 1 shows the estimated coefficients of the predictors of the means of practical (η_{GP}) and financial (η_{GF}) help given by respondents to parents. There is little evidence that the respondent's partnership status or the presence or age of their children are associated with the tendency to give help. Women tend to give more practical help than men, but there is no gender difference in giving financial help. Indicators of lower socioeconomic status or a more difficult economic situation of the respondent (lower education, not being a homeowner, lower household income, and not being employed) are associated with a higher tendency to give practical help, while having more education and higher household income predict a higher tendency to give financial help. These results are consistent with a pattern where children give help to the best of their ability, with the less well-off children giving, on average, relatively more practical support and less financial support. However, the results for household tenure and employment status (where home owners and the employed also tend to give less financial help) deviate from this pattern, after controlling for education and income. There is also some evidence that respondents with one sibling give less help than those with none, which could suggest some sharing of support between the siblings (although there is no similar reduction for those with more siblings).

Having a parent who lives alone and older parental age are positively associated with giving both forms of help, with the positive association with financial help emerging when the oldest parent reaches their early 70s. These findings are consistent with children giving help according to parental need. After controlling for parental age, respondent's age has an inverse U-shaped relationship with giving help, with highest levels of giving by respondents aged in their 40s. Finally, respondents who live more than an hour away from the nearest parent have a lower tendency to give practical help, but a higher tendency to give financial help. As for the effects of socioeconomic status, the different directions of these associations suggest differences in the mix of the two types of help related to the giver's circumstances, in this case according to how feasible it is to provide practical help.

Covariate effects on levels of practical and financial help that the respondents receive from their parents (variables η_{RP} and η_{RF}) are shown in Table 2. Women tend to receive more of both types of support than men. Expected levels of help from parents are also higher for respondents who are not employed, have less education, or have no coresident partner, all of which can be taken to indicate higher levels of need for support. Respondents with two or more siblings tend to receive less of either form of help than those from one or two-child families, which may reflect greater competition for parental resources in larger families. For financial help, the tendency to receive it is higher for respondents who have lower household income or who rent rather than own their homes, as well as for those with very young or secondary school age children. These associations are also consistent with parents providing more financial assistance to children who are most in need.

Levels of both practical and financial help received decline with the respondent's age, which is consistent with reduced need by respondents. As a function of the oldest parent's age, receipt of practical help also declines from age 67 onwards, but the tendency to receive financial help increases with parental age. This may be interpreted as another instance of the balance of different types of help depending on the giver's capacities, in this case with older parents being more able to give financial than practical support. Finally, longer travel time between the respondent and their nearest parent is associated with less practical and more financial help, as it was also for help from respondents to parents.

7.5. *Models for the correlations.* Estimated coefficients ($\hat{\boldsymbol{\alpha}}$) of model (8)–(9) for the residual correlations of $\boldsymbol{\eta} = (\eta_{GP}, \eta_{GF}, \eta_{RP}, \eta_{RF})$ are shown in Table 3. Here we included as covariates the respondent's age, age squared, gender, household income, and travel time to the nearest parent. Whereas the models in Section 7.4 concern the expected level of each helping tendency separately, these correlations focus on their joint distribution for a given child-parent dyad, over and above the levels predicted by the mean models.

For ease of interpretation we focus on some fitted correlations from these models, as shown in Table 4. The figures on its first row are the fitted correlations (for each of the six pairs of helping tendencies) averaged over the parameter values in the MCMC sample and over the respondents in the analysis sample. The other fitted values in the table are obtained similarly, except that one covariate at a time is fixed at specific value (e.g. age at 35 years) while leaving the other covariates at their sample values.

The four correlations between the tendencies to give and receive help (of the same or different type) are measures of reciprocity or symmetry in exchanges between children and their parents (research question (b) above). Results for them are given in the first four columns of each table. Focusing on Table 4, consider first the fitted correlations on its first row, averaged over the sample distribution of all the covariates. There is a moderate positive correlation of 0.38 between giving and receiving practical help (GP $\leftrightarrow$ RP). In other words, when a child has a high tendency to give practical help to their parent(s), relative to what would be predicted by their own and the parents' characteristics, they also tend to receive a relatively high level of support from the parents. This suggests a fair amount of reciprocity in practical help. The other three correlations are weaker, indicating little dyad-level reciprocity in anything other than practical help. What is not observed here are any substantial negative correlations. They would indicate that when the tendency to help is high in one direction it is low in the other, as would happen for example if help was given only in the direction of greater need. This is not seen here even for giving and receiving financial help, even though we might expect financial exchanges to be largely unidirectional. A possible explanation of this is that the single financial support item covers also small sums of money, which may be exchanged more frequently and symmetrically than large ones.

The (GP $\leftrightarrow$ RP) correlation is also the one for which we see the clearest covariate effects, as illustrated by the other rows of Table 4. It declines sharply with age, and is significantly higher for men than for women and among parents and children who live farther apart. Reciprocity in practical support is highest at younger ages of the adult children. This captures a different aspect of the effects of age than the mean models in Section 7.4. There respondent's age was negatively associated with tendency to give practical help and positively associated (up to age around 43) with receiving it. Thus younger individuals tend to give less practical help and receive more of it, and the expected balance of support is more toward help from parents to children, than is the case at older ages (comparable conclusions were reached in a different way by Mudrazija 2016, who considered net financial values of the differences between these two directions). The residual correlations, however, show that, around these expected levels, for younger respondents the level of practical help that they do (or do not) give is particularly strongly predictive of how much support they receive. Similarly, the gender difference in the correlation suggests that men are more likely than women to engage in two-way exchanges or not exchange practical help at all.

TABLE 1

Estimated parameters of the linear model for the expected value of the tendency to give practical help (η_{GP}) and to give financial help (η_{GF}) to individuals' non-coresident parents. The estimates are posterior means from MCMC samples (with posterior standard deviations in parentheses).

	Giving practical help Estimate	(s.d.)	Giving financial help Estimate	(s.d.)
Estimated coefficients:	$\hat{\boldsymbol{\beta}}_{GP}$		$\hat{\boldsymbol{\beta}}_{GF}$	
Intercept	-0.70^{***}	(0.18)	-2.35^{***}	(0.31)
Respondent (child) characteristics				
Age† (×10)	0.03	(0.03)	0.12^{***}	(0.04)
Age squared† ($\times 10^3$)	-0.60^{***}	(0.12)	-0.70^{***}	(0.19)
Gender				
Female (vs. Male)	0.41^{***}	(0.03)	0.03	(0.04)
Partnership status				
Partnered (vs. Single)	−0.04	(0.03)	0.01	(0.05)
Age of youngest coresident child (vs. No children):				
0–1 years	−0.08	(0.06)	−0.05	(0.09)
2–4 years	0.01	(0.05)	0.03	(0.08)
5–10 years	0.02	(0.04)	0.09	(0.07)
11–16 years	−0.04	(0.05)	−0.10	(0.07)
17– years	0.03	(0.04)	−0.03	(0.06)
Number of siblings (vs. None)				
1	-0.08^{*}	(0.04)	-0.12^{*}	(0.07)
2 or more	0.00	(0.04)	0.06	(0.07)
Longstanding illness (vs. No)	0.07^{*}	(0.04)	0.07	(0.06)
Employment status (vs. Employed)				
Not employed	0.21^{***}	(0.03)	0.11^{**}	(0.05)
Education (vs. Secondary or less)				
Post-secondary	-0.05^{**}	(0.03)	0.12^{***}	(0.04)
Household tenure (vs. Renter)				
Own home outright or with mortgage	-0.17^{***}	(0.03)	-0.19^{***}	(0.05)
Logarithm of household equivalised income	-0.04^{**}	(0.02)	0.09^{***}	(0.03)
Parent characteristics				
Age of the oldest living parent† (×10)	0.28^{***}	(0.02)	−0.02	(0.04)
Age of the oldest parent squared† ($\times 10^3$)	0.52^{***}	(0.11)	0.63^{***}	(0.17)
At least one parent lives alone (vs. No)	0.33^{***}	(0.03)	0.24^{***}	(0.04)
Child-parent characteristics				
Travel time to the nearest parent				
More than 1 hour (vs. 1 hour or less)	-0.43^{***}	(0.04)	0.14^{**}	(0.05)
Residual s.d.:	$\hat{\sigma}_{GP}$			
	0.73	(0.01)	1	

The posterior credible interval excludes zero at level 90% (*), 95% (**) or 99% (***).
† Age of respondent is centered at 40, and age of oldest living parent at 70.

TABLE 2

Estimated parameters of the linear model for the expected value of the tendency to receive practical help (η_{RP}) and to receive financial help (η_{RF}) from individuals' non-coresident parents. The estimates are posterior means from MCMC samples (with posterior standard deviations in parentheses).

	Receiving practical help Estimate	(s.d.)	Receiving financial help Estimate	(s.d.)
Estimated coefficients:	$\hat{\boldsymbol{\beta}}_{RP}$		$\hat{\boldsymbol{\beta}}_{RF}$	
Intercept	-2.17^{***}	(0.23)	1.03^{***}	(0.34)
Respondent (child) characteristics				
Age† ($\times 10$)	-0.26^{***}	(0.03)	-0.28^{***}	(0.04)
Age squared† ($\times 10^3$)	-0.16	(0.18)	-0.46^{*}	(0.23)
Gender				
Female (vs. Male)	0.27^{***}	(0.03)	0.15^{***}	(0.04)
Partnership status				
Partnered (vs. Single)	-0.35^{***}	(0.04)	-0.30^{***}	(0.05)
Age of youngest coresident child (vs. No children):				
0–1 years	0.02	(0.05)	0.14^{*}	(0.07)
2–4 years	-0.03	(0.05)	0.07	(0.06)
5–10 years	-0.09^{**}	(0.04)	-0.02	(0.06)
11–16 years	-0.11^{*}	(0.06)	0.18^{**}	(0.07)
17– years	-0.12	(0.07)	0.02	(0.09)
Number of siblings (vs. None)				
1	0.00	(0.05)	-0.07	(0.07)
2 or more	-0.14^{***}	(0.05)	-0.25^{***}	(0.06)
Longstanding illness (vs. No)	0.03	(0.05)	0.06	(0.06)
Employment status (vs. Employed)				
Not employed	0.21^{***}	(0.04)	0.14^{**}	(0.05)
Education (vs. Secondary or less)				
Post-secondary	-0.06^{**}	(0.03)	-0.06	(0.04)
Household tenure (vs. Renter)				
Own home outright or with mortgage	0.08^{**}	(0.03)	-0.34^{***}	(0.05)
Logarithm of household equivalised income	0.01	(0.02)	-0.14^{***}	(0.03)
Parent characteristics				
Age of the oldest living parent† ($\times 10$)	-0.03	(0.03)	0.22^{***}	(0.04)
Age of the oldest parent squared† ($\times 10^3$)	-0.44^{***}	(0.15)	0.28	(0.19)
At least one parent lives alone (vs. No)	-0.05	(0.03)	0.04	(0.04)
Child-parent characteristics				
Travel time to the nearest parent				
More than 1 hour (vs. 1 hour or less)	-0.42^{***}	(0.05)	0.27^{***}	(0.06)
Residual s.d.:	$\hat{\sigma}_{RP}$			
	0.68	(0.02)	1	

The posterior credible interval excludes zero at level 90% (*), 95% (**) or 99% (***).
† Age of respondent is centered at 40, and age of oldest living parent at 70.

