LATENT VARIABLE MODELS FOR MULTIVARIATE DYADIC DATA WITH ZERO INFLATION: ANALYSIS OF INTERGENERATIONAL EXCHANGES OF FAMILY SUPPORT

BY JOUNI KUHA[1,a], SILIANG ZHANG[2,c] AND FIONA STEELE[1,b]

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

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

Understanding the help and support that is exchanged between family members of different generations is of increasing importance, with research questions in sociology and social policy focusing on both predictors of the levels of help given and received, and on reciprocity between them. We propose general latent variable models for analysing such data, when helping tendencies in each direction are measured by multiple binary indicators of specific types of help. The model combines two continuous latent variables, which represent the helping tendencies, with two binary latent class variables which allow for high proportions of responses where no help of any kind is given or received. This defines a multivariate version of a zero inflation model. The main part of the models is estimated using MCMC methods, with a bespoke data augmentation algorithm. We apply the models to analyse exchanges of help between adult individuals and their non-coresident parents, using survey data from the UK Household Longitudinal Study.

1. Introduction. In this article we propose and apply latent variable models for the joint distribution of variables within a dyad of two interacting units. This is motivated by research questions in sociology and social policy about exchanges of help and support between adult individuals and their non-coresident parents. In all societies such intergenerational transfers have major implications for individual, family, and societal wellbeing (Mason and Lee, 2018). Transfers between adult children and their parents are an important element of intergenerational linkages and a means of providing support to those in need (Künemund et al., 2005), especially in a context of shrinking social services (Pickard, 2015). Increases in life expectancy imply an increase in the volume of help needed by older people with age-related functional limitations. At the same time, there may be an increased need for assistance in younger age groups as a result of delayed transitions to adulthood, precarious employment, and increasingly diverse and complex family life courses (Lesthaeghe, 2014; Henretta et al., 2018). Analysis of the factors associated with exchanges of support between generations is important in order to anticipate which population sub-groups may be at risk from lack of support in the future, either because of an increased unmet need for help or a reduced capacity to provide help among potential donors.

We consider two broad research questions on such intergenerational support: what characteristics of individuals and their parents are associated with different levels of help given and received between them, and what is the extent and nature of reciprocity of these exchanges (i.e. to what extent do children with a high tendency to give help to parents also have a high or low tendency to receive help). Previous research suggests that reciprocity, either contemporaneous or over the life course, is an important motivating factor in intergenerational exchanges of support (e.g. Grundy 2005; Silverstein et al. 2002). For example, studies

Keywords and phrases: Item response theory models, latent class analysis, mixture models, non-equivalence of measurement, two-step estimation.

that have analysed contemporaneous reciprocity have found a positive association between support to parents and the receipt of support by their adult children in the U.S. (Cheng et al. 2015) and Britain (Grundy 2005; Steele and Grundy 2021), and studies that have examined reciprocity across the lifecourse have found that a higher level of parental support during childhood is associated with an increased propensity to help parents in later life (e.g. Silverstein et al. 2002). Another reason for considering the extent to which child-parent exchanges are balanced is that being unable to reciprocate support may have negative consequences for the mental health and wellbeing of older people (e.g. Davey and Eggebeen 1998).

Research on intergenerational support is framed by theoretical perspectives from sociology, social psychology and economics (see e.g. the discussions in Silverstein et al. 2002, Grundy 2005 and Kalmijn 2014, and references therein). A prominent distinction is between explanations which focus on altruism and ones which focus on the costs and benefits of giving support, although these motivations do not need to be mutually exclusive. The theories in turn inform considerations of possible explanatory variables for levels of support. Many of them can be seen as instances of two broad kinds of factors: *capacity* (financial and time resources) of the provider of help, and the *needs* of the recipient (Fingerman et al., 2015). A wide range of such predictors have been examined for exchanges of support between generations in different contexts (see the studies cited in this section, and references therein).

Our goal is to improve the methodology of analysing these questions. We consider the case of Britain, using cross-sectional survey data from the Family Network module of the UK Household Longitudinal Study (UKHLS). These data are described in Section 2. They include sixteen questions ('items') about exchanges of help within the dyad defined by a survey respondent and their non-coresident parent or parents. Eight of the items indicate whether the respondent gives each of eight specific types of help to the parents (for example, helping them with housework), and eight indicate types of help that they may receive from the parents. These items are regarded as measures of two latent variables, which we interpret as the general tendencies to give and to receive help. We thus have 'doubly multivariate' data, with two sets of observed binary items measuring two latent variables. The substantive research questions correspond to questions about the joint distribution of the latent variables, both on their means conditional on covariates and on the association between the latent variables.

The analysis of this situation should be handled with an appropriate form of *latent variable modelling*. Further, the data have two peculiar features which should be allowed for. First, they display a multivariate form of *zero inflation*, where the proportion of respondents who give a zero response to all eight items for a latent variable (i.e. no help of any kind given, or none received) is larger than can be accounted for by basic models. Second, the signal value of specific types of help may be different for different types of respondents, especially for men and women (because of gendered patterns of helping) or for people who live at longer or shorter distances from their parents (because of different practicalities of different kinds of help). This can be seen as an instance of *non-equivalence of measurement* in the items.

We propose a general latent variable modelling framework for the analysis of such data. Its starting point is a conventional model for two continuous latent variables given covariates, measured by binary items. Non-equivalence of measurement is represented by letting the measurement component of the model depend on some covariates. Zero inflation is allowed for by supplementing the bivariate continuous latent variable with a bivariate latent class variable which accounts for the excess of all-zero responses in one or both of the sets of items. This specification combines and extends several modelling elements, and draws on the corresponding literatures (we discuss this further in Section 3.2, after the models have been defined in Section 3.1). Beyond the analysis of exchanges of help between generations, the models could also be applied to other questions with comparable doubly multivariate data elsewhere, for instance studies of family members' perceptions of relationship quality in family psychology or of cooperation between coworkers in organisational behaviour.

We contribute to the literature on intergenerational exchanges of support by addressing two methodological limitations of previous research: the measurement of support given or received, and estimation of reciprocity of exchanges. Most previous studies using data similar to those collected in UKHLS have reduced the multivariate data on different types of support in a given direction to a single binary variable indicating whether any support was given or received. Another approach has been to analyse the sum score of the items using linear regression (e.g. Cheng et al. 2015), which assigns equal weight to each item and ignores zero inflation. It is common to focus on one direction of exchange only, for example just support given to elderly parents (e.g. Silverstein et al. 2002), or to analyse receipt and provision of support separately. The disadvantage of both approaches is that they preclude investigation of reciprocity of exchanges, the importance of which has been widely acknowledged. Among the few studies that have investigated reciprocity, one approach has been to treat helping tendencies as categorical, by using first latent class analysis to identify a typology of exchanges and then modelling class membership using multinomial regression (Hogan et al. 1993; Chan 2008), and another has been to treat them asymmetrically, by including the receipt of support as a predictor of provision of support, and vice versa (Grundy 2005; Cheng et al. 2015). A recent study by Steele and Grundy (2021) models bidirectional exchanges between adult children and their parents jointly, interpreting the residual correlation as a measure of reciprocity, but it treats support given and received as a bivariate binary response.

We estimate the models using a two-step approach where the parameters of the measurement model for the items given the latent variables are estimated first, and their values are then held fixed in the second step where the structural model for the latent variables is estimated. The second step is carried out using Markov Chain Monte Carlo (MCMC) estimation. It was implemented using a tailored algorithm written for these models, which is made available as an R package. This substantially speeds up the estimation, compared to implementation with general-purpose MCMC packages. The algorithm has a convenient data augmentation structure which alternates between sampling the latent variables given the model parameters and the observed data, and sampling the parameters given the observed and latent data. These methods of estimation are described in Section 4, with details of the MCMC algorithm given in the Appendix.

Our analysis of intergenerational exchanges of help is then described in Section 5. The results suggest that parents and childen with some characteristics and family circumstances that are likely to be associated with higher levels of capacity do indeed have a greater tendency to give help, and those with characteristics associated with higher levels of need have a greater tendency to receive help. Helping tendencies in the two directions are positively correlated, conditional on the explanatory variables, suggesting a substantial amount of contemporaneous reciprocity in helpfulness between the generations.

2. Data on exchanges of support between generations. We use data from the UK Household Longitudinal Study (UKHLS), also known as ‘Understanding Society’ (University of Essex et al. 2018; see Knies 2018 for more information on the study). This is a longitudinal survey of the members of approximately 40,000 households (at Wave 1 of UKHLS) drawn from the residential population living in private households in the United Kingdom. UKHLS started in 2009, but it also subsumes the smaller British Household Panel Survey (BHPS) which began in 1991.

Information on exchanges of help with parents living outside a respondent’s household was collected in the Family Network module which was administered in 2001, 06, 11/12, 13/14, and 15/16 (BHSP Waves 11 and 16 and UKHLS Waves 3, 5, and 7). We carry out cross-sectional analysis which uses only the last of these, Wave 7 (with the exception that the development and estimation of the measurement models was done using pooled data across all five waves; this is explained separately in Section 5.1). In this module, respondents with at least one non-coresident parent were asked whether they 'nowadays' gave 'regularly or frequently' the following eight types of help to their parent(s), each with a yes–no response: 'giving them lifts in your car (if you have one)' [referred to as *lifts* below], 'shopping for them' (*shopping*), 'providing or cooking meals' (*meals*), 'helping with basic personal needs like dressing, eating or bathing' (*personal care*), 'washing, ironing or cleaning' (*housework*), 'dealing with personal affairs e.g. paying bills, writing letters' (*personal affairs*), 'decorating, gardening or house repairs' (*diy*), and 'financial help' (*financial*). The same questions were asked about receipt of support from the parents, but with the personal care item replaced by 'looking after your children' (*childcare*).

Although respondents were asked to report on giving parents a lift in their car if they had one, the recorded variable had only 'yes' or 'no' responses. We therefore 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 dependent children aged 16 or under. For the item on receiving lifts from parents, we do not have information about whether the parents have access to a car, so responses of 'no' to this item will include also cases where they do not.

A set of covariates (explanatory variables) was considered to capture factors that may be associated with help given or received between individuals and their parents. They are gender, age, number of siblings, partnership status and employment status of the respondent, the presence and age of children in their family, their household income, age of the oldest non-coresident parent, whether any parent lived alone, and the travel time between the respondent and the parent living closest.

