An example of multilevel analysis of network dynamics using `sienaBayes`

Tom A.B. Snijders

University of Groningen
University of Oxford

June, 2023

Example: data Andrea Knecht

This is an example of the use of the function `sienaBayes` for the estimation of multilevel longitudinal network models.

The data set used is about 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, dependent variables friendship and delinquent behavior.

Model specification

In addition to the regular effects, for multilevel models
we should think about group-level effects.

1. The groups may have different numbers n of actors.
   Snijders (2005; Section 11.13(B)) derives that for the empty model the outdegree parameter will have a component approximately $-\frac{1}{2}\log(n)$.
   Therefore it is recommendable to include $\log(n)$ as a covariate;
   the expected regression coefficient is something like -0.5,
   but for non-empty models the value will differ and is unknown.
   An alternative is to use a creation effect of $\log(n)$;
   then the expected coefficient is something like –1.
2. Furthermore, it may be advisable to include interactions of $\log(n)$ with reciprocity and transitivity.
3. Other group-level variables may also be relevant.

However, the number of group-level variables should not be too large!
The same considerations apply as for the number of covariates,
given sample size, in linear regression models;
sample size here is number of groups, which usually is small.

Within- and between-group regressions

Similar to the Hierarchical Linear Model of multilevel analysis, we should be aware that within-group regression coefficients may differ from between-group coefficients.
The group mean of covariates, or dependent behavior variables, may be included in the model to account for this.

By implication, cross-level interactions may be included.

In this case, we include the group mean of delinquency (ego) and the interaction of this group mean with delinquency alter.

Effects of delinquency on network evolution

From Snijders & Lomi (*Network Science*, 2019),
we know that actor variables (here: delinquency)
may have a variety of effects on networks,
because such effects imply a level transition monadic ⇒ dyadic.

In a first analysis, the five-parameter model was used:
$V(\text{ego})$, $V(\text{alter})$, $V^2(\text{ego})$, $V^2(\text{alter})$, $\big(V(\text{ego}) - V(\text{alter})\big)^2$.

From a provisional analysis it seemed that delinquency alter,
for given ego, has approximately a linear effect.
Therefore the model was reduced to four parameters:
$V(\text{ego})$, $V(\text{alter})$, $V^2(\text{ego})$, $V(\text{ego}) \times V(\text{alter})$.

Summary of model specification

Network dynamics:

outdegree; reciprocity; transitive triplets; transitive reciprocated triplets; indegree popularity; outdegree activity; reciprocal degree-activity; old friends; same sex; $\log(n)$;

dependence on delinquency V:
V(ego), V(alter); V^2(ego); $V(\text{ego}) \times V(\text{alter})$;
V(group mean ego); $V(\text{group mean ego}) \times V(\text{alter})$;

Delinquency dynamics:

linear shape; quadratic shape; sex; average alter.

For 21 groups, 2 is a rather high number of group-level variables.
Therefore a prior distribution (mean 0, variance 1) was assumed for the variables $\log(n)$ and V(group mean ego) .

Random / fixed

The choice for which parameters to define as random was based on a preliminary multi-group analysis by MoM (`siena07`)
where all parameters were assumed fixed,
followed by `sienaTimeTest` to test parameter homogeneity;
parameters with the largest test statistics were defined as random.

The following effects were considered to vary randomly:

Network dynamics:
outdegree; reciprocity; transitive triplets; indegree popularity; outdegree activity; reciprocal degree-activity; delinquency ego; same sex;

Delinquency dynamics:
linear shape; quadratic shape.

Prior distributions

Prior for rates: data dependent, calculated internally;
other prior means for μ: outdegree –2, same sex 0.4, others 0;
other prior variances for μ: 0.01;
prior covariances: all 0;
prior variances for η: infinite, but for group-level variables 1;
prior κ: 0.01.

This means that the between-group differences of parameters $\theta_j^{(1)}$
are thought to be in the order of magnitude of $\sqrt{0.01} = 0.1$,
and the uncertainty about the value of the prior means of $\theta_j^{(1)}$
is of the order of $\sqrt{0.01/0.01} = 1$.