TABLE 3

Estimated coefficients ($\hat{\alpha}$) of the model for the residual correlations of the tendencies to give and receive practical help (GP and RP) and to give and receive financial help (GF and RF). The estimates are posterior means from MCMC samples (with posterior standard deviations in parentheses).

	Correlation					
	GP↔RP	GP↔RF	GF↔RP	GF↔RF	GP↔GF	RP↔RF
Intercept	0.087	0.166	−0.133	−0.126	0.475***	0.148
	(0.171)	(0.186)	(0.220)	(0.159)	(0.174)	(0.197)
Age of respondent†	−0.014***	0.004*	0.003	−0.001	−0.002	−0.009***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Age squared† ($\times 10^3$)	−0.277**	−0.137	0.001	0.159	−0.112	−0.251*
	(0.124)	(0.149)	(0.178)	(0.184)	(0.129)	(0.133)
Female	−0.151***	−0.025	−0.119*	−0.103*	−0.080*	0.044
	(0.044)	(0.047)	(0.063)	(0.062)	(0.043)	(0.046)
Travel time to nearest parent > 1hr	0.141***	−0.206***	−0.119	−0.226***	−0.273***	−0.252***
	(0.051)	(0.058)	(0.080)	(0.076)	(0.056)	(0.055)
Log(household income)	0.044***	0.007	0.025	0.017	0.003	0.017
	(0.017)	(0.019)	(0.022)	(0.016)	(0.018)	(0.020)

The posterior credible interval excludes zero at level 90% (*), 95% (**) or 99% (***).
† Age of respondent is centered at 40.

TABLE 4

Fitted residual correlations calculated using the parameter estimates in Table 3, averaged over parameter values in the MCMC samples and over covariate values in the analysis sample. The 'Overall' values are averaged over sample values of all the covariates, and the other fitted values over the sample values of all the covariates except for the one fixed at the specified value.

Covariate	Correlation					
setting	GP↔RP	GP↔RF	GF↔RP	GF↔RF	GP↔GF	RP↔RF
Overall	0.38	0.16	0.02	−0.06	0.36	0.20
Age of respondent						
35 years	0.53	0.14	0.00	−0.07	0.39	0.31
45 years	0.39	0.18	0.03	−0.08	0.37	0.22
55 years	0.20	0.19	0.06	−0.06	0.32	0.08
Gender						
Female	0.31	0.14	−0.03	−0.10	0.32	0.22
Male	0.47	0.17	0.09	0.00	0.40	0.18
Travel time to the nearest parent						
> 1 hr	0.48	0.01	−0.06	−0.22	0.16	0.02
≤ 1 hr	0.34	0.21	0.05	0.00	0.43	0.27
Logarithm of household equivalised income						
25th percentile	0.37	0.15	0.02	−0.06	0.36	0.20
50th percentile	0.38	0.16	0.02	−0.06	0.36	0.20
75th percentile	0.39	0.16	0.03	−0.05	0.36	0.21

The only other clearly significant covariate effects that relate to reciprocity are those between within-dyad distance and the (GP ↔ RF), (GF ↔ RF) and (GP ↔ RP) correlations. Recall that the models for the means showed that the balance of the expected levels of different types of help moves towards more financial and less practical support when the child and the parent(s) live further apart. Of the residual correlations here, (GP $\leftrightarrow$ RP) is quite strongly positive when the distance is longer vs. less positive when it is shorter, while (GP $\leftrightarrow$ RF) is near zero vs. moderately positive and (GF $\leftrightarrow$ RF) moderately negative vs. near zero similarly (and GF $\leftrightarrow$ RP is always small). One possible interpretation of these different patterns is that among children and parents who live further apart providing practical support requires a greater effort and the tendency to give such support may be higher when reciprocated. For dyads at a longer distance, financial help may also more often involve one-way (and perhaps larger) transfers which are less often and less easily reciprocated by practical help.

The two remaining correlations, between the tendencies to give financial and practical help, and between the tendencies to receive them, relate to whether one form of help that a person may give serves as a substitute for the other or whether they are complementary, and whether this varies according to individual characteristics (research question (c)). The mean models in Section 7.4 also give information about one version of this question, when they show that the expected balance of the two types of help is, on average, different for different types of dyads. This is most obvious when the coefficient of a covariate has different signs for practical and financial help, as it does for example for the distance between respondent and their parents (a similar result for expected levels of financial vs. time assistance given distance was found by Bonsang 2007 in a cross-national European study). However, this is again not the same as the question of substitution for a person, i.e. whether the level of one kind of help that he or she gives predicts higher or lower levels of the other kind of help.

Results for the correlations that address this question are given in the final two columns of Tables 3 and 4. The fitted correlations are positive overall and in all sub-groups defined by the covariates. This indicates clearly that within a person the types of help are complementary rather than substitutes of each other: a child or parent (or parents) who has a high tendency to give one kind of help (relative to what would be expected given the characteristics of their dyad) also has a high tendency to give the other kind of help. The most noticeable covariate effect that holds for both children and the parents is that the degree of complementarity in practical and financial help is greater when the child-parent distance is small. For help received from the parents, complementarity also declines with the respondent's (and thus in effect also the parents') age. This suggests that at older ages the parents more often tend to limit the support that they give to one of these types (most often financial help, in light of the results in Table 2) rather than both of them.

In conclusion, we return to the research questions that were stated in Section 7.1. The first question was addressed by the models for the mean levels of helping tendencies in Section 7.4. Their results may be summarised in terms of two broad types of characteristics: the capacities of a giver of support and the level of need of the recipient. The model results indicate clearly that recipients with higher level of need (such as children with less privileged socioeconomic status or parents who are older or live alone) tend to receive more support. For capacities of giving, the results are more subtle. There is no strong evidence that lower capacity is associated with less help given in some overall sense. Instead, different types of individuals tend to give the types of help that they are best able to give, e.g. with less wealthy children giving, on average, relatively more practical than financial help to their parents, and older parents providing relatively more financial help to their children.

The other two research questions correspond to the models for residual correlations in this section. The results show evidence of reciprocity between children and parents in practical help, and of within-person complementarity in giving different types of help. A prominent covariate effect here was that the patterns of correlations between helping tendencies of different types and directions where somewhat different for children who live far from rather than close to their parents.

8. Conclusions. We have proposed methods for modelling joint distributions of multivariate continuous variables, including models for how their correlations depend on covariates. A linear model was specified for each correlation, and we developed an estimation procedure that ensures that the estimated model implies positive definite correlation matrices over a relevant set of values of the covariates. This builds on literature on such 'constrained' methods of estimation for models for correlations, which are here extended to include unit-level covariates. The estimation is carried out using a tailored MCMC algorithm which includes an efficient Metropolis-Hastings sub-procedure for estimating the correlation model.

These methods were motivated by substantive research questions on the levels and correlations of intergenerational family support. There the model was defined for the joint distribution of latent variables which represent individuals' tendencies of giving and receiving different types of support. We applied it to study exchanges of support between adult individuals and their non-coresident parents in the UK, using survey data from the UK Household Longitudinal Study. We modelled the conditional means and correlations of different helping tendencies. The mean levels are positively associated with many characteristics of the recipients that indicate higher need, and with characteristics of givers that indicate their higher capacity to give help. These results are, arguably, fairly encouraging about patterns of intergenerational support in this population. Less positively, large proportions of both adult individuals and their parents do not typically give any of the kinds of help considered here. The estimated correlations indicate reciprocity, where those who tend to give high levels of practical help also tend to receive much of it, and complementarity, where those who tend to give high levels of one kind of help (practical or financial) also tend to give much of the other kind. This suggests a picture of a general culture of helpfulness within some families, and general lack of it in others, rather than a sort of zero-sum game where help would flow only in one direction at a time and one kind of help would reduce the amount of other kinds.

This work could be extended in a number of ways. Methodologically, the proposed modelling approach for the correlation matrix could be embedded into other covariance modelling tasks, such as the copula model. (Hoff, 2007; Murray et al., 2013). The computational efficiency and mixing rates of the simple element-wise Metropolis-Hastings MCMC sampler that was used here could perhaps be improved by using other approaches, for example adaptive MCMC (Haario, Saksman and Tamminen, 2001; Andrieu and Thoms, 2008) which proposes multiple parameters from an adaptive proposal in each iteration.

Substantively, the choices of this analysis were constrained by the available data. Although we considered practical and financial support separately, the single indicator of financial support leaves us unable to examine varieties of it in more detail. Because we analyse data collected from the adult children only, we have limited information about their parents. Another promising direction would be to extend the models to longitudinal data. This would allow us to examine reciprocity and complementarity of help not only contemporaneously, using models for correlations as described in this paper, but also over time, using predictive models for types of help at one time given help at earlier times. These areas of further research remain to be pursued.

Acknowledgements This research was supported by a UK Economic and Social Research Council (ESRC) grant "Methods for the Analysis of Longitudinal Dyadic Data with an Application to Inter-generational Exchanges of Family Support" (ref. ES/P000118/1). Additional funding for Siliang Zhang was provided by Shanghai Science and Technology Committee Rising-Star Program (22YF1411100) and National Natural Science Foundation of China (12301373).

SUPPLEMENTARY MATERIAL

Pseudodata and code for data analysis. The supplement includes a representative pseudo version of the data and R package and code for its analysis, together with information about access to the actual data used in the paper.

Additional results. The supplementary materials also include supplementary Appendices A–G which provide some additional tables and theoretical results. They are referred to in appropriate places in the main text of the paper above.

REFERENCES

ALBERTINI, M., KOHLI, M. and VOGEL, C. (2007). Intergenerational transfers of time and money in European families: common patterns - different regimes? *Journal of European Social Policy* **17** 319-334.

ANDERSON, T. W. (1973). Asymptotically Efficient Estimation of Covariance Matrices with Linear Structure. *The Annals of Statistics* **1** 135–141.

ANDRIEU, C. and THOMS, J. (2008). A Tutorial on Adaptive MCMC. *Statistics and Computing* **18** 343–373.

ARCHAKOV, I. and HANSEN, P. R. (2021). A new parametrization of correlation matrices. *Econometrica* **89** 1699–1715.

ATTIAS-DONFUT, C., OGG, J. and WOLFF, F. C. (2005). European patterns of intergenerational financial and time transfers. *European Journal of Ageing* **2** 161-173.

BAKK, Z. and KUHA, J. (2018). Two-Step Estimation of Models Between Latent Classes and External Variables. *Psychometrika* **83** 871–892.

BARNARD, J., MCCULLOCH, R. and MENG, X.-L. (2000). Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, with Application to Shrinkage. *Statistica Sinica* **10** 1281–1311.