These family network data have several limitations, which are shared by other large-scale studies of intergenerational exchanges. First, for practical reasons, data on exchanges with non-coresident relatives were collected only from the perspective of the survey respondent, in our case the child. As noted by Chan and Ermisch (2015), studies with matched pair data which are collected from both non-coresident parents and children in the same family are rare and tend to be for small selective samples. For UKHLS this means that we have rich data on children, but much less information on their parents. Child responses may also suffer from reporting bias, for example over-reporting of help given and under-reporting of help received. Second, respondents were asked to report on exchanges with both parents collectively. so it is not possible to distinguish between exchanges with the mother and with the father, even when they are living apart. Where a respondent had both biological and step/adoptive parents alive, the recorded responses refer to the ones that the respondent had most contact with. The data also include some respondents who are siblings to each other, and thus refer to the same parents. This can happen when respondents who are currently adults in their own households were originally sampled as children in the same household. The number of such cases is small in our analysis data (635 respondents have a sibling in the dataset), and we do not include a further adjustment for dependencies among them. The number of siblings that a respondent has is included in the analysis as a covariate for every respondent.

The sample was first restricted to the 19,052 respondents in UKHLS Wave 7 who were aged 16 or over and who had at least one non-coresident parent but no coresident parent. Respondents living with a parent were mainly younger individuals who had not left the parental home; they were excluded because their exchanges with their non-coresident parent are likely to differ from those of respondents who do not live with either parent. Also excluded were respondents whose closest parent lived or worked abroad (n=2590) and those with missing data on any covariate or on *all* of the help items (n=1719). This gives our main analysis sample of 14,743 respondents. Most of the omissions from missing data (1226 cases) were due to the indicator of whether either parent lived alone, while nonresponse on the other covariates and the help items was much rarer.

TABLE 1
Percentage of respondents giving help to their non-coresident parents and receiving help from the parents, by item.

Item	Help given to parents	Help received from parents
Lifts in car (*lifts*)	29.1	11.2
Shopping (*shopping*)	20.8	8.0
Providing or cooking meals (*meals*)	12.3	14.0
Basic personal needs (*personal care*)	3.6	–
Looking after children (*childcare*)	–	40.3
Washing, ironing or clearning (*housework*)	7.7	5.6
Personal affairs (*personal affairs*)	16.9	2.4
Decorating, gardening or house repairs (*diy*)	17.9	8.3
Financial help (*financial*)	6.3	13.1
At least one of the eight kinds of help:	43.3	40.0

Data from UKHLS, Wave 7. Valid percentages, excluding cases with missing data. The sample sizes are n =14,736 for items on help to parents, and n =14,738 for help from parents. Of these, the *lifts* item is missing for the 19.0% of respondents who have no access to a car, and *childcare* is missing for the 52.3% who have no coresident dependent children.

The percentages of respondents in the analysis sample who reported giving and receiving each type of help are shown in Table 1, and descriptive statistics and coding for the covariates in Table 2. Less than half of the respondents report giving (43.3%) or receiving (40.0%) even one of these kinds of help. This large proportion of all-No responses is a feature that we will want allow for in modelling these data. The specification and estimation of the models is described in Sections 3 and 4. We will then return to the analysis of the data in Section 5.

3. Latent variable models for dyadic data. Here we define, in Section 3.1, the latent variable models that we propose for analysing dyadic data like those introduced in Section 2. In Section 3.2 we discuss how different elements of this specification draw on previous literature. For ease of exposition the dyads and variables are mostly introduced with reference to their meaning in the application to intergenerational support, but we note that the models are also applicable to any data with a similar doubly multivariate structure.

3.1. *Model specification.* Consider data on variables $(\mathbf{X}_i, \mathbf{Y}_{Gi}, \mathbf{Y}_{Ri})$ for a sample of n dyads $i = 1, \dots, n$, where $\mathbf{X}_i$ is a $q \times 1$ vector of covariates and $\mathbf{Y}_{Gi} = (Y_{G1i}, \dots, Y_{Gp_Gi})'$ and $\mathbf{Y}_{Ri} = (Y_{R1i}, \dots, Y_{Rp_Ri})'$ are two vectors of binary indicator variables (*items*). In our application, a dyad is composed of an individual (survey respondent) and their non-coresident parents, and the items are the questions on the $p_G = p_R = 8$ specific types of help the respondent gives to the parents ($\mathbf{Y}_{Gi}$) or receives from the parents ($\mathbf{Y}_{Ri}$). Each item is coded 1 if that kind of help is given or received, and 0 if it is not. We treat $\mathbf{Y}_{Gi}$ as multiple indicators of a latent variable η_{Gi} which describes an individual's tendency to give help to their parents, and $\mathbf{Y}_{Ri}$ as indicators of another latent variable η_{Ri} which describes the parents' tendency to give help to the individual. We take η_{Gi} and η_{Ri} to be continuous variables. The goal is to estimate their joint distribution and how it depends on the covariates.

Our data contain a substantial number of respondents for whom all of the items in $\mathbf{Y}_{Gi}$ or $\mathbf{Y}_{Ri}$ are 0 (see Table 1). The frequencies of such all-zero response patterns are higher than what would be expected under standard models with continuous latent variables. To allow for this, we introduce for each of the two sets of items a second, binary latent variable. It defines two latent classes, denoted by 0 and 1, where class 0 accounts for the excess zeros. For help to parents, this latent class variable is denoted ξ_{Gi}. The *measurement model* for how

TABLE 2
Descriptive statistics for the covariates used in the analysis (n=14,743).

Variable	n	Percent
Respondent (child) characteristics		
Age (years)	Mean=42.7	SD=11.4
Gender		
Female	8489	57.6
Male	6254	42.4
Number of siblings respondent has		
0	1383	9.4
1–2	9096	61.7
3 or more	4264	28.9
Partnership status		
Partnered	11,255	76.3
Single	3488	23.7
Employment status		
Employed	11,423	77.4
Unemployed	583	4.0
Economically inactive	2737	18.6
Annual household income (pounds, log-transformed)*	Mean=9.9	SD=0.7
Age of youngest child		
No children	5982	40.6
$0-1$ years	1264	8.6
$2-4$ years	1681	11.4
$5-10$ years	2278	15.5
$11-16$ years	1707	11.6
>16 years	1831	12.4
Parent characteristics		
Age of oldest parent (years)	Mean=71.1	SD=11.3
At least one parent lives alone		
Yes	5586	37.9
No	9157	62.1
Child-parent characteristics		
Travel time to nearest parent		
1 hour or less	10,675	72.4
More than 1 hour	4068	27.6

* Equivalized and adjusted for inflation using the Consumer Price Index for the year of interview within the two-year survey field period.

$\mathbf{Y}_{Gi}$ measures the latent variables is then specified by

(1) $$p(\mathbf{Y}_{Gi} = \mathbf{0} | \xi_{Gi} = 0, \mathbf{X}_i) = 1 \qquad \text{and}$$

(2) $$p(\mathbf{Y}_{Gi} | \xi_{Gi} = 1, \eta_{Gi}, \mathbf{X}_i; \boldsymbol{\phi}_G) \equiv p_1(\mathbf{Y}_{Gi} | \eta_{Gi}, \mathbf{X}_i; \boldsymbol{\phi}_G) = \prod_j p_1(Y_{Gji} | \eta_{Gi}, \mathbf{Z}_i; \boldsymbol{\phi}_G)$$

where $p(\cdot|\cdot)$ denotes a conditional distribution and $p_1(\cdot|\cdot)$ that a distribution is also conditional on $\xi_{Gi} = 1$, $\mathbf{Z}_i$ are a subset of $\mathbf{X}_i$, and $\boldsymbol{\phi}_G$ are parameters. When $\xi_{Gi} = 0$, a respondent is thus certain to answer 'No' to all the items in $\mathbf{Y}_{Gi}$. When $\xi_{Gi} = 1$, the probabilities of the responses depend on the latent helping tendency η_{Gi} and covariates $\mathbf{Z}_i$, and the different items Y_{Gji} are taken to be conditionally independent of each other; this is a conventional latent variable model for the binary items, with the extension that the measurement may be non-equivalent with respect to some covariates $\mathbf{Z}_i$. Together, (1) and (2) define a zero-inflation model where the class $\xi_{Gi} = 0$ allows for that part of the probabilities of $\mathbf{Y}_{Gi} = \mathbf{0}$ which are not accounted for by the distribution of η_{Gi} and the measurement model given η_{Gi}.

The measurement model for $\mathbf{Y}_{Ri}$ given $(\xi_{Ri}, \eta_{Ri}, \mathbf{Z}_i)$ is defined similarly, with parameters $\boldsymbol{\phi}_R$. We assume that $\mathbf{Y}_{Gi}$ do not depend on (ξ_{Ri}, η_{Ri}), $\mathbf{Y}_{Ri}$ do not depend on (ξ_{Gi}, η_{Gi}), and $\mathbf{Y}_{Gi}$ and $\mathbf{Y}_{Ri}$ are conditionally independent of each other, and define $\boldsymbol{\phi} = (\boldsymbol{\phi}_G, \boldsymbol{\phi}_R)$. Some of the items in $\mathbf{Y}_{Gi}$ and/or $\mathbf{Y}_{Ri}$ may be missing, in which case the products over j in (2) and the corresponding model for $\mathbf{Y}_{Ri}$ are over only those items which are observed for that respondent. This implies that these missing data are assumed to be missing at random. We assume here that there are no missing data in the covariates $\mathbf{X}_i$.

The model for the latent variables given the explanatory variables is specified by the distributions $p(\xi_{Gi} = j, \xi_{Ri} = k|\mathbf{X}_i; \boldsymbol{\psi}_\xi) \equiv \pi_{jk}(\mathbf{X}_i; \boldsymbol{\psi}_\xi)$ and $p(\eta_{Gi}, \eta_{Ri}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)$, where $\boldsymbol{\psi} = (\boldsymbol{\psi}_\xi, \boldsymbol{\psi}_\eta)$ are parameters, and (η_{Gi}, η_{Ri}) and (ξ_{Gi}, ξ_{Ri}) are taken to be independent of each other given $\mathbf{X}_i$. We refer to this as the *structural model* for the latent variables. It will be the focus of interest for the substantive research questions.