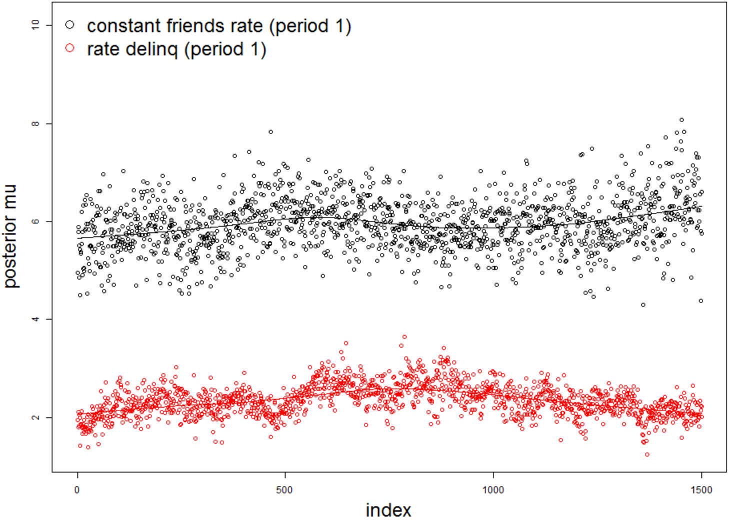
For the MCMC algorithm, we used:

1. groupwise number of MH iterations for sampling mini-steps varies between 100–600 depending on distance between observed networks;
2. $\Rightarrow$ 500 iterations sampling $\theta_j^{(1)}, \eta, \mu, \Sigma$ for warmup

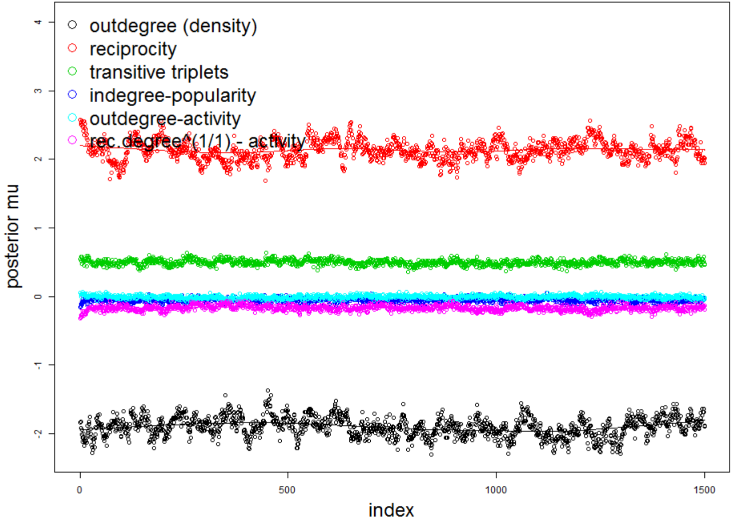
   $\Rightarrow$ 1000 iterations sampling $\theta_j^{(1)}, \mu, \Sigma$
   and 3000 for sampling η for estimation, with a thinning ratio of 1:40.

The results are provisional,
because a good convergence check was not carried out.

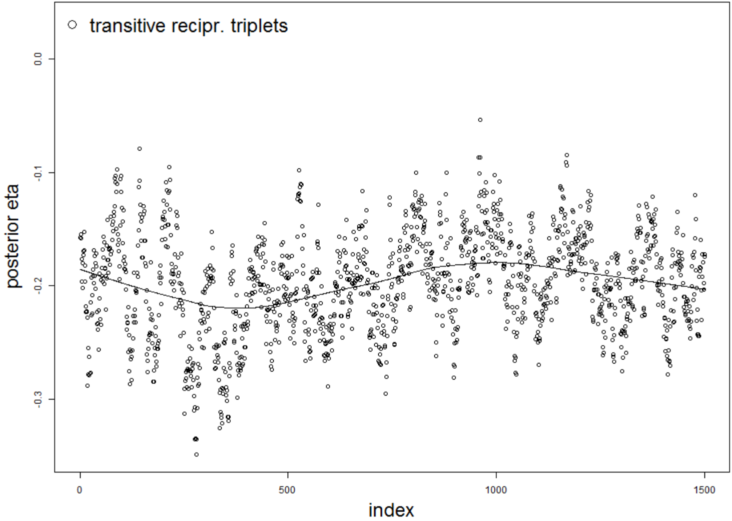
Trace plots for rate parameters:



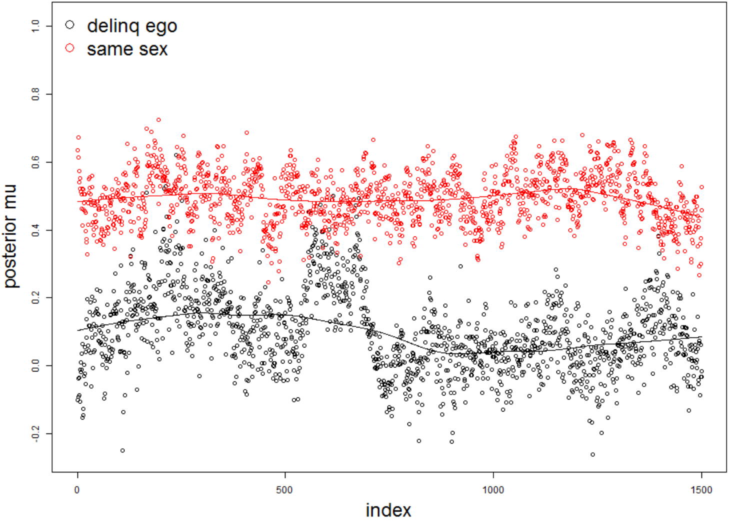
Trace plots for means (μ) of structural network effects



