The Multilevel p_2 Model

A Random Effects Model for the Analysis of Multiple Social Networks

Bonne J. H. Zijlstra, Marijtje A. J. van Duijn, and Tom A. B. Snijders

University of Groningen, the Netherlands

Abstract. The p_2 model is a random effects model with covariates for the analysis of binary directed social network data coming from a single observation of a social network. Here, a multilevel variant of the p_2 model is proposed for the case of multiple observations of social networks, for example, in a sample of schools. The multilevel p_2 model defines an identical p_2 model for each independent observation of the social network, where parameters are allowed to vary across the multiple networks. The multilevel p_2 model is estimated with a Bayesian Markov Chain Monte Carlo (MCMC) algorithm that was implemented in free software for the statistical analysis of complete social network data, called StOCNET. The new model is illustrated with a study on the received practical support by Dutch high school pupils of different ethnic backgrounds.

Keywords: social network analysis, random effects, multilevel modeling, p2 model

Introduction

An important goal of social network analysis is to find structure in relations within groups. Usually, one social network at a time is analyzed. This article deals with the analysis of multilevel network data where multiple observations of the same binary relation in different groups are available. In an educational setting, for example, networks may be observed in multiple schools or school classes, forming the higher-level units. Analyzing multiple networks simultaneously provides greater generalizability of research results compared to analyses of single-network data. For multilevel network data, an interesting question is whether all social networks show a common structure. If this turns out not to be the case, subsequent questions are on which aspects the networks differ, and whether there are network attributes to which these differences can be ascribed. The multilevel p_2 model has been designed to answer these types of questions. So far, a statistical analysis of multiple networks with the p_2 model involved a two-stage estimating procedure, where a meta-analysis was performed on the parameters obtained in separate p_2 analyses of each network (Baerveldt, Van Duijn, Vermeij, & Van Hemert, 2004). The multilevel p_2 model estimates the parameters more efficiently and, moreover, quantifies the differences between networks by modeling the variability of parameters over networks. For continuous social network data, the social relations model for multiple groups can be used that can be estimated with standard multilevel software (Snijders & Kenny, 1999).

The data under consideration consist of K networks. Each single network is defined by a set of social actors and a relation (e.g., friendship, collaboration) defined on this set. Each relation is expressed by a collection of tie variables: Y_{ij} equals 1 if there is a tie from actor i to actor j, and 0 otherwise (where the notation temporarily omits to indicate which of the networks is being referred to). It is assumed that the sets of actors for the K networks are disjoint (e.g., different school classes), and that the content, or meaning, of the relation is the same in each network (e.g., friendship). It should be noted that the actor sets are not required to have the same size for every network.

In the p_2 model (Van Duijn, Snijders, & Zijlstra, 2004), the tie variables are regressed on explanatory variables, while the dependence of ties from and to the same actor is modeled using random effects. The multilevel p_2 model defines an identically specified p_2 model with varying parameters for multiple independent social networks. One of the Bayesian Markov Chain Monte Carlo (MCMC) algorithms developed for the single-network p_2 model by Zijlstra, Van Duijn, and Snijders (2005a) can be expanded for the multilevel p_2 model. This hybrid Metropolis-Hastings algorithm will be briefly described in the third section, after defining the multilevel p_2 model in the next section. Software for the multilevel p_2 model is available in the p_2 module of StOCNET (Boer, Huisman, Snijders, & Zeggelink, 2003), an open software system for the statistical analysis of social networks.

The multilevel p_2 model is applied to network data collected by Chris Baerveldt in 20 Dutch high schools (Baerveldt, 2000; Snijders & Baerveldt, 2003). The central question dealt with in this article is whether these pupils tend to report more practical support from other pupils with the same ethnic background.

DOI 10.1027/1614-1881.2.1.42

The Multilevel p_2 Model

The multilevel p_2 model is an extension of the p_2 model (Van Duijn et al., 2004) that originates from the p_1 model proposed by Holland and Leinhardt (1981). These models are defined by specifying the probability of observing one of the four possible outcomes of the pair of two directed ties between each pair of actors, which is called a dyad. Let a dependent network with n actors be denoted by the tie indicator variables Y_{ij} and let the actors i and j have numbers $1, \ldots, n$. Then the p_1 model for the probabilities of the two observed ties between actors i and j is defined by

$$\begin{aligned} P(Y_{ij} = y_1, Y_{ji} = y_2) &= \exp(y_1\mu_{ij} + y_1 y_2 \rho)/c, \\ c &= \sum_{y_1 y_2 \in \{0,1\}, i \neq j} \exp(y_1\mu_{ij} + y_2\mu_{ji} + y_1 y_2 \rho), \\ \mu_{ij} &= \mu + \alpha_i + \beta_j, \end{aligned} \tag{1}$$

where α_i is a sender parameter for actor i, β_j a receiver parameter for actor j, μ the density parameter, and ρ the reciprocity parameter.

