Multilevel Longitudinal Analysis of Social Networks

Tom A.B. Snijders and Johan Koskinen

University of Oxford
University of Groningen
University of Manchester

June 29, 2012

Traditionally, network analysis tended to consist of cases studies of single networks.

However, it is preferable to generalize to a population of networks.

Traditionally, network analysis tended to consist of cases studies of single networks.

However, it is preferable to generalize to a population of networks.

This is permitted, in principle, by multilevel network analysis in the sense of analyzing multiple similar networks, mutually independent.

This was proposed by Snijders & Baerveldt (*J. Math. Soc. 2003*).

Also see Entwisle, Faust, Rindfuss, & Kaneda (*AJS*, 2007) who also gave on overview of empirical work involving multiple networks.

Sample from Population of Networks

Suppose we have a sample indexed by $j = 1, \ldots, N$
from a population of networks,
where the networks are ‘replications’ of each other
in the following sense:

Sample from Population of Networks

Suppose we have a sample indexed by $j = 1, \ldots, N$
from a population of networks,
where the networks are 'replications' of each other
in the following sense:

they all are regarded as realizations of
processes obeying the same model,
but having different parameters $\theta_1, \ldots, \theta_j, \ldots, \theta_N$.

.

Several approaches are possible for combining such data:

Several approaches are possible for combining such data:

1. Meta analysis: no population assumption:
   Fisher combination of independent tests.

Several approaches are possible for combining such data:

1. Meta analysis: no population assumption:
   Fisher combination of independent tests.
2. Meta analysis: population assumption $\sim$ random effects, no distributional assumptions:
   two-stage meta analysis.

Several approaches are possible for combining such data:

1. Meta analysis: no population assumption:
   Fisher combination of independent tests.
2. Meta analysis: population assumption $\sim$ random effects, no distributional assumptions:
   two-stage meta analysis.
3. Meta analysis: population assumption $\sim$ random effects, assume multivariate normal distribution for θ_j :
   integrated hierarchical approach.

Meta-Analysis $\sim$ Fixed Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are arbitrary values,
no assumption about a population is made.

Meta-Analysis ∼ Fixed Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are arbitrary values,
no assumption about a population is made.

two-stage procedure:
estimate each θ_j separately, combine the results by *Fisher's procedure for combining independent tests*:

'*is there any evidence for a hypothesized effect?*'

Meta-Analysis ∼ Fixed Effects Model (contd.):

For coordinate k of the parameter, test null hypothesis

$$H_0 : \theta_{kj} = 0 \text{ for all } j$$

against alternative hypothesis

$$H_1 : \theta_{kj} = 0 \text{ for at least one } j \, .$$

(Two-sided variants also are possible; SIENA manual.)

Procedure: see, e.g., Snijders & Bosker Section 3.7.

Mercken, Snijders, Steglich, & de Vries (2009)
applied this in a study of smoking initiation:
7704 adolescents in 70 schools in 6 countries.

Meta-Analysis ∼ Random Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are drawn randomly
from a *population* $\mathcal{P}$[net] of networks,
no further distributional assumptions are made.

Meta-Analysis ~ Random Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are drawn randomly
from a *population* $\mathcal{P}$[net] of networks,
no further distributional assumptions are made.

Two-stage procedure:
estimate each θ_j separately,
combine the results in a meta-analysis (Cochran 1954),
(*'V-known problem* in multilevel analysis)
which allows testing hypotheses about $\mathcal{P}$[net]
such as, for a coordinate k ,

$$\begin{aligned} H_0^{\text{total}}: &\quad \text{all } \theta_{kj} = 0; \\ H_0^{\text{mean}}: &\quad \mathrm{E}\{\theta_{kj}\} = 0; \\ H_0^{\text{spread}}: &\quad \mathrm{var}\{\theta_{kj}\} = 0. \end{aligned}$$

The input for the meta-analysis consists of estimates $\hat{\theta}_j$ and their standard errors s.e.$_j$.

The meta analysis is constructed based on the model

$$\hat{\theta}_j = \mu + U_j + E_j \ ,$$

where μ is the population mean,
U_j is the true effect of group j,
and E_j is the statistical error of estimation.

The input for the meta-analysis consists of estimates $\hat{\theta}_j$ and their standard errors s.e.$_j$.

