Generalized Linear Latent and Mixed Models with Composite Links and Exploded Likelihoods

Anders Skrondal[1] and Sophia Rabe-Hesketh[2]

[1] Norwegian Institute of Public Health, Oslo (anders.skrondal@fhi.no)
[2] University of California, Berkeley

Abstract: Applications of composite links and exploded likelihoods for generalized linear latent and mixed models are explored.

Keywords: Generalized linear latent and mixed models; Composite link; Exploded likelihood.

1 Introduction

Instead of linking the expectation of each observation with a single linear predictor as in generalized linear models, it is often useful to link it with a composite function of several linear predictors. Moreover, each likelihood contribution can sometimes be exploded into a product of terms.
We explore how these tools can be used to extend 'Generalized Linear Latent And Mixed Models' or GLLAMMs (Rabe-Hesketh, Skrondal and Pickles, 2004a; Skrondal and Rabe-Hesketh, 2004). Applications considered include discrete time frailty models, item response models for ordinal items, unfolding models for attitudes, small area estimation with census information, measurement models combining discrete and continuous latent variables, ability testing with guessing, sensitivity analysis of the assumption of normal random effects, and zero-inflated Poisson models.

2 Generalized Linear Models

Let y_i be the response and $\mathbf{x}_i$ explanatory variables for unit i, and define the conditional expectation of the response given the covariates as μ_i, i.e. $\mu_i \equiv \mathrm{E}[y_i|\mathbf{x}_i]$. Generalized linear models can be specified as

$$\mu_i \;=\; g^{-1}(\nu_i),$$

where $g^{-1}(\cdot)$ is an inverse link function, $\nu_i = \mathbf{x}_i'\boldsymbol{\beta}$ is a linear predictor and $\boldsymbol{\beta}$ are fixed effects. The specification is completed by choosing a conditional distribution for the responses y_i given the conditional expectations μ_i, $f(y_i|\mu_i)$, from the exponential family.

3 Exploded likelihoods and composite links

3.1 Exploded likelihoods

Generalized linear models can be extended to handle *multivariate responses* y_{it}, $t=1,\ldots,T$, for each unit. The responses may be of mixed types combining different links and families, for instance a Poisson distributed count and a logistically distributed dichotomous response. Dependence can be modelled by including latent variables (random effects and/or factors) in the linear predictors; see Section 4. Given the corresponding vectors of conditional means $\boldsymbol{\mu}_i$ (which depend on the latent variables), the joint conditional distribution of the vector of responses $\mathbf{y}_i$ is

$$\Pr(\mathbf{y}_i|\boldsymbol{\mu}_i) = \prod_{t=1}^{T} f_t(y_{it}|\mu_{it}). \tag{1}$$

We now distinguish between two types of artificial multivariate responses where the response is univariate but individual likelihood contributions are nevertheless 'exploded' into product terms:

Phantom responses A univariate response y_i can in some cases be represented by S phantom responses y_{it} entering the likelihood (1) as if they were truly multivariate responses.
Phantom responses can be used for the Luce-Plackett model for rankings where the likelihood contribution of a ranking is the product of successive multinomial logit choice probabilities among remaining alternatives (e.g. Skrondal and Rabe-Hesketh, 2003). Another example is survival analysis based on data exploded into risk sets, for instance the Cox proportional hazard model implemented via Poisson regression and the complementary log-log model for discrete time hazards (e.g. Skrondal and Rabe-Hesketh, 2004, Ch.2).

Mutually exclusive responses A univariate response y_i can sometimes be represented by one of S mutually exclusive responses y_{it} having distributions $f_t(y_{it}|\mu_{it})$ from generalized linear models. For the case of $T=2$ the likelihood can be written as

$$\Pr(\mathbf{y}_i|\boldsymbol{\mu}_i) = f_1(y_{i1}|\mu_{i1})^{1-\delta_i} f_2(y_{i2}|\mu_{i2})^{\delta_i},$$