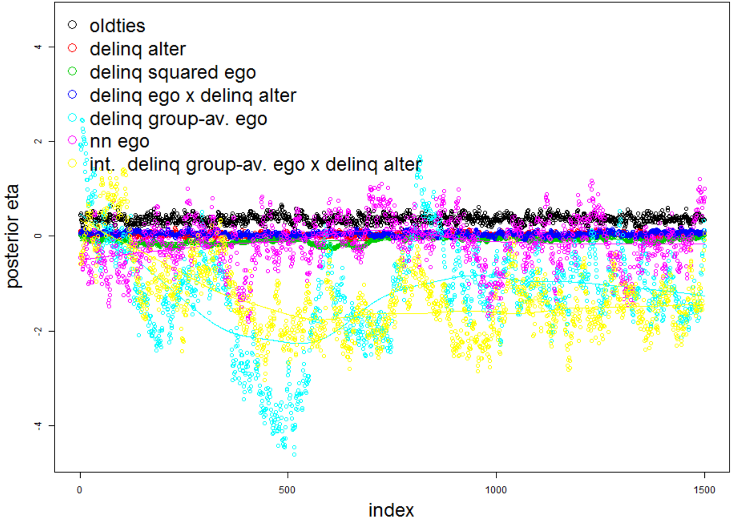
Trace plots for fixed parameters (η) of structural network effects



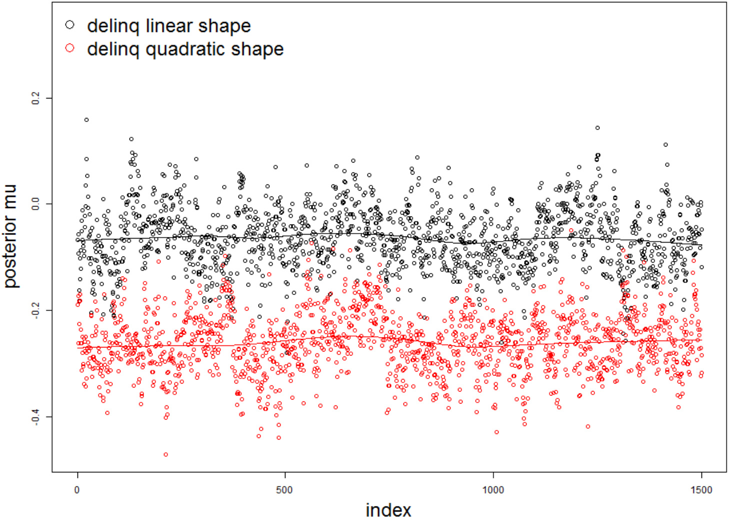
Trace plots for means (μ) of covariate network effects



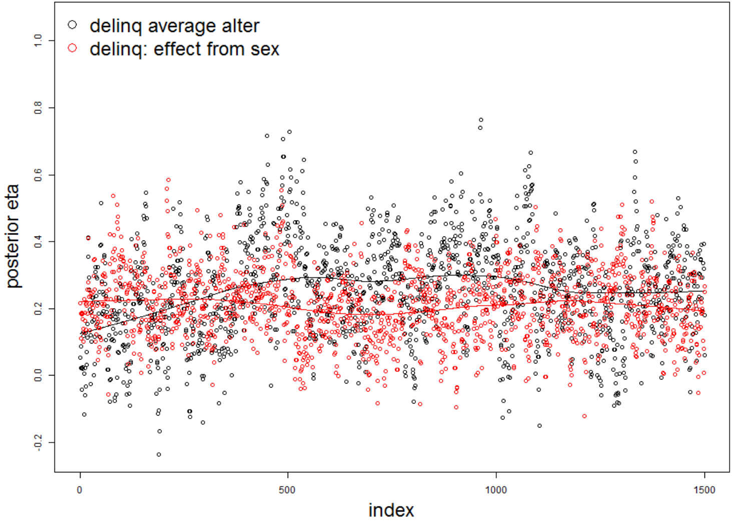
Trace plots for fixed parameters (η) of covariate network effects



Trace plots for means (μ) of behavior effects



Trace plots for fixed parameters (η) of behavior effects



Conclusion: non-stationarity mainly in warming phase,
but up to run 800 there still seems some non-stationarity.
A longer run is necessary!