BENGTSON, V. L. (2001). Beyond the nuclear family: the increasing importance of multigenerational bonds. *Journal of Marriage and the Family* **63** 1-16.

BONSANG, E. (2007). How do middle-aged children allocate time and money transfers to their older parents in Europe? *Empirica* **34** 171-188.

CHEN, Z. and DUNSON, D. B. (2003). Random effects selection in linear mixed models. *Biometrics* **59** 762–769.

CHENG, Y. P., BIRDITT, K. S., ZARIT, S. H. and FINGERMAN, K. L. (2015). Young adults' provision of support to middle-aged parents. *Journals of Gerontology, Series B* **70** 407-416.

CHIB, S. and GREENBERG, E. (1998). Analysis of Multivariate Probit Models. *Biometrika* **85** 347–361.

CHIU, T. Y., LEONARD, T. and TSUI, K.-W. (1996). The matrix-logarithmic covariance model. *Journal of the American Statistical Association* **91** 198–210.

DAVEY, A. and EGGEBEEN, D. J. (1998). Patterns of intergenerational exchange and mental health. *Journal of Gerontology: Psychological Sciences* **53B** P86-P95.

DEINDL, C. and BRANDT, M. (2011). Financial support and practical help between older parents and their middle-aged children. *Ageing & Society* **31** 645–62.

EVANDROU, M., FALKINGHAM, J., GOMEZ-LEON, M. and VLACHANTONI, A. (2018). Intergenerational flows of support between parents and adult children in Britain. *Ageing & Society* **38** 321-351.

FAN, J., LIAO, Y. and LIU, H. (2016). An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal* **19** C1–C32.

GELMAN, A., GILKS, W. R. and ROBERTS, G. O. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability* **7** 110–120.

GHOSH, R. P., MALLICK, B. and POURAHMADI, M. (2021). Bayesian Estimation of Correlation Matrices of Longitudinal Data. *Bayesian Analysis* **16** 1039–1058.

GRUNDY, E. (2005). Reciprocity in relationships: socio-economic and health influences on intergenerational exchanges between third age parents and their adult children in Great Britain. *British Journal of Sociology* **56** 233-255.

HAARIO, H., SAKSMAN, E. and TAMMINEN, J. (2001). An Adaptive Metropolis Algorithm. *Bernoulli* **7** 223–242.

HENRETTA, J. C., VAN VOORHIS, M. F. and SOLDO, B. J. (2018). Cohort differences in parental financial help to adult children. *Demography* **55** 1567-1582.

HOFF, P. D. (2007). Extending the Rank Likelihood for Semiparametric Copula Estimation. *The Annals of Applied Statistics* **1** 265–283.

HOFF, P. D. and NIU, X. (2012). A Covariance Regression Model. *Statistica Sinica* **22** 729–753.

HOGAN, D. P., EGGEBEEN, D. J. and CLOGG, C. C. (1993). The structure of intergenerational exchanges in American families. *American Journal of Sociology* **98** 1428–1458.

HU, J., CHEN, Y., LENG, C. and TANG, C. Y. (2021). Regression Analysis of Correlations for Correlated Data. *arXiv preprint arXiv:2109.05861*.

HUININK, J., BRUDERL, J., NAUCK, B., WALPER, S., CASTIGLIONI, L. and FELDHAUS, M. (2011). Panel analysis of intimate relationships and family dynamics (pairfam): conceptual framework and design. *Zeitschrift für Familienforschung (Journal of Family Research)* **23** 77-101.

KIM, K., ZARIT, S. H., EGGEBEEN, D. J., BIRDITT, K. S. and FINGERMAN, K. L. (2011). Discrepancies in reports of support exchanges between aging parents and their middle-aged children. *Journals of Gerontology Series B* **66** 527-537.

KUHA, J. and BAKK, Z. (2023). Two-step estimation of latent trait models. arXiv preprint arXiv: 2303.16101.

KUHA, J., ZHANG, S. and STEELE, F. (2023). Latent variable models for multivariate dyadic data with zero inflation: Analysis of intergenerational exchanges of family support. *Annals of Applied Statistics* **17** 1521–1542.

LESTHAEGHE, R. (2014). The second demographic transition: A concise overview of its development. *Proceedings of the National Academy of Sciences* **111** 18112-18115.

LIECHTY, J. C., LIECHTY, M. W. and MÜLLER, P. (2004). Bayesian Correlation Estimation. *Biometrika* **91** 1–14.

LITWIN, H. (2004). Intergenerational exchange and mental health in later-life - the case of older Jewish Israelis. *Aging and Mental Health* **8** 196-200.

LITWIN, H., VOGEL, C., KÜNEMUND, H. and KOHLI, M. (2008). The balance of intergenerational exchange: correlations of net transfers in Germany and Israel. *European Journal of Ageing* **5** 91-102.

LUO, R. and PAN, J. (2022). Conditional Generalized Estimating Equations of Mean-Variance-Correlation for Clustered Data. *Computational Statistics and Data Analysis* **168** 107386.

MANDERMAKERS, J. J. and DYKSTRA, P. A. (2008). Discrepancies in parents' and adult child's reports of support and contact. *Journal of Marriage and Family* **70** 495-506.

MUDRAZIJA, S. (2016). Public transfers and the balance of intergenerational family support in Europe. *European Societies* **18** 336-35.

MURRAY, J. S., DUNSON, D. B., CARIN, L. and LUCAS, J. E. (2013). Bayesian Gaussian Copula Factor Models for Mixed Data. *Journal of the American Statistical Association* **108** 656–665.

MUTHÉN, L. K. and MUTHÉN, B. (2010). *Mplus User's Guide (Sixth Edition)*. Muthén & Muthén, Los Angeles, CA.

PAN, J. and MACKENZIE, G. (2006). Regression Models for Covariance Structures in Longitudinal Studies. *Statistical Modelling* **6** 43–57.

PAN, J. and PAN, Y. (2017). jmcm: An R Package for Joint Mean-Covariance Modeling of Longitudinal Data. *Journal of Statistical Software* **82** 1–29.

PINHEIRO, J. C. and BATES, D. M. (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing* **6** 289–296.

POURAHMADI, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika* **86** 677–690.

POURAHMADI, M. (2007). Cholesky decompositions and estimation of a covariance matrix: orthogonality of variance–correlation parameters. *Biometrika* **94** 1006–1013.

POURAHMADI, M. (2011). Covariance Estimation: The GLM and Regularization Perspectives. *Statistical Science* **26** 369–387.

ROSSEEL, Y. and LOH, W. W. (2022). A Structural After Measurement Approach to Structural Equation Modeling. *Psychological Methods*. Advance online publication. https://dx.doi.org/10.1037/met0000503.

ROUSSEEUW, P. J. and MOLENBERGHS, G. (1994). The Shape of Correlation Matrices. *The American Statistician* **48** 276–279.

SHAPIRO, A. (2004). Revisiting the generation gap: exploring the relationships of parent/adult-child dyads. *International Journal of Aging and Human Development* **58** 127–146.

SILVERSTEIN, M. and BENGTSON, V. L. (1997). Intergenerational solidarity and the structure of adult child-parent relationships in American families. *American Journal of Sociology* **103** 429-460.

SILVERSTEIN, M., CONROY, S. J., WANG, H., GIARRUSSO, R. and BENGTSON, V. L. (2002). Reciprocity in parent-child relations over the adult life course. *Journal of Gerontology: Social Sciences* **57B** S3-S13.

STEELE, F. and GRUNDY, E. (2021). Random effects dynamic panel models for unequally spaced multivariate categorical repeated measures: an application to child-parent exchanges of support. *Journal of the Royal Statistical Society C (Applied Statistics)* **70** 3-23.

STEELE, F., ZHANG, S., GRUNDY, E. and BURCHARDT, T. (2024). Longitudinal analysis of exchanges of support between parents and children in the UK. *Journal of the Royal Statistical Society, Series A* **187** 279–304.

STONE, J., BERRINGTON, A. and FALKINGHAM, J. (2011). The changing determinants of UK young adults' living arrangements. *Demographic Research* **25** 629-666.

SUITOR, J. J., GILLIGAN, M., PILLEMER, K., FINGERMAN, K. L., KIM, K., SILVERSTEIN, M. and BENGTSON, V. L. (2017). Applying within-family differences approaches to enhance understanding of the complexity of intergenerational relations. *The Journals of Gerontology: Series B* **73** 40-53.

TIERNEY, L. (1994). Markov Chains for Exploring Posterior Distributions. *The Annals of Statistics* **22** 1701–1728.

TIERNEY, L. (1996). Introduction to general state-space Markov chain theory. *Markov chain Monte Carlo in Practice* 59–74.

UNIVERSITY OF ESSEX, INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH, NATCEN SOCIAL RESEARCH AND KANTAR PUBLIC (2019). Understanding Society: Waves 1-9, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection], 12th ed. University of Essex, Institute for Social and Economic Research. UK Data Service. SN: 6614.

WANG, Y. and DANIELS, M. J. (2013). Bayesian Modeling of the Dependence in Longitudinal Data via Partial Autocorrelations and Marginal Variances. *Journal of Multivariate Analysis* **116** 130–140.

WILDING, G. E., CAI, X., HUTSON, A. and YU, Z. (2011). A Linear Model-Based Test for the Heterogeneity of Conditional Correlations. *Journal of Applied Statistics* **38** 2355–2366.

WONG, F., CARTER, C. K. and KOHN, R. (2003). Efficient Estimation of Covariance Selection Models. *Biometrika* **90** 809–830.

YAN, J. and FINE, J. (2007). Estimating Equations for Association Structures. *Statistics in Medicine* **23** 859–874.

ZHANG, W. and LENG, C. (2012). A Moving Average Cholesky Factor Model in Covariance Modelling for Longitudinal Data. *Biometrika* **99** 141–150.

ZHANG, W., LENG, C. and TANG, C. Y. (2015). A Joint Modelling Approach for Longitudinal Studies. *Journal of the Royal Statistical Society, Series B* **77** 219–238.

ZOU, T., LAN, W., WANG, H. and TSAI, C.-L. (2017). Covariance Regression Analysis. *Journal of the American Statistical Association* **112** 266–281.

ZOU, T., LAN, W., LI, R. and TSAI, C.-L. (2022). Inference on Covariance-Mean Regression. *Journal of Econometrics* **230** 318–338.

ZWIERNIK, P., UHLER, C. and RICHARDS, D. (2017). Maximum Likelihood Estimation for Linear Gaussian Covariance Models. *Journal of the Royal Statistical Society, Series B* **79** 1269–1292.