Let $\mathbf{Y} = (\mathbf{Y}_G, \mathbf{Y}_R)$ denote all of the observed data on $\mathbf{Y}_i = (\mathbf{Y}_{Gi}, \mathbf{Y}_{Ri})$, and $\mathbf{X}$ all the $\mathbf{X}_i$. Define $G_i = 1$ if $\mathbf{Y}_{Gi} \neq \mathbf{0}$ and $G_i = 0$ if $\mathbf{Y}_{Gi} = \mathbf{0}$, and define R_i similarly for $\mathbf{Y}_{Ri}$. If we take the observations i to be independent, the log-likelihood for the model is

$$
\begin{aligned}
(3)\quad & \log p(\mathbf{Y}|\mathbf{X}; \boldsymbol{\phi}, \boldsymbol{\psi}) = \\
& \sum_{i=1}^{n} \log \Bigg[\pi_{11}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \iint p_1(\mathbf{Y}_{Gi}|\eta_{Gi}, \mathbf{Z}_i; \boldsymbol{\phi}_G)\, p_1(\mathbf{Y}_{Ri}|\eta_{Ri}, \mathbf{Z}_i; \boldsymbol{\phi}_R)\, p(\eta_{Gi}, \eta_{Ri}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{Gi} d\eta_{Ri} \\
& \qquad + (1 - R_i)\, \pi_{10}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \int p_1(\mathbf{Y}_{Gi}|\eta_{Gi}, \mathbf{Z}_i; \boldsymbol{\phi}_G)\, p(\eta_{Gi}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{Gi} \\
& \qquad + (1 - G_i)\, \pi_{01}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \int p_1(\mathbf{Y}_{Ri}|\eta_{Ri}, \mathbf{Z}_i; \boldsymbol{\phi}_R)\, p(\eta_{Ri}|\mathbf{X}_i; \boldsymbol{\psi}_\eta)\, d\eta_{Ri} \\
& \qquad + (1 - G_i)(1 - R_i)\, \pi_{00}(\mathbf{X}_i; \boldsymbol{\psi}_\xi) \Bigg].
\end{aligned}
$$

We further specify the structural model for each $i = 1, \ldots, n$ as

$$
(4)\qquad p(\eta_{Gi}, \eta_{Ri}|\mathbf{X}_i; \boldsymbol{\psi}_\eta) \sim N\left(\begin{bmatrix} \boldsymbol{\beta}'_G \mathbf{X}_i \\ \boldsymbol{\beta}'_R \mathbf{X}_i \end{bmatrix}, \begin{bmatrix} \sigma_G^2 & \\ \rho_{GR}\sigma_G\sigma_R & \sigma_R^2 \end{bmatrix} \right) \qquad \text{and}
$$

$$
(5)\qquad \log \left[\frac{\pi_{jk}(\mathbf{X}_i; \boldsymbol{\psi}_\xi)}{\pi_{00}(\mathbf{X}_i; \boldsymbol{\psi}_\xi)} \right] = \boldsymbol{\gamma}'_{jk} \mathbf{X}_i
$$

for $j, k = 0, 1$ with $\boldsymbol{\gamma}_{00} = \mathbf{0}$, i.e. as a bivariate normal linear model for (η_{Gi}, η_{Ri}) and a multinomial logistic model for (ξ_{Gi}, ξ_{Ri}). Thus here $\boldsymbol{\psi}_\eta$ includes $(\boldsymbol{\beta}_G, \boldsymbol{\beta}_R, \sigma_G^2, \sigma_R^2, \rho_{GR})$ in (4), and $\boldsymbol{\psi}_\xi$ includes $(\boldsymbol{\gamma}_{01}, \boldsymbol{\gamma}_{10}, \boldsymbol{\gamma}_{11})$ in (5). Finally, the measurement models given the continuous latent variables are specified as

$$
(6)\qquad \text{logit}[p_1(Y_{Gji} = 1|\eta_{Gi}, \mathbf{Z}_i; \boldsymbol{\phi}_G)] = \tau_{Gj} + \boldsymbol{\delta}'_{Gj}\mathbf{Z}_i + (\lambda_{Gj} + \boldsymbol{\zeta}'_{Gj}\mathbf{Z}_i)\, \eta_{Gi}
$$

for $j = 1, \ldots, p_G$, so that $\boldsymbol{\phi}_G$ consists of all the τ, $\boldsymbol{\delta}$, λ and $\boldsymbol{\zeta}$ parameters for $\mathbf{Y}_{Gi}$, and the models for the items in $\mathbf{Y}_{Ri}$ are specified similarly, with parameters $\boldsymbol{\phi}_R$. The baseline parameters of these models are the intercepts (the τs) and the loadings of the latent η variables (the λs). These are then further modified by the covariates if any of the $\boldsymbol{\delta}$ or $\boldsymbol{\zeta}$ parameters are non-zero, in which case the measurement model for that item is non-equivalent with respect to the corresponding variables in $\mathbf{Z}_i$. For simplicity, we consider only models where any non-equivalence in an item affects both the intercept and the loading, so that for that item the elements of $\boldsymbol{\delta}$ and $\boldsymbol{\zeta}$ corresponding to the same variable in $\mathbf{Z}_i$ are either both zero or both non-zero. The motivation and choice of $\mathbf{Z}_i$ in our application are discussed in Section 5.1.

3.2. *Previous literature on the elements of the models.* The model in Section 3.1 combines several existing modelling elements, and draws on the corresponding literatures. The starting point is the conventional general framework for latent variable modelling with covariates (see e.g. Skrondal and Rabe-Hesketh 2004 and Bartholomew et al. 2011). If (η_G, η_R) were the only latent variables, this would be a standard model for the joint distribution of two continuous latent variables given covariates $\mathbf{X}$. When, as here, all the measures $\mathbf{Y}$ of the latent variables are binary and the measurement models for them are logistic models, this is a common instance of what is known, especially in psychometrics and educational testing, as Item Response Theory (IRT) modelling (see e.g. de Ayala 2009 and van der Linden 2016).

Including covariates $\mathbf{Z}$ in a measurement model, as we do in (6), allows the measurement of a latent variable to be non-equivalent with respect to these covariates. This is also a standard approach in applications where such non-equivalence may be of concern, such as in cross-national survey research and other 'multigroup' situations, and in many applications of IRT, where non-equivalence of measurement is commonly known as differential item functioning (DIF). For overviews of these ideas and methods, see Kankaraš et al. (2011) and Millsap (2011).

The least familiar element of the model is the way we allow for the large number of all-zero responses by adding the latent class variables (ξ_R, ξ_R). To motivate this, consider first models for a single non-negative variable Y with excess zeros, meaning that the observed probability $P(Y = 0)$ is greater than can be expected under an assumed distribution $p(Y)$ for Y. There are, broadly, three ways of representing this situation, depending on how many of the zero values are taken to be accounted for by $p(Y)$: (1) all of them—*censoring* models where it is assumed that Y could actually be negative but that all such values are recorded as 0, so that $P(Y = 0) = p(Y \leq 0)$; (2) none of them—*hurdle* models where we model separately $P(Y = 0)$ and $p(Y|Y > 0)$; or (3) some of them—*zero-inflated* models where $P(Y = 0) = \pi + (1 - \pi)p(Y = 0)$ with an additional probability parameter π for the proportion of zeros which is not accounted for by $p(Y)$ (see e.g. Tobin 1958, Cragg 1971, Mullahy 1986, Lambert 1992, and Min and Agresti 2005 for introductions and comparisons of these possibilities).

We are interested in latent-variable models for multivariate items. Denote for the moment a generic continuous latent variable by η, and its indicators by $\mathbf{Y}$, omitting covariates, so that the model is specified by $p(\mathbf{Y}|\eta)p(\eta)$. Suppose that the observed proportion of $\mathbf{Y} = \mathbf{0}$ is so high that we want to allow for it specially. Here the basic model for $p(\mathbf{Y}|\eta)$ is in effect already a censoring model, in that estimates of its parameters will be determined so that they accommodate these zeros. This, however, can distort the parameters in a way which leaves the model as a whole badly specified to account for the non-zero patterns of responses (see Wall et al. 2015 for a discussion of the biases which can arise when a latent-variable model is poorly specified in this way). A hurdle model is also unappealing here, because it would involve conditioning on the observed items $\mathbf{Y}$. This leads us to consider zero-inflated models, extended to multivariate $\mathbf{Y}$.

These models can be seen as an instance of *finite mixture models*. The general form of them is here $\sum_g p_g(\mathbf{Y}|\eta)\, p_g(\eta)\, \pi_g$, where $\pi_g = P(\xi = g)$ are probabilities of a latent-class variable ξ. One type of such models is obtained when $p_g(\mathbf{Y}|\eta) = p(\mathbf{Y}|\eta)$, i.e. when the measurement model is the same in every class g. Then the model becomes $p(\mathbf{Y}|\eta)\, p^*(\eta)$ where $p^*(\eta) = \sum_g p_g(\eta)\, \pi_g$ is a finite mixture distribution. This provides a way of specifying the basic latent variable model with a more flexible distribution for η than is possible with a single parametric (for example normal) distribution. Mixture modelling with this purpose is discussed by Wall et al. (2012) and Wall et al. (2015). Here, in contrast, we particularly need models where the measurement models do depend on the class. This represents a situation where the latent classes correspond to individuals with different *response styles*, i.e. different relationships between the latent variable η and its measures $\mathbf{Y}$. This idea has been used in various contexts of measurement, especially in applications of psychological and educational testing; see Wall et al. (2015), Huang (2016), and references therein.

A zero-inflation model for multivariate $\mathbf{Y}$ involves two response styles: one where $\mathbf{Y} = \mathbf{0}$ always, and one where $\mathbf{Y}$ follows an IRT model given η. This has been proposed for models where the items are binary (Muthén and Asparouhov 2006; Finkelman et al. 2011; Wall et al. 2015), ordinal (Magnus and Liu, 2017) or count variables (Magnus and Thissen, 2017), sometimes with extensions such as separate classes for all-0 and all-1 response patterns or more than one class for general response patterns. Our model is similar to the previous ones for binary items (different versions of them use ostensibly different specifications for η in the all-zero class for ξ, but these are all equivalent). To accommodate the dyadic data, however, we have extended them to include two latent variables (η_G and η_R), with separate zero-inflation classes for each of them.