where the indicator δ_i picks out the appropriate component.
A simple example is a log-normal survival model with right-censoring. Let $\mathbf{x}_i'\boldsymbol{\beta}$ be the linear predictor, y_{i1} the log survival time if the event is observed for i ($\delta_i = 0$) and y_{i2} the censoring time if the event is censored ($\delta_i = 1$). The likelihood contribution then becomes either a normal distribution with identity link and linear predictor $\mathbf{x}_i'\boldsymbol{\beta}$, $f_1(y_{i1}|\mu_{i1})=\phi(y_{i1};\mu_i,\sigma^2)$, or a Bernoulli distribution with a (scaled) probit link and linear predictor $\mathbf{x}_i'\boldsymbol{\beta}$, $f_2(y_{i2}|\mu_{i2}) = \Phi(\frac{\mathbf{x}_i'\boldsymbol{\beta}-y_{i2}}{\sigma})$. Here, $\Phi(\cdot)$ is the cumulative standard normal distribution and $-y_{i2}$ is treated as an offset.

3.2 Composite links

Thompson and Baker (1981) suggested linking the expectation μ_i with a composite function of several linear predictors instead of a function of a single linear predictor as in generalized linear models.

Simple composite links In this case the expectation μ_i is a weighted sum of inverse links with known weights w_{ir},

$$\mu_i \;=\; \sum_r w_{ir}\, g_r^{-1}(\nu_{ir}),$$

where ν_{ir} is the rth linear predictor for unit i and $g_r^{-1}(\cdot)$ an inverse link function.
A simple example of composite links are cumulative models for categorical responses with S ordered response categories $s = 1, \ldots, S$, which can be expressed as

$$\Pr(y_i > s|\mathbf{x}_i) \;=\; g^{-1}(\nu_i - \kappa_s), \quad s = 1, \ldots, S-1$$

where κ_s are threshold parameters and the inverse link function is a cumulative distribution function such as the standard normal, logistic or extreme value distributions. The response probabilities can be written as a composite link,

$$\Pr(y_i = s|\mathbf{x}_i) \;=\; g^{-1}(\nu_{i,s-1}) - g^{-1}(\nu_{is}), \quad \nu_{is} = \nu_i - \kappa_s, \quad s = 1, \ldots, S, \tag{2}$$

where $\kappa_0 = -\infty$ and $\kappa_S = \infty$ so that $g^{-1}(\nu_{i0}) = 1$ and $g^{-1}(\nu_{iS}) = 0$. An advantage of the composite link formulation is that left and right-censoring, or even interval censoring of an ordinal response are easily accommodated. This is particularly useful for discrete time survival data.

Bilinear composite links A first extension is to consider unknown linear functions of inverse links, replacing the known constants w_{ir} with products of the constants and unknown parameters α_r, giving

$$\mu_i \;=\; \sum_r \alpha_r w_{ir}\, g_r^{-1}(\nu_{ir}).$$

A second extension is to let the expectation be some (not necessarily linear) function $h\{\cdot\}$ of the above sum,

$$\mu_i \;=\; h\{\sum_r \alpha_r w_{ir}\, g_r^{-1}(\nu_{ir})\}.$$

General composite links In this case general functions $f_{ir}[g_r^{-1}(\nu_{ir})]$ replace $w_{ir}\, g_r^{-1}(\nu_{ir})$ in the above expressions.

4 Generalized Linear Latent and Mixed Models

4.1 Generalized Linear Mixed Models (GLMMs)

A crucial assumption of generalized linear models is that the responses of different units i are independent given the covariates $\mathbf{x}_i$. This assumption is often unrealistic since data are frequently of a multilevel nature with units i nested in clusters j, for instance repeated measurements (units) nested in subjects (clusters) or subjects (units) nested in families (clusters). There will often be unobserved heterogeneity at the cluster level inducing dependence among the units, even after conditioning on covariates. In generalized linear mixed models (e.g. Breslow and Clayton, 1993) unobserved heterogeneity is modeled by including random effects $\eta_{mj}^{(2)}$ in the linear predictor,

$$g(\mu_{ij}) \;=\; \nu_{ij} = \underbrace{\mathbf{x}_{ij}'\boldsymbol{\beta}}_{\text{Fixed part}} \;+\; \underbrace{\sum_{m=1}^{M} \eta_{mj}^{(2)} z_{mij}^{(2)}}_{\text{Random part}} . \tag{3}$$

Here, $\mu_{ij} \equiv \mathrm{E}[y_{ij}|\mathbf{x}_{ij}, \mathbf{z}_{ij}^{(2)}, \boldsymbol{\eta}_j^{(2)}]$ where $\boldsymbol{\eta}_j^{(2)} = (\eta_{1j}^{(2)}, \cdots, \eta_{M,j}^{(2)})'$ are random effects varying at level 2 and $\mathbf{z}_{ij}^{(2)}$ corresponding covariates. Specifically, $\eta_{mj}^{(2)}$ is a random effect of covariate $z_{mij}^{(2)}$ for cluster j, a random intercept if $z_{mij}^{(2)} = 1$. It is typically assumed that the random effects are multivariate normal.