MODELLING CORRELATION MATRICES IN MULTIVARIATE DATA, WITH APPLICATION TO RECIPROCITY AND COMPLEMENTARITY OF CHILD-PARENT EXCHANGES OF SUPPORT

BY SILIANG ZHANG[1,*], JOUNI KUHA[2] AND FIONA STEELE[2,†]

[1]Key Laboratory of Advanced Theory and Application in Statistics and Data Science-MOE, School of Statistics, East China Normal University, [*]slzhang@fem.ecnu.edu.cn

[2]Department of Statistics, London School of Economics and Political Science, j.kuha@lse.ac.uk; [†]f.a.steele@lse.ac.uk

SUPPLEMENTARY APPENDICES

These supplementary materials include the following supplementary appendices, which are referred to in the main text:

A: Summary statistics for the analysis sample of data from UK Household Longitudinal Study (UKHLS)
B: Properties of set C_{α,S_X} — Proposition 1 discussed in Section 5, with discussion
C: Feasible intervals for individual parameters of the model for correlations — Proof of Proposition 2 discussed in Section 5
D: MCMC algorithm for the latent variables and for the parameters except for the correlation parameters $\boldsymbol{\alpha}$
E: Further information on the MCMC algorithm for the correlation parameters $\boldsymbol{\alpha}$ discussed in Section 6.2 — Remarks 1–4
F: Estimated measurement models for the analysis in Section 7
G: Estimated multinomial logistic model for the latent class variables $\boldsymbol{\xi}$ in the analysis in Section 7

A Summary statistics for the analysis sample of data from UK Household Longitudinal Study (UKHLS)

TABLE A1
Estimated parameters (measurement loadings and intercepts) of the measurement models for survey items on help given by respondents to their parents and on help received from the parents.

Item	Giving practical help loading	Giving practical help intercept	Receiving practical help loading	Receiving practical help intercept
Lifts in car	1.12	0.83	1.14	1.54
Shopping	2.38	1.02	1.70	2.08
Providing or cooking meals	1.24	-0.28	1.15	1.57
Basic personal needs (to parent only)	1.32	-1.32	–	–
Looking after children (from parents only)	–	–	0.89	2.25
Washing, ironing or cleaning	1.32	-0.77	1.15	0.82
Personal affairs	1.00	0.00	1.00	0.00
Decorating, gardening or house repairs	0.57	-0.22	0.74	0.37

TABLE A2
Descriptive statistics of the covariates used in the analysis.

Variable	n	%
Respondent (child) characteristics:		
Age (years)	Mean=43.7	SD=11.4
Gender		
Female	7060	57.9
Male	5143	42.1
Partnership status		
Partnered	9373	76.8
Single	2830	23.2
Age of youngest coresident child		
No children	5002	41.0
$0-1$ years	910	7.5
$2-4$ years	1231	10.1
$5-10$ years	1910	15.7
$11-16$ years	1548	12.7
$17-$ years	1602	13.1
Number of siblings		
None	1235	10.1
1	4325	35.4
2 or more	6643	54.4
Longstanding illness		
Yes	1533	12.6
No	10670	87.4
Employment status		
Employed	9688	79.4
Not employed	2515	20.6
Education (highest qualification)		
Secondary or less	6024	49.4
Post-secondary	6179	50.6
Household tenure		
Own home outright or with mortgage	8817	72.3
Other (private or social renter)	3386	27.7
Logarithm of household equivalised income	Mean=9.9	SD=0.79
Parent characteristics:		
Age of the oldest living parent (years)	Mean=72.1	SD=11.2
At least one parent lives alone		
Yes	4641	38.0
No	7562	62.0
Child–parent characteristics:		
Travel time to the nearest parent		
1 hour or less	8851	72.5
More than 1 hour	3352	27.5

Data from UKHLS, 2017-19 (Wave 9). The sample size for all covariates is $n = 12,203$.

B Properties of set C_{α,S_X} — Proposition 1 discussed in Section 5, with discussion

Denote $S_X = \mathbf{X}(S_Z) = \{\mathbf{X}(\mathbf{Z}) \mid \mathbf{Z} \in S_Z\}$ and let

$$C_{\alpha,S_X} = \{\boldsymbol{\alpha} \in \mathbb{R}^{L\times q} \mid \boldsymbol{\rho} = \boldsymbol{\alpha}^\mathsf{T}\mathbf{X} \in C_\rho \text{ for all } \mathbf{X} \in S_X\}$$

be the set of values of $\boldsymbol{\alpha}$ which are feasible when combined with any $\mathbf{X}$ in S_X. Proposition 1 gives basic properties of C_{α,S_X}.

PROPOSITION 1. *Properties of C_{α,S_X}:*

(i) *If $S_{X_2} \subseteq S_{X_1}$, then $C_{\alpha,S_{X_1}} \subseteq C_{\alpha,S_{X_2}}$.*
(ii) $C_{\alpha,\mathit{Conv}(S_X)} = C_{\alpha,S_X}$, *where Conv$(S_X)$ denotes the convex hull of S_X.*
(iii) $\mathbf{0} \in C_{\alpha,S_X}$.
(iv) *Suppose further that there exist q linearly independent elements in S_X. Then C_{α,S_X} is bounded.*
(v) C_{α,S_X} *is a convex set.*

PROOF OF PROPOSITION 1.

(i) Let $\boldsymbol{\alpha} \in C_{\alpha,S_{X_1}}$, so that $\boldsymbol{\alpha}^\mathsf{T}\mathbf{X} \in C_\rho$ for all $\mathbf{X} \in S_{X_1}$. Since $S_{X_2} \subseteq S_{X_1}$, in particular, for all $\mathbf{X} \in S_{X_2} \subseteq S_{X_1}$, $\boldsymbol{\alpha}^\mathsf{T}\mathbf{X} \in C_\rho$, and thus $\boldsymbol{\alpha} \in C_{\alpha,S_{X_2}}$.
(ii) Since $S_X \subseteq \text{Conv}(S_X)$, we have $C_{\alpha,\text{Conv}(S_X)} \subseteq C_{\alpha,S_X}$ by (i). So we just need to prove the other direction. Suppose that $\boldsymbol{\alpha} \in C_{\alpha,S_X}$, for any $\mathbf{X}' \in \text{Conv}(S_X)$, there exist a finite number of points $\mathbf{X}_1,...,\mathbf{X}_r \in S_X$ and $\lambda_1,\ldots,\lambda_r \geq 0$, $\sum_j \lambda_j = 1$, such that $\mathbf{X}' = \sum_j \lambda_j \mathbf{X}_j$. We then have $\boldsymbol{\alpha}^\mathsf{T}\mathbf{X}' = \boldsymbol{\alpha}^\mathsf{T}(\sum_j \lambda_j \mathbf{X}_j) = \sum_j \lambda_j\,(\boldsymbol{\alpha}^\mathsf{T}\mathbf{X}_j) \in C_\rho$, i.e., $\boldsymbol{\alpha} \in C_{\alpha,\text{Conv}(S_X)}$, which holds because $\boldsymbol{\alpha}^\mathsf{T}\mathbf{X}_j \in C_\rho$ for all $j = 1,\ldots,r$, and C_ρ is a convex set.
(iii) $\boldsymbol{\alpha} = \mathbf{0}$ gives $\boldsymbol{\rho} = \mathbf{0}$. This implies the identity correlation matrix, which is in C_ρ.
(iv) Under the further assumption stated in (iv), we can find a set $S_{X_*} = \{\mathbf{X}_1,\ldots,\mathbf{X}_q\} \subseteq S_X$ such that the matrix $\mathbf{X}_* = [\mathbf{X}_1,\ldots,\mathbf{X}_q]$ is non-singular. Suppose that $\boldsymbol{\alpha} \in C_{\alpha,S_{X_*}}$, and let $\boldsymbol{\alpha}^\mathsf{T}\mathbf{X}_* = [\boldsymbol{\rho}_1,\ldots,\boldsymbol{\rho}_q]$. Then $\boldsymbol{\alpha}^\mathsf{T} = [\boldsymbol{\rho}_1,\ldots,\boldsymbol{\rho}_q]\mathbf{X}_*^{-1}$. This is bounded, because all elements of $\boldsymbol{\rho}_1,\ldots,\boldsymbol{\rho}_q$ are bounded (moreover, $\boldsymbol{\rho} \in [-1,1]^L$). Finally, since $S_{X_*} \subseteq S_X$, we have $C_{\alpha,S_X} \subseteq C_{\alpha,S_{X_*}}$ by (ii), and thus C_{α,S_X} is also bounded.
(v) Suppose that $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in C_{\alpha,S_X}$ and that $0 \leq \lambda \leq 1$. Then $(\lambda\,\boldsymbol{\alpha}_1 + (1-\lambda)\,\boldsymbol{\alpha}_2)^\mathsf{T}\mathbf{X} = \lambda\,\boldsymbol{\alpha}_1^\mathsf{T}\mathbf{X} + (1-\lambda)\,\boldsymbol{\alpha}_2^\mathsf{T}\mathbf{X} = \lambda\,\boldsymbol{\rho}_1 + (1-\lambda)\,\boldsymbol{\rho}_2 \in C_\rho$, where the last equation holds since C_ρ is a convex set. Thus $\lambda\,\boldsymbol{\alpha}_1 + (1-\lambda)\,\boldsymbol{\alpha}_2 \in C_{\alpha,S_X}$.

□

Parts (i) and (ii) of Proposition 1 explain how C_{α,S_X} depends on the set S_X of values for $\mathbf{X}$. When step (3) (as discussed in Section 5) is discussed, we know that $\boldsymbol{\alpha} \in C_{\alpha,S_{XT}}$ for all $\boldsymbol{\alpha} \in S_\alpha$. Then also $\boldsymbol{\alpha} \in C_{\alpha,\text{Conv}(S_{XT})}$ by (ii). Even though feasibility was checked only for a finite number of values of $\mathbf{X}$, we thus know that it holds also for the infinite set of their convex hull.

We then need to translate this result for $\mathbf{X}$ back to $\mathbf{Z}$. This is simple if $\mathbf{X} = \mathbf{Z}$, so that we can denote $S_{XT} = S_{ZT}$. Here S_{ZT} should be chosen so that $S_Z \subseteq \text{Conv}(S_{ZT})$, i.e. that its the convex hull covers S_Z. Then, for any $\boldsymbol{\alpha} \in S_\alpha$, we have $\boldsymbol{\alpha} \in C_{\alpha,\text{Conv}(S_{ZT})}$ by (ii) and $\boldsymbol{\alpha} \in C_{\alpha,S_Z}$ by (i), as required. In terms of the possible target sets S_Z defined in Section 5, the test set S_{ZT} could be S_{ZN}, which ensures feasibility also for all $\mathbf{Z}$ in S_{Zh}, or S_{ZT} could consist of the vertices of S_{Zr}, which ensures feasibility in all of S_{Zr}, S_{Zh} and S_{ZN}.