The meta analysis is constructed based on the model

$$\hat{\theta}_j = \mu + U_j + E_j \,,$$

where μ is the population mean,
U_j is the true effect of group j,
and E_j is the statistical error of estimation.

U_j and E_j are independent residuals with mean 0,
the U_j are *i.i.d.* with unknown variance,
and $\text{var}(E_j) = \text{s.e.}_j^2$ ('V–known').

Implemented in MLwiN, HLM, R package *Metafor*,
RSiena function *siena08*.

Meta-Analysis ~ Integrated Random Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are drawn randomly
from a *population* $\mathcal{P}$[net] of networks,
and are assumed to have
a common multivariate normal $\mathcal{N}(\mu, \Sigma)$ distribution,
perhaps conditionally on network-level covariates.

Meta-Analysis ~ Integrated Random Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are drawn randomly
from a *population* $\mathcal{P}$[net] of networks,
and are assumed to have
a common multivariate normal $\mathcal{N}(\mu, \Sigma)$ distribution,
perhaps conditionally on network-level covariates.

Integrated procedure:
Estimate μ and Σ and consider the
'posterior' distribution of θ_j given the data.

Meta-Analysis ∼ Integrated Random Effects Model:

$\theta_1, \ldots, \theta_j, \ldots, \theta_N$ are drawn randomly
from a *population* $\mathcal{P}$[net] of networks,
and are assumed to have
a common multivariate normal $\mathcal{N}(\mu, \Sigma)$ distribution,
perhaps conditionally on network-level covariates.

Integrated procedure:
Estimate μ and Σ and consider the
'posterior' distribution of θ_j given the data.

Advantage:
The analysis of the separate networks draws strength from the total sample of networks by regression to the mean.

Useful especially for many small networks.

Meta-Analysis ∼ Integrated Random Effects Model (contd.)

New developments
for the stochastic actor-oriented model for network dynamics,
implemented in the SIENA program.

Recall that this is a model for network dynamics,
where the dynamics is
an unobserved sequence of 'micro steps'
and the parameters are estimated from network panel data.

Meta-Analysis ∼ Integrated Random Effects Model (contd.)

New developments
for the stochastic actor-oriented model for network dynamics,
implemented in the SIENA program.

Recall that this is a model for network dynamics,
where the dynamics is
an unobserved sequence of 'micro steps'
and the parameters are estimated from network panel data.

This is elaborated following a likelihood-based approach;
see Koskinen & Snijders (*JSPI* 2007),
Snijders, Koskinen & Schweinberger (*AAS* 2010),
Schweinberger (PhD thesis 2007, Chapters 4 and 5).

Here we discuss a Bayesian approach, where the parameters μ, Σ have a prior distribution. We assume the conjugate prior,

- $\Sigma^{-1} \sim \text{wishart}_p(\Lambda_0^{-1}, \nu_0)$, and conditionally on Σ
- $\mu \mid \Sigma \sim N_p(\mu_0, \Sigma/\kappa_0)$.

Thus, the parameters of the prior are $\Lambda_0, \nu_0, \kappa_0$.

For the 'basic rate parameters' ρ, normal distributions are assumed after transforming to $\sqrt{\rho}$,
representing the greater relative uncertainty at higher levels (with pragmatic truncation to ensure positivity).

.

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$, is

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$, is

$$f_{\mathrm{InvWish}}\big(\Sigma \mid \Lambda_0^{-1}, \nu_0\big)\, \phi_p\big(\mu \mid \mu_0, \Sigma/\kappa_0\big) \quad \text{prior}$$

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$, is

$$f_{\text{InvWish}}\Big(\Sigma \mid \Lambda_0^{-1}, \nu_0\Big)\, \phi_p\Big(\mu \mid \mu_0, \Sigma/\kappa_0\Big) \quad \text{prior}$$

$$\times \prod_{j=1}^{N} \phi_p(\theta_j \mid \mu, \Sigma) \quad \text{hierarchical model}$$

The joint p.d.f., for data $y_1, \dots, y_j, \dots, y_N$, is

$$f_{\text{InvWish}}\left(\Sigma \mid \Lambda_0^{-1}, \nu_0\right) \phi_p\left(\mu \mid \mu_0, \Sigma/\kappa_0\right) \quad \text{prior}$$

$$\times \prod_{j=1}^{N} \phi_p(\theta_j \mid \mu, \Sigma) \quad \text{hierarchical model}$$

$$\times \prod_{j=1}^{N} p_{\text{SAOM}}(y_j \mid \theta_j) \quad \text{network model}$$

Since $p_{\text{SAOM}}(y_j \mid \theta_j)$ cannot be calculated directly,
we employ *data augmentation* (Tanner & Wong, 1987):
augment the network panel data by the sequence v_j
of all microsteps connecting the consecutive observations.