4. Estimation of the models.

4.1. *Two-step estimation.* We employ a two-step approach to estimate these models. What this means is that the measurement model is first selected and estimated separately, and its parameters $\boldsymbol{\phi}$ are then fixed at their estimated values for all subsequent exploration and estimation of the structural model. These two steps for our models are described separately in Sections 4.2 and 4.3 below.

This idea of two-step estimation of latent variable models goes back to Burt (1976, 1973), and the implementation of it has been developed more recently by Xue and Bandeen-Roche (2002) and Bakk and Kuha (2018). Our motivation for using it here is twofold. First, it substantially reduces the computational demands compared to the 'one-step' method of estimating all parts of the models together. This is beneficial here, where the estimation of even the structural model alone is demanding. Second, a conceptual advantage of the two-step approach is that fixing the measurement models in the first step also fixes the exact operational definition of the latent variables. This then remains fixed in subsequent analyses, and do not change when the specification of the structural model is changed, for example when covariates are added or removed. In our work, this extends also to other analyses of intergenerational exchanges of family support outside this paper, where we also want to keep the definitions of the latent variables unchanged in this sense.

4.2. *Estimation of the measurement models.* In the first step of the estimation, the measurement models for $\mathbf{Y}_G$ and $\mathbf{Y}_R$ are estimated separately and conditional on $\mathbf{Z}$ alone. This means that for $\mathbf{Y}_G$ we consider the log likelihood

$$\log p(\mathbf{Y}_G|\mathbf{Z};\boldsymbol{\phi}_G,\boldsymbol{\psi}^*_G) = \sum_{i=1}^{n} \log \Bigg[\pi_G(\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\xi}) \int p_1(\mathbf{Y}_{Gi}|\eta_{Gi},\mathbf{Z}_i;\boldsymbol{\phi}_G)\, p(\eta_{Gi}|\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\eta})\, d\eta_{Gi}$$
$$+ (1-G_i)\,\big(1-\pi_G(\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\xi})\big)\Bigg] \tag{7}$$

where $p_1(\mathbf{Y}_{Gi}|\eta_{Gi},\mathbf{Z}_i;\boldsymbol{\phi}_G)$ is as before. The structural model here consists of $\pi_G(\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\xi}) = P(\xi_i = 1|\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\xi})$ and $p(\eta_{Gi}|\mathbf{Z}_i;\boldsymbol{\psi}^*_{G\eta})$, specified as a binary logistic and a normal linear model, and $\boldsymbol{\psi}^*_G = (\boldsymbol{\psi}^*_{G\xi},\boldsymbol{\psi}^*_{G\eta})$ are the parameters of these models. Here (7) is obtained by integrating (3) over $p(\mathbf{Y}_R,\mathbf{X}_*|\mathbf{Z})$, where $\mathbf{X}_*$ denotes the variables in $\mathbf{X}$ but not in $\mathbf{Z}$. This is actually only approximately true, because if (4)–(5) holds given $\mathbf{X}$, then the structural models given $\mathbf{Z}$ only are generally not exactly of binary logistic and normal linear form. We ignore this small approximation and maximize (7) to estimate $\boldsymbol{\phi}_G$. This step can be carried out using standard latent variable modelling software such as Mplus. The parameters $\boldsymbol{\phi}_R$ are estimated similarly from a model like (7) for $\mathbf{Y}_R$. Denote these estimates by $\tilde{\boldsymbol{\phi}} = (\tilde{\boldsymbol{\phi}}_G,\tilde{\boldsymbol{\phi}}_R)$. The estimates of $\boldsymbol{\psi}^*_G$ and $\boldsymbol{\psi}^*_R$ from this step are discarded.

4.3. *Estimation of the structural models.* In the second step of estimation, the structural models are then estimated, treating the estimated measurement parameters $\tilde{\boldsymbol{\phi}}$ from the first step as known numbers. In other words, the log-likelihood for the second step is (3) but in the form $\log p(\mathbf{Y}|\mathbf{X}; \tilde{\boldsymbol{\phi}}, \boldsymbol{\psi})$ where only $\boldsymbol{\psi}$ are unknown parameters. We omit below the fixed $\tilde{\boldsymbol{\phi}}$ from the notation for simplicity. We further write $\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\eta})$, where $\boldsymbol{\xi}$ denotes all the values of the latent (ξ_{Gi}, ξ_{Ri}) for the dyads i in the sample, and $\boldsymbol{\eta}$ all the values of (η_{Gi}, η_{Ri}).

In our analyses, this step was carried out in the Bayesian framework and using MCMC methods of estimation. The estimation algorithm has a data augmentation structure, which alternates between sampling the latent variables and sampling the model parameters:

- **Imputation step**: Given the observed data $(\mathbf{Y}, \mathbf{X})$ and the most recently sampled value of the parameters $\boldsymbol{\psi}$, sample a value for the latent variables $\boldsymbol{\zeta}$ from the conditional distribution

$$p(\boldsymbol{\zeta}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\psi}) \propto p(\mathbf{Y}|\boldsymbol{\zeta}, \mathbf{X}) p(\boldsymbol{\zeta}|\mathbf{X}; \boldsymbol{\psi}).$$

  This is further split into sampling $\boldsymbol{\xi}$ from $p(\boldsymbol{\xi}|\boldsymbol{\eta}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi})$, using $\boldsymbol{\eta}$ from the previous iteration, and then $\boldsymbol{\eta}$ from $p(\boldsymbol{\eta}|\boldsymbol{\xi}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi})$.

- **Posterior step**: Given the observed data $(\mathbf{Y}, \mathbf{X})$ and the most recently sampled value of the latent variables $\boldsymbol{\zeta}$, sample a value for the parameters $\boldsymbol{\psi}$ from the conditional distribution

$$p(\boldsymbol{\psi}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\zeta}) = p(\boldsymbol{\psi}|\mathbf{X}, \boldsymbol{\zeta}) \propto p(\boldsymbol{\zeta}|\mathbf{X}; \boldsymbol{\psi}) p(\boldsymbol{\psi})$$

  where $p(\boldsymbol{\psi}) = p(\boldsymbol{\psi}_\xi) p(\boldsymbol{\psi}_\eta)$ is the prior distribution of the parameters, taking $\boldsymbol{\psi}_\xi$ and $\boldsymbol{\psi}_\eta$ to be independent a priori. The conditional distribution then further splits into

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

  which can be sampled separately and in parallel for $\boldsymbol{\psi}_\xi$ and $\boldsymbol{\psi}_\eta$. This does not depend on the measurement items $\mathbf{Y}$, because they are not in the 'Markov blanket' of $\boldsymbol{\psi}$ (in the directed acyclic graph for the model, $\mathbf{Y}$ are not parents, children or co-parents of children of $\boldsymbol{\psi}$). The posterior step thus involves sampling the parameters of two regression models given $\mathbf{X}$, a multinomial logistic model for $\boldsymbol{\xi}$ and a bivariate linear model for $\boldsymbol{\eta}$, exactly as if from their posterior distributions if the most recently imputed values of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ were real observed data.

We note that since the measurement models are fixed, the structural model is straightforwardly identified here. In particular, 'label switching', where the numbering of the latent classes changes between MCMC iterations, cannot occur.

The details of these steps are described in the Appendix. We wrote bespoke code for them, implemented in an R package. Because this is designed specifically for these models, it can achieve substantially higher speeds of estimation than general MCMC packages. The sampling procedure can be tailored to the distributions that are needed here, for some of them sampling from standard distributions and for others using adaptive rejection sampling which is enabled by log-concavity of the target distributions. Parts of the sampling within each MCMC iteration can also be implemented in parallel and using multiple processors.

5. Analysis of intergenerational exchanges of help.

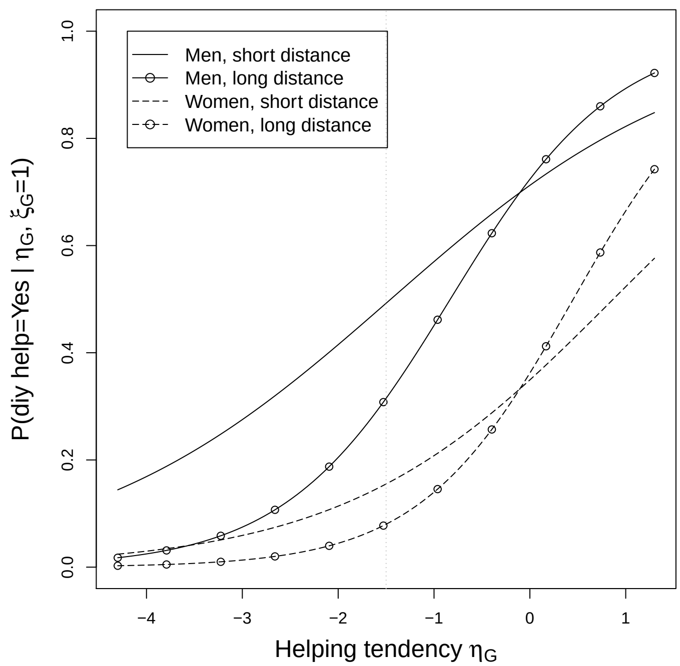
5.1. *Measurement models.* We use measurement models which are non-equivalent with respect to two covariates $\mathbf{Z}$: the gender of the respondent, and the distance between where the respondent and their parents live. It is substantively to be expected, and empirically confirmed for these data, that the patterns of what kinds of help a person gives and receives may vary by these covariates. Some types of help are strongly gendered among the generations considered here, with men and women expressing their helpfulness in different ways and receiving different kinds of help from their parents. Similarly, for obvious practical reasons a longer distance between the parties may affect some types of help more than others. This being the case, the expected levels and patterns of different kinds of help may be different between men and women and between respondents at different distances from their parents, even for individuals who actually have a similar latent tendency to give or receive help. The non-equivalent measurement models allow for this possibility. We thus define $\mathbf{Z}_i = (Z_{gi}, Z_{di})$, where Z_{gi} is an indicator variable for a female respondent and Z_{di} an indicator for a respondent who lives at a distance of an hour or more's travel time from their parents.