The following page shows
posterior means and standard deviations
with 95 % credibility interval of $E(\theta_j)$,
computed from runs 801-1500. for Bayesian estimation
of friendship and delinquency dynamics
in 21 classrooms (data Andrea Knecht).

Effect	posterior mean	(s.d.)	interval from	to	varying
Friendship dynamics					
outdegree (density)	–1.932	(0.145)	–2.208	–1.648	+
reciprocity	2.132	(0.133)	1.895	2.405	+
transitive triplets	0.493	(0.037)	0.421	0.565	+
transitive recipr. triplets	–0.186	(0.038)	–0.262	–0.117	–
indegree-popularity	–0.052	(0.030)	–0.113	0.004	+
outdegree-activity	–0.004	(0.024)	–0.051	0.042	+
recipr. degree - activity	–0.175	(0.037)	–0.251	–0.105	+
old ties	0.368	(0.083)	0.200	0.534	–
delinq alter	0.035	(0.037)	–0.035	0.111	–
delinq ego	0.053	(0.088)	–0.114	0.234	+
delinq squared ego	–0.043	(0.045)	–0.127	0.046	–
delinq ego × delinq alter	0.045	(0.044)	–0.042	0.127	–
delinq group-average ego	–0.863	(0.690)	–1.857	0.727	–
del gr-av ego × del alter	–1.532	(0.535)	–2.522	–0.463	–
same sex	0.498	(0.078)	0.356	0.646	+
log(n) ego	–0.152	(0.539)	–1.264	0.833	–

Post. means, standard dev.s, and 95 % credibility intervals for μ, η; varying between classrooms: + = "yes", – = "no".

Effect	posterior		interval		varying
	mean	(s.d.)	from	to	
Delinquency dynamics					
linear shape	–0.068	(0.059)	–0.178	0.047	+
quadratic shape	–0.264	(0.054)	–0.368	–0.155	+
average alter	0.268	(0.147)	–0.030	0.546	–
effect from sex	0.212	(0.104)	0.007	0.421	–

Posterior means, standard deviations, and 95 % credibility intervals for μ, η;

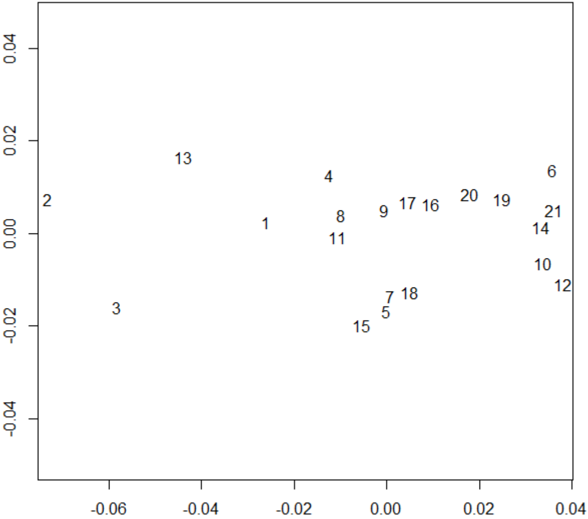
varying between classrooms: + = "yes", – = "no".

The hierarchical multilevel approach gives
much more information:
for each group, the posterior distribution of the parameters
(which are constant across groups for the fixed parameters).

In the following plots, for the varying parameters,
note the difference between
the posterior density of μ
and that of average θ_j;

the average θ_j has a bearing on this sample only;
for μ, there is the extra uncertainty
due to generalisation from sample to population.

