The PDF of this article has been modified from its original version.

On page 140, Equation (1) has been changed to

$$g(\gamma_{crij}) = \alpha_{cr} + \boldsymbol{\beta}_r^T \mathbf{x}_{rij} + \boldsymbol{\lambda}_r^{(1)^T} \mathbf{v}_{ij}^{(1)} + \boldsymbol{\lambda}_r^{(2)^T} \mathbf{v}_j^{(2)} + u_{rj}, \quad c = 0, \ldots, C_r - 1 \; .$$

On page 141, the unnumbered equation has been changed to

$$\lambda_r^{(l)*} = \frac{\lambda_r^{(l)}}{\sqrt{Var(y_{rij}^*)}} \; .$$

On page 143, the equation that appears five lines from the top of the pages as been changed to

$$\boldsymbol{\beta}_r^T \mathbf{x}_{rij} + \boldsymbol{\lambda}_r^{(1)^T} \mathbf{v}_{ij}^{(1)} + \boldsymbol{\lambda}_r^{(2)^T} \mathbf{v}_j^{(2)} + u_{rj} \; .$$

Sociological Methods & Research
Volume 35 Number 1
August 2006 137-153

10.1177/0049124106289112
http://smr.sagepub.com
hosted at
http://online.sagepub.com

A Multilevel Factor Model for Mixed Binary and Ordinal Indicators of Women's Status

Fiona Steele
Harvey Goldstein
University of Bristol, UK

The authors present a factor model for the analysis of categorical responses with a two-level hierarchical structure. The model allows for multiple, potentially correlated factors at each level as well as covariate effects on the responses. Estimation using Markov chain Monte Carlo (MCMC) methods is described. The methodology is applied in an analysis of women's status in Bangladesh. Two dimensions of women's status are considered—social independence and decision-making power—and the factor structure of the responses at the woman and district levels is explored.

Keywords: *multilevel factor analysis; ordinal response data; women's status*

The low status of women has been cited as a major obstacle to demographic change in Bangladesh, where women's lives are constrained by the norms of patriarchal society and poverty (Balk 1994). Previous research has found that low status, characterized by a lack of social independence and autonomy within the household, is associated with low uptake of family planning and health services, which has led to high fertility and poor health outcomes (Amin, Diamond, and Steele 1997; Balk 1994; Schuler, Hashemi, and Riley 1997; Steele, Amin, and Naved 2001). The social constraints faced by women have long been recognized in the design of health and social programs. Health initiatives, such as the doorstep delivery of family planning and mobile immunization clinics, have been designed to take into account women's lack of mobility, while interventions such as micro-credit schemes have sought to improve women's position within their household and in the wider public sphere (see, e.g., Schuler et al. 1997; Steele et al. 2001).

Women's status is an abstract multidimensional concept. Balk (1994) distinguishes between four aspects of women's empowerment: the extent to which a woman can move about freely and autonomously in public, her perception of what her family will permit her to do, her participation in household decisions, and her attitudes concerning women's rights. Demographic surveys, such as those carried out under the Demographic and Health Surveys (DHS) program, measure women's status indirectly using a set of questions that map loosely onto the above dimensions. For example, a woman might be asked who in the household makes decisions about children's schooling: she alone, the woman and her husband jointly, or her husband alone. The responses to these questions are usually analyzed by constructing an index for each dimension, most commonly by assigning a numeric code to each response alternative and taking the total of the responses across the different questions. The questions may be given equal weight (e.g., Balk 1994), or greater weight may be assigned to indicators thought to have particularly strong discriminatory power (Schuler et al. 1997). Rather than arbitrarily choosing weights a priori, a preferred approach is to use some form of factor analysis in which weights are estimated from the data.

In factor analysis, it is assumed that the correlation between responses on a set of observed variables is due to their shared dependency on one or more latent variables or factors. In an analysis of women's status, a separate factor might represent each underlying dimension. As women's status indicators are categorical, most often a mixture of binary and ordinal measures, it is inappropriate to use standard factor models that assume multivariate normality. Jöreskog and Moustaki (2001) describe alternative factor models for ordinal responses, and Moustaki (2003) proposes a generalized latent trait model for the analysis of mixed binary and ordinal responses. A review of factor analysis for categorical response data can be found in Bartholomew et al. (2003). While Amin et al. (1997) and Steele et al. (2001) analyzed binary status indicators using latent class analysis and latent trait analysis to obtain, respectively, categorical and continuous indices, no previous study has properly accounted for ordinal indicators of women's status.