In the p_1 model, all ties coming from or directed toward the same actor are mutually related through—or conditionally independent given—the $2n$ parameters α_i and β_i. The number of parameters in the p_1 model increases linearly with the number of actors, an undesirable property for any model. The p_2 model extends the p_1 model by including covariate effects for α, β, μ, and ρ, while the total number of parameters in this model is reduced by assuming random (instead of fixed) sender and receiver effects A_i and B_i. The latter are assumed to be independent, identically bivariate normally distributed variables with zero means, and covariance matrix $\boldsymbol{\Sigma}$ with diagonal elements σ_A^2 (sender variance), and σ_B^2 (receiver variance) and off-diagonal elements σ_{AB} (sender-receiver covariance). In the p_2 model, the sender and receiver parameters are regressed on (actor) covariates $\mathbf{X_1}$ and $\mathbf{X_2}$ with fixed regression parameters $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$,

$$\begin{aligned} \alpha_i &= \mathbf{X}_{1i}\boldsymbol{\gamma}_1 + A_i, \\ \beta_i &= \mathbf{X}_{2i}\boldsymbol{\gamma}_2 + B_i. \end{aligned}$$

The density and reciprocity parameters μ and ρ are regressed on $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ with fixed regression parameters $\boldsymbol{\delta_1}$ and $\boldsymbol{\delta_2}$,

$$\begin{aligned} \mu_{ij} &= \mu + \boldsymbol{Z}_{1ij}\boldsymbol{\delta}_1, \\ \rho_{ij} &= \rho + \boldsymbol{Z}_{2ij}\boldsymbol{\delta}_2. \end{aligned}$$

Note the added subscripts i and j for the density and reciprocity parameters, which are now assumed to be dyad-specific. The variables $\boldsymbol{Z_1}$ and $\boldsymbol{Z_2}$ specify dyadic covariates depending on the ordered pair of actors (i, j), where $\boldsymbol{Z_2}$ is a symmetric matrix, expressing the mutuality-by-definition of reciprocity, $\rho_{ij} = \rho_{ji}$.

The multilevel p_2 model specifies an identical p_2 model for K independent observations of a network relation. To account for differences between the K networks, the multilevel p_2 model includes random coefficients for the fixed p_2 regression parameters at the network level. The fixed regression parameters become random coefficients by adding random effects $\boldsymbol{G_1}$ and $\boldsymbol{G_2}$ to the regression parameters $\boldsymbol{\gamma_1}$ and $\boldsymbol{\gamma_2}$ for the sender and receiver effects. The density and reciprocity parameters μ and ρ obtain random effects M and R, their regression parameters random effects $\boldsymbol{D_1}$ and $\boldsymbol{D_2}$. The parameters in the p_1 model (Equation 1) for the kth network are thus substituted by

$$\begin{aligned} \alpha_{ik} &= \boldsymbol{X}_{1ik}(\boldsymbol{\gamma}_1 + G_{1k}) + A_{ik}, \\ \beta_{ik} &= \boldsymbol{X}_{2ik}(\boldsymbol{\gamma}_2 + G_{2k}) + B_{ik}, \\ \mu_{ijk} &= \mu + \mathrm{M}_k + \boldsymbol{Z}_{1ijk}(\boldsymbol{\delta}_1 + D_{1k}), \\ \rho_{ijk} &= \rho + \mathrm{R}_k + \boldsymbol{Z}_{2ijk}(\boldsymbol{\delta}_2 + D_{2k}). \end{aligned} \tag{2}$$

The vector with random effects for each network k is denoted by $\boldsymbol{T}_k' = (M_k, R_k, \boldsymbol{G}_{1k}', \boldsymbol{G}_{2k}', \boldsymbol{D}_{1k}', \boldsymbol{G}_{2k}')$, and is assumed to be normally distributed with zero means and covariance matrix $\boldsymbol{\Omega}$. A further assumption, common in multilevel modeling, is that the random effects at the network level (k) are independent from the random effects at the actor level (i).