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$,
using data augmentation, is the sum over v_j of

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$,
using data augmentation, is the sum over v_j of

$$f_{\text{InvWish}}\Big(\Sigma \mid \Lambda_0^{-1}, \nu_0\Big)\, \phi_p\Big(\mu \mid \mu_0, \Sigma/\kappa_0\Big) \quad \text{prior}$$

$$\times \prod_{j=1}^{N} \phi_p(\theta_j \mid \mu, \Sigma) \quad \text{hierarchical model}$$

$$\times \prod_{j=1}^{N} p_{\text{SAOM}}(v_j \mid \theta_j, y_j)\ . \quad \text{network model}$$

The joint p.d.f., for data $y_1, \ldots, y_j, \ldots, y_N$,
using data augmentation, is the sum over v_j of

$$
\begin{array}{ll}
f_{\text{InvWish}}\Big(\Sigma \mid \Lambda_0^{-1}, \nu_0\Big)\, \phi_p\Big(\mu \mid \mu_0, \Sigma/\kappa_0\Big) & \text{prior} \\
\times \prod_{j=1}^{N} \phi_p(\theta_j \mid \mu, \Sigma) & \text{hierarchical model} \\
\times \prod_{j=1}^{N} p_{\text{SAOM}}(v_j \mid \theta_j, y_j)\ . & \text{network model}
\end{array}
$$

The posterior distribution can be sampled
by Markov chain Monte Carlo (MCMC).
The unknown random variables are

$$\mu, \Sigma;\ \theta_1, \ldots, \theta_N;\ v_1, \ldots, v_N$$

and these are sampled in turn, as follows.

1. For all j make some Metropolis Hastings steps
   sampling $v_j \mid y_j, \theta_j$,
   as in Snijders, Koskinen & Schweinberger (2010).
   This is implemented already for the Maximum Likelihood estimation procedure in SIENA.
   Works well, but time consuming.

1. For all j make some Metropolis Hastings steps
   sampling $v_j \mid y_j, \theta_j$,
   as in Snijders, Koskinen & Schweinberger (2010).
   This is implemented already for the Maximum Likelihood estimation procedure in SIENA.
   Works well, but time consuming.
2. For all j make one or more Metropolis Hastings steps
   sampling $\theta_j \mid v_j, \mu, \Sigma$,
   using a random walk proposal distribution
   (Schweinberger 2007, Ch. 5.4;
   Koskinen & Snijders 2007, Sect. 4.4).
   Covariance matrix for proposals obtained as
   covariance matrix of groupwise MoM estimators
   approximated at quick and easy initial values,
   and scaled to obtain $\sim 40\%$ acceptance rates.

3. Sample $(\mu, \Sigma) \mid \theta_1, \ldots, \theta_N, \Lambda_0, \nu_0, \kappa_0$ from the full conditional distribution (Gibbs sampling step).

This requires tuning to obtain good mixing – as usual.

Still time-consuming.

Example: data Andrea Knecht

As an example,
we use friendship networks in 21 school classes
from the study by Andrea Knecht (PhD thesis Utrecht, 2008);
see Knecht, Snijders, Baerveldt, Steglich, & Raub,
'Friendship and Delinquency:
Selection and Influence Processes in Early Adolescence',
Social Development, 2010.

We consider a model for a longitudinal study with 2 waves,
and with 9 parameters:
rate of change; outdegree; reciprocity; transitive triplets;
3-cycles; delinquency ego, alter, ego × alter; sex similarity.

The Bayesian MCMC procedure produces, if there is convergence
(i.e., hopefully, after a burn-in period),
a sample from the posterior distribution of all the parameters,
both the θ_j referring to the individual sampled networks,
and μ and Σ referring to the population of networks.