4.2 Extending GLMMs to GLLAMMs

Multilevel factor structures The basic idea of factor or IRT models is that one or more unobserved variables, latent traits or factors 'explain' the dependence between different observed measurements for a subject, in the sense that the measurements are conditionally independent given the factor(s).

A simple example of a unidimensional factor model is the two-parameter logistic item response model often used in ability testing. Examinees j answer test items i, $i = 1, \ldots, I$, giving responses y_{ij} equal to 1 if the answer is correct and 0 otherwise. The probability of a correct response is modelled as a function of the examinee's latent ability η_j,

$$\Pr(y_{ij} = 1|\eta_j) \;=\; \frac{\exp(\nu_{ij})}{1 + \exp(\nu_{ij})}, \qquad \nu_{ij} = \beta_i + \lambda_i \eta_j . \tag{4}$$

The latent ability η_j is assumed to have a normal distribution, λ_i are factor loadings or discrimination parameters (with $\lambda_1 = 1$) signifying how well the items discriminate between examinees with different abilities, and -β_i/λ_i are item 'difficulties'.

We can specify models of this form by extending the two-level generalized linear mixed model in (3) to allow each random effect to be multiplied not just by a single variable but by a linear combination of variables. To obtain the two-parameter logistic item response model, we stack the dichotomous responses y_{ij} into a single response vector and define dummy variables

$$d_{pi} = \begin{cases} 1 & \text{if } p = i \\ 0 & \text{otherwise} \end{cases}$$

The linear predictor of the item response model can then be written as

$$\nu_{ij} = \sum_p d_{pi}\beta_p + \eta_j \sum_p d_{pi}\lambda_p = \beta_i + \eta_j \lambda_i.$$

The linear predictor for a three-level multidimensional factor model can be expressed as

$$\nu_{ijk} = \underbrace{\mathbf{x}'_{ijk}\boldsymbol{\beta}}_{\text{Fixed part}} + \underbrace{\sum_{m_2=1}^{M_2} \eta^{(2)}_{m_2 jk} \boldsymbol{\lambda}^{(2)\prime}_{m_2} \mathbf{z}^{(2)}_{m_2 ijk}}_{\text{Level-2 random part}} + \underbrace{\sum_{m_3=1}^{M_3} \eta^{(3)}_{m_3 k} \boldsymbol{\lambda}^{(3)\prime}_{m_3} \mathbf{z}^{(3)}_{m_3 ijk}}_{\text{Level-3 random part}},$$

where $\mathbf{z}^{(2)}_{m_2 ijk}$ and $\mathbf{z}^{(3)}_{m_3 ijk}$ are vectors of dummy variables with corresponding vectors of factor loadings, $\boldsymbol{\lambda}^{(2)}_m$ and $\boldsymbol{\lambda}^{(3)}_m$. See Rabe-Hesketh, Skrondal and Pickles (2004a) for an application of a multilevel factor model with dichotomous responses.

Discrete latent variables The response model can be further generalized by allowing the latent variables η_j to have discrete distributions. This is useful if the level 2 units are believed to fall into a number of groups or 'latent classes' within which the latent variables do not vary.
If the number of latent classes, or masses, is chosen to maximize the likelihood the nonparametric maximum likelihood estimator (NPMLE) can be achieved (e.g. Rabe-Hesketh, Pickles and Skrondal, 2003), relaxing the assumption of multivariate normal latent variables.

Multilevel structural equations Continuous latent variables (random coefficients and/or factors) can be regressed on covariates (see Section 6) and other latent variables at the same or higher levels, generalizing conventional structural models to a multilevel setting. If the latent variables are discrete, the masses, component weights or latent class probabilities can depend on covariates via multinomial logit models. See Skrondal and Rabe-Hesketh (2004, Ch.4).