Another issue that needs to be considered in the analysis of women's status indicators is the potential for geographical clustering in individual responses. Clustering of responses may occur due to the presence of unobserved area characteristics, such as cultural norms and religiosity, which influence the position of women. In Bangladesh, Amin et al. (1997) suggest that the shared cultural identity of people living in the same district is manifest in attitudes and behavioral patterns. Research has found that districts vary in their levels of social conservatism and their attitudes toward nongovernmental organizations (NGOs) that work with women, with eastern districts showing resistance toward NGO activity (Amin, Basu, and Stephenson 2001). Multilevel factor analysis (e.g., Goldstein and Browne 2002) can be used to control for and explore area effects on women's empowerment. In a multilevel factor model, there may be multiple factors at both the individual and area levels. Thus, it is possible to explore contextual effects and the dimensionality of women's status at each hierarchical level. Multilevel factor models have recently been developed for the analysis of binary and ordinal responses (Goldstein, Bonnet, and Rocher forthcoming).

In this article, we present a practically useful methodology for simultaneously handling multiple factor structures at two hierarchical levels with covariates, where the responses are categorical. The methodology is applied in an analysis of women's status in Bangladesh. Using data from a nationally representative survey, we study two aspects of a woman's lifestyle: her social independence and her role in decisions on family matters. We extend previous research on women's status in several ways. First, of the few studies that have used factor analysis to analyze women's status indicators, separate models are fitted for each dimension; we analyze two dimensions jointly and estimate their correlation. Second, no study has used multilevel modeling to explore area effects on women's status. Third, using a multilevel factor model, we explore the factor structure at both the individual and district levels. For example, one question of interest is whether there exists just one district-level factor, representing "overall" status, or one factor for each dimension, as found in previous individual-level analyses. Finally, we consider the effects of observed background characteristics on the status indicators, after controlling for the effects of the common factors.

Method

The Multilevel Factor Model for Binary and Ordinal Responses

A two-level factor model is described in the context of our application where women (Level 1 units) are nested within districts (Level 2).

We denote by y_{rij} the response on variable r for woman i in district j ($r = 1, \ldots, R$; $i = 1, \ldots, n_j$; $j = 1, \ldots, J$). Variable r has $C_r + 1$ response categories (coded 0, 1, ..., C_r) and may be binary ($C_r = 1$) or ordinal ($C_r > 1$). The probability of being in response category c on variable r is denoted by π_{crij}. We expect a woman's response on variable r to depend on a vector of observed covariates $\mathbf{x}_{rij}$, her unobserved overall status (represented by a vector of individual-level factors $\mathbf{v}_{ij}^{(1)}$), and the unobserved status of women living in the same district (represented by a vector of district-level factors $\mathbf{v}_{j}^{(2)}$). All vectors are of arbitrary length, although, as in any factor model, the number of factors that can be identified will depend on the number of items R (see Bartholomew et al. 2003:155). A factor model for the cumulative response probabilities $\gamma_{crij} = P(y_{rij} \leq c)$ may be written as follows:

$$g(\gamma_{crij}) = \alpha_{cr} + \boldsymbol{\beta}_r^T \mathbf{x}_{rij} + \boldsymbol{\lambda}_r^{(1)^T} \mathbf{v}_{ij}^{(1)} + \boldsymbol{\lambda}_r^{(2)^T} \mathbf{v}_{j}^{(2)} + u_{rj}, \quad c = 0, \ldots, C_r - 1 \tag{1}$$

where $g(.)$ is the link function. The model for binary responses is a special case of (1) with $C_r = 1$, where the category coded 1 is taken as the reference category. In this article, we use a probit link, which has certain computational advantages and also has a useful interpretation in terms of an underlying normal "propensity" distribution for the responses. For example, for a response variable r with $C_r = 2$, we can suppose that there is an underlying continuous latent variable y_{rij}^* with thresholds given by $\beta_{0r} + \alpha_{cr}$, such that

$$y_{rij} = \begin{cases} 0 & \text{if } y_{rij}^* \leq \beta_{0r} + \alpha_{0r} \\ 1 & \text{if } \beta_{0r} + \alpha_{0r} < y_{rij}^* \leq \beta_{0r} + \alpha_{1r} \\ 2 & \text{if } y_{rij}^* > \beta_{0r} + \alpha_{1r}, \end{cases}$$

where β_{0r} is an overall intercept, the first element of $\boldsymbol{\beta}_r$, and $\alpha_{0r} = 0$.