The multilevel p_2 model can be regarded as a three-level random effects model where Level 1 is formed by the tie observations, cross-nested in the actors (Level 2), who are nested in the networks (Level 3). Just like the p_2 model allows for covariates at the actor level, the multilevel model allows for explanatory variables at the network level, adding a regression model to the network-specific parts of Equation 2. For instance, the term $\mu + M_k$ can be replaced by $\mu + M_k + \boldsymbol{W}_k\eta$, where $\boldsymbol{W}_k$ denotes a vector of network-level covariates such as network size and aggregated actor characteristics.

Outline of the MCMC Estimation Algorithm

For random effects models, maximum likelihood estimation of the fixed effects and of the variances of the random effects requires integration over the random parameters. In the case of the multilevel p_2 model, these are practically intractable integrals, which is why an MCMC algorithm is applied.

In MCMC estimation methods, prior distributions need to be specified for the model parameters, which, together with the data, determine the posterior distributions for the parameters, as follows from Bayes's theorem. In these algorithms, a Markov chain generates a sample of all parameters. If convergence of the chain can be postulated after an initial burn-in period, then this is a sample from the posterior distribution and Monte Carlo estimates can be calculated from this sample. This section provides a succinct overview of how the posterior distributions of the multilevel p_2 model parameters are obtained and how these are sampled from.

The multivariate distribution of the data and the fixed and random parameters has a probability density that is proportional to the joint density

$$P(\boldsymbol{Y}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\theta}, \boldsymbol{T}, \boldsymbol{\Omega}), \tag{3}$$

where $\boldsymbol{Y} = (\boldsymbol{Y}_1', \ldots, \boldsymbol{Y}_K')'$ contains the observed data for all networks, with $\boldsymbol{Y}_k'$ the vector of all dyads $(\mathrm{Y}_{ijk}, \mathrm{Y}_{jik})$ in network k. $\boldsymbol{C}$ contains all pairs of random actor effects $(\mathrm{A}_{ik}, \mathrm{B}_{jk})$, and $\boldsymbol{\Sigma}$ denotes their 2×2 covariance matrix. Vector $\boldsymbol{\theta}' = (\mu, \rho, \boldsymbol{\gamma}_1', \boldsymbol{\gamma}_2', \boldsymbol{\delta}_1', \boldsymbol{\delta}_2')$ is the vector of fixed parameters, and $\boldsymbol{T} = (\boldsymbol{T}_1', \ldots, \boldsymbol{T}_K')'$ is the vector of random effects of the parameters in $\boldsymbol{\theta}$ with covariance matrix $\boldsymbol{\Omega}$.

A convenient way to obtain a sample from the multivariate distribution of all variables in Equation 3 is by means of the Gibbs sampler, which involves drawing subsequent random variables from each distribution formed by a separate set of parameters, conditional on all other parameters and $\boldsymbol{Y}$ (see, e.g., Chib and Greenberg, 1995).

For some of the parameters in Equation 3, the posterior distributions are easy to sample from due to conditional independence and convenient conjugate priors. Assuming that $\boldsymbol{\Sigma}$, $\boldsymbol{\theta}$ and $\boldsymbol{\Omega}$ are mutually independent, the multivariate density can be factorized as

$$P(\boldsymbol{Y}, \boldsymbol{C}, \boldsymbol{\Sigma}, \boldsymbol{\theta}, \boldsymbol{T}, \boldsymbol{\Omega}) = P_Y(\boldsymbol{Y}|\boldsymbol{C}, \boldsymbol{\theta}, \boldsymbol{T})P_C(\boldsymbol{C}|\boldsymbol{\Sigma})P_\Sigma(\boldsymbol{\Sigma})P_\theta(\boldsymbol{\theta})P_T(\boldsymbol{T}|\boldsymbol{\Omega})P_\Omega(\boldsymbol{\Omega}). \quad (4)$$

Here, $P_Y(\boldsymbol{Y}|\boldsymbol{C}, \boldsymbol{\theta}, \boldsymbol{T})$ is the conditional likelihood of the multilevel p_2 model given the fixed parameters and random actor and network parameters—that is, the probability of network k according to Equation 1, with the substitutions defined in Equation 2, multiplied over all networks. As can be seen from Equation 4, the density of the random actor effects, $\boldsymbol{C}$, depends only on $\boldsymbol{\Sigma}$, and the density of the random effects of the fixed parameters, $\boldsymbol{T}$, depends only on $\boldsymbol{\Omega}$.