5 Composite links and exploded likelihoods in GLLAMMs

An outline is given of some extensions of GLLAMMs arising from plugging in linear predictors with latent variables from GLLAMMs into composite links and exploded likelihoods.

Discrete time frailty models If we let the linear predictor in (2) be $\nu_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \eta_j$ and use a logit link we can obtain a proportional odds model with frailty (see Skrondal and Rabe-Hesketh, 2004, Ch.12).

Item response models for ordinal items Letting the linear predictor in (2) be $\nu_{ij} = \beta_i + \lambda_i \eta_j$ as in the two parameter IRT model (4) and the thresholds be item-specific, we obtain Samejima's graded response model for ordinal items (see Skrondal and Rabe-Hesketh, 2004, Ch.10).

Unfolding or ideal point models In standard item response models the probability of a positive response for an item is a monotonic function of the latent trait η_j. This assumption may be violated for attitude items where respondents are asked to rate their agreement as 'disagree' or 'agree', or more generally in terms of $s=1,\ldots,S$ ordered categories.
For instance, as sentiments favouring capital punishment increase from negative infinity, the probability of agreeing with the statement 'capital punishment seems wrong but is sometimes necessary' initially increases from 0, reaches a maximum when the latent trait is in the 'ambiguous' zone (at the 'ideal point') and then declines as the latent trait goes to infinity.
It has been argued (e.g. Roberts and Laughlin, 1996) that a respondent may give a particular rating of an attitude item for two reasons. Considering 'disagree', he can 'disagree from below' because his latent trait is below the position of the item or 'disagree from above' because it exceeds the position. These two possibilities can be expressed in terms of 'subjective ratings' z_{ij}; such that $z_{ij}=s$ if the respondent 'disagrees from below' and $z_{ij}=2S+1-s$ if he 'disagrees from above'.
Since the z_{ij} are not observed, the probabilities of the observed rating y_{ij}, given the latent trait η_j, can be written as the sum of the probabilities of the two disjunct 'subjective ratings' corresponding to the observed rating. We propose using a cumulative model (2) for the subjective ratings

$$\Pr(y_{ij}=s|\eta_j) \;=\; \Pr(z_{ij}=s|\eta_j) + \Pr(z_{ij}=2S+1-s|\eta_j) \;= \qquad (5)$$
$$\left[g^{-1}(\nu_{ij}-\kappa_{s-1}) - g^{-1}(\nu_{ij}-\kappa_s)\right] + \left[g^{-1}(\nu_{ij}-\kappa_{2S-s}) - g^{-1}(\nu_{ij}-\kappa_{2S-s+1})\right],$$

where $\nu_{ij} = \beta_i + \lambda_i \eta_j$ as in (4). For identification, the thresholds must be constrained as for instance $\kappa_s = -\kappa_{2S-s},\ s=1,\ldots,S$, and $\kappa_S = 0$.
Importantly, embedding the models in the GLLAMM framework produce a wide range of novel unfolding models. The latent trait can for instance be regressed on same or higher level latent variables and/or regressed on covariates as demonstrated in Section 6.

Small area estimation Rindskopf (1992) emphasizes that composite link functions are useful for modelling count data where some observed counts represent sums of counts for different groups of units, due to different kinds of missing or partially observed categorical variables. These ideas have been used by Tranmer et al. (2004) in random effects modeling and empirical Bayes prediction of area specific odds-ratios, for instance for the association between ethnicity and unemployment. They make use of one-way marginal tables from the census 'tabular output', e.g. unemployment rate and ethnic composition, in addition to borrowing strength from other areas as usual in empirical Bayes prediction.