The effects of the covariates, as well as of the Level 1 and 2 factors (the factor "loadings"), are denoted by $\boldsymbol{\beta}_r$, $\boldsymbol{\lambda}_r^{(1)}$, and $\boldsymbol{\lambda}_r^{(2)}$. The factors are assumed to follow multivariate normal distributions, with covariances at Levels 1 and 2 denoted by $\boldsymbol{\Omega}_{v(1)}$ and $\boldsymbol{\Omega}_{v(2)}$. Factors at the same hierarchical level may be correlated. The model also includes district effects u_{rj}, sometimes referred to as "unique" or "specific" factors, that are specific to each response r. We assume $u_{rj} \sim N(0, \sigma_{ur}^2)$ with $\mathrm{Cov}(u_{rj}, u_{r'j}) = 0$ for $r \neq r'$. Thus, it is assumed that the correlation between a pair of responses y_{rij} and $y_{r'ij}$ is due to their mutual dependence on the individual- and district-level "common" factors $\mathbf{v}_{ij}^{(1)}$ and $\mathbf{v}_{j}^{(2)}$; conditional on these factors, the responses are independent.

As is usual in factor analysis, constraints on either the factor loadings or the factor variances are required to fix the scale of the factors at each level. Here, we constrain the factor variances to equal 1. An alternative is to fix one loading for each factor to 1, which constrains each factor to have the same scale as one of the responses.

Standardized Loadings and Covariate Effects

If all factors are constrained to have the same scale, with unit variance in our case, the loadings for a given response *r* can be compared across factors at the same or different levels. However, because the underlying variables y_{rij}^{*} have different scales, loadings cannot be compared across responses. Specifically, if a probit link is used in (1), the variance of the underlying response *r* is given by

$$\text{Var}(y_{rij}^{*}) = \boldsymbol{\lambda}_r^{(1)^T} \boldsymbol{\Omega}_{v(1)} \boldsymbol{\lambda}_r^{(1)} + \boldsymbol{\lambda}_r^{(2)^T} \boldsymbol{\Omega}_{v(2)} \boldsymbol{\lambda}_r^{(2)} + \boldsymbol{\sigma}_{ur}^2 + 1. \tag{2}$$

The last two terms in (2) are, respectively, the residual Level 2 and Level 1 variances in y_{rij}^{*}, after accounting for the common factors at each level. A unit residual Level 1 variance (i.e., the within-district, between-woman variance in an underlying response) arises from use of the probit link. To obtain loadings that are comparable across responses, standardized loadings may be calculated as

$$\lambda_r^{(l)*} = \frac{\lambda_r^{(l)}}{\sqrt{Var(y_{rij}^{*})}},$$

where $\lambda_r^{(l)}$ is the vector of loadings for factors at Level l (l = 1, 2) for response *r*. These standardized loadings can be interpreted as (partial) pairwise correlations between y_{rij}^{*} and each of the Level l factors. Standardized covariate effects may be computed in a similar way.

Estimation

Multilevel factor models may be estimated in a number of ways. We begin with a brief review of maximum likelihood methods. Until recently, multilevel factor models were usually estimated in two stages. The advantage of two-step procedures is that they are computationally efficient and easily implemented in a range of software packages. Goldstein (2003) describes one such a method: At the first stage, a multivariate multilevel model is fitted to the responses to obtain estimates of the within- and between-cluster covariance matrices, possibly adjusting for covariate effects; the second step of the procedure is to analyze these matrices separately using any structural equation modeling (SEM) software. A further advantage of this approach is that missing responses can be handled straightforwardly, under a missing-at-random assumption. Another two-step approach (Muthén 1994) involves first estimating the within- and between-community covariance matrices, which are then analyzed simultaneously using procedures for multigroup analysis available in standard SEM software. The two-stage maximum likelihood methods outlined above assume multivariate normality of the response variables or of the latent continuous variables underlying binary or ordered responses. A more general model for mixed response types, estimated via adaptive quadrature in a single step and implemented in the Stata program GLLAMM, is proposed by Rabe-Hesketh, Skrondal, and Pickles (2004).

In this article, the multilevel factor model for binary and ordinal responses in (1) is estimated in a single step using Markov chain Monte Carlo (MCMC) methods. MCMC methods are increasingly being used for the estimation of multilevel models (see Browne 2003 for an introduction). The major reason for their popularity is that MCMC algorithms consist of a set of distinct steps that can be more readily extended to handle complex structures than can maximum likelihood methods. Other advantages of MCMC methods are the facility to generate the posterior distribution of any function of the unknown parameters (e.g., standardized factor loadings) and the ability to incorporate prior beliefs about parameter values. Goldstein and Browne (2002) outline estimation of a factor model for normal response data using Gibbs sampling. Estimation of a model for binary responses is described in Goldstein and Browne (2005), and their method is extended by Goldstein et al. (forthcoming) to include ordinal responses.