Some more care is needed when $\mathbf{X} = \mathbf{X}(\mathbf{Z})$ includes non-linear functions of $\mathbf{Z}$. If $S_Z = S_{ZN}$ is finite, a simple pragmatic choice is to set $S_{XT} = \mathbf{X}(S_Z) = \{\mathbf{X}(\mathbf{Z}) \mid \mathbf{Z} \in S_Z\}$, and check all of their values. Otherwise, the forms of these functions should be considered. This can be seen even for a single correlation ρ, e.g. for a quadratic model $\rho = \alpha_0 + \alpha_1 Z + \alpha_2 Z^2$ given a single Z. Here $\mathbf{X}(Z) = (X_1, X_2, X_3)^\mathsf{T} = (1, Z, Z^2)^\mathsf{T}$. Suppose that $S_Z = [Z_1, Z_2]$. It is not enough to check just $\mathbf{X}_1 = \mathbf{X}(Z_1)$ and $\mathbf{X}_2 = \mathbf{X}(Z_2)$, because we may still have $\rho \notin (-1, 1)$ for some values between Z_1 and Z_2. It is sufficient to consider one more point $\mathbf{X}_3 = (1, X_{23}, X_{33})^\mathsf{T}$ such that the convex hull of $S_{XT} = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ covers $\mathbf{X}(S_Z)$. For example, this can be the intersection point of the tangents of $f(Z) = Z^2$ drawn at Z_1 and Z_2, i.e. $X_{23} = (Z_1 + Z_2)/2$ and $X_{33} = Z_1 Z_2$. Note that such a choice depends only on the forms of S_Z and $\mathbf{X}(\mathbf{Z})$, so it can be used with any value of $\boldsymbol{\alpha}$ and for any number of correlations ρ_l.

At this point, after the MCMC step (3), we know that $\boldsymbol{\alpha} \in C_{\alpha,\mathbf{X}(S_Z)}$ for all $\boldsymbol{\alpha} \in S_\alpha$, i.e. that all the values in the MCMC sample $S_\alpha = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M\}$ are feasible when combined with any value of $\mathbf{Z}$ in the target set S_Z. But we still need to extend this conclusion to other values of $\boldsymbol{\alpha}$ that were not sampled, specifically to the convex hull $\text{Conv}(S_\alpha)$ of S_α. This is justified by parts (iii)–(v) of Proposition 1, which concern the values of $\boldsymbol{\alpha}$ in C_{α,S_X} given a fixed S_X. Part (iii) shows that this set is non-empty, so some feasible $\boldsymbol{\alpha}$ always exist, and (iv) states that feasible $\boldsymbol{\alpha}$ will not drift away, as long as S_X is not degenerate. Finally, part (v) shows that $\boldsymbol{\alpha} \in C_{\alpha,\mathbf{X}(S_Z)}$ for all $\boldsymbol{\alpha} \in \text{Conv}(S_\alpha)$ as required, thus completing step (4) in Section 5.

We note that part (v) of Proposition 1 would not necessarily hold if the individual correlations were modelled using a nonlinear transformation, for example Fisher's z transformation where $\boldsymbol{\rho} = \tanh(\boldsymbol{\alpha}^\mathsf{T}\mathbf{X})$. Such transformations are used to ensure that the fitted correlations will be in the range $(-1, 1)$. That, however, is not needed here, because positive definiteness of the matrix as a whole already implies that all of the correlations are in the valid range.

C Feasible intervals for individual parameters of the model for correlations — Proof of Proposition 2 discussed in Section 5

The proof of Proposition 2 builds on the key ideas of Barnard, McCulloch and Meng (2000), extended to the case of models with covariates that we consider.

LEMMA 1. *Let $\mathbf{R}(\boldsymbol{\rho}) = \mathbf{R}(\rho_l, \boldsymbol{\rho}_{-l})$ be the positive definite correlation matrix defined by distinct correlations $\boldsymbol{\rho} = (\rho_l, \boldsymbol{\rho}_{-l}^{\mathsf{T}})^{\mathsf{T}}$. Consider $f_l(\rho'_l) = |\mathbf{R}(\rho'_l, \boldsymbol{\rho}_{-l})|$ as a univariate function of $\rho'_l \in [-1, 1]$. Then $f_l(\rho'_l)$ is a quadratic function of ρ'_l with negative second order coefficient. The matrix $\mathbf{R}_l = \mathbf{R}(\rho'_l, \boldsymbol{\rho}_{-l})$ is positive definite if and only if $f_l(\rho'_l) > 0$.*

PROOF OF LEMMA 1. $\mathbf{R}_l$ is a symmetric matrix where ρ'_l appears once in both its upper and lower triangles, so $f_l(\rho'_l)$ is a quadratic function. Suppose that $\mathbf{R}_l$ is a $K \times K$ matrix. Without loss of generality, assume that ρ'_l is in its Kth row, first column (and first row, Kth column), as we can always swap both row and column without changing the positive definiteness and determinant value. Thus, the coefficient of $(\rho'_l)^2$ in $f_l(\rho'_l)$ is $c_l = (-1)^{2K+1}|\mathbf{R}_{(l)}|$, where $\mathbf{R}_{(l)}$ is the submatrix of $\mathbf{R}_l$ obtained by deleting the first and last rows and columns. Here $\mathbf{R}_{(l)}$ is a correlation matrix, obtained by deleting from $\mathbf{R}(\boldsymbol{\rho})$ all those correlations which involve either of the two variables whose correlation is ρ_l. Thus $\mathbf{R}_{(l)}$ is positive definite, $|\mathbf{R}_{(l)}| > 0$, and $c_l < 0$.

$\mathbf{R}_l$ is positive definite if and only if $|\mathbf{R}_{lk}| > 0$ for all $k = 1, \ldots, K$, where $\mathbf{R}_{lk}$ is the kth primary submatrix of $\mathbf{R}_l$ (Sylvester's criterion). Here ρ'_l only affects $|\mathbf{R}_{lK}| = |\mathbf{R}_l|$. Because $\mathbf{R}_{l1}, \ldots, \mathbf{R}_{l,K-1}$ are equal to the corresponding submatrices of the positive definite correlation matrix $\mathbf{R}(\boldsymbol{\rho})$, we have $|\mathbf{R}_{lk}| > 0$, for $k = 1, \ldots, K-1$. So $\mathbf{R}_l = \mathbf{R}(\rho'_l, \boldsymbol{\rho}_{-l})$ is positive definite if and only if $f_l(\rho'_l) = |\mathbf{R}_l| > 0$. □

PROOF OF PROPOSITION 2.

From Lemma 1 we know that $\mathbf{R}_{jl} = \mathbf{R}(\rho'_l, \boldsymbol{\rho}_{-l}^{(j)})$ is positive definite if and only if $f_{jl}(\rho'_l) = |\mathbf{R}_{jl}| > 0$. We can write $f_{jl}(\rho'_l) = c_{jl}(\rho'_l)^2 + d_{jl}\rho'_l + e_{jl}$, where $c_{jl} = [f_{jl}(1) + f_{jl}(-1) - 2f_{jl}(0)]/2$, $d_{jl} = [f_{jl}(1) - f_{jl}(-1)]/2$ and $e_{jl} = f_{jl}(0)$. The set of values for ρ'_l for which $f_{jl}(\rho'_l) > 0$ is a finite interval because $c_{jl} < 0$, $f_{jl}(0) = |\mathbf{R}(0, \boldsymbol{\rho}_{-l})| > 0$, and $f_{jl}(\rho'_l)$ is a continuous function. Let us denote the roots of $f_{jl}(\rho'_l) = 0$ by $x_{jl1} > x_{jl2}$, and define

$$\begin{aligned} g_{jl} &= \frac{x_{jl1} + x_{jl2}}{2} = -\frac{d_{jl}}{2c_{jl}}, \\ h_{jl} &= \frac{x_{jl1} - x_{jl2}}{2} = \sqrt{\frac{d_{jl}^2 - 4c_{jl}e_{jl}}{4c_{jl}^2}}. \end{aligned} \tag{C1}$$

$\mathbf{R}_{jl}$ is positive definite when $\rho'_l \in (g_{jl} - h_{jl}, g_{jl} + h_{jl})$.

Consider now $\boldsymbol{\rho}_j = \boldsymbol{\alpha}^{\mathsf{T}}\mathbf{X}_j$ as specified by model (9), as functions of coefficients $\boldsymbol{\alpha}$ and covariates $\mathbf{X}_j$. Consider $\rho'_l = \alpha'_{lm}X_{jm} + \sum_{k \neq m}\alpha_{lk}X_{jk}$ as implied by this model, treating α'_{lm} for a single $m = 1, \ldots, q$ as the argument of the function and fixing all the other elements of $\boldsymbol{\alpha}$ and $\mathbf{X}_j$ at the values which defined $\boldsymbol{\rho}_j$. Solving the end points of the feasible interval of ρ'_l for α'_{lm}, and taking into account the sign of X_{jm} gives the feasible interval for α'_{lm} with end points $a_{lm}^{(j)}$ and $b_{lm}^{(j)}$ as shown in equation (10) in Proposition 2, when $X_{jm} \neq 0$. When $X_{jm} = 0$, $f_{jl}(\rho'_l)$ does not depend on α'_{lm} and the interval can be taken to be infinite. The interval for α_{lm} which is feasible for all of the $\mathbf{X}_j \in S_{XT}$ is then $(a_{lm}, b_{lm}) = \cap_j(a_{lm}^{(j)}, b_{lm}^{(j)})$. □

Computationally the most demanding part of using this result is the calculation of the necessary determinants. Efficient methods for obtaining them, and other elements of the computations, are described in Section 6.2 and Appendix E.

D MCMC algorithm for the latent variables and for the parameters except for the correlation parameters $\boldsymbol{\alpha}$

Here we describe the MCMC sampling algorithm for estimating the structural-model parameters of the model which was introduced in Section 3 of the paper. The general idea of this estimation was outlined in Section 6.2. As discussed there, the steps for the other elements of the model other than the parameters $\boldsymbol{\alpha}$ of the correlation model (which is considered separately in Section 6.2 and Appendix E) are the same or very similar to the ones proposed in Kuha, Zhang and Steele (2022). Their details are given here in order to keep the description of the full MCMC algorithm self-contained in this paper.

The algorithm has been packed into an R (R Core Team, 2020) package [which will be included in the supplementary materials and made available open source on an author's GitHub page]. The algorithm was programmed in R with core functions implemented in C++, where two techniques are used to speed up the procedure. First, for sampling steps with non-standard distributions, adaptive rejection sampling (Gilks, Best and Tan, 1995) is used, exploiting log-concavity of the posterior density functions. Second, parallel sampling is used within each MCMC iteration where possible. The parallelization is implemented through the OpenMP C++ API (Dagum and Menon, 1998).

Different elements of $\boldsymbol{\zeta}$ and $\boldsymbol{\psi}$ are sampled one at a time, as scalars or vectors as appropriate. In the notation below, those quantities that are not being sampled in a given step are taken to be observed and fixed at their most recently sampled values.

Sampling the latent variables: Generate values for the latent variables $\boldsymbol{\zeta}_i = (\boldsymbol{\xi}_i^\mathsf{T}, \boldsymbol{\eta}_i^\mathsf{T})^\mathsf{T}$, given the observed data and current values of the parameters $\boldsymbol{\psi}$. This can be parallelised, because $\boldsymbol{\zeta}_i$ for different units i are conditionally independent.