Models combining discrete and continuous latent variables Latent class models can be specified by modeling the 'complete' data (including latent class membership) using log linear models. Since latent class membership is unknown, we must sum over the latent classes to obtain expected counts for the observed response patterns. For a two-class model with three dichotomous observed responses y_i, $i = 1, \ldots, 3$, a log-linear model with conditionally independent responses given latent class membership can be written as

$$\log \mu_{y_1 y_2 y_3 c} = \nu_{y_1 y_2 y_3 c} = \beta_0 + c\alpha_0 + \sum_i y_i \beta_i + \sum_i y_i c \alpha_i,$$

where $c = 0, 1$ is the latent class indicator, $\mu_{y_1 y_2 y_3 c}$ is the expected count for response pattern y_1, y_2, y_3 and latent class c, and β_p and α_p, $p = 0, \ldots, 3$ are parameters. The expected values $\mu_{y_1 y_2 y_3}$ of the observed counts are modeled as the sum of the class-specific expected counts,

$$\mu_{y_1 y_2 y_3} = \exp(\nu_{y_1 y_2 y_3 0}) + \exp(\nu_{y_1 y_2 y_3 1}).$$

Qu, Tan and Kutner (1996) include continuous random effects η_j within a latent class model to relax conditional independence among the responses given latent class membership. To incorporate subject-specific random effects in the model, we expand the data to obtain counts (0 or 1) for each response and latent class pattern for each subject j. The model can then be written as

$$\begin{aligned} \log \mu_{y_1 y_2 y_3 c j} = \nu_{y_1 y_2 y_3 c j} = \quad & \beta_0 + c\alpha_0 + \sum_i y_i \beta_i + \sum_i y_i c \alpha_i \\ & + \eta_j (\sum_i y_i (1-c) \lambda_{i0} + \sum_i y_i c \lambda_{i1}), \end{aligned}$$

where η_j can be interpreted as subject j's propensity to have a '1' (e.g. score positively on a diagnostic test, have a symptom, be diagnosed by a rater), with item-specific effects λ_{i0} for those who are healthy and λ_{i1} for those who have the disease. Since the total count for each person j is fixed at 1, we can estimate the multinomial logit version of this model

$$\Pr(y_1 y_2 y_3 c | j) = \frac{\exp(\nu_{y_1 y_2 y_3 c j})}{\sum_{y_1 y_2 y_3 c} \exp(\nu_{y_1 y_2 y_3 c j})}$$

Again, we do not know c, so the likelihood contribution for subject j becomes

$$\Pr(y_1 y_2 y_3|j) = \frac{\exp(\nu_{y_1 y_2 y_3 0j}) + \exp(\nu_{y_1 y_2 y_3 1j})}{\sum_{y_1 y_2 y_3 c} \exp(\nu_{y_1 y_2 y_3 cj})}.$$

This is a composite link model if each multinomial logit term is viewed as an inverse link. Note that this set-up makes it easy to relax conditional independence among pairs of items by including interaction effects of the form $\beta_{12} y_1 y_2$ in the linear predictors.

Item response models accommodating guessing If it is possible to guess the right answer of an 'item' in ability testing, as when multiple choice questions are used, the two-parameter logistic item response model in (4) is sometimes replaced by the three-parameter model

$$\Pr(y_{ij}=1|\eta_j) \;=\; c_i + (1-c_i)\frac{\exp(\nu_{ij})}{1+\exp(\nu_{ij})}.$$

The c_i are often called 'guessing parameters' and can be interpreted as the probability of a correct answer on item i for an examinee with ability minus infinity.

If we fix the guessing parameters to some common constant w, the response model can be expressed as a generalized linear model with a composite link

$$\Pr(y_{ij}=1|\eta_j) \;=\; w g_1^{-1}(1) + (1-w) g_2^{-1}(\nu_{ij}),$$

where g_1 is the identity link and g_2 is the logit link. If we let $\alpha_1 = w$ be a free parameter, we have a simple example of a bilinear composite link model.

The above kind of model (without latent variables) is said to have 'natural responsiveness' or 'nonzero background' in quantal response bioassay.

Log-normal random effects If the random effects distribution is skewed, we may want to specify a linear mixed model with log-normal random effects

$$\mu_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \exp(\eta_{1j}) + \exp(\eta_{2j}) z_{ij},$$

which can be accomplished using the composite link

$$\mu_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \exp(\eta_{1j}) + \exp(\eta_{2j} + \log(z_{ij})).$$

This is also a useful way of conducting a sensitivity analysis of the conventional normality assumption for the random effects. Using the GLLAMM formulation, we can also have log-normal common factors.