Estimation of the factor model for mixed binary and ordinal responses is described in detail in Goldstein et al. (2004). We give a brief outline of the procedure here. Albert and Chibb (1993) propose an MCMC algorithm for estimation of a single-level probit model for binary and polychotomous data, which was later adapted by Goldstein and Browne (2005) to estimate a multilevel factor model for binary responses. Using their approach, a binary or ordinal response y_{rij} is viewed as a threshold from an underlying latent continuous response y^*_{rij}. If we knew the value of y^*_{rij}, we could use the standard Gibbs sampling algorithm for normal response models. We therefore add an extra step to the Gibbs sampling algorithm for continuous data to generate y^*_{rij} at each iteration from its posterior distribution conditional on current parameter values, which is a truncated normal distribution with mean $\boldsymbol{\beta}_r^T \mathbf{x}_{rij} + \boldsymbol{\lambda}_r^{(1)^T} \mathbf{v}_{ij}^{(1)} + \boldsymbol{\lambda}_r^{(2)^T} \mathbf{v}_j^{(2)} + u_{rj}$ and variance 1. A further step is included to sample the threshold parameters α_{cr}, conditional on current parameter values, ensuring that for ordinal responses, the order of relationships among the thresholds is preserved, that is, $\alpha_{cr} \geq \alpha_{c-1,r}$ for $c = 1, \ldots, C_r - 1$. A lower bound for α_{0r} is the maximum value of y^*_{rij} for all individuals with $y_{rij} = 0$. For $c = 1, \ldots, C_r - 2$, a lower bound for α_{cr} is $\max\{\alpha_{c-1,r}, \max(y^*_{rij}$ for $y_{rij} = c)\}$, and an upper bound is $\min\{\alpha_{c+1,r}, \min(y^*_{rij}$ for $y_{rij} = c+1)\}$. An upper bound for $\alpha_{C_r-1,r}$ is simply $\min(y^*_{rij}$ for $y_{rij} = C_r)$. Albert and Chibb (1993) show that sampling of each α_{cr} is from a uniform distribution with intervals defined by these bounds. Having generated y^*_{rij} and α_{cr}, the remaining steps broadly follow those outlined by Goldstein and Browne (2002) for continuous response data.

Gibbs sampling is used, except for the case of correlated factors where Metropolis-Hastings sampling is used to obtain estimates of the factor covariances. Some authors (e.g., Lee and Zhu 2000) have noted that Metropolis-Hastings sampling typically produces chains with better mixing properties for the threshold parameters. These parameters are not the principal concern of this article, and we have not pursued this further. Default starting values are to set factor scores to a random sample from N(0,1) and factor loadings to zero. When assessing convergence in factor models estimated using MCMC, it is especially important to examine the factor loading chains for evidence of "flipping states," where the factor values and loadings switch signs between iterations, and convergence is therefore never reached (Browne 2003; Goldstein and Browne 2002). It is also advisable to try different starting values for the loadings. Fixed coefficient starting values are estimated from overall response proportions, assuming a model with intercept terms only. As we have no prior information on likely parameter values, we have incorporated suitable "diffuse" prior distributions in the model. The basic code for the algorithm is written in MATLAB (Mathworks 2004) and will be incorporated into MLwiN (Rasbash et al. 2004; Browne 2003) by extending the existing factor-fitting procedures.

Analysis of Women's Status in Bangladesh

Data

We analyze data from the Bangladesh Fertility Survey of 1989 (Huq and Cleland 1990), a nationally representative survey of 11,905 ever-married women ages 15 to 49. As several of the woman's status indicators used in the analysis refer to a husband's participation in household decisions, the analysis is restricted to the 10,871 women who were married at the time of interview. Six women with missing responses and 50 whose religion was neither Muslim nor Hindu were also excluded. The final analysis sample contains 10,815 women who are nested in 60 districts.

We consider two dimensions of women's status, which we refer to as *social independence* and *decision making*. Social independence is measured by seven binary variables, based on a woman's responses to questions about whether she is able to do the following activities alone: go to any part of her village/town/city, go outside the village/town/city, talk to a man she does not know, go to the cinema or a cultural show, go shopping, attend a cooperative or a social club, or visit a health center. A woman's role in family decision making is measured by five categorical variables. Women were asked who made decisions on the following matters: whether children go to school, visits to relatives or friends, household purchases, use of family planning, and seeking treatment for a sick family member. Each variable has three response alternatives, which we treat as ordinal: Decisions may be made by the husband alone, the woman jointly with her husband, or the woman alone. Table 1 shows the distribution of all 12 women's status indicators. The responses reveal a general lack of social mobility and autonomy. While most women can move freely around their locality and talk to men they do not know, few can go shopping or visit a health center unaccompanied. The majority of women participate to some extent in family decisions, but the proportion that can make decisions on their own is low.