(1) Sampling $\boldsymbol{\xi}$ from $p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi})$: Draw $\boldsymbol{\xi}_i = (\xi_{Gi}, \xi_{Ri})^\mathsf{T}$ independently for $i = 1, \ldots, n$, from multinomial distributions with probabilities

$$\text{(D1)}\qquad \begin{aligned} &p(\xi_G = j, \xi_R = k|\boldsymbol{\eta}, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}) \\ &\propto p(\mathbf{Y}_{Gi}|\xi_G = j, \eta_{Gi})\, p(\mathbf{Y}_{Ri}|\xi_R = k, \eta_{Ri})\, p(\xi_G = j, \xi_R = k|\mathbf{X}_i; \boldsymbol{\psi}_\xi) \end{aligned}$$

for $j, k = 0, 1$, where the measurement model is specified by equations (1)–(4) in the paper for $\mathbf{Y}_{Gi}$ and similarly for $\mathbf{Y}_{Ri}$, and the structural model for $\boldsymbol{\xi}_i$ is specified by (5).

(2) Sampling $\boldsymbol{\eta}$ from $p(\boldsymbol{\eta}|\boldsymbol{\xi}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi})$: Draw $\boldsymbol{\eta}_i = (\eta_{GPi}, \eta_{RPi}, \eta_{GFi}, \eta_{RFi})^\mathsf{T}$ independently for $i = 1, \ldots, n$, from

$$\text{(D2)}\qquad p(\eta_{GP}|\boldsymbol{\eta}_{-GPi}, \boldsymbol{\xi}_i, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}) \propto p(\mathbf{Y}_{GPi}|\xi_{Gi}, \eta_{GP})\, p(\eta_{GP}|\boldsymbol{\eta}_{-GPi}, \mathbf{X}_i; \boldsymbol{\psi}_\eta)$$

$$\text{(D3)}\qquad p(\eta_{GF}|\boldsymbol{\eta}_{-GFi}, \boldsymbol{\xi}_i, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}) \propto p(Y_{GFi}|\xi_{Gi}, \eta_{GF})\, p(\eta_{GF}|\boldsymbol{\eta}_{-GFi}, \mathbf{X}_i; \boldsymbol{\psi}_\eta)$$

$$\text{(D4)}\qquad p(\eta_{RP}|\boldsymbol{\eta}_{-RPi}, \boldsymbol{\xi}_i, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}) \propto p(\mathbf{Y}_{RPi}|\xi_{Ri}, \eta_{RP})\, p(\eta_{RP}|\boldsymbol{\eta}_{-RPi}, \mathbf{X}_i; \boldsymbol{\psi}_\eta)$$

$$\text{(D5)}\qquad p(\eta_{RF}|\boldsymbol{\eta}_{-RFi}, \boldsymbol{\xi}_i, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}) \propto p(Y_{RFi}|\xi_{Ri}, \eta_{RF})\, p(\eta_{RF}|\boldsymbol{\eta}_{-RFi}, \mathbf{X}_i; \boldsymbol{\psi}_\eta).$$

Here $\boldsymbol{\eta}_{-GPi}$ denotes $(\eta_{GFi}, \eta_{RPi}, \eta_{RFi})$ and $\boldsymbol{\eta}_{-GFi}$, $\boldsymbol{\eta}_{-RPi}$ and $\boldsymbol{\eta}_{-RFi}$ are defined similarly. The conditional distributions for the η-variables on the right hand sides of (D2)–(D5) are the univariate conditional normal distributions implied by the joint normal distribution given by (6)–(8) in the paper. The sampling distributions depend on the values of the ξ-variables. When $\xi_{Gi} = 0$, in which case always $\mathbf{Y}_{Gi} = \mathbf{0}$, we have $p(\mathbf{Y}_{GPi}|\xi_{Gi}, \eta_{GP}) = p(Y_{GFi}|\xi_{Gi}, \eta_{GF}) = 1$ and η_{GPi} and η_{GFi} are drawn directly from the conditional normal distributions. When $\xi_{Gi} = 1$, adaptive rejection sampling is used for η_{GPi} and truncated normal sampling for η_{GFi}. The sampling of η_{RPi} and η_{RFi} is analogous.

Sampling the parameters of the structural model: Generate values for the parameters $\boldsymbol{\psi}$ from their distributions given the observed variables and current imputed values of the latent variables $\boldsymbol{\zeta}$. These have the form of posterior distributions of these structural parameters when both $\boldsymbol{\zeta}$ and $\mathbf{X}$ are taken to be observed data (this step does not depend on $\mathbf{Y}$). The prior distributions are taken to be of the form $p(\boldsymbol{\psi}) = p(\boldsymbol{\psi}_\xi)p(\boldsymbol{\beta})p(\boldsymbol{\sigma})p(\boldsymbol{\alpha})$, i.e. independent for different blocks of parameters; their specific forms are given below. The sampling steps for $\boldsymbol{\psi}_\xi$ and $\boldsymbol{\psi}_\eta$ do not depend on each other, so they can be carried out in either order or in parallel.

(3) Sampling $\boldsymbol{\psi}_\xi = (\boldsymbol{\gamma}_{01}^\mathsf{T}, \boldsymbol{\gamma}_{10}^\mathsf{T}, \boldsymbol{\gamma}_{11}^\mathsf{T})^\mathsf{T}$ from $p(\boldsymbol{\psi}_\xi|\mathbf{X}, \boldsymbol{\xi}) \propto p(\boldsymbol{\xi}|\mathbf{X}; \boldsymbol{\psi}_\xi)\, p(\boldsymbol{\psi}_\xi)$. This is the posterior distribution of the coefficients of the multinomial logistic model (5) for $\boldsymbol{\xi}_i$ given $\mathbf{X}_i$. Define $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_{00}^\mathsf{T}, \boldsymbol{\gamma}_{01}^\mathsf{T}, \boldsymbol{\gamma}_{10}^\mathsf{T}, \boldsymbol{\gamma}_{11}^\mathsf{T})^\mathsf{T}$, where $\boldsymbol{\gamma}_{00} = \mathbf{0}$. Let γ_{jkr} demote the coefficient of X_{jkr} in the model for $p(\xi_\chi = j, \xi_{Ri} = k|\mathbf{X}_i; \boldsymbol{\psi}_\xi)$, and $\boldsymbol{\gamma}_{-jkr}$ denote the vector obtained by omitting γ_{jkr} from $\boldsymbol{\gamma}$. We take the prior distributions of each non-zero γ_{jkr} to be independent of each other, with $p(\gamma_{jkr}) \sim N(0, \sigma_\gamma^2)$ with $\sigma_\gamma^2 = 100$. The sampling is done using conditional Gibbs sampling, one parameter at a time. We cycle over all $r = 1 \ldots, Q$ and over $(j,k) = (0,1),\ (1,0),\ (1,1)$ to draw γ_{jkr} from

(D6) $$p(\gamma_{jkr}|\boldsymbol{\gamma}_{-jkr}, \mathbf{X}, \boldsymbol{\xi}) \propto \left[\prod_{i=1}^n \frac{\prod_{r,s=0,1} \exp(\boldsymbol{\gamma}_{rs}^\mathsf{T}\mathbf{X}_i)^{\delta_{ijk}}}{\sum_{r,s=0,1} \exp(\boldsymbol{\gamma}_{rs}^\mathsf{T}\mathbf{X}_i)}\right] p(\gamma_{jkr})$$

where $\delta_{ijk} = \mathbb{1}(\xi_{Gi} = j, \xi_{Ri} = k)$. These are sampled using adaptive rejection sampling.

(4) Sampling $\boldsymbol{\psi}_\eta = (\mathrm{vec}(\boldsymbol{\beta})^\mathsf{T}, \boldsymbol{\sigma}^\mathsf{T}, \mathrm{vec}(\boldsymbol{\alpha})^\mathsf{T})^\mathsf{T}$ from $p(\boldsymbol{\psi}_\eta|\mathbf{X}, \boldsymbol{\eta}) \propto p(\boldsymbol{\eta}|\mathbf{X}; \boldsymbol{\psi}_\eta)\, p(\boldsymbol{\psi}_\eta)$. Here the sampling of $\boldsymbol{\alpha}$ will be described in Appendix E below. For $\boldsymbol{\beta}$, the sampling is from the posterior distribution $p(\mathrm{vec}(\boldsymbol{\beta})|\mathbf{X}, \boldsymbol{\eta}) \propto p(\boldsymbol{\eta}|\mathbf{X}; \boldsymbol{\psi}_\eta)\, p(\mathrm{vec}(\boldsymbol{\beta}))$ where $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ are regarded as known. This means that the conditional covariance matrices $\boldsymbol{\Sigma}_i = \mathrm{cov}(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\sigma}, \boldsymbol{\alpha})$ are also known here. We specify $p(\mathrm{vec}(\boldsymbol{\beta})) \sim N(\mathbf{0}, \sigma_\beta^2\, \mathbf{I}_{4Q})$ with $\sigma_\beta^2 = 100$. The sampling is done separately for each of the four subvectors of $\boldsymbol{\beta}$. Let $\boldsymbol{\beta}_1$ denote one of them, say $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_{GP}$, and $\boldsymbol{\beta}_2$ the rest of them, say $\boldsymbol{\beta}_2 = [\boldsymbol{\beta}_{RP}, \boldsymbol{\beta}_{GF}, \boldsymbol{\beta}_{RF}]$, and let $\boldsymbol{\psi}_{\eta(\beta_1)}$ denote all the elements of $\boldsymbol{\psi}_\eta$ other than $\boldsymbol{\beta}_1$. Let $\boldsymbol{\eta}_i$ be partitioned correspondingly into η_{1i} and $\boldsymbol{\eta}_{2i}$, and $\boldsymbol{\Sigma}_i$ into the blocks Σ_{11i}, $\boldsymbol{\Sigma}_{12i}$ and $\boldsymbol{\Sigma}_{22i}$. The conditional distribution $p(\eta_{1i}|\boldsymbol{\eta}_{2i}, \mathbf{X}_i; \boldsymbol{\psi}_\eta)$ is then univariate normal with mean $\boldsymbol{\beta}_1^\mathsf{T}\mathbf{X}_i + d_{2i}$, where $d_{2i} = \boldsymbol{\Sigma}_{12i}\boldsymbol{\Sigma}_{22i}^{-1}(\boldsymbol{\eta}_{2i} - \boldsymbol{\beta}_2^\mathsf{T}\mathbf{X}_i)$, and variance $\sigma_{1i}^2 = \Sigma_{11i} - \boldsymbol{\Sigma}_{12i}\boldsymbol{\Sigma}_{22i}^{-1}\boldsymbol{\Sigma}_{12i}^\mathsf{T}$. Let $\mathbf{V}_1 = \mathrm{diag}(\sigma_{1i}^2, \ldots, \sigma_{1n}^2)$ and $\mathbf{e}_1 = (\eta_{11} - d_{21}, \ldots, \eta_{1n} - d_{2n})^\mathsf{T}$. The value of $\boldsymbol{\beta}_1$ is then sampled from $p(\boldsymbol{\beta}_1|\mathbf{X}, \boldsymbol{\eta}, \boldsymbol{\psi}_{\eta(\beta_1)}) \sim N(\mathbf{V}_{\beta_1}(\mathbf{X}^\mathsf{T}\mathbf{V}_1^{-1}\mathbf{e}_1),\, \mathbf{V}_{\beta_1})$ where $\mathbf{V}_{\beta_1} = (\mathbf{X}^\mathsf{T}\mathbf{V}_1^{-1}\mathbf{X} + \mathbf{I}_Q/\sigma_\beta^2)^{-1}$. This is repeated with each of the four subvectors of $\boldsymbol{\beta}$ in turn in the role of $\boldsymbol{\beta}_1$.