If we use a bilinear composite link, we can include log-normal random effects in generalized linear mixed (and item response) models as well,

$$\mu_{ij} = h[\mathbf{x}'_{ij}\boldsymbol{\beta} + \exp(\eta_{1j}) + \exp(\eta_{2j} + \log(z_{ij}))].$$

Zero-inflated Poisson (ZIP) models The likelihood of ZIP models can be expressed using a combination of composite links and exploded likelihoods.

The ZIP model is a finite mixture model for counts where the population is assumed to consist of two components, a component $c=0$ where the count can only be zero and a component $c=1$ where the count has a Poisson distribution. The probability of belonging to the zero-count component is modelled as

$$\pi_{i0} = \frac{\exp(\mathbf{z}_i'\boldsymbol{\gamma})}{1+\exp(\mathbf{z}_i'\boldsymbol{\gamma})} \tag{6}$$

and the Poisson distribution for the other component is

$$\Pr(y_i=k|\mathbf{x}_i, c_i=1) = \exp(-\mu_i)\mu_i^k/k!, \quad \mu_i = \exp(\mathbf{x}_i'\boldsymbol{\beta}). \tag{7}$$

The probability of a non-zero count becomes

$$\begin{aligned}\Pr(y_i=k>0|\mathbf{z}_i,\mathbf{x}_i) &= \Pr(y_i=k>0, c_i=1) = (1-\pi_{i0})\exp(-\mu_i)\mu_i^k/k! \\ &= \left(\frac{1}{1+\exp(\mathbf{z}_i'\boldsymbol{\gamma})}\right)\left[\exp(-\mu_i)\mu_i^k/k!\right]\end{aligned}$$

and the probability of a zero count

$$\begin{aligned}\Pr(y_i=0|\mathbf{z}_i,\mathbf{x}_i) &= \Pr(y_i=0, c_i=0|\mathbf{z}_i,\mathbf{x}_i) + \Pr(y_i=0, c_i=1|\mathbf{z}_i,\mathbf{x}_i) \\ &= \pi_{i0} + (1-\pi_{i0})\exp(-\mu_i) \\ &= \left(\frac{1}{1+\exp(\mathbf{z}_i'\boldsymbol{\gamma})}\right)\left[\exp(\mathbf{z}_i'\boldsymbol{\gamma}) + \exp(-\exp(\mathbf{x}_i'\boldsymbol{\beta}))\right].\end{aligned}$$

For a non-zero count, the probability is the product of the probability of 0 in a logistic regression model with linear predictor $\mathbf{z}_i'\boldsymbol{\gamma}$ and the Poisson probability of a count k with a log link and linear predictor $\mathbf{x}_i'\boldsymbol{\beta}$. Therefore, for non-zero counts, we obtain the correct likelihood by creating two responses, 0 and k and specifying a mixed response (logistic and Poisson) model.

For a zero count, we again create a 0 response, modelled as a logistic regression, for the first term. For the second term, we specify a composite link,

$$[\exp(\mathbf{z}_i'\boldsymbol{\gamma}) + \exp(-\exp(\mathbf{x}_i'\boldsymbol{\beta}))] = g_1^{-1}(\mathbf{z}_i'\boldsymbol{\gamma}) + g_2^{-1}(\mathbf{x}_i'\boldsymbol{\beta}),$$

where g_1 is the log link and g_2 the log-log link. If we create a 1 response and specify a Bernoulli distribution with this composite link, we obtain the required term.

This set-up also makes it fairly straightforward to include random effects in ZIP models to capture dependence induced by clustered data. For instance, in modeling the number of alcoholic drinks consumed by respondents nested in regions, we could include region-specific random effects in both (6) and (7) to model variations in the prevalence of non-drinking and in the amount consumed among drinkers, with possible correlations between these random effects.