We also consider the effects on different aspects of women's lifestyles of several background characteristics: current age, level of education (no schooling versus some education), religion (Hindu or Muslim), and type of region of residence (urban or rural). Previous demographic research in Bangladesh has found that these variables are often predictors of women's status and a range of behaviors. For example, Balk (1994) finds that younger women are treated more leniently by their families than older women but have a more limited role in household

Table 1
Percentage Distribution for Each Response ($N = 10{,}815$)

Decision Making	Husband Only	Joint	Woman Only
Children's education	28.6	66.7	4.8
Visits to family/friends	45.1	46.6	8.3
Household purchases	43.4	43.5	13.0
Family planning	23.0	70.8	6.2
Medical treatment	46.3	44.5	9.2

Social Independence	No	Yes
Go out locally	21.3	78.7
Go outside village/town	66.6	33.4
Talk to unknown man	23.2	76.8
Go to cinema/show	57.3	42.7
Go shopping	87.4	12.6
Go to cooperative/club	84.7	15.3
Go to health center	84.8	15.2

decision making. Balk also reports that Muslim women have lower status than Hindus on all dimensions.

Model Selection

We considered three specifications of the multilevel factor model, with different assumptions made about the factor structure at each level. All models include covariate effects on the response variables. The first model (Model 1) includes a single factor at both the woman and district levels. Model 2 is an extension of this model with a second factor at the woman level. In the two-factor model, a simple structure was imposed on the factors. For one of the factors, the loadings for the social independence items were constrained to equal zero, while the loadings for the decision-making items were estimated freely; for the other factor, this pattern of constraints was reversed. The two factors were permitted to have a nonzero correlation. In the third model considered, a second factor was included at the district level. As for the two factors at the woman level, the district-level factors were assumed to have simple structure, and the between-factor covariance was estimated.

The three models were compared using the deviance information criterion (DIC) (Spiegelhalter et al. 2002). The DIC statistic is a Bayesian

Table 2
Mean Deviance ($\bar{D}$), Effective Number of Parameters (p_D), and Deviance Information Criterion (DIC) Goodness-of-Fit Statistic for Two-Level Factor Models

Model	Description	$\bar{D}$	p_D	DIC
1	One factor at each level	123,288	9204	132,492
2	Two correlated factors at woman level; one factor at district level	99,275	16,382	115,657
3	Two correlated factors at each level	99,260	16,371	115,631

analogue of the likelihood-based Akaike information criterion (AIC). Like the AIC, the DIC measures the trade-off between model fit and model complexity. The DIC is calculated as $\bar{D} + p_D$, where $\bar{D}$ is the mean deviance of the chain, the average of the values of the deviance obtained at each iteration, and p_D is the "effective" number of parameters. p_D is calculated as $\bar{D} - D(\bar{\theta})$, where $D(\bar{\theta})$ is the deviance calculated at the mean values of the unknown parameters θ. The model with the smallest DIC is deemed the "best" fitting. Table 2 shows the DIC values, together with $\bar{D}$ and p_D, for the three models described above. Model 3 has the smallest DIC value, although it should be noted that the drop in deviance is substantially larger when adding a second factor at the woman level (Model 2 vs. 1) compared to adding a factor at the district level (Model 3 vs. 2). We therefore select this model as the best fit to the data and conclude that there is evidence of two dimensions of women's status at both the woman and district levels, with particularly strong evidence of a two-factor structure at Level 1.

Results

We focus on the interpretation of Model 3 above, with two correlated factors at each level. The standardized factor loadings from this model are presented in Table 3. All unconstrained loadings differ significantly from zero, suggesting that each factor is correlated with all indicators of the dimension it represents. We begin by examining the loadings associated with the decision-making variables. At Level 1, the decision-making factor is most strongly correlated with the "visits to family/friends" and "household purchases" indicators, suggesting that these variables are the best at

Table 3
Standardized Factor Loadings $\lambda_r^{(l)*}$ (Standard Errors)

	Woman Level				District Level			
Response (r)	$\lambda_{1r}^{(l)*}$	(SE)	$\lambda_{2r}^{(l)*}$	(SE)	$\lambda_{1r}^{(2)*}$	(SE)	$\lambda_{2r}^{(2)*}$	(SE)
Decision making								
Children's education	0.70	(0.01)	0	—	0.16	(0.05)	0	—
Visits to family/friends	0.81	(0.01)	0	—	0.22	(0.04)	0	—
Household purchases	0.80	(0.01)	0	—	0.13	(0.03)	0	—
Family planning	0.60	(0.01)	0	—	0.14	(0.05)	0	—
Medical treatment	0.78	(0.01)	0	—	0.11	(0.04)	0	—
Social independence								
Go out locally	0	—	0.65	(0.02)	0	—	0.26	(0.07)
Go outside village/town	0	—	0.73	(0.01)	0	—	0.24	(0.05)
Talk to unknown man	0	—	0.56	(0.02)	0	—	0.21	(0.08)
Go to cinema/show	0	—	0.84	(0.01)	0	—	0.31	(0.04)
Go shopping	0	—	0.86	(0.01)	0	—	0.30	(0.04)
Go to cooperative/club	0	—	0.76	(0.02)	0	—	0.34	(0.06)
Go to health center	0	—	0.82	(0.01)	0	—	0.22	(0.04)