Prior probability densities $P_\Sigma(\boldsymbol{\Sigma})$ and $P_\Omega(\boldsymbol{\Omega})$ are assumed to have inverse Wishart$(\boldsymbol{\Sigma}_\Sigma, \nu_\sigma)$ and inverse Wishart$(\boldsymbol{\Sigma}_\omega, \nu_\omega)$ distributions, respectively, the natural conjugate priors for the covariance matrices of normally distributed random variables (see, e.g., Press, 1989, p. 141). The covariance matrices of the Wishart prior distributions are chosen as $\boldsymbol{\Sigma}_\sigma = \boldsymbol{I}$ and $\boldsymbol{\Sigma}_\omega = \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix. The degrees of freedom, ν_σ and ν_ω, are chosen as the number of dimensions plus one, representing little prior information. Consequently, the posterior inverse Wishart distribution of $\boldsymbol{\Sigma}^{-1}$ has $3 + \sum_{k=1}^{K} n_k$ degrees of freedom and covariance matrix $(\boldsymbol{C}'\boldsymbol{C}+\boldsymbol{I})^{-1}$. The number of degrees of freedom for $\boldsymbol{\Omega}^{-1}$ is ν_ω with covariance matrix $(\boldsymbol{T}'\boldsymbol{T}+\boldsymbol{I})^{-1}$ (see, e.g., Box & Tiao, 1973, p. 427).

A priori, the parameters in $\boldsymbol{\theta}$ are assumed to follow independent normal distributions with zero means and variances for μ and ρ equal to 100. The variances of the regression parameters γ_1, γ_2, $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are set to 100 divided by the observed variance of the corresponding covariate. Thus, the variance of a parameter for a "standardized" covariate is equal to 100 as well. Since parameters in $\boldsymbol{\theta}$ are on a logistic scale, a standard deviation of 10 implies that 33% of the observations are larger than the absolute value of 10. This prior for $\boldsymbol{\theta}$ may thus be regarded as very lightly informative and reflects the fact that almost any statistician will be surprised when seeing log odds ratios larger than 10.

With all distributions in Equation 3 defined, the conditional posterior distributions for each of the model parameters can be derived. By repeatedly sampling from these distributions, eventually a sample from the multivariate posterior distribution of all parameters is obtained. Draws from the conditional distributions of the covariance matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ can be obtained directly from their inverse Wishart distributions, whereas sampling of the random effects $\boldsymbol{C}$ and $\boldsymbol{T}$ and the fixed parameters $\boldsymbol{\theta}$ requires using a Metropolis-Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller & Teller, 1953). For this purpose, we used a similar type of random walk algorithm as in Zijlstra et al. (2005a), in which the variance of the proposals is based on its estimate obtained from a normal approximation to the conditional distributions.

A complication in the case of the p_2 model, which is proliferated in the multilevel p_2 model, is that parameters are not independent of each other. For instance, as can be seen from its definition in Equations 1 and 2, the random actor effects appear at the same place as the density parameter μ and the random network effects M_k. This parameter dependence reduces the efficiency of the MCMC algorithm. For a higher efficiency, the MCMC algorithm could be adjusted using a different parameterization that partly controls for these dependencies in the spirit of Gelfand, Sahu, and Carlin (1995) and Hoff (2005).

Application: Reported Practical Support Between Dutch High School Pupils

In the Dutch Social Behavior Study (Baerveldt, 2000; Baerveldt et al., 2004; Snijders and Baerveldt, 2003), social network data were collected among 16–18-year-old pupils belonging to the same year group. The data are from 20 urban high schools with a total of 1,337 pupils. All high schools were so-called MAVO schools, which educate children of medium intellectual ability. One of the questions asked was "Which pupils help you with practical problems, such as doing homework, organizing a party or completing a difficult form?" and was intended to measure social support. This type of question is typical for social network studies.

The pupils are from different ethnic backgrounds. Ethnicity is determined by the country of birth of both parents, where for pupils from parents with different countries of birth ethnicity is treated as a missing value. One of the research questions of the study was whether more social support relations are found between pupils from the same ethnic background. In this example, we consider this as our main question, taking into account that emotional support relations have been found to be more prevalent among pupils of the same gender (see Baerveldt et al., 2004; Zijlstra et al., 2005b).