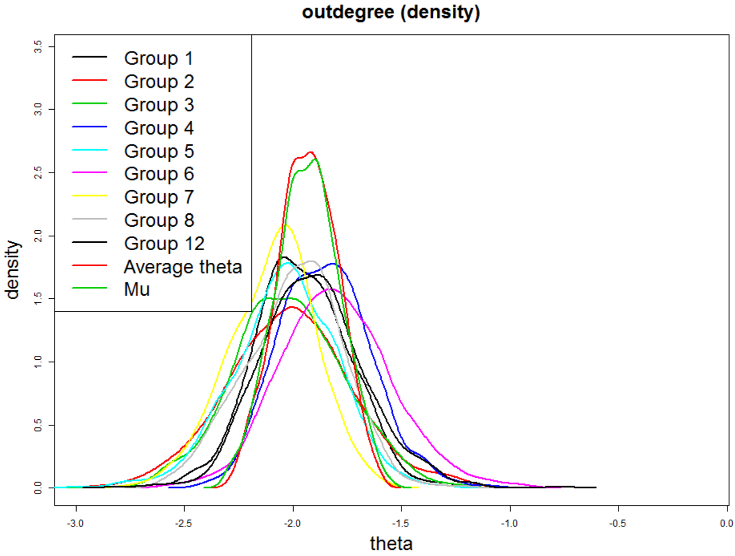
MDS plot of posterior means



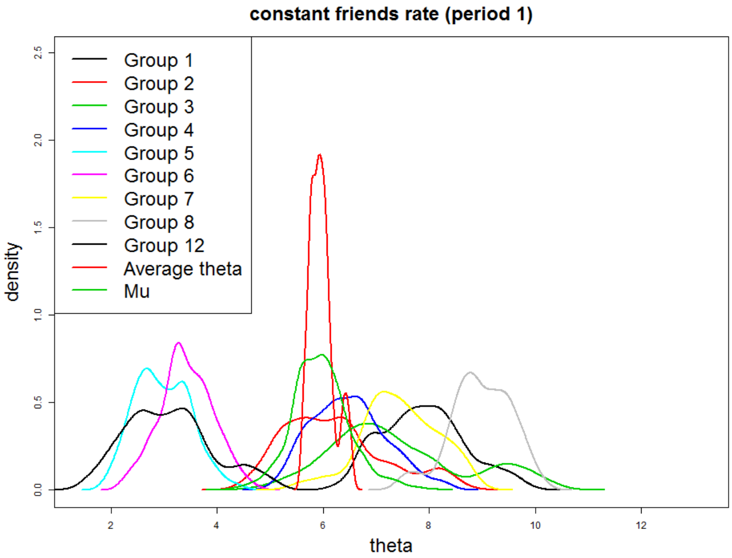
The MDS plot of the previous page
can be used as a diagnostic.

Based on this, we select groups 1–8 and 12 for plotting.