The measurement models were estimated using more of the UKHLS data than are used for the second step of estimation discussed in Section 5.2 below. This was because these models were intended for use in multiple analyses of the data from the Family Networks module, and were developed prior to the analysis described here. The models were first explored for data from the 2001 Wave 11 of BHPS, to identify items for which non-equivalence with respect to gender and/or distance was substantial enough that it should be allowed for. This selection was done using a combination of likelihood ratio tests and the AIC and BIC statistics. The selected models were then re-estimated using pooled data from all the five available waves of UKHLS/BHPS, to maximize the amount of data which contributed to these estimates. (Longitudinal observations for a respondent in these pooled data are not independent; ignoring this, however, affects only the standard errors of the parameter estimates, which are not needed for what follows. We have also repeated the estimation of these measurement models using only the data for the main analysis sample from Wave 7 of UKHLS; this would give very similar estimated parameters of the measurement model.) The models were estimated using maximum likelihood estimation, with the Mplus 6.12 software (Muthén and Muthén, 2010). More information about the estimated measurement models is given in the online supplementary materials.

The selected measurement models include some non-equivalence in most items, especially with respect to distance. Of the items on help to parents ($\mathbf{Y}_G$), *financial*, *lifts* and *diy* are non-equivalent with respect to gender, and all but *personal affairs*, *personal care* and *financial help* with respect to distance. Of the items on help from parents ($\mathbf{Y}_R$), *financial* and *meals* are non-equivalent with respect to gender, and all but *personal affairs* and *financial* with respect to distance. The intercept and loading parameters of *personal care*, which was fully equivalent for both η_G and η_R, were fixed at 0 and 1 respectively in both measurement models, to fix the measurement scales of the two latent variables. For each of η_G and η_R at least two items which measure them are equivalent with respect to gender, and at least two with respect to distance. This means that the coefficients of gender and distance in the structural models for (η_G, η_R) are also identified, separately from the measurement models. However, information that is available for estimating these associations is clearly reduced, especially for distance for which the non-equivalent measurement models account for much of the observed association between distance and the items $\mathbf{Y}_G$ and $\mathbf{Y}_R$. The associations between other variables in $\mathbf{X}$ and (η_G, η_R) in the structural model are then conditional on gender and distance in this sense, i.e. they refer to the latent variables as they are defined by these measurement models with this adjustment for non-equivalence.

An illustrative example of the estimated measurement models is shown in Figure 1, for the *diy* item (decorating, gardening or house repairs) on help given to the parents. The model for this item is non-equivalent with respect to both covariates. The plot shows the estimated probabilities of giving such help as a function of the latent tendency of helpfulness η_G, separately for each combination of gender and distance. Considering the genders, it can be seen that, at the same level of this tendency, men are more likely to give this kind of help. The non-equivalence with respect to distance shows most clearly in the loading (or 'discrimination')

FIG 1. *Item response curves for the item* diy *(decorating, gardening or house repairs) for help that respondents give to their parents, for probability of giving help conditional on the latent variable η_G (tendency to give help), separately for the combinations of gender of the respondent and distance between respondent and their parents. The dotted vertical line is approximate mean of η_G.*



parameters. These are larger — and the probability curves thus steeper — when a respondent lives further away from their parents, so that giving such help is a more discriminating signal of helpfulness for such respondents than for those who live near their parents.

Another implication of non-equivalent measurement models, which is less often pointed out, is that they imply that the marginal associations between different items will also depend on covariates, here gender and distance. For example, when the measurement loadings are stronger at a longer distance, as they are for *diy* and several other items, the associations between the items are also stronger, both within and between items in $\mathbf{Y}_G$ and $\mathbf{Y}_R$. This could arise, for instance, if children who live further from their parents tend to give multiple types of help on the occasions when they visit the parents.

5.2. *Models for help between respondents and their parents.* Fixing the parameters of the measurement models at their estimated values from Section 5.1, we then estimated the structural models which are our focus of interest. The MCMC algorithm described in Section 4.3 and the Appendix was run for two MCMC chains of 110,000 iterations each, from different starting values. Discarding the first 10,000 iterations of each, conventional convergence diagnostics indicated that the chains had converged. The two chains were then combined, so the estimates are based on 200,000 draws from the posterior distributions of the parameters.

The estimated parameters of the bivariate linear model (4) for the continuous latent variables (η_G, η_R) are shown in Table 3. The coefficients of the multinomial logistic model (5) for the categorical latent variables (ξ_G, ξ_R) are less convenient for interpretation, because they express comparisons of the probabilities in the joint distribution of the variables, relative to $(\xi_G, \xi_R) = (0, 0)$ (these coefficients are given in the supplementary materials). Instead, in Table 4 we summarise this model with a focus on the marginal distributions of

ξ_G and ξ_R, using comparisons of fitted probabilities. We first calculated the probabilities $p(\xi_G = j, \xi_R = k|\mathbf{X}_i; \boldsymbol{\psi}_\xi)$ for $j, k = 0, 1$ given selected values of $\mathbf{X}_i$ for each of the $n = 14,743$ respondents i and for each of the MCMC draws of $\boldsymbol{\psi}_\xi$. Table 4 shows these fitted probabilities, averaged over respondents and parameter draws. It also shows the odds ratios between ξ_G and ξ_R calculated from these averages, and the average marginal probabilities $p(\xi_G = 1)$ and $p(\xi_R = 1)$ that a respondent belongs to the class 1 where they may give and receive help respectively. Different choices are considered for the n values of $\mathbf{X}_i$. In the first row of the table, these are the actual covariate values in the observed sample. On the other rows, one covariate in turn is fixed at a single value for every respondent, while the rest are left at their sample values. For example, the second row of the table shows the results for a hypothetical sample where every respondent is aged 35. For cases with different fixed values of the same covariate, we also show the differences of the marginal probabilities between them, and posterior standard deviations of these differences over the parameter draws.

As discussed in Section 3, we interpret (η_G, η_R) as continuous latent tendencies to give and to receive help. Although ξ_G and of ξ_R were introduced primarily to account for zero inflation, on the face of it they can also be interpreted in terms of binary helping tendencies, with class 0 of each being the class of firm 'non-givers' or 'non-receivers' of help. In this sense we can interpret both higher conditional means of η_G and η_R, and higher conditional probabilities of class 1 of ξ_G and of ξ_R, as indications of higher levels of helpfulness of the respondent to the parents or vice versa; here we refer to both of these as 'positive associations' between a covariate and helpfulness. We note first that the average marginal probability of class 0 is here 0.27 for ξ_G and 0.32 for ξ_R. Each of these accounts for about half of the proportions of all-zero responses to the corresponding items (which were 0.57 and 0.60 respectively, as shown in Table 1).

Considering first the models for help given by respondents to the parents, covariates which are strongly and positively associated with it in the linear model for η_G are higher age of the oldest parent, at least one parent living alone, the respondent having no siblings or 3 or more siblings, lower household income of the respondent, and the respondent being single or not employed. Respondents who have no young children at home (i.e. have no children, or only older children) also tend to help more, although this association is less clear. In the model for ξ_G, significant positive associations are also found for age of oldest parent, a parent living alone and lower household income, and additionally for younger age of the respondent. Considering then help received from parents, characteristics which are positively associated with it in the model for η_R are the respondent having a low household income, no siblings or no children at home, the respondent being younger, single or not employed, and the parents not living alone. Similar associations are seen in the model for ξ_R for younger, less wealthy and single respondents, and, in addition, there is a negative association between help from parents and the respondent having only older children at home.

For help received from parents, associations with η_R and ξ_R are in different directions for employment status and for having no children at home. Respondents who are unemployed or economically inactive, rather than employed, are more likely to be in the no-help-received class $\xi_R = 0$ but, if they are not in this class, the level of help they do receive (η_R) tends to be higher. Similarly, respondents with no children at home have higher probability of $\xi_R = 0$, but otherwise tend to receive more help. These diverging findings for the categorical and continuous parts of the model are intriguing, but it is not clear what substantive interpretations we can draw from them.

Here we omitted comments about associations involving gender of the respondent and the distance between them and the nearest parent. As discussed in Section 4.2, the interpretation for these covariates is somewhat different because they are also included in the measurement models for the non-equivalent items in $\mathbf{Y}_G$ and $\mathbf{Y}_R$. This means, in effect, that the estimated

TABLE 3

Estimated parameters of the linear models for the tendency to give help to (η_G) and to receive help from (η_R) individuals' non-coresident parents, from the estimated model for data from Wave 7 of the UK Household Longitudinal Study described in Section 5.2. The estimates are posterior means from MCMC samples (with posterior standard deviations in parentheses).

	Help to parents		Help from parents	
	Estimate	(s.d.)	Estimate	(s.d.)
Coefficients of explanatory variables ($\hat{\boldsymbol{\beta}}_G$ and $\hat{\boldsymbol{\beta}}_R$):				
Intercept	-1.69^{***}	(0.34)	-3.37^{***}	(0.37)
Respondent (child) characteristics				
Age ($\times$10 years)	0.04	(0.04)	-0.55^{***}	(0.06)
Gender				
Female (vs. Male)	0.86^{***}	(0.05)	0.73^{***}	(0.06)
Number of siblings respondent has (vs. 0)				
1–2	-0.18^{**}	(0.07)	-0.16	(0.10)
3 or more	0.16^{**}	(0.08)	-0.48^{***}	(0.11)
Partnership status				
Partnered (vs. Single)	-0.19^{***}	(0.06)	-0.67^{***}	(0.07)
Employment status (vs. Employed)				
Unemployed	0.25^{**}	(0.12)	0.45^{***}	(0.14)
Economically inactive	0.37^{***}	(0.06)	0.25^{***}	(0.08)
Household income (log-transformed)	-0.08^{**}	(0.03)	-0.08^{**}	(0.04)
Age of youngest child in the respondent's own household (vs. 0–1 years):				
No children	0.20^{*}	(0.10)	0.18^{*}	(0.10)
0–1 years	0		0	
2–4 years	-0.05	(0.11)	-0.00	(0.10)
5–10 years	0.19^{*}	(0.11)	0.04	(0.10)
11–16 years	0.13	(0.11)	0.09	(0.13)
>16 years	0.15	(0.12)	-0.13	(0.15)
Parent characteristics				
Age of oldest parent ($\times$ 10 years)	0.34^{***}	(0.04)	0.05	(0.05)
At least one parent lives alone (vs. No)	0.55^{***}	(0.05)	-0.31^{***}	(0.06)
Child-parent characteristics				
Travel time to nearest parent				
More than 1 hour (vs. 1 hour or less)	-1.01^{***}	(0.08)	-0.46^{***}	(0.08)
Residual variances ($\hat{\sigma}^2_G$ and $\hat{\sigma}^2_R$):	2.08^{***}	(0.08)	2.45^{***}	(0.11)
Residual correlation ($\hat{\rho}_{GR}$):	0.51^{***}	(0.02)		

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

associations for gender and distance in Tables 3 and 4 are informed only by those items which are equivalent with respect to them. Even so, these associations are strong for η_G and η_R, with women tending to both give and to receive more help than men, and the level of help in both directions being lower when children and parents live far apart.