Note: Estimates and standard errors are the means and standard deviations of 50,000 chains, with a burn-in of 4,000. Each factor is constrained to have zero loadings on either the decision-making responses or the social independence responses.

discriminating between women with different levels of participation in family decisions. The "decisions about family planning" indicator has relatively weak discriminatory power. The loadings at the district level are all considerably smaller than those at the woman level. As would be expected, although there is evidence of a contextual district effect on some of the responses, it is a woman's own overall decision-making score (the Level 1 factor) that has the greater influence. The "visits to family/friends" indicator is the most strongly correlated with the district-level factor.

Turning to the loadings for the social independence indicators, we again find stronger effects of the individual factor. However, for these indicators, the district-to-woman ratios of the factor loadings are all larger than those for the decision-making indicators, which implies that where a woman lives has a greater relative influence on her mobility than on her decision making. This difference in the relative impact of district-level measures of status is expected. Social independence concerns women's activities in public life, which are more likely to be constrained by local

Table 4
Covariate Effects β_r (Standard Errors) on Each Response

Response (*r*)	Age	(*SE*)	Urban	(*SE*)	Educated	(*SE*)	Hindu	(*SE*)
Decision making								
Children's education	0.008	(0.002)	0.25	(0.04)	0.50	(0.04)	0.08	(0.05)
Visits to family/friends	0.010	(0.002)	0.32	(0.05)	0.61	(0.05)	0.10	(0.07)
Household purchases	0.013	(0.002)	0.39	(0.05)	0.51	(0.04)	0.00	(0.06)
Family planning	–0.002	(0.002)	0.18	(0.04)	0.28	(0.03)	0.04	(0.05)
Medical treatment	0.013	(0.002)	0.34	(0.05)	0.57	(0.04)	0.04	(0.06)
Social independence								
Go out locally	0.023	(0.002)	–0.03	(0.05)	0.16	(0.05)	–0.16	(0.07)
Go outside village/town	0.020	(0.002)	0.35	(0.05)	0.61	(0.05)	–0.08	(0.07)
Talk to unknown man	0.009	(0.002)	0.34	(0.05)	0.58	(0.04)	0.14	(0.06)
Go to cinema/show	0.025	(0.003)	1.26	(0.08)	0.97	(0.07)	–0.09	(0.09)
Go shopping	–0.005	(0.005)	1.75	(0.12)	1.59	(0.11)	0.47	(0.12)
Go to cooperative/club	0.012	(0.004)	0.99	(0.08)	0.97	(0.07)	0.03	(0.09)
Go to health center	0.011	(0.004)	1.41	(0.08)	1.08	(0.07)	0.21	(0.09)

Note: A positive coefficient implies that a variable is positively associated with being in a high status category. See text for details. Age is in years; the reference categories for Urban, Educated, and Hindu are, respectively, Rural, Uneducated, and Muslim. Estimates and standard errors are the means and standard deviations of 50,000 chains, with a burn-in of 4,000.

cultural norms than is participation in decisions about more private family matters. Of the social independence indicators, shopping, going to a cinema or show, and attending a cooperative or club are good discriminators at both the woman and district levels.

In standard factor analysis, it is usual to assume that the factors are independent. In the present case, however, we might expect that women who are socially independent will tend also to participate more in household decisions. Similarly, districts with high scores on one dimension would be expected to have high scores on the other. Previous research, based on simple summated scales and not taking into account area effects, has found pairwise correlations between scales ranging from 0.18 to 0.29 (Balk 1994; Schuler et al. 1997). We find moderate, positive correlations between the social independence and decision-making factors at each level. The correlation between woman-level factors is estimated as 0.35 (with 95 percent credible interval 0.33-0.38), while at the district level, the estimated correlation is 0.40 (0.02-0.74).

Table 4 shows the effects of covariates on each of the women's status variables. In our parameterization of the multilevel factor model, the covariate coefficients (and factor loadings) are interpreted as effects on the probability of being in a lower response category. All indicators have been coded so that a lower value corresponds to higher status. For example, the ordinal decision-making items are coded as follows: 0 = woman only, 1 = woman and husband jointly, and 2 = husband alone. Therefore, a positive coefficient (loading) implies that a variable is positively associated with being in a high status category. On all but two indicators (decisions about family planning and the freedom to go shopping alone), older women have significantly higher status than younger women. Living in an urban rather than a rural area is associated with having a say in household decisions and greater social freedom. There are particularly strong effects of type of region on the last four social independence indicators, which are all to do with engagement in public life. For all items, having some education is associated with higher status; as for region, the effects of education are stronger the more public an activity. The effects of religious denomination, where significant, are weak. There is little evidence to suggest differences between Muslim and Hindu women in their participation in family decisions. However, in general, Hindus have more social independence than Muslims; they are more likely than Muslims to be able to talk to an unknown man, go shopping, or visit a health center alone. The one exception to this pattern is the finding that Hindus are less likely than Muslims to be able to go outside within their locality.