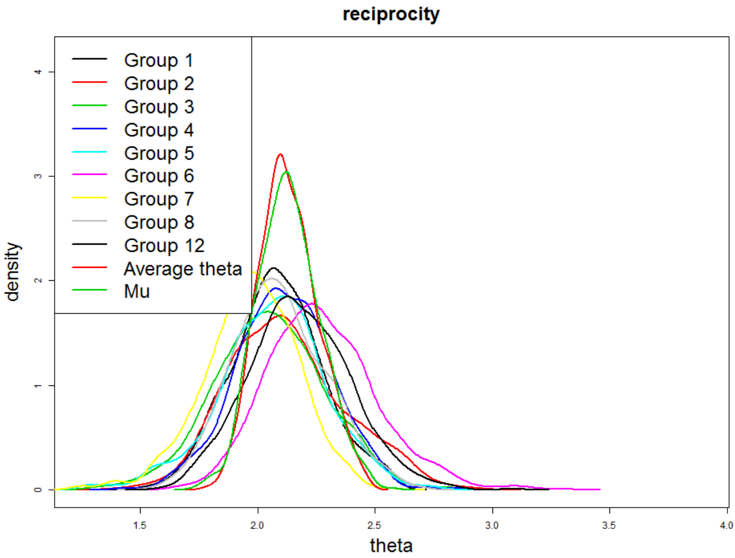
Density plot for outdegree



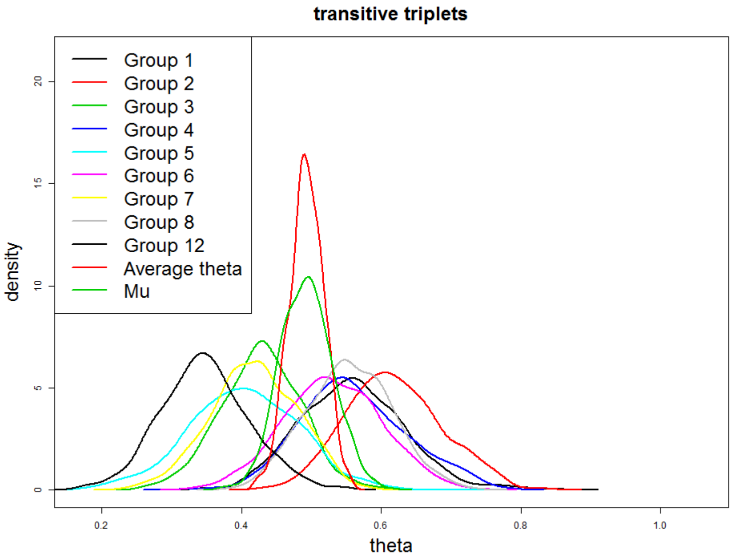
Density plot for network rate



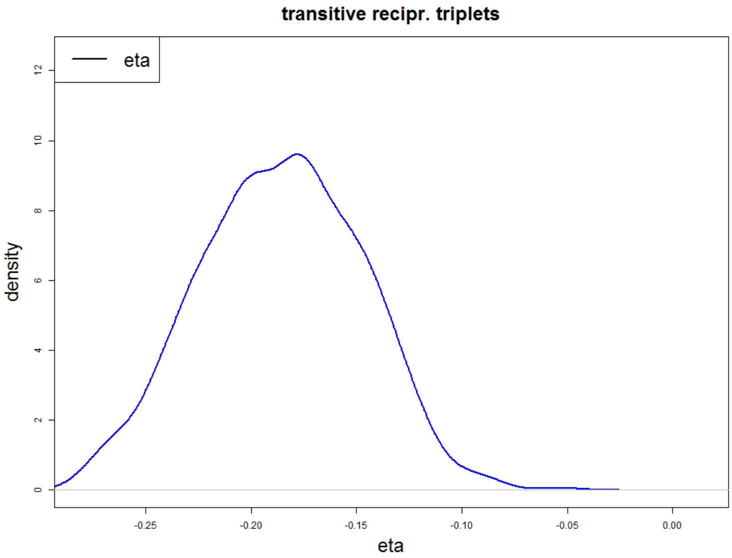
Density plot for reciprocity



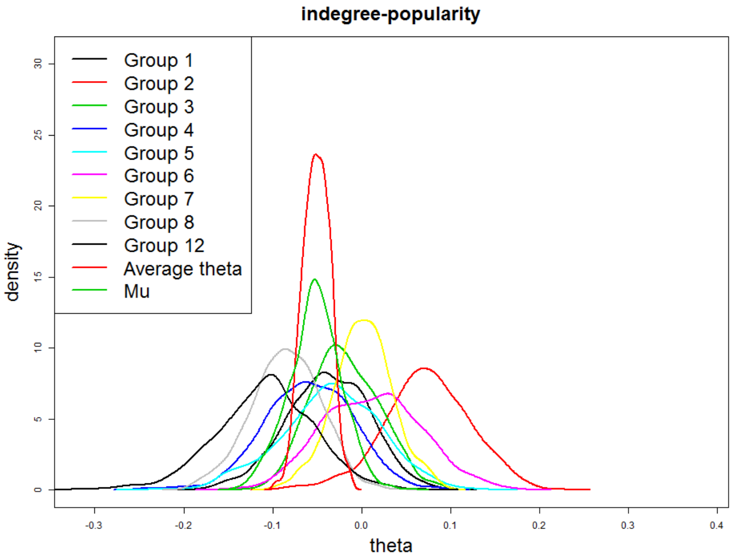
Density plot for transitive triplets



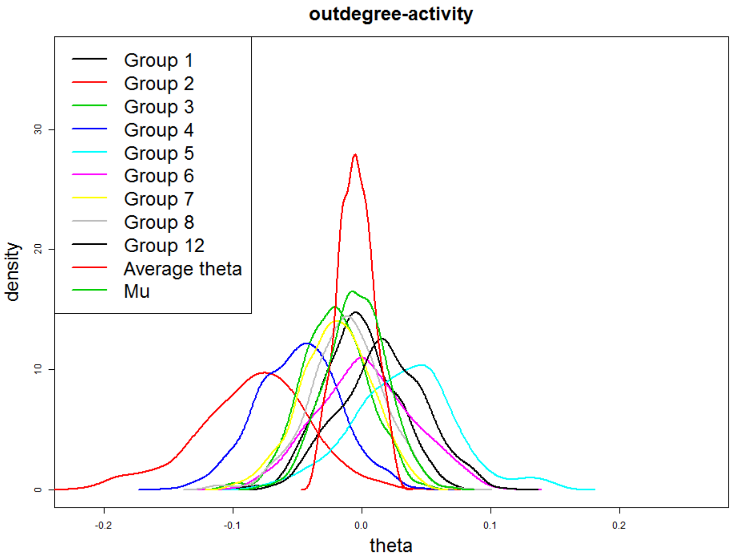
Density plot for transitive reciprocated triplets