How, then, should we summarise these results? One way to do so is to think of the covariates as different instances of two broad categories of characteristics: an actor's (the child's or the parents') *capacity to give* help, and the other actor's *need to receive* help. Considering the models for help received by the respondents from the parents in this light, the lower levels of helpfulness when the parent(s) live alone or when the respondent has siblings may be taken to reflect the parents' reduced capacity to help, while the positive associations with the child being younger, single or not employed, and having a lower household income, may be interpreted as instances of higher need by the child.

TABLE 4

Fitted probabilities of the zero-inflation latent classes (ξ_G, ξ_R), from the estimated model described in Section 5.2, averaged over parameter values in MCMC samples and over covariate values $\mathbf{X}_i$ in the sample of dyads i of a respondent and their parent(s), and odds ratios (OR) calculated from these averages. On the first row the values of $\mathbf{X}_i$ are all as in the observed data, while on the other rows one covariate is set to the same value for every dyad as indicated, while the rest keep their sample values. The last six columns show the average marginal probabilities of classes $\xi_G = 1$ and $\xi_R = 1$ (those who may give or receive help), their differences between different covariate settings, and standard deviations of these differences across the MCMC samples.

Covariate setting	$p(\xi_G = j, \xi_R = k)$ (0,0)	(0,1)	(1,0)	(1,1)	OR	Marginal probabilities of helper classes [with difference (and its SD)] $p(\xi_G = 1)$			$p(\xi_R = 1)$		
Sample	.17	.10	.14	.59	7.4	.73			.68		
Respondent (child) characteristics											
Age											
35 years	.13	.11	.09	.67	8.9	.76			.78		
45 years	.22	.09	.12	.56	11.0	.69	−.07***	(.01)	.66	−.12***	(.02)
Gender											
Male	.20	.06	.13	.61	14.2	.74			.67		
Female	.16	.12	.15	.57	5.0	.72	−.02	(.02)	.69	+.02	(.02)
Number of siblings respondent has											
0	.13	.11	.16	.61	4.6	.76			.71		
1–2	.16	.11	.13	.60	6.4	.73	−.03	(.02)	.71	−.00	(.02)
3 or more	.22	.06	.16	.56	13.2	.72	−.04*	(.03)	.62	−.10***	(.03)
Partnership status											
Single	.15	.12	.09	.64	8.6	.73			.76		
Partnered	.18	.09	.16	.57	7.2	.73	−.00	(.02)	.66	−.10***	(.02)
Employment status											
Employed	.16	.10	.13	.60	7.3	.74			.70		
Unemployed	.21	.08	.18	.54	8.1	.72	−.02	(.04)	.61	−.09**	(.04)
Inactive	.22	.09	.16	.52	8.2	.69	−.05	(.02)	.61	−.09***	(.02)
Household income											
25th percentile	.17	.08	.14	.60	8.8	.74			.69		
75th percentile	.18	.11	.14	.57	6.2	.71	−.03***	(.01)	.68	−.01	(.01)
Age of youngest child in the respondent's own household											
No children	.21	.06	.14	.59	14.0	.73	+.00	(.03)	.65	−.13***	(.04)
0–1 years	.12	.15	.09	.64	5.7	.73			.79		
2–4 years	.11	.14	.10	.66	5.3	.75	+.02	(.04)	.80	+.01	(.05)
5–10 years	.15	.15	.12	.59	4.9	.70	−.03	(.03)	.74	−.05	(.04)
11–16 years	.22	.09	.20	.49	6.5	.69	−.04	(.04)	.58	−.21***	(.04)
>16 years	.18	.07	.14	.61	11.2	.74	+.01	(.04)	.68	−.11**	(.05)
Parent characteristics											
Age of oldest parent											
70 years	.22	.10	.07	.61	19.3	.68			.71		
80 years	.13	.11	.16	.60	4.3	.76	+.09***	(.01)	.71	−.00	(.01)
At least one parent lives alone											
No	.20	.12	.12	.56	7.6	.68			.69		
Yes	.14	.05	.17	.64	11.7	.81	+.13***	(.02)	.69	+.00	(.02)
Child–parent characteristics											
Travel time to nearest parent											
1 hour or less	.15	.12	.16	.57	4.3	.73			.69		
>1 hour	.24	.04	.09	.63	45.2	.72	−.02	(.03)	.67	−.02	(.03)

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

Conversely, in the models for help to parents, there is a clear positive association between the parents' need and help given: older parents and ones who live alone tend to receive more help from their children. In terms of the children's capacity to help, the respondent characteristics which are positively associated with helping — being single, not being employed, and not having young children at home — can perhaps be interpreted in these terms if we take 'capacity' to mean 'opportunity' in the sense of having fewer other commitments. The finding that higher household income is negatively associated with helping is likely to be a reflection of the fact that the types of help covered by these items are, with one exception, practical rather than financial (it also emphasises the importance of more detailed analysis of forms of financial help in future studies). When it comes to the number of siblings, respondents who have 1–2 siblings tend to help less than those with no siblings, which is consistent with some amount of sharing of helping between the siblings. On the other hand, respondents who have 3 or more siblings tend to help their parents about as much as those who have none. We might perhaps speculate that this could reflect family dynamics and expectations of helping which are particular to large families. Finally, we also observe a reduction of helping behaviour for respondents who live with a partner and/or have young children at home. Other things being equal, such respondents are less likely both to give help to their non-coresident parents and to receive help from them. We could perhaps think of this situation as one of a self-contained family unit whose support activities may be more likely to be directed within the family rather than outside of it.

The models also give estimates of the associations between the levels of help in the two directions, allowing us to examine reciprocity of support between children and their parents. These estimated associations are strong and positive. For the categorical part of the model, the odds ratios between ξ_G and ξ_R (which depend on the covariates) are typically between 5 and 10. For the continuous part, the conditional correlation between η_G and η_R is 0.51. This is in fact substantially higher than their marginal correlation, estimated from a model without covariates (not shown here), which is 0.23. This difference is mainly due to the age variables. The ages of the respondent and their oldest parent are strongly associated, with a sample correlation of 0.87. If we include either one of them alone in the models, the conditional correlation between η_G and η_R is already about 0.50, and the age variable has a strong positive association with help given and a negative one with help received (the models where both are included, as in Tables 3 and 4, further indicate, more specifically, that older respondents tend to receive less help, and older parents tend to receive more help). The two age effects thus naturally pull in different directions, so that the marginal correlation of help given and received is somewhat suppressed. If, however, we condition on the ages, i.e. account for the different levels of help we would expect on average from children and parents of given ages, the correlation between them is substantially higher. In this sense, the results suggest a high level of reciprocity in helpfulness between the generations.

6. Conclusions. In this paper we have developed methods for analysing intergenerational help and support that is exhanged between individuals and their non-coresident parents. This involved specifying latent variable models for data where the help given and received are measured by multiple items for different types of help. The models include a multivariate zero inflation component to allow for the fact that a large proportion of people in our data gave or received no help of any kind. Estimation of the models was done in a two-step fashion, where the measurement model for the items given latent helping tendencies was estimated and fixed first, before the structural model for the joint distribution of these latent variables given explanatory variables was then estimated. The estimation of the structural model was carried out using an MCMC algorithm implemented specifically for these models.

We analysed data from the UK Household Longitudinal Study, where the respondents (the children in the parent-child dyads) are aged around 40 on average, and their parents around 70. The results of the analysis indicated some characteristics of individuals and their circumstances which were strong predictors of helping, and which we interpreted in terms of the capacities and needs of the two parties. For example, parents who are older or living alone had a higher tendency to be receiving help from their children, while children who are single or have no children of their own had a higher tendency to receive help from their parents — but also a higher tendency to give help. The levels of help given by the parents and by the children were positively correlated, suggesting substantial contemporaneous reciprocity of help between the generations.

The survey data that we have used is extensive and rich in many respects, but it also has some limitations. In particular, because only one member of the dyad — here the children — were interviewed, the data on their parents is limited. It would be preferable to survey both parties directly, but this data collection design is difficult and costly to implement on a large scale.

The model proposed here is immediately applicable also to other applications with the same structure, that is 'doubly multivariate' data with two sets of observed binary items measuring two latent variables. For example, it could be used to analyse attitudes among couples, when the interest was also on the concordance between the partners. Further, the model could be extended in different ways, both for this and for other applications. This would involve, in essence, combining the kinds of structural models that would be used in each situation if the variables of interest were directly observed with the kinds of measurement models considered here when they are latent rather than observed. For example, data where the dyads are grouped in natural clusters could be accommodated in this way by including random effects (higher-order latent variables) to allow for within-cluster associations. An important instance of this is longitudinal data on dyads, which will be needed for questions about levels and reciprocity of intergenerational help over time. Models for longitudinal data can also be specified in other ways; for example, Steele and Grundy (2021) consider a dynamic (autoregressive) panel model that allows for unequal spacing between the measurements, but simplifying the analysis in another way by reducing giving and receiving help each to a binary variable.