For sampling of the standard deviation parameters $\boldsymbol{\sigma}$, denote here $\sigma_1 = \sigma_{GP}$ and $\sigma_2 = \sigma_{RP}$. For both of them we use the prior distribution Inv-Gamma(α_0, β_0) with $\alpha_0 = \beta_0 = 10^{-5}$, independently for σ_1^2 and σ_2^2. This implies the priors $p(\sigma_k) \propto \sigma_k^{-2\alpha_0 - 1}\exp(\beta_0/\sigma_k^2)$ for $k = 1, 2$. Denote by $\boldsymbol{\psi}_{\eta(\sigma)}$ all other parameters in $\boldsymbol{\psi}_\eta$ apart from σ_k. Recall that this means that in $\boldsymbol{\Sigma}_i = \mathbf{S}\,\mathbf{R}_i\,\mathbf{S}$, where $\mathbf{S} = \mathrm{diag}(\sigma_1, \sigma_2, 1, 1)$, the correlation matrix $\mathbf{R}_i = \mathbf{R}(\mathbf{X}_i; \boldsymbol{\alpha})$ is also treated as known here. Let $\mathbf{e}_i = (e_{i1}, e_{i2}, e_{i3}, e_{i4})^\mathsf{T} = \boldsymbol{\eta}_i - \boldsymbol{\beta}^\mathsf{T}\mathbf{X}_i$. The parameter σ_k is then drawn from

(D7)
$$\begin{aligned} p(\sigma_k|\mathbf{X}, \boldsymbol{\eta}, \boldsymbol{\psi}_{\eta(\sigma)}) &\propto \prod_{i=1}^n p(\boldsymbol{\eta}_i|\mathbf{X}_i; \boldsymbol{\psi}_\eta)p(\sigma_k) \propto \prod_{i=1}^n \sigma_k^{-1}\exp\left(-\frac{1}{2}\mathbf{e}_i^\mathsf{T}\,\boldsymbol{\Sigma}_i^{-1}\,\mathbf{e}_i\right)p(\sigma_k) \\ &\propto \sigma_k^{-\alpha-1}\exp\left(-\beta_1/\sigma_k^2 - 2\beta_2/\sigma_k\right), \end{aligned}$$

where $\alpha = n + 2\alpha_0$, $\beta_1 = \beta_0 + (\sum_{i=1}^n e_{ik}^2 w_{kki})/2$, $\beta_2 = \sum_{i=1}^n e_{ik}(\sum_{j\neq k} w_{kji}e_{ij}/\sigma_j)/2$, and w_{kji} is the (k,j)th element of $\mathbf{R}_i^{-1}$. Then random-walk Metropolis sampler or the adaptive rejection Metropolis sampler (ARMS, Gilks, Best and Tan, 1995) can be used to sample σ_1 and σ_2.

E Further information on the MCMC algorithm for the correlation parameters α discussed in Section 6.2 — Remarks 1–4

Remark 1: *Calculating feasible interval for α_{lm} by Cholesky decomposition.*

Proposition 2 (in Section 5) and Lemma 1 (in Appendix D) descibe one way of calculating the interval (a_{lm}, b_{lm}). This requires the calculation of $3T$ determinants of correlation matrices. An alternative, more efficient procedure for its first steps can be obtained by adapting a method proposed by Wong, Carter and Kohn (2003). Let $\mathbf{R}_j = \mathbf{\Gamma}_j^{\mathsf{T}}\mathbf{\Gamma}_j$ as defined in the Input statement of Algorithm 1, where $\mathbf{\Gamma}_j = (\gamma_{k_1,k_2}^{(j)})$. Recall that K denotes the dimension of $\mathbf{R}_j$, and assume that ρ_l in Lemma 1 corresponds to the $(K, K-1)$th element of $\mathbf{R}_j$. We then have $g_{jl} = \sum_{k=1}^{K-2}\gamma_{k,K-1}^{(j)}\gamma_{k,K}^{(j)}$ and $h_{jl} = \gamma_{K-1,K-1}^{(j)}(1-\sum_{k=1}^{K-2}(\gamma_{k,K}^{(j)})^2)^{1/2}$, and (a_{lm}, b_{lm}) can be obtained by plugging in g_{jl} and h_{jl} into equation (10) in Proposition 2 as before. If ρ_l is not the $(K, K-1)$th element of $\mathbf{R}_j$, we can permute the indices with a permutation matrix $\mathbf{P}$ so that it is the $(K, K-1)$th element of the matrix $\mathbf{P}^{\mathsf{T}}\mathbf{R}_j\mathbf{P} = (\mathbf{\Gamma}_j\mathbf{P})^{\mathsf{T}}(\mathbf{\Gamma}_j\mathbf{P})$, followed by a Givens rotation by an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{\Gamma}_j\mathbf{P} = \tilde{\mathbf{\Gamma}}_j$, where $\tilde{\mathbf{\Gamma}}_j$ is upper-triangular and $\mathbf{P}^{\mathsf{T}}\mathbf{R}_j\mathbf{P} = \tilde{\mathbf{\Gamma}}_j^{\mathsf{T}}\tilde{\mathbf{\Gamma}}_j$. Then apply the calculation above to $\tilde{\mathbf{\Gamma}}_j$.

Remark 2: *Generating proposal values for α_{lm}.*

We have used a simple random walk Metropolis sampler. It generates the proposal through an independent Gaussian increment to the previous value, as $\alpha'_{lm} = \alpha_{lm} + \gamma_m\delta$, where δ is drawn from the standard normal distribution. Thus $\alpha_{lm}|\alpha'_{lm} \sim N(\alpha'_{lm}, \gamma_m^2)$. The step size γ_m should be chosen to achieve a good balance between rejection rate and mixing efficiency. We have used $\gamma_m = C(\sqrt{n}||(X_{1m}, \ldots, X_{nm})^{\mathsf{T}}||_\infty)^{-1}$, where $||\cdot||_\infty$ is the infinity norm and C is a chosen constant, the same for all γ_m, which is used to control rejection rates in the range 0.7–0.8. The sampler was efficient enough in our real data analysis when the step sizes were chosen appropriately.

An alternative would be to use the ARMS algorithm (Gilks, Best and Tan, 1995) to adaptively construct the proposal function of α_{lm} in (a_{lm}, b_{lm}). This can improve the acceptance rate but the algorithm may require the likelihood function $p(\boldsymbol{\epsilon}|\boldsymbol{\alpha}, \mathbf{X})$ to be evaluated multiple times based on the rejection condition, whereas in the random walk Metropolis method it needs to be calculated at most once in each iteration. Other methods for improving the acceptance rate exist (Chib and Greenberg, 1998), but their implementation is more complex and relies heavily on tuning.

Remark 3: *Updating the determinant and inverse of correlation matrix.*

Here we want to update the determinant and inverse of $\mathbf{R}_i(\alpha_{lm})$ to those of $\mathbf{R}_i(\alpha'_{lm})$. Suppose that the correlation parameter ρ_l corresponds to the (k_1, k_2)th element of $\mathbf{R}_i(\alpha_{lm})$. Let $\varepsilon_{ilm} = (\alpha'_{lm} - \alpha_{lm})X_{im}$, and denote by $\mathbf{w}_1$ and $\mathbf{w}_2$ the $K \times 1$ vectors which are zero except that the k_1th element of $\mathbf{w}_1$ and the k_2th element of $\mathbf{w}_2$ are $\sqrt{|\varepsilon_{ilm}|}$. Then

$$\mathbf{R}_i(\alpha'_{lm}) = \Big[\mathbf{R}_i(\alpha_{lm}) + (\mathrm{sgn}(\varepsilon_{ilm})\mathbf{w}_1)\mathbf{w}_2^{\mathsf{T}}\Big] + \mathbf{w}_2(\mathrm{sgn}(\varepsilon_{ilm})\mathbf{w}_1)^{\mathsf{T}}.$$

Since this is of the form $(\mathbf{A}+\mathbf{u}\mathbf{v}^{\mathsf{T}})+\mathbf{v}\mathbf{u}^{\mathsf{T}}$, $\mathbf{R}_i(\alpha'_{lm})^{-1}$ can be computed efficiently with two rank-1 updates by applying twice the Sherman-Morrison formula

$$(\mathbf{A}+\mathbf{u}\mathbf{v}^{\mathsf{T}})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^{\mathsf{T}}\mathbf{A}^{-1}}{1+\mathbf{v}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{u}}$$

and $|\mathbf{R}_i(\alpha'_{lm})|$ can be calculated by updating $|\mathbf{R}_i(\alpha_{lm})|$ through two applications of $|\mathbf{A} + \mathbf{u}\mathbf{v}^\mathsf{T}| = (1 + \mathbf{v}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u})\,|\mathbf{A}|$, the second of which employs the first update of the inverse. These steps reduce the computation complexity of $p(\boldsymbol{\epsilon}_i|\boldsymbol{\alpha}, \mathbf{X})$ from $O(K^3)$ to $O(K^2)$.

Remark 4: *Updating the Cholesky decomposition of a correlation matrix*.

Let $\varepsilon_{jlm} = (\alpha'_{lm} - \alpha_{lm})X_{jm}$. Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be defined as in Remark 3, and define $\mathbf{w}$ as the $K \times 1$ vector where the k_1th and k_2th elements are $\sqrt{|\varepsilon_{jlm}|}$ and the other elements are zero. Then we can write

$$\mathbf{R}_j(\alpha'_{lm}) = \left[\left(\mathbf{R}_j(\alpha_{lm}) + \text{sgn}(\varepsilon_{jlm})\mathbf{w}\mathbf{w}^\mathsf{T}\right) - \text{sgn}(\varepsilon_{jlm})\mathbf{w}_1\mathbf{w}_1^\mathsf{T}\right] - \text{sgn}(\varepsilon_{jlm})\mathbf{w}_2\mathbf{w}_2^\mathsf{T}.$$

The Cholesky decomposition of $\mathbf{R}_j(\alpha'_{lm})$ can be computed efficiently from this, with three rank-1 updates for the Cholesky decomposition of the form $\mathbf{A} + \mathbf{u}\mathbf{u}^\mathsf{T}$ or $\mathbf{A} - \mathbf{u}\mathbf{u}^\mathsf{T}$(Seeger, 2008); built-in functions for this are available in Matlab and linear algebra libraries like Eigen (Guennebaud et al., 2010). This updating rule reduces the computation complexity of the Cholesky decomposition $\mathbf{R}_j = \boldsymbol{\Gamma}_j^\mathsf{T}\boldsymbol{\Gamma}_j$ from $O(K^3)$ to $O(K^2)$.