Table 5 shows the residual district-level variances, or the variances of the "specific factors," for each response variable. The specific factors represent unobserved district characteristics that are unique to each response, whereas the common factors represent characteristics that affect all responses. The greatest amount of residual between-district variation is found for three of the social independence measures: going out locally, talking to an unknown man, and attending a cooperative or club. The responses on each of these indicators are influenced by unobserved district-level variables that do not affect the other responses.

Discussion

Multilevel modeling and factor analysis are now well-established members of the quantitative social researcher's methods toolkit. Since the late 1980s, there have been efforts to synthesize these two powerful techniques,

Table 5
Residual District-Level Variances (95 Percent Credible Intervals)

Item (r)	σ^2_{ur}	(95 Percent Credible Interval)
Decision making		
Children's education	0.13	(0.08, 0.21)
Visits to family/friends	0.07	(0.00, 0.15)
Household purchases	0.06	(0.03, 0.10)
Family planning	0.08	(0.04, 0.13)
Medical treatment	0.09	(0.05, 0.14)
Social independence		
Go out locally	0.56	(0.35, 0.86)
Go outside village/town	0.18	(0.11, 0.29)
Talk to unknown man	0.52	(0.34, 0.78)
Go to cinema/show	0.08	(0.00, 0.18)
Go shopping	0.13	(0.03, 0.28)
Go to cooperative/club	0.51	(0.28, 0.85)
Go to health center	0.08	(0.04, 0.15)

Note: Estimates are the means of 50,000 chains, with a burn-in of 4,000. The limits of the 95 percent credible intervals are the 2.5 percent and 97.5 percent points of the distribution of the chains for each variance parameter.

but until recently, most research has focused on the multivariate normal case. In this article, we describe a two-level probit factor model for mixed binary and ordinal responses, with covariate effects on each response. We illustrate the potential of this new methodology in an analysis of two aspects of women's status in Bangladesh: social independence and autonomy in making household decisions. The analysis reveals strong evidence of two correlated factors at the woman level and somewhat weaker evidence of a second factor at the district level. Among the interesting substantive findings is the result that district effects are stronger on the social independence indicators, which measure participation in the public sphere, than on the decision-making indicators, which concern more private family matters. These effects persist after accounting for the effects of age, education level, religion, and type of region of residence, whose effects on the responses are in the expected directions.

Several extensions to the method presented here may be fruitful in this and other applications. One extension would be to generalize the factor model to a full SEM, consisting of the measurement model in (1) combined with a structural model. In the structural component of the model, one or more factors at either level may depend on covariates and possibly other factors. An SEM would be useful if, for example, the goal of the analysis was to identify women with little involvement in family decisions; in that case, it would be of greater interest to estimate covariate effects on the decision-making factor (i.e., in the structural part of the model) than on the individual indicators (in the measurement model). The SEM could be further extended to a random coefficient model, in which covariate effects on the factors vary randomly across Level 2 units. For instance, in our application, the effect of education on decision making might vary across districts. Further work is planned in this area.

The factor model described in this article may be applied to a mixture of binary and ordinal responses, but it is straightforward to incorporate normally distributed responses (see Goldstein et al. forthcoming). Goldstein et al. (2004) also describe how missing response data can be handled under a missing-at-random assumption. Another possible generalization is to allow inclusion of nominal (unordered) responses. Following Albert and Chibb (1993), a multinomial probit model can be fitted to the nominal responses. Each nominal response y_r is replaced by a vector of binary responses $\{y_{cr}\}, c = 0, \cdots, C_r$, where $y_{cr} = 1$ if $y_r = c$ and $y_{cr} = 0$ otherwise. We assume that underlying y_{cr} is a continuous latent variable y^*_{cr}, such that we observe $y_r = c$ if and only if $y^*_{cr} > y^*_{c'r}$ for all $c \neq c'$. (See Aitchison and Bennett 1970 for further details of the multiple indicator approach to modeling nominal response data.) The extension to nominal responses would be useful in the present application to test our assumption that the three-category decision-making responses are indeed ordinal in terms of the status they imply.

Our approach may also be generalized to handle partially ordered responses that may arise, for example, if there are indicators with "don't know" response alternatives. In this case, we can define two variables: a binary variable $y_{DK,r}$, coded 1 for a "don't know" response and 0 otherwise, and an ordered categorical variable y_r, which is defined only when $y_{DK,r} = 0$. We then specify probit models for $y_{DK,r}$ and $y_r | y_{DK,r} = 0$, allowing for correlation between the underlying normal responses $y^*_{DK,r}$ and $y^*_r | y_{DK,r} = 0$ at Level 2.