Some descriptive statistics for the data in the application are given in Table 1. The data set on which the analyses are performed is reduced to 1,232 after discarding those pupils for whom either their gender or ethnic background was missing. Sample sizes for the different schools are then between 38 and 96, with varying composition. The range of percentages are 36–68% boys, 12–92% Dutch, 0–19% Moroccan, 0–23% Turkish, and 0–28% Surinamese. Two schools have no pupils from a Moroccan, Turkish, or Surinamese background. Seven schools have no Moroccan pupils, four no Turkish, and four no Surinamese.

The mean degree reported in Table 1 indicates the average number of ties per pupil (reporting having received support or being reported by others to have helped) in the 20 networks. The mean reciprocal degree is the average number of reciprocal ties per pupil. The variation of the mean degree and the reciprocal degree between networks is small. This is reflected by the results in Table 2, where quantiles and

Table 1. Descriptive statistics for the 20 schools in the Dutch Social Behavior Study data.

School	1	2	3	4	5	6	7	8	9	10
Number of pupils	50	43	55	38	73	96	62	39	91	42
Mean degree	2.98	2.30	1.47	1.95	3.22	2.16	2.71	3.33	2.44	1.52
Mean reciprocal degree	1.56	1.12	0.62	0.89	1.70	0.96	0.90	1.54	1.10	0.67
School	11	12	13	14	15	16	17	18	19	20
Number of pupils	91	73	59	42	52	62	55	56	77	76
Mean degree	3.01	2.71	2.12	1.07	3.00	2.24	2.18	3.21	2.49	2.55
Mean reciprocal degree	1.60	1.23	1.12	0.52	1.77	0.97	1.56	1.75	1.43	1.08

Table 2. Parameter estimates of a multilevel p_2 analysis of the Dutch Social Behavior Study data.

			Estimates					
			Posterior		Quantiles			
Fixed Effect	Covariate	Parameter	Mean	SE	0.5	2.5	97.5	99.5
Sender	Boy	γ_{11}	−0.11	0.08	−0.32	−0.28	0.05	0.10
Receiver	Boy	γ_{21}	−0.12	0.07	−0.29	−0.25	0.01	0.05
Density		μ	−3.58	0.11	−3.83	−3.77	−3.36	−3.28
	Gender	δ_{11}	1.26	0.06	1.09	1.14	1.40	1.44
	Dutch	δ_{12}	0.50	0.06	0.33	0.37	0.62	0.66
	Moroccan	δ_{13}	0.40	0.13	0.06	0.14	0.64	0.70
	Turkish	δ_{14}	0.53	0.12	0.23	0.30	0.77	0.85
	Surinamese	δ_{15}	0.30	0.04	0.10	0.10	0.52	0.60
	Girl	δ_{16}	−0.14	0.06	−0.29	−0.26	−0.02	0.02
Reciprocity		ρ	4.19	0.18	3.71	3.84	4.54	4.64
	Gender	δ_{21}	−0.45	0.17	−0.89	−0.80	−0.11	0.02
	Dutch	δ_{22}	−0.33	0.15	−0.71	−0.62	−0.03	0.06
	Moroccan	δ_{23}	−0.32	0.30	−1.07	−0.89	0.27	0.48
	Turkish	δ_{24}	−0.03	0.31	−0.79	−0.60	0.66	0.83
	Surinamese	δ_{25}	0.02	0.24	−0.60	−0.47	0.48	0.63
	Girl	δ_{26}	0.16	0.14	−0.20	−0.12	0.43	0.50
Actor-level random effects								
Sender variance		σ^2_A	1.13	0.09	0.92	0.97	1.31	1.37
Receiver variance		σ^2_B	0.44	0.05	0.33	0.36	0.54	0.57
Sender receiver covariance		σ_{AB}	−0.57	0.05	−0.71	−0.67	−0.47	−0.44
Network-level random effects of fixed parameters								
Density variance		$\Omega_{1,1}$	0.21	0.08	0.08	0.10	0.41	0.52
Reciprocity variance		$\Omega_{2,2}$	0.24	0.11	0.08	0.10	0.52	0.68
Density-reciprocity covariance		$\Omega_{1,2}$	−0.13	0.08	−0.45	−0.33	−0.02	0.00

Note. The covariate effects for density and reciprocity are similarity effects; burn-in length and sample size of the MCMC estimation algorithm were 8,000 and 40,000 respectively.

estimates of the mean and standard error of the posterior distributions of the parameters are reported. The variances of the density and reciprocity parameters are rather low, especially compared to the sender variance. Next, the network-level covariate “percentage of Dutch pupils” was included in the model, but this turned out to be nonsignificant. Moreover, the quality of the MCMC sample deteriorated, probably because there was little information in the data about the effect of the additional model parameter.