6 Unfolding attitudes to female work participation

In the 1988 and 2002 General Social Surveys respondents in the USA were presented with the following attitude statements regarding female work participation:

[famhapp] A woman and her family will all be happier if she goes to work

[twoincs] Both the husband and wife should contribute to the family income

[warmrel]: A working mother can establish just as warm and secure a relationship with her children as a woman who does not work

[jobindep] Having a job is the best way for a woman to be an independent person

[housewrk] Being a housewife is just as fulfilling as working for pay

[homekid] A job is alright, but what most women really want is a home and children

[famsuff] All in all, family life suffers when the woman has a full-time job

[kidsuff] A pre-school child is likely to suffer if his or her mother works

[hubbywrk] A husband's job is to earn money; a wife's job is to look after the home

The respondents rated each statement as either 'disagree completely' (1), 'disagree' (2), 'agree somewhat' (3), 'agree' (4), or 'agree completely' (5). In 2002, the 'disagree completely' and 'disagree' response options were collapsed into a single 'disagree' option.
We use the unfolding model proposed in Section 5, with g as scaled probit links with item-specific scale parameters σ_i (estimated on the log-scale),

$$g^{-1}(\nu_{ijs}) \;=\; \Phi^{-1}\left(\frac{\beta_i + \lambda_i\eta_j - \kappa_s}{\sigma_i}\right).$$

In 2002, the composite link for 'disagree' is the sum of the composite links for 'disagree' and 'disagree completely'.
To investigate if sentiments in favour of female work participation η_j (loosely referred to as 'feminism') have changed from 1988 to 2002, we specify the structural model

$$\eta_j \;=\; \gamma_1 w_j + \zeta_j, \qquad \zeta_j \sim \mathrm{N}(0,\psi),$$

where w_j is a dummy variable for year being [2002].
Maximum likelihood estimates based on data from 1462 respondents are given in Table 1 where the items have been ordered from the most positive to the most negative according to their estimated scale values $\widehat{\beta}_i$. Since the magnitude of $\widehat{\gamma}_1$ is negligible, mean 'feminism' does not appear to have changed.

TABLE 1. Estimates for scaled probit unfolding model

	Item parameters					
	β_i		λ_i		$\ln \sigma_i$	
Item i	Est	SE	Est	SE	Est	SE
[famhapp]	-2.32	0.08	0.30	0.04	-0.24	0.05
[twoincs]	-1.60	0.07	0.29	0.05	-0.06	0.05
[warmrel]	-0.99	0.07	1	–	0	–
[jobindep]	-0.27	0.14	1.15	0.15	0.64	0.05
[housewrk]	1.29	0.08	0.54	0.08	0.22	0.06
[homekid]	2.11	0.07	0.76	0.06	-0.06	0.04
[famsuff]	2.19	0.08	1.43	0.09	-0.29	0.05
[kidsuff]	2.24	0.08	1.49	0.09	-0.46	0.06
[hubbywrk]	2.42	0.09	1.14	0.09	-0.11	0.05
	Thresholds $-\kappa_s = \kappa_{2S-s}$					
s (categories)			Est	SE		
1 (‘disagree completely’/‘disagree’)			3.43	0.11		
2 (‘disagree’/‘agree somewhat’)			2.36	0.08		
3 (‘agree somewhat’/‘agree’)			1.67	0.06		
4 (‘agree/‘agree completely’)			0.72	0.03		
	Latent trait regression					
			Est	SE		
		[2002] γ_1	-0.04	0.04		
		Variance ψ	0.62	0.08		

Following Roberts and Laughlin (1996) we assess model fit graphically. First, we estimate the position or ‘dominance’ $\widetilde{\nu}_{ij}$ of respondent j relative to item i (how much more ‘feminist’ the respondent is than the item) by plugging in the empirical Bayes prediction $\widetilde{\eta}_j$ of the latent trait and the parameter estimates into the linear predictor. Substituting this into the unfolding model, we obtain the expected response category for each person-item pair. Grouping the $\widetilde{\nu}_{ij}$ into approximately homogeneous groups of size 30 for each item and plotting the corresponding average observed and expected frequencies versus the average $\widetilde{\nu}_{ij}$ for each item gives Figure 1. Our unfolding model appears to fit quite well.

Although the expected response takes the form of a single-peaked function consistent with an unfolding process when all items are considered together, none of the individual items exhibit single-peaked behaviour with the possible exception of [jobindep]. Using conventional item response models that assume monotonicity might therefore be appropriate if either (1) reversing the coding of the appropriate items can be based on a priori information or (2) the model accommodates negative factor loadings.