Another straightforward extension of the model is obtained by allowing multiple latent variables for each member of the dyad, each measured by their own multiple indicators. This would be needed, for example, if we wanted to consider different kinds of financial and practical help separately from each other. It is also possible for the same individuals to appear in multiple dyads. For some such cases the structural model would be an obvious extension of the models considered in this paper, for example if we analysed data where survey respondents were asked about help that they exchanged with their children as well as with their parents. In more complex situations, such as for 'round-robin' data where each individual is paired with more than one other individual, the models should include further role-specific latent variables for 'actors' and 'partners' (or 'givers' and 'receivers'), as well as group (e.g. family) effects. This would define multivariate extensions of different versions of the Social Relations Model (Kenny and LaVoie 1984; Snijders and Kenny 1999). Gin et al. (2020) have recently proposed latent-variable formulations for such situations, and our measurement models would add to them the element of zero inflation. These combinations remain to be explored in future research.

APPENDIX: DETAILS OF THE MCMC ALGORITHM

Here we describe the details of the tailored MCMC sampling algorithm for the estimation of the structral model parameters ψ which was outlined in Section 4.3. The algorithm has been packed into an R (R Core Team 2020) package which is included in the supplementary materials; source code for it is available on Siliang Zhang's GitHub page at `https://github.com/slzhang-fd/jsem-ukhls`. The main part of the algorithm was programmed in C++, where two techniques are used to speed up the sampling. First, for sampling steps with non-standard distributions, adaptive rejection sampling (Gilks and Wild 1992) is used, exploiting log-concavity of the posterior density functions. This is used for sampling $\boldsymbol{\psi}_\xi$ and (some of) $\boldsymbol{\eta}$, while $\boldsymbol{\psi}_\eta$ and $\boldsymbol{\xi}$ can be drawn very efficiently from standard distributions. Second, parallel sampling techniques, spread out across multiple processors, are used within each MCMC iteration when there is no dependence between the quantities being sampled. This can be done when sampling the latent variables $\boldsymbol{\zeta}_i$ for different units i, and also when sampling the two subsets of structural parameters $\boldsymbol{\psi}_\xi$ and $\boldsymbol{\psi}_\eta$ separately from each other. This parallelisation is implemented through OpenMP C++ API (Dagum and Menon 1998).

Let $\boldsymbol{\zeta}^{(t)} = (\boldsymbol{\xi}^{(t)}, \boldsymbol{\eta}^{(t)})$ and $\boldsymbol{\psi}^{(t)} = (\boldsymbol{\psi}_\xi^{(t)}, \boldsymbol{\psi}_\eta^{(t)})$ denote the values of the latent variables and the structural parameters sampled in iteration $t = 0, 1, 2, \ldots$, where 0 denotes the initial values. Given $\boldsymbol{\zeta}^{(t-1)}$, $\boldsymbol{\psi}^{(t-1)}$ and the observed data $(\mathbf{Y}, \mathbf{X})$, the values for the next iteration t are sampled as follows:

Imputation step: Generating values for the latent variables $\boldsymbol{\zeta}$, given the observed data and current values of the parameters $\boldsymbol{\psi}$:

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

$$\text{(A1)}\quad p(\xi_G = j, \xi_R = k|\boldsymbol{\eta}^{(t-1)}, \mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\psi}^{(t-1)}) \propto p(\mathbf{Y}_{Gi}|\xi_G = j, \eta_{Gi}^{(t-1)}, \mathbf{X}_i; \tilde{\boldsymbol{\phi}}_G) \times p(\mathbf{Y}_{Ri}|\xi_R = k, \eta_{Ri}^{(t-1)}, \mathbf{X}_i; \tilde{\boldsymbol{\phi}}_R)\, p(\xi_G = j, \xi_R = k|\mathbf{X}_i; \boldsymbol{\psi}_\xi^{(t-1)})$$

for $j, k = 0, 1$, where the structural model for $\boldsymbol{\xi}_i$ is specified by (5), the measurement model is specified as in (1)–(2) for $\mathbf{Y}_{Gi}$ and similarly for $\mathbf{Y}_{Ri}$, and the parameters of the measurement models are fixed at their estimated values $\tilde{\boldsymbol{\phi}}_G$ and $\tilde{\boldsymbol{\phi}}_R$ from the first step of the two-step estimation (as described in Section 4.2) throughout. Note that here the probabilities which involve $\xi_G = 0$ are zero when $\mathbf{Y}_{Gi} \neq \mathbf{0}$, and the ones which involve $\xi_R = 0$ are zero when $\mathbf{Y}_{Ri} \neq \mathbf{0}$. Conversely, when $\mathbf{Y}_{Gi}$ and/or $\mathbf{Y}_{Gi}$ is $\mathbf{0}$, the imputation assigns such a unit i either to the corresponding zero-inflation class 0 or to class 1 for the duration of iteration t.

(2) Sampling $\boldsymbol{\eta}^{(t)}$ from $p(\boldsymbol{\eta}|\boldsymbol{\xi}^{(t)}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi}^{(t-1)})$: Draw $\boldsymbol{\eta}_i^{(t)} = (\eta_{Gi}^{(t)}, \eta_{Ri}^{(t)})$ independently for $i = 1, \ldots, n$, as follows. First, draw $\eta_{Gi}^{(t)}$ from

$$\text{(A2)}\quad p(\eta_G|\eta_{Ri}^{(t-1)}, \boldsymbol{\xi}_i^{(t)}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi}^{(t-1)}) \propto p(\mathbf{Y}_{Gi}|\xi_{Gi}^{(t)}, \eta_G, \mathbf{X}_i; \tilde{\boldsymbol{\phi}}_G)\, p(\eta_G|\eta_{Ri}^{(t-1)}, \mathbf{X}_i; \boldsymbol{\psi}_\eta^{(t-1)})$$

and then $\eta_{Ri}^{(t)}$ from

$$\text{(A3)}\quad p(\eta_R|\eta_{Gi}^{(t)}, \boldsymbol{\xi}_i^{(t)}, \mathbf{Y}, \mathbf{X}, \boldsymbol{\psi}^{(t-1)}) \propto p(\mathbf{Y}_{Ri}|\xi_{Ri}^{(t)}, \eta_R, \mathbf{X}_i; \tilde{\boldsymbol{\phi}}_R)\, p(\eta_R|\eta_{Gi}^{(t)}, \mathbf{X}_i; \boldsymbol{\psi}_\eta^{(t-1)})$$

where the conditional distributions for η_G and η_R on the right-hand sides are the univariate normal distributions implied by (4). When $\xi_{Gi}^{(t)} = 0$, in which case always $\mathbf{Y}_{Gi} = \mathbf{0}$, the probability for $\mathbf{Y}_{Gi}$ in (A2) is 1 by (1), and $\eta_{Gi}^{(t)}$ is generated directly from this normal distribution, whereas adaptive rejection sampling is used when $\xi_{Gi}^{(t)} = 1$; the procedure for $\eta_{Ri}^{(t)}$ is analogous, depending on whether or not $\mathbf{Y}_{Ri}$ is $\mathbf{0}$.

Posterior step: Drawing values for the model parameters $\boldsymbol{\psi}$ from their distributions given the observed data and current imputed values of the latent variables $\boldsymbol{\zeta}$. These are standard posterior distributions of the parameters of regression models for $\boldsymbol{\zeta}$ given $\mathbf{X}$, the bivariate linear model (4) for $\boldsymbol{\eta}$ and the multinomial logistic model (5) for $\boldsymbol{\xi}$. These do not depend on each other, so these sampling steps can be carried out in either order or in parallel.

(3) Sampling $\boldsymbol{\psi}_\eta$ from its posterior distribution $p(\boldsymbol{\psi}_\eta|\mathbf{X}, \boldsymbol{\eta}^{(t)}) \propto p(\boldsymbol{\eta}^{(t)}|\mathbf{X}; \boldsymbol{\psi}_\eta)\, p(\boldsymbol{\psi}_\eta)$. These parameters are handled in two blocks, $\boldsymbol{\beta} = (\boldsymbol{\beta}'_G, \boldsymbol{\beta}'_R)' = \text{vec}(\mathbf{B})$ where $\mathbf{B} = [\boldsymbol{\beta}_G\, \boldsymbol{\beta}_R]$, and $(\sigma^2_G, \sigma^2_R, \rho_{GR})$ which define the conditional covariance matrix in (4), which we denote $\boldsymbol{\Sigma}_\eta$. Here we define the notation specifically as $\mathbf{X} = [\mathbf{X}_1\ \ldots\ \mathbf{X}_n]'$, $\boldsymbol{\eta}_G = (\eta_{G1}, \ldots, \eta_{Gn})'$, $\boldsymbol{\eta}_R = (\eta_{G1}, \ldots, \eta_{Rn})'$ and $\boldsymbol{\eta} = [\boldsymbol{\eta}_G\ \boldsymbol{\eta}_R]$. The bivariate linear model (4) can then be written as $\text{vec}(\boldsymbol{\eta}) \sim N(\text{vec}(\mathbf{XB}), \boldsymbol{\Sigma}_\eta \otimes \mathbf{I}_n)$, where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix.