The inference is based on these sampled posterior distributions.

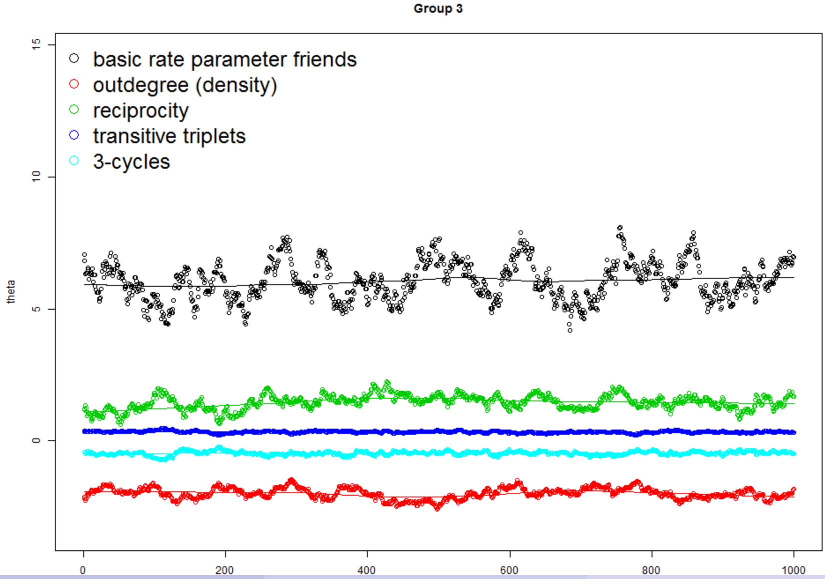
Two kinds of plot will be given:

1. *trace plots*, representing successive draws from the posterior distribution (after *thinning*),
2. *density plots*, representing the plausible values of the parameters, given the observed data.

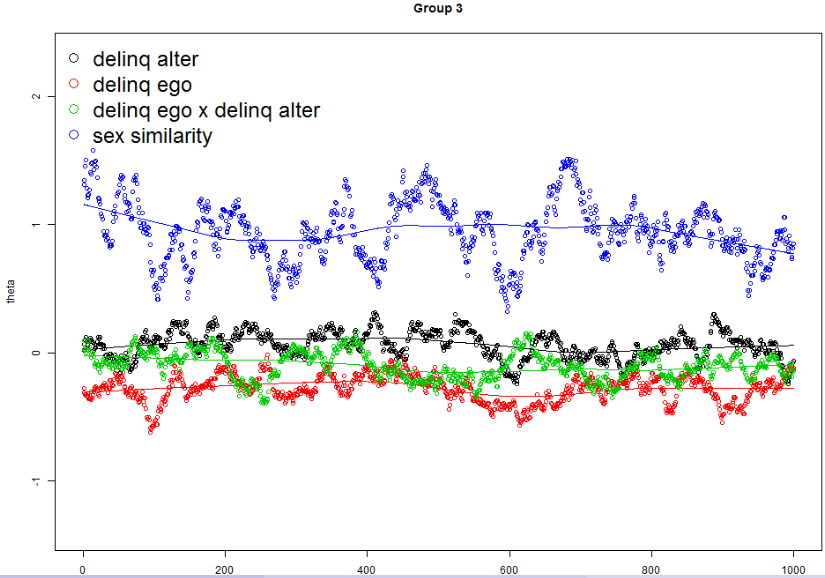
For the MCMC algorithm, we used:

1. groupwise number of MH iterations for sampling micro-steps varies between 75–500 depending on distance between observed networks;
2. 2,000 iterations sampling θ_j, μ, Σ for warmup
3. 20,000 iterations sampling θ_j, μ, Σ for estimation, with a thinning ratio of 1:20.