The reported results in Table 2 are based on a burn-in sequence of 8,000 iterations and a sampling sequence of 40,000 iterations. By inspecting the trace plots, one can investigate how well the sampled parameters have converged to a stable distribution. In Figure 1 the traces appear stable, although some sudden jumps are noticeable, which demonstrates the mutual dependence of the parameters discussed above.

The model in Table 2 can be regarded as a so-called random intercept model (cf. Goldstein, 2003) with random effects M and R for the intercepts μ and ρ on the network level. The parameter estimates show that reported practical support relations are more prevalent between pupils with the same gender. This is pointed out by the strong positive density effect of “similarity gender.” The interpretation of this effect is moderated by the “similarity girl” effect, which shows that the increased practical support within the same

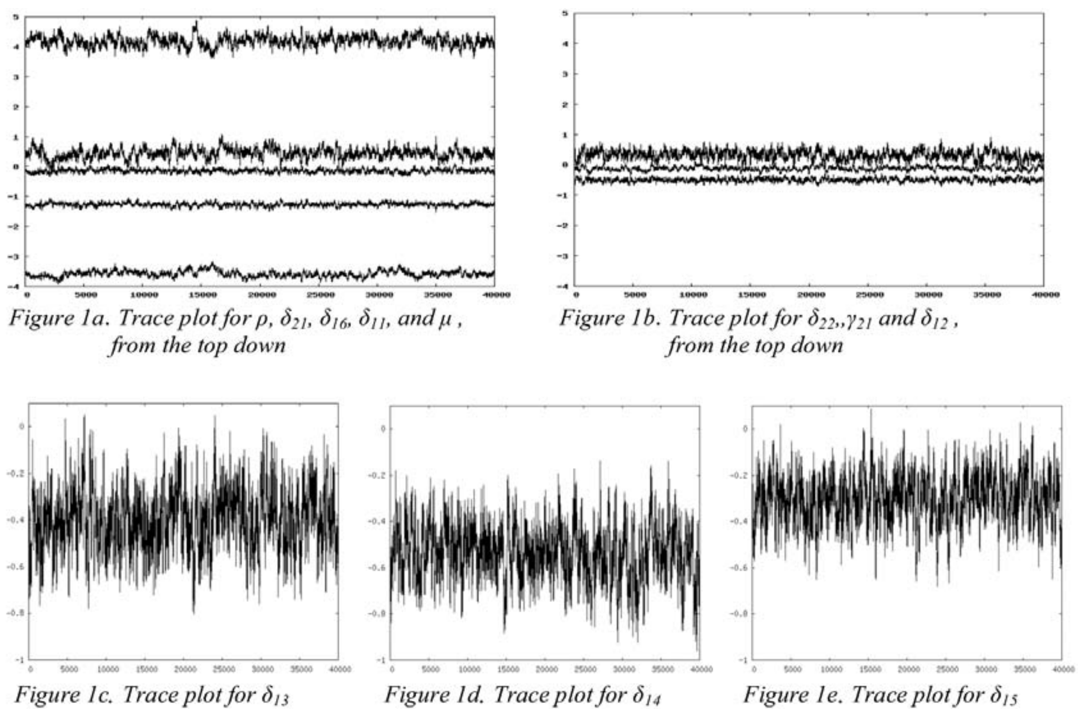

Figure 1. (a) Trace plot for ρ, δ_{21}, δ_{16}, δ_{11}, and μ, from the top down. (b) Trace plot for δ_{22}, γ_{21}, and δ_{12}, from the top down. (c) Trace plot for δ_{13}. (d) Trace plot for δ_{14}. (e) Trace plot for δ_{15}.

gender is stronger for boys than for girls. The negative similarity gender effect for reciprocity shows that reported practical support relations within the same sex is not a doubled density effect but slightly less, making reciprocal same-gender dyads more likely than asymmetrical dyads. Furthermore, boys and girls do not differ strongly with respect to sending and receiving tendencies.