We specify the prior distribution as $p(\boldsymbol{\psi}_\eta) = p(\boldsymbol{\beta})p(\boldsymbol{\Sigma}_\eta)$, where $p(\boldsymbol{\beta}) \sim N(\mathbf{0}, \sigma^2_\beta \mathbf{I}_{2q})$ with $\sigma^2_\beta = 100$, and $p(\boldsymbol{\Sigma}_\eta) \sim \mathcal{W}^{-1}(\mathbf{I}_2, 2)$, an inverse Wishart prior for $\boldsymbol{\Sigma}_\eta$. This is a 'semi-conjugate' prior for $\boldsymbol{\psi}_\eta$, meaning that conditional on $\boldsymbol{\Sigma}_\eta$, the posterior distribution of $\boldsymbol{\beta}$ is also multivariate normal, and conditional on $\boldsymbol{\beta}$ the posterior of $\boldsymbol{\Sigma}_\eta$ is inverse Wishart. Specifically, $\boldsymbol{\beta}^{(t)}$ is then sampled from the distribution

$$p(\boldsymbol{\beta}|\boldsymbol{\eta}^{(t)}, \mathbf{X}, \boldsymbol{\Sigma}_\eta^{(t-1)}) \sim N(\boldsymbol{\mu}_\beta^{(t)}, \mathbf{V}_\beta^{(t)}) \tag{A4}$$

where

$$\mathbf{V}_\beta^{(t)} = \left(\mathbf{I}_{2q}/\sigma^2_\beta + (\boldsymbol{\Sigma}_\eta^{(t-1)})^{-1} \otimes (\mathbf{X}'\mathbf{X})\right)^{-1} \qquad \text{and} \tag{A5}$$

$$\boldsymbol{\mu}_\beta^{(t)} = \mathbf{V}_\beta^{(t)} \left((\boldsymbol{\Sigma}_\eta^{(t-1)})^{-1} \otimes \mathbf{X}'\right) \text{vec}(\boldsymbol{\eta}) \tag{A6}$$

where $\otimes$ denotes the Kronecker product, and $\boldsymbol{\Sigma}_\eta^{(t)}$ is sampled from

$$p(\boldsymbol{\Sigma}_\eta|\boldsymbol{\beta}^{(t)}, \mathbf{X}, \boldsymbol{\eta}^{(t)}) \sim \mathcal{W}^{-1}(\mathbf{I}_2 + (\boldsymbol{\eta}^{(t)} - \mathbf{XB}^{(t)})'(\boldsymbol{\eta}^{(t)} - \mathbf{XB}^{(t)}), n+2). \tag{A7}$$

(4) Sampling $\boldsymbol{\psi}_\xi^{(t)} = \boldsymbol{\gamma}^{(t)} = (\boldsymbol{\gamma}_{00}^{(t)\prime}, \boldsymbol{\gamma}_{01}^{(t)\prime}, \boldsymbol{\gamma}_{10}^{(t)\prime}, \boldsymbol{\gamma}_{11}^{(t)\prime})'$, where $\boldsymbol{\gamma}_{00}^{(t)} = \mathbf{0}$, from the posterior distribution $p(\boldsymbol{\psi}_\xi|\mathbf{X}, \boldsymbol{\xi}^{(t)}) \propto p(\boldsymbol{\xi}^{(t)}|\mathbf{X}; \boldsymbol{\psi}_\xi)\, p(\boldsymbol{\psi}_\xi)$. This is done using conditional Gibbs sampling, one parameter at a time. We specify the prior distribution as $p(\boldsymbol{\psi}_\xi) \sim N(\mathbf{0}, \sigma^2_\gamma \mathbf{I}_{3q})$, with $\sigma^2_\gamma = 100$. Letting γ_{jkr} denote the rth element of $\boldsymbol{\gamma}_{jk}$, we cycle over all $r = 1 \ldots, q$ and over $(j,k) = (0,1),\ (1,0),\ (1,1)$ to draw $\gamma_{jkr}^{(t)}$ from

$$p(\gamma_{jkr}|\boldsymbol{\gamma}_{(jkr)}^{(t-1)}, \mathbf{X}, \boldsymbol{\xi}^{(t)}) \propto \left[\prod_{i=1}^{n} \frac{\prod_{u,v=0,1} \exp(\boldsymbol{\gamma}_{uvr}^{(t-1)\prime}\mathbf{X}_i)^{\delta_{iuv}^{(t)}}}{\sum_{u,v=0,1} \exp(\boldsymbol{\gamma}_{uvr}^{(t-1)\prime}\mathbf{X}_i)}\right] p(\gamma_{jkr}) \tag{A8}$$

where $\boldsymbol{\gamma}_{uvr}^{(t-1)}$ are vectors where all the γ-parameters except for γ_{jkr} are fixed at their most recently sampled values (from iteration $t-1$ or t), $\boldsymbol{\gamma}_{(jkr)}^{(t-1)}$ denotes all of these fixed parameter values, $\delta_{iuv}^{(t)} = I(\xi_{Gi}^{(t)} = u, \xi_{Ri}^{(t)} = v)$, and $p(\gamma_{jkr})$ is the prior density of γ_{jkr} implied by $p(\boldsymbol{\psi}_\xi)$, in our case $p(\gamma_{jkr}) \sim N(0, 100)$. These $\gamma_{jkr}^{(t)}$ are generated using adaptive rejection sampling.

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

Acknowledgement of the data used in the paper: Understanding Society (UKHLS) is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service.

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 a note which gives additional information on two topics: (1) estimated regression coefficients for the multinomial logistic model from which fitted probabilities are shown in Table 4 of the paper, and (2) some further information about the estimation of the measurement model, to supplement the information in Sections 4.2 and 5.1 of the paper.

REFERENCES

Bakk, Z. and J. Kuha (2018). Two-step estimation of models between latent classes and external variables. *Psychometrika 83*, 871–892.

Bartholomew, D., M. Knott, and I. Moustaki (2011). *Latent Variable Models and Factor Analysis: A Unified Approach* (Third ed.). Chichester: Wiley.

Burt, R. S. (1973). Confirmatory factor-analytic structures and the theory construction process. *Sociological Methods & Research 2*, 131–190.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods & Research 5*, 3–52.

Chan, T. W. (2008). The structure of intergenerational exchange in the UK. University of Oxford, Department of Sociology Working Paper Series 2008-05.

Chan, T. W. and J. Ermisch (2015). Residential proximity of parents and their adult offspring in the United Kingdom, 2009–10. *Population Studies 69*, 355–372.

Cheng, Y.-P., K. S. Birditt, S. H. Zarit, and K. L. Fingerman (2015). Young adults' provision of support to middle-aged parents. *Journals of Gerontology, Series B 70*, 407–416.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica 39*, 829–844.

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

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

de Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. New York: The Guilford Press.

Fingerman, K. L., K. Kim, E. M. Davis, F. F. Furstenberg, Jr, K. S. Birditt, and S. H. Zarit (2015). "I'll give you the world": Socioeconomic differences in parental support of adult children. *Journal of Marriage and Family 77*, 844–865.

Finkelman, M. D., J. G. Green, M. J. Gruber, and A. M. Zaslavsky (2011). A zero- and K-inflated mixture model for health questionnaire data. *Statistics in Medicine 30*, 1028–1043.

Gilks, W. R. and P. Wild (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics 41*, 337–348.

Gin, B., N. Sim, A. Skrondal, and S. Rabe-Hesketh (2020). A dyadic IRT model. *Psychometrika 85*, 815–836.

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. *The British Journal of Sociology 56*, 233–255.

Henretta, J., M. Van Voorhis, and B. Soldo (2018). Cohort differences in parental financial help to adult children. *Demography 55*, 1567–1582.

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

Huang, H.-Y. (2016). Mixture random-effect IRT models for controlling extreme response style on rating scales. *Frontiers in Psychology 7*(1706).

Kalmijn, M. (2014). Adult intergenerational relationships. In J. Treas, J. Scott, and M. Rochards (Eds.), *The Wiley Blackwell Companion to the Sociology of Families*, pp. 385–403. Chichester: Wiley.

Kankaraš, M., J. K. Vermunt, and G. Moors (2011). Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class approaches. *Sociological Methods and Research 40*, 279–310.

Kenny, D. A. and L. LaVoie (1984). The social relations model. *Advances in Experimental Social Psychology 18*, 141–182.

Knies, G. (Ed.) (2018). *Understanding Society: The UK Household Longitudinal Study Waves 1-8. User Guide*. Colchester: Institute for Social and Economic Research, University of Essex.

Künemund, H., A. Motel-Lingebiel, and M. Kohli (2005). Do intergenerational transfers from elderly parents increase social inquality among their middle-aged children? Evidence from the German Ageing Survey. *Journal of Gerontology: Social Sciences 60B*, S30–S36.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics 34*, 1–14.

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

Magnus, B. E. and Y. Liu (2017). A zero-inflated Box-Cox normal unipolar item response model for measuring constructs of psychopathology. *Journal of Educational and Behavioral Statistics 42*, 531–558.

Magnus, B. E. and D. Thissen (2017). Item response modeling of multivariate count data with zero inflation, maximum inflation, and heaping. *Journal of Educational and Behavioral Statistics 42*, 531–558.

Mason, A. and R. Lee (2018). Intergenerational transfers and the older population. In M. Hayward and M. Majmundar (Eds.), *Future directions for the demography of aging: Proceedings of a workshop*. Washington DC: The National Academies Press.

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York: Routledge.

Min, Y. and A. Agresti (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling 5*, 1–19.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics 33*, 341–365.

Muthén, B. and T. Asparouhov (2006). Item response mixture modeling: Application to tobacco dependence criteria. *Addictive Behaviors 31*, 1050–1066.

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

Pickard, L. (2015). A growing care gap? The supply of unpaid care for older people by their adult children in England to 2032. *Ageing and Society 35*, 96–123.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Silverstein, M., S. J. Conroy, H. Wang, R. Giarrusso, and V. L. Bengtson (2002). Reciprocity in parent-child relations over the adult life course. *Journal of Gerontology: Social Sciences 57B*, S3–S13.

Skrondal, A. and S. Rabe-Hesketh (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal,and Structural Equation Models*. Boca Raton, FL: Chapman & Hall / CRC.

Snijders, T. A. B. and D. A. Kenny (1999). The social relations model for family data: A multilevel approach. *Personal Relationships 6*, 471–486.

Steele, F. and E. Grundy (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, Series C 70*, 3–23.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica 26*, 24–36.

University of Essex, Institute for Social and Economic Research, NatCen Social Research, and Kantar Public (2018). Understanding Society: Waves 1–8, 2009–2017 and Harmonised BHPS: Waves1–18, 1991–2009. [data collection]. 11th edition. UK Data Service. SN: 6614. `http://doi.org/10.5255/UKDA-SN-6614-12`.

van der Linden, W. (Ed.) (2016). *Handbook of Item Response Theory*. New York: Chapman and Hall/CRC.

Wall, M. M., J. Guo, and Y. Amemiya (2012). Mixture factor analysis for approximating a nonnormally distributed continuous latent factor with continuous and dichotomous observed variables. *Multivariate Behavioral Research 47*, 276–313.

Wall, M. M., J. Y. Park, and I. Moustaki (2015). IRT modeling in the presence of zero-inflation with application to psychiatric disorder severity. *Applied Psyhological Measurement 39*, 583–597.

Xue, Q.-L. and K. Bandeen-Roche (2002). Combining complete multivariate outcomes with incomplete covariate information: A latent class approach. *Biometrics 58*, 110–120.