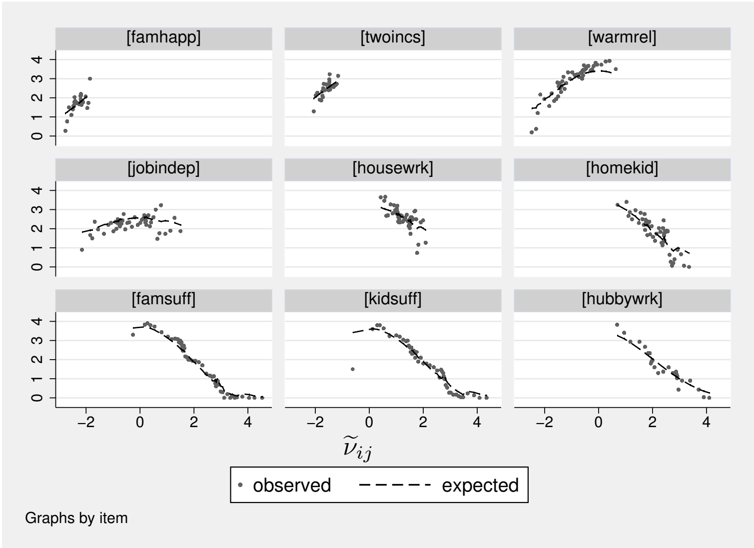

FIGURE 1. Mean expected and observed responses as a function of 'dominance' $\widetilde{\nu}_{ij}$ of person j over item i

7 Conclusions

Although simple to implement, composite links and exploded likelihoods have been demonstrated to be remarkably powerful tools for specifying novel GLLAMMs. Indeed, we do not purport to exhaust potential applications in this paper.

A further useful extension would be to generalize the traditional composite links suggested by Thompson and Baker (1981) to accommodate products of inverse links. A simple variant is of the form

$$\mu_i \;=\; \sum_r \alpha_r \prod_t g_{rt}^{-1}(\nu_{irt}).$$

A composite link with products can be used for additive relative risk models with random effects. The risk or rate parameter μ_{ij} in the Poisson distribution is specified as

$$\mu_{ij} \;=\; \exp(\beta_0 + \eta_j)[1 + \mathbf{x}_{ij}'\boldsymbol{\beta}],$$

where $\mathbf{x}_{ij}$ does not include a 1 and $\boldsymbol{\beta}$ correspondingly not a constant. Note that the baseline risk when $\mathbf{x}_{ij} = \mathbf{0}$ becomes $\exp(\beta_0 + \eta_j) > 0$. It follows that the 'relative risk' RR_{ij}, the risk when the covariate vector is $\mathbf{x}_{ij}$ relative to the baseline risk, is

$$\mathrm{RR}_{ij} \;=\; 1 + \mathbf{x}_{ij}'\boldsymbol{\beta},$$

an additive function of the covariates.

Maximum likelihood estimation and of GLLAMMs and empirical Bayes prediction using adaptive quadrature (e.g. Rabe-Hesketh, Skrondal and Pickles, 2004b) are implemented in the `gllamm` software running in `Stata`. See `http://www.gllamm.org` for further information.

References

Breslow, N.E., and Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**, 9-25.

Qu, Y., Tan, M., and Kutner, M.H. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, **52**, 797-810

Rabe-Hesketh, S., Pickles, A., and Skrondal, A. (2003). Correcting for covariate measurement error in logistic regression using nonparametric maximum likelihood estimation. *Statistical Modelling*, **3**, 215-232.

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004a). Generalized multilevel structural equation modeling. *Psychometrika*, in press.

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004b). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics*, in press.

Rindskopf, D. (1992). A general approach to categorical data analysis with missing data, using generalized linear models with composite links. *Psychometrika*, **57**, 29-42.

Roberts, J.S., and Laughlin, J.E. (1996). A unidimensional item response model for unfolding responses from a graded disagree-agree response scale. *Applied Psychological Measurement*, **20**, 231-255.

Skrondal, A., and Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, **68**, 267-287.

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

Thompson, R., and Baker, R.J. (1981). Composite link functions in generalized linear models. *Journal of the Royal Statistical Society, Series C*, **30**, 125-131.

Tranmer, M., Pickles, A., Fieldhouse, E., et al. (2004). The case for small area microdata. *Journal of the Royal Statistical Society, Series A*, in press.