From fellow pupils with the same ethnic background, more practical support is received, following from the positive density effects for the covariates indicating similarity in ethnic background. For Dutch pupils only a clear effect of reciprocity was found, modifying the double density effect for reciprocal dyads. The posterior distributions for the greater reciprocity effects among pupils with the same ethnic backgrounds are much wider for the other ethnic groups, containing large positive as well as negative values, which may be the result of the fact that for some of these groups our sample contains only a small number of children.

Concluding Remarks

With the multilevel p_2 model multiple parallel network observations can be analyzed. Such data are likely to be gathered in, for instance, educational settings. Compared to the two-stage meta-analytic approach used by Baerveldt et al. (2004), the advantage is clear: All data can be analyzed in a single model, resulting in an increase in power to detect possible actor or dyadic covariate effects. Moreover, it is possible to investigate whether these effects differ over networks and whether these differences can be explained by network covariates. In the application, a small amount of between-network variability was found. In general, however, one should be careful not to include too many random effects for the fixed parameters at the network level. Including t random effects implies estimating $t(t+1)/2$ parameters in $\mathbf{\Omega}$. This means that the number of parameters soon becomes large compared to the number of networks. It also implies that the multilevel p_2 model is more easily applied in data sets with a larger number of networks.

The results obtained in the example show that reported practical support is more prevalent among pupils with the same ethnic background. Further research is intended to combine multiple types of relations with a multilevel data structure in a single model. Then the current model can be extended with, for example, the data on emotional support that were analyzed in Baerveldt et al. (2004).

Software for the multilevel p_2 model is available in StOCNET (Boer et al., 2003).

Acknowledgments

This research was supported by the Netherlands Organisation for Scientific Research, grant NWO-425-10-0003-P. The authors would like to thank Chris Baerveldt for his permission to use the data.

References

Baerveldt, C. (2000, March). *Pupil's networks in high schools. Network sampling, program and some results from a theory-oriented research on petty crime of pupils*. Paper presented at the

2nd International Network Sampling Workshop, Maastricht, the Netherlands.

Baerveldt, C., Van Duijn, M. A. J., Vermeij, L., & Van Hemert, D. A. (2004). Ethnic boundaries and personal choice. Assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils' networks. *Social Networks*, *26,* 55–74.

Boer, P., Huisman, M., Snijders, T. A. B., & Zeggelink, E. P. H. (2003). *StOCNET: An Open Software System for the Advanced Analysis of Social Networks (Version 1.7)* [Computer Software]. Groningen, the Netherlands: ProGAMMA/ICS. Retrieved February 8, 2006, from http://stat.gamma.rug.nl/stocnet.

Box, G. E. P., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading, MA: Addison-Wesley.

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician*, *49,* 327–335.

Gelfand, A. E., Sahu, S. K., & Carlin, B. P. (1995). Efficient parametrizations for normal linear mixed models. *Biometrika*, *82,* 479–488.

Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Edward Arnold.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Assosciation*, *100,* 286–295.

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, *77,* 33–50.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, *21,* 1087–1092.

Press, J. S. (1989). *Bayesian statistics*. New York: Wiley.

Snijders, T. A. B, & Baerveldt, C. (2003). A multilevel network study of the effects of delinquent behavior on friendship evolution. *Journal of Mathematical Sociology*, *27,* 123–151.

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

Van Duijn, M. A. J., Snijders, T. A. B., & Zijlstra, B. J. H. (2004). p_2: A random effects model with covariates for directed graphs. *Statistica Neerlandica*, *58,* 234–254.

Zijlstra, B. J. H., Van Duijn, M. A. J., & Snijders, T. A. B. (2005a). *MCMC estimation of the p_2 model for directed graphs with covariates*. Manuscript submitted for publication.

Zijlstra, B. J. H., Van Duijn, M. A. J., & Snijders, T. A. B. (2005b). Model selection in random effects models for directed graphs using approximated Bayesian factors. *Statistica Neerlandica*, *59*, 107–118.

Bonne J. H. Zijlstra

University of Groningen
Faculty of Behavioural and Social Sciences
Grote Rozenstraat 31
NL-9721 TG Groningen
The Netherlands
Tel. +31 50 363 6199
Fax +31 50 363 6226
E-mail b.j.h.zijlstra@rug.nl