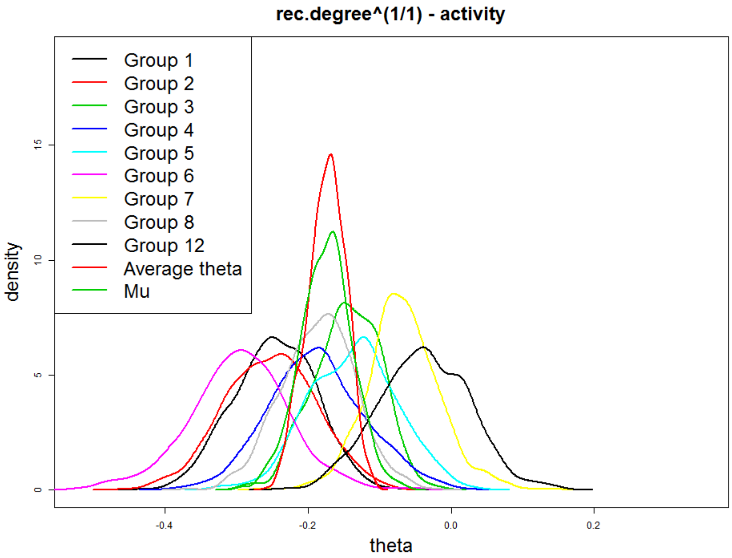
Density plot for indegree-popularity



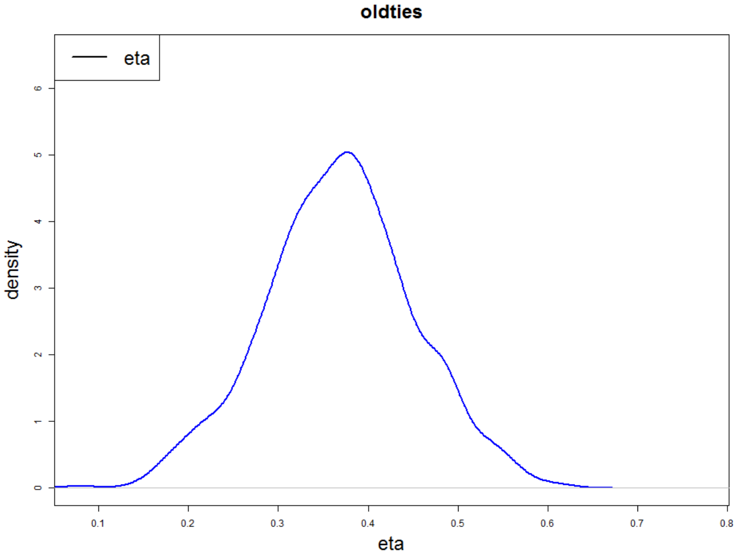
Density plot for outdegree-activity



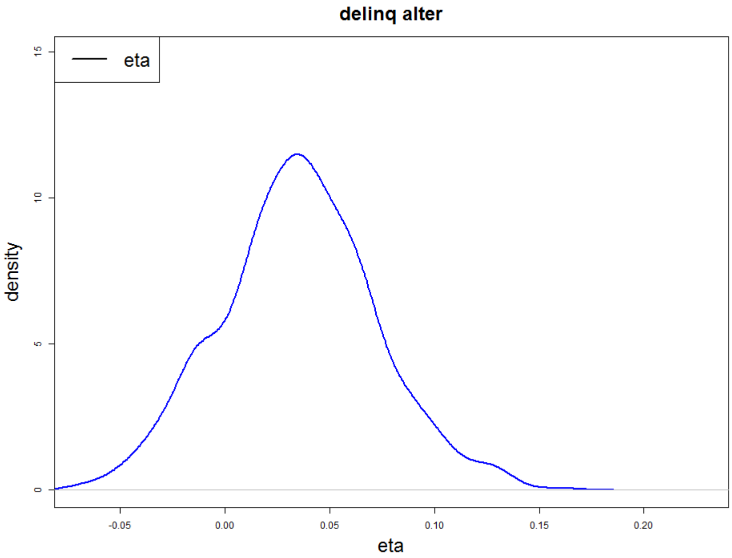
Density plot for reciprocal degree-activity