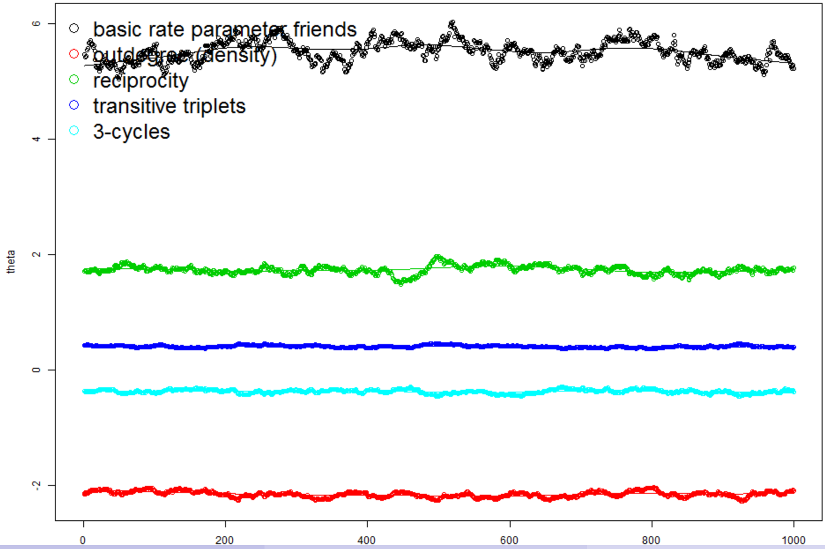
Trace plots for (e.g.) group 3, structural effects:



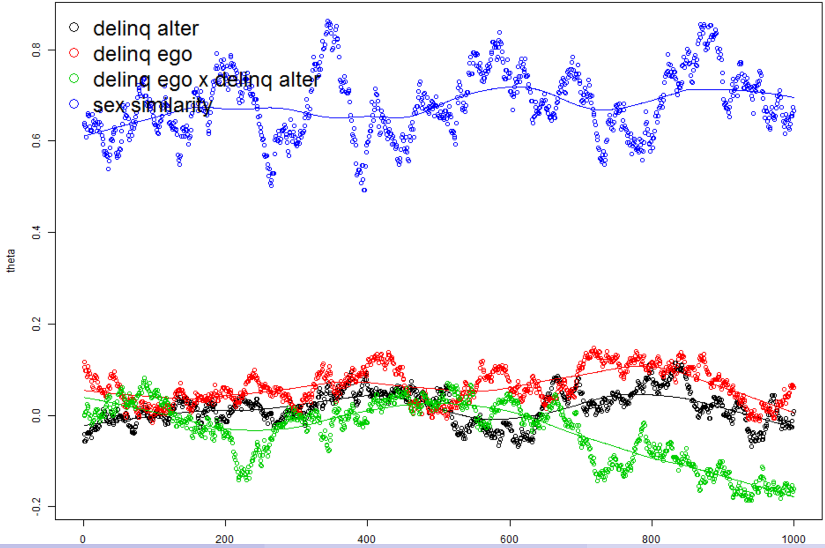
Trace plots for group 3, covariate effects:



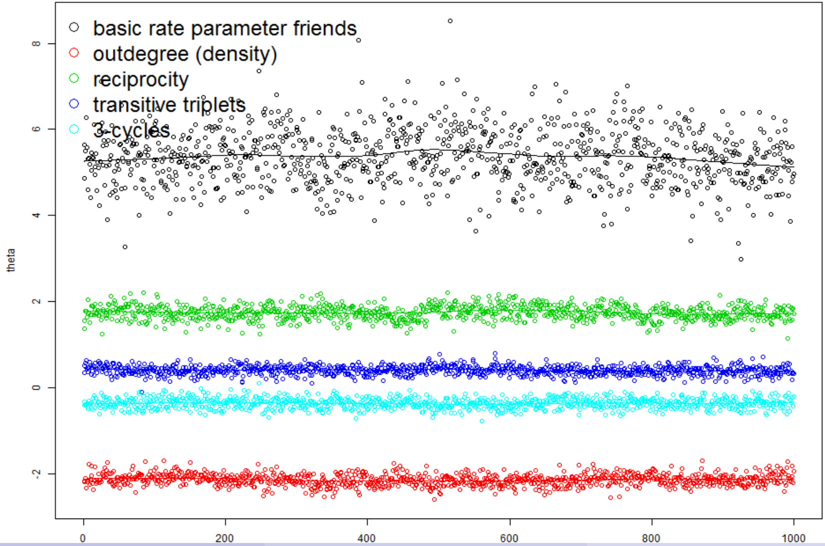
Trace plots average posterior $\bar{\theta}_{\cdot}$: structural effects