Alternatives to Algorithm 1 in Section 6.2 could also be considered. In cases when $\mathbf{X} = \mathbf{X}(\mathbf{Z})$ is a complex function such as a cubic spline, for better efficiency the element-wise MH algorithm could be replaced with a blockwise algorithm where subvectors of $\boldsymbol{\alpha}$ can be proposed and rejected together. Apart from the MH algorithm we use in this paper, we note that the "griddy Gibbs" sampler discussed in Barnard, McCulloch and Meng (2000) also works here in principle, where the feasible intervals for each α_{lm} can be discretized into grids. However, the computational efficiency for evaluating posterior function over these grids may suffer.

F Estimated measurement models for the analysis in Section 7

TABLE F1

Estimated parameters (measurement loadings and intercepts) of the measurement models for survey items on help given by respondents to their parents and on help received from the parents.

	Giving practical help		Receiving practical help	
Item	loading	intercept	loading	intercept
Lifts in car	1.12	0.83	1.14	1.54
Shopping	2.38	1.02	1.70	2.08
Providing or cooking meals	1.24	-0.28	1.15	1.57
Basic personal needs (to parent only)	1.32	-1.32	–	–
Looking after children (from parents only)	–	–	0.89	2.25
Washing, ironing or cleaning	1.32	-0.77	1.15	0.82
Personal affairs	1.00	0.00	1.00	0.00
Decorating, gardening or house repairs	0.57	-0.22	0.74	0.37

G Estimated multinomial logistic model for the latent class variables ξ in the analysis in Section 7

TABLE G1

Estimated coefficients of the multinomial logistic model for the zero-inflation latent classes (ξ_G, ξ_R). The coefficients $\boldsymbol{\gamma}_{00}$ are fixed at $\mathbf{0}$ for identification. The estimates are posterior means from MCMC samples (with posterior standard deviations in parentheses).

Covariate	$\gamma_{jk}(\xi_G = j, \xi_R = k)$ γ_{01}		γ_{10}		γ_{11}	
Intercept	-3.98^{***}	(0.76)	-1.73^{*}	(1.08)	1.50^{***}	(0.52)
Respondent (child) characteristics						
Age (centered at 40) ($\times 10$)	-0.13	(0.21)	-0.44^{*}	(0.24)	-0.44^{***}	(0.09)
Age squared ($\times 10^3$)	-2.63	(1.72)	1.34	(0.93)	1.54^{***}	(0.46)
Gender						
Female (vs. Male)	1.47^{***}	(0.30)	0.21	(0.17)	0.06	(0.09)
Partnership status						
Partnered (vs. Single)	0.13	(0.22)	0.71^{***}	(0.20)	-0.05	(0.10)
Age of youngest coresident child (vs. No children):						
0–1 years	0.51	(0.35)	-0.60	(0.53)	-0.01	(0.17)
2–4 years	0.50	(0.32)	0.10	(0.40)	0.45^{***}	(0.16)
5–10 years	0.43^{*}	(0.26)	-1.69^{***}	(0.66)	0.16	(0.14)
11–16 years	-0.59^{*}	(0.30)	-0.03	(0.22)	-0.32^{**}	(0.14)
17– years	-0.24	(0.39)	-0.20	(0.24)	-0.11	(0.17)
Number of siblings (vs. None)						
1	-0.11	(0.28)	-0.49^{**}	(0.23)	0.02	(0.15)
2	-0.70^{**}	(0.27)	-0.41^{*}	(0.22)	-0.23	(0.15)
Long standing illness (vs. No)	0.00	(0.23)	-0.37^{*}	(0.21)	-0.27^{**}	(0.11)
Employment status (vs. Employed)						
Not employed	-0.25	(0.21)	0.36^{*}	(0.19)	-0.34^{***}	(0.10)
Education (vs. Secondary or less)						
Post-secondary	0.63^{***}	(0.18)	-0.05	(0.16)	0.20^{**}	(0.09)
Household tenure (vs. Renter)						
Own home outright or by mortgage	-0.23	(0.21)	0.36^{*}	(0.22)	0.06	(0.10)
Logarithm of household equivalised income	0.37^{***}	(0.09)	0.01	(0.11)	-0.01	(0.05)
Parent characteristics						
Age of the oldest living parent (centered at 70) ($\times 10$)	0.37^{*}	(0.20)	1.58^{***}	(0.42)	0.17^{**}	(0.07)
Squared Age of the oldest parent ($\times 10^3$)	-12.72^{***}	(2.41)	-2.91^{**}	(1.50)	-0.37	(0.37)
At least one parent lives alone (vs. No)	-0.84^{***}	(0.21)	1.13^{***}	(0.17)	0.25^{***}	(0.09)
Child-parent characteristics						
Travel time to the nearest parent						
More than 1 hour (vs. 1 hour or less)	-1.65^{***}	(0.24)	-1.73^{***}	(0.20)	-1.99^{***}	(0.10)

The posterior credible interval excludes zero at level 90% (*), 95% (**) or 99% (***).

TABLE G2

Fitted membership probabilities of the zero-inflation latent classes (ξ_G, ξ_R), from the estimated model in Table S2. The fitted probabilities are averaged over parameter values in MCMC samples and over covariate values in the observed sample (for all covariates for the "Overall" figures, and for all but the fixed covariate for the rest. The odds ratios (OR) calculated from these averages are also shown.

Covariate setting	$p(\xi_G = j, \xi_R = k)$ (0,0)	(0,1)	(1,0)	(1,1)	OR	Marginal probabilities [with difference (and its SD)] $p(\xi_G = 1)$			$p(\xi_R = 1)$		
Overall	.24	.09	.14	.53	10.6	.67			.62		
Respondent (child) characteristics											
Age											
35 years	.22	.09	.16	.54	9.1	.70			.63		
45 years	.28	.10	.14	.48	9.5	.62	$-.07^{***}$	(.02)	.58	$-.05^{**}$	(.02)
55 years	.31	.07	.14	.48	18.2	.62	$-.07^{**}$	(.03)	.55	$-.08^{*}$	(.04)
Gender											
Female	.23	.13	.14	.50	6.5	.65			.63		
Male	.26	.04	.13	.56	30.0	.70	$+.05^{***}$	(.02)	.60	$-.03$	(.02)
Partnership status											
Single	.25	.08	.09	.57	19.7	.66			.65		
Partnered	.24	.09	.15	.52	9.1	.67	$+.01$	(.02)	.61	$-.05^{**}$	(.02)
Age of youngest coresident child											
No children	.24	.08	.15	.52	10.5	.68	$+.05$	(.03)	.60	$-.05$	(.04)
0-1 years	.25	.12	.10	.53	13.1	.63			.65		
2-4 years	.19	.10	.13	.59	9.7	.72	$+.09^{***}$	(.03)	.68	$+.03$	(.04)
5-10 years	.24	.11	.04	.60	51.0	.64	$+.02$	(.03)	.71	$+.06^{*}$	(.04)
11-16 years	.29	.06	.18	.48	13.8	.65	$+.02$	(.03)	.54	$-.11^{***}$	(.04)
17– years	.26	.07	.14	.52	14.6	.66	$+.03$	(.04)	.59	$-.06$	(.05)
Number of siblings											
No sibling	.21	.11	.17	.51	6.3	.68			.62		
1 sibling	.23	.10	.12	.55	10.3	.67	$-.01$	(.02)	.65	$+.03$	(.03)
2 or more	.26	.07	.14	.52	13.0	.66	$-.01$	(.02)	.59	$-.03$	(.03)
Longstanding illness											
Yes	.28	.10	.13	.49	11.4	.62			.60		
No	.24	.09	.14	.53	10.4	.67	$+.05^{***}$	(.02)	.62	$+.02$	(.02)
Employment status											
Not employed	.27	.09	.18	.46	8.3	.65			.55		
Employed	.24	.09	.13	.55	11.8	.67	$+.03$	(.02)	.64	$+.09^{***}$	(.02)
Education											
Secondary or less	.26	.07	.15	.52	13.4	.67			.59		
Post-secondary	.23	.10	.13	.54	9.6	.67	$-.00$	(.01)	.64	$+.05^{***}$	(.02)
Household tenure											
Own home	.24	.08	.15	.53	10.7	.68			.61		
Renter	.25	.10	.12	.53	11.5	.64	$-.03^{*}$	(.02)	.63	$+.02$	(.02)
Logarithm of household equivalised income											
25 percentile	.25	.08	.14	.53	12.0	.67			.61		
50 percentile	.24	.09	.14	.53	10.7	.67	$-.01^{**}$	(.00)	.62	$+.00$	(.00)
75 percentile	.24	.10	.14	.52	9.6	.66	$-.01^{**}$	(.01)	.62	$+.00$	(.01)
Parent characteristics											
Age of the oldest living parent											
65 years	.29	.11	.04	.56	46.3	.60			.67		
70 years	.25	.15	.08	.52	12.6	.60	.00	(.01)	.67	$+.00$	(.01)
80 years	.22	.06	.20	.52	9.7	.72	$+.12^{***}$	(.02)	.58	$-.10^{***}$	(.03)
At least one parent lives alone											
Yes	.22	.05	.19	.55	14.5	.74			.59		
No	.27	.11	.09	.53	14.0	.62	$-.11^{***}$	(.01)	.64	$+.05^{**}$	(.02)
Child-parent characteristics											
Travel time to the nearest parent											
$>$ 1 hour	.51	.07	.11	.31	22.7	.42			.38		
$\leq$ 1 hour	.15	.10	.15	.61	6.2	.76	$+.34^{***}$	(.02)	.70	$+.32^{***}$	(.02)

The posterior credible interval excludes zero at level 90% (*), 95% (**) or 99% (***).

REFERENCES

BARNARD, J., MCCULLOCH, R. and MENG, X.-L. (2000). Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, with Application to Shrinkage. *Statistica Sinica* **10** 1281–1311.

CHIB, S. and GREENBERG, E. (1998). Analysis of Multivariate Probit Models. *Biometrika* **85** 347–361.

DAGUM, L. and MENON, R. (1998). OpenMP: An industry standard API for shared-memory programming. *Computational Science & Engineering, IEEE* **5** 46–55.

GILKS, W. R., BEST, N. G. and TAN, K. K. C. (1995). Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **44** 455–472.

GUENNEBAUD, G., JACOB, B. et al. (2010). Eigen v3. http://eigen.tuxfamily.org.

KUHA, J., ZHANG, S. and STEELE, F. (2022). Latent variable models for multivariate dyadic data with zero inflation: Analysis of intergenerational exchanges of family support. *Annals of Applied Statistics*. To appear.

SEEGER, M. (2008). Low Rank Updates for the Cholesky Decomposition Technical Report, University of California at Berkeley.

R CORE TEAM (2020). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

WONG, F., CARTER, C. K. and KOHN, R. (2003). Efficient Estimation of Covariance Selection Models. *Biometrika* **90** 809–830.