References

Aitchison, J. and J. A. Bennett. 1970. "Polychotomous Quantal Response by Maximum Indicant." *Biometrika* 57:253-62.

Albert, James H. and Siddhartha Chibb. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88:669-79.

Amin, Sajeda, Alaka M. Basu, and Rob Stephenson. 2001. "Spatial Variation in Contraceptive Use in Bangladesh." *Demography* 39:251-67.

Amin, Sajeda, Ian Diamond, and Fiona Steele. 1997. "Contraception and Religiosity in Bangladesh." Pp. 268-89 in *The Continuing Demographic Transition*, edited by G. W. Jones, R. M. Douglas, J. C. Caldwell, and R. M. D'Souza. Oxford, UK: Clarendon.

Balk, Deborah. 1994. "Individual and Community Aspects of Women's Status and Fertility in Rural Bangladesh." *Population Studies* 48:21-45.

Bartholomew, David J., Fiona Steele, Irini Moustaki, and Jane I. Galbraith. 2003. *The Analysis and Interpretation of Multivariate Data for Social Scientists*. London: CRC Press.

Browne, William J. 2003. *MCMC Estimation in MLwiN, Version 2.0*. London: Institute of Education.

Goldstein, Harvey. 2003. *Multilevel Statistical Models*. 3rd ed. London: Arnold.

Goldstein, Harvey, Gerard Bonnet, and Thierry Rocher. Forthcoming. "Multilevel Structural Equation Models for the Analysis of Comparative Data on Educational Performance." *Journal of Educational and Behavioral Statistics*.

Goldstein, Harvey and William J. Browne. 2002. "Multilevel Factor Analysis Modelling Using Markov Chain Monte Carlo (MCMC) Estimation." Pp. 225-43 in *Latent Variable and Latent Structure Models*, edited by G. Marcoulides and I. Moustaki. Englewood Cliffs, NJ: Lawrence Erlbaum.

———. 2005. "Multilevel Factor Analysis for Continuous and Discrete Data." Pp. 453-75 in *Contemporary Psychometrics: A Festschrift for Roderick P. McDonald*, edited by A. Maydeu-Olivares and J. J. McArdle. Mahwah, NJ: Lawrence Erlbaum.

Huq, M. Najmul and John Cleland. 1990. *Bangladesh Fertility Survey, 1989*. Dhaka, India: National Institute of Population Research and Training (NIPORT).

Jöreskog, Karl G. and Irini Moustaki. 2001. "Factor Analysis of Ordinal Variables: A Comparison of Three Approaches." *Multivariate Behavioral Research* 36:347-87.

Lee, S.-Y. and H.-T. Zhu. 2000. "Statistical Analysis of Nonlinear Structural Equation Models With Continuous and Polytomous Data." *British Journal of Mathematical and Statistical Psychology* 53:209-32.

Mathworks. 2004. MATLAB program, Version 7.0. Natick, MA: Mathworks.

Moustaki, Irini. 2003. "A General Class of Latent Variables Models for Ordinal Manifest Variables With Covariate Effects on the Manifest and Latent Variables." *British Journal of Mathematical and Statistical Psychology* 56:337-57.

Muthén, Bengt O. 1994. "Multilevel Covariance Structure Analysis." *Sociological Methods & Research* 22:376-98.

Rabe-Hesketh, Sophia, Anders Skrondal, and Andrew Pickles. 2004. "Generalized Multilevel Structural Equation Modelling." *Psychometrika* 69:167-90.

Rasbash, Jon, Fiona Steele, William J. Browne, and Bob Prosser. 2004. *A User's Guide to MLwiN, Version 2.0*. London: Institute of Education.

Schuler, Sidney R., Syed M. Hashemi, and Ann P. Riley. 1997. "The Influence of Women's Changing Roles and Status in Bangladesh's Fertility Transition: Evidence From a Study of Credit Programs and Contraceptive Use." *World Development* 25:563-75.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (With Discussion)." *Journal of the Royal Statistical Society, Series B* 64:583-639.

Steele, Fiona, Sajeda Amin, and Ruchira T. Naved. 2001. "Savings/Credit Group Formation and Change in Contraception." *Demography* 38:267-82.

Fiona Steele is a reader in social statistics in the Graduate School of Education, University of Bristol. Her research interests include multilevel modeling, event history analysis, and structural equation modeling, with applications in demography.

Harvey Goldstein is a professor of social statistics in the Graduate School of Education, University of Bristol. His research interests include multilevel modeling and educational assessment.