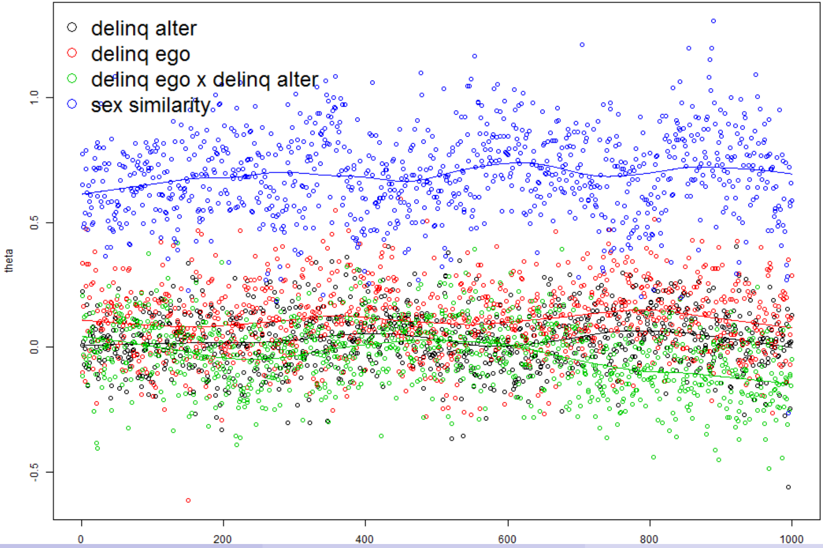
Trace plots average posterior $\bar{\theta}_{.}$: covariate effects



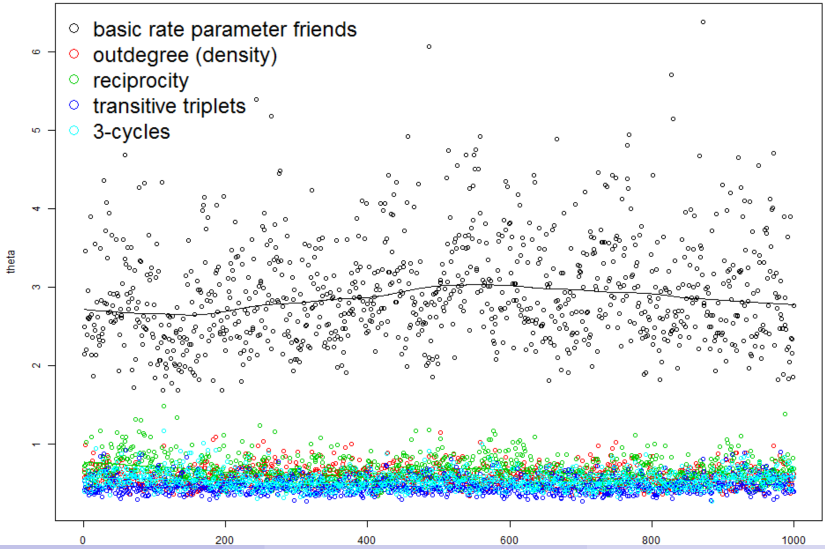
Trace plots of posterior μ: structural effects



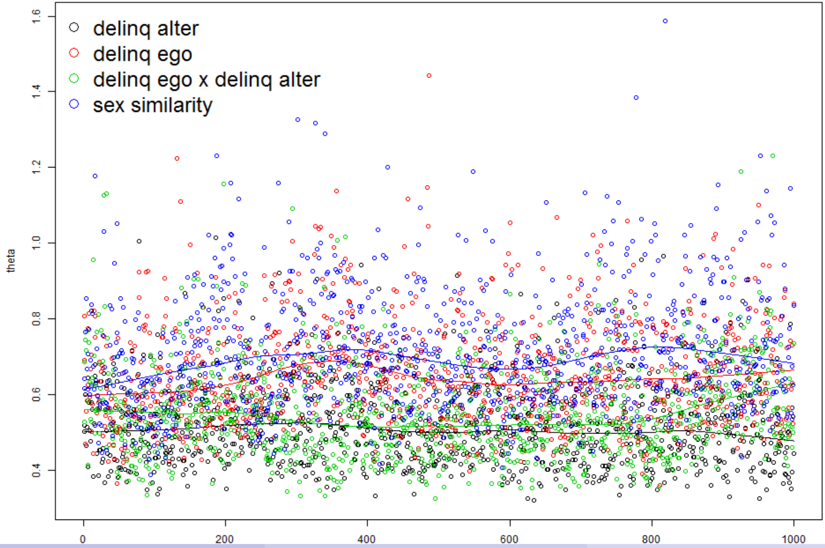
Trace plots of posterior μ: covariate effects



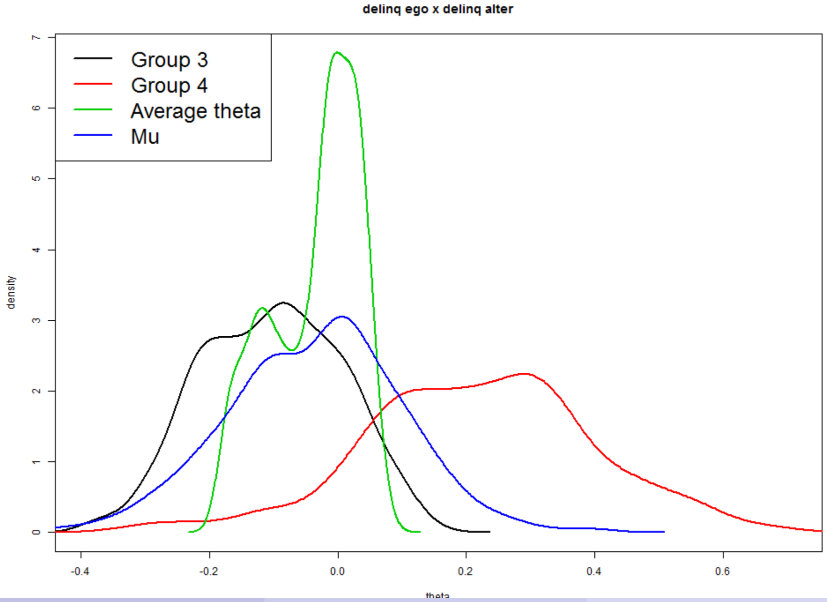
Trace plots of posterior σ_k: structural effects



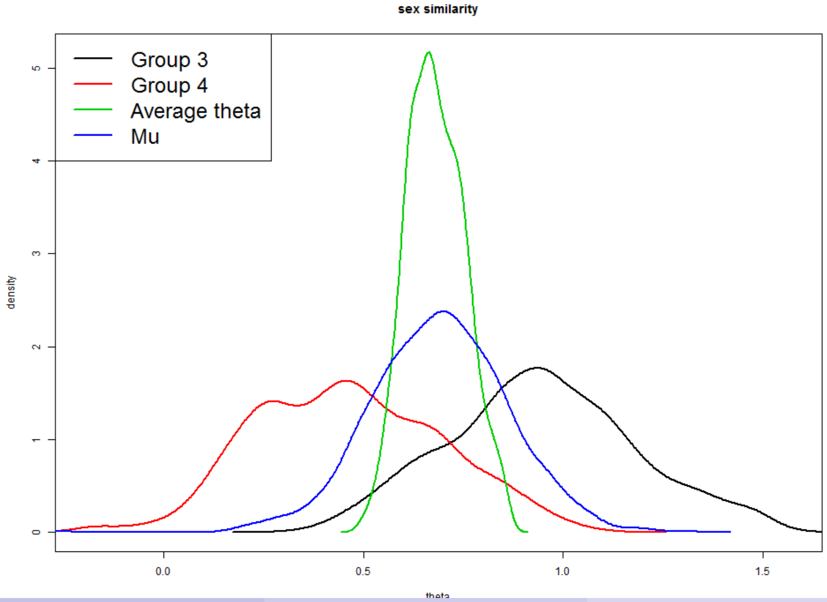
Trace plots of posterior σ_k: covariate effects



Density plots for del. ego × del. alter; groups 3,4



Density plots for sex similarity; groups 3,4



Conclusion

The method seems to work well.

It is promising for analyzing collections of small networks; however, time-consuming.

Note the much larger posterior uncertainty for μ
compared to $\bar{\theta}_{.}$;
this is a general feature of multilevel modeling,
more apparent for small numbers of highest-level units.

.