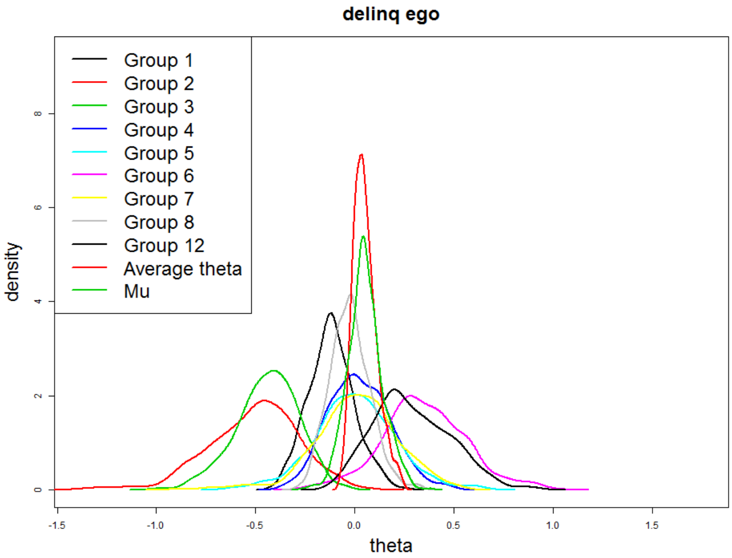
Density plot for old ties



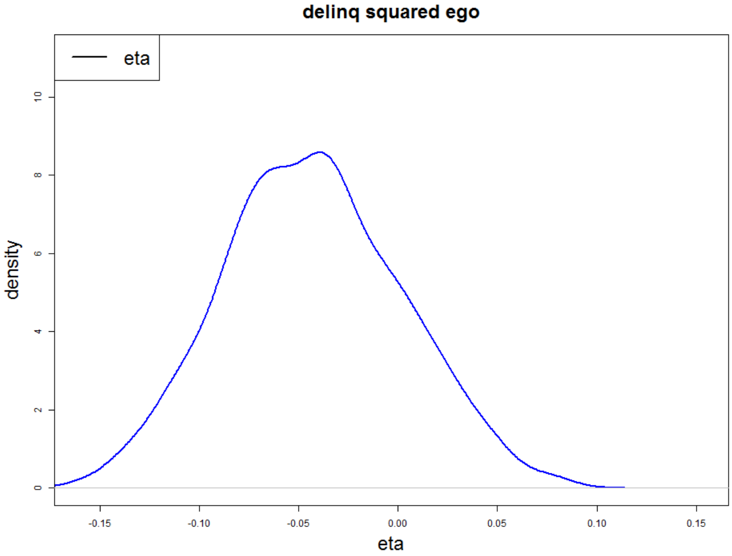
Density plot for delinquency alter



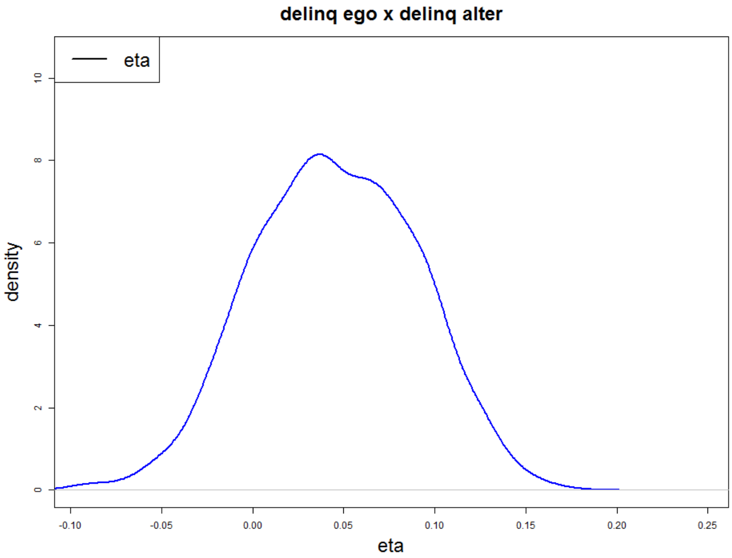
Density plot for delinquency ego



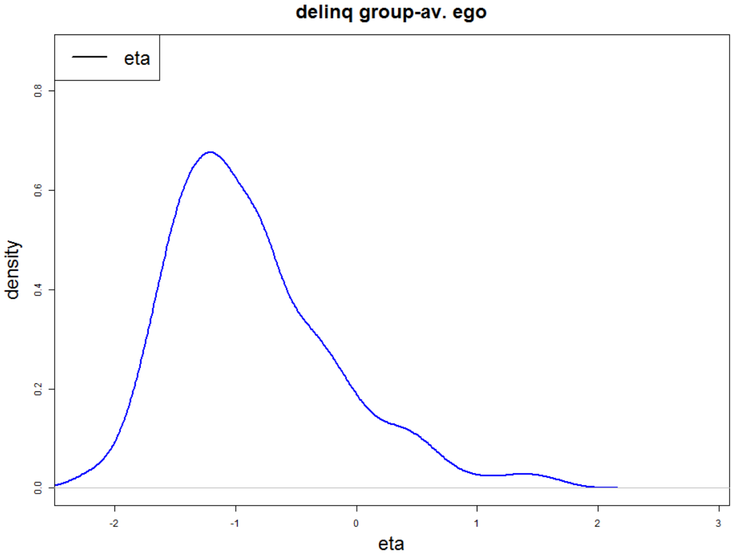
Density plot for delinquency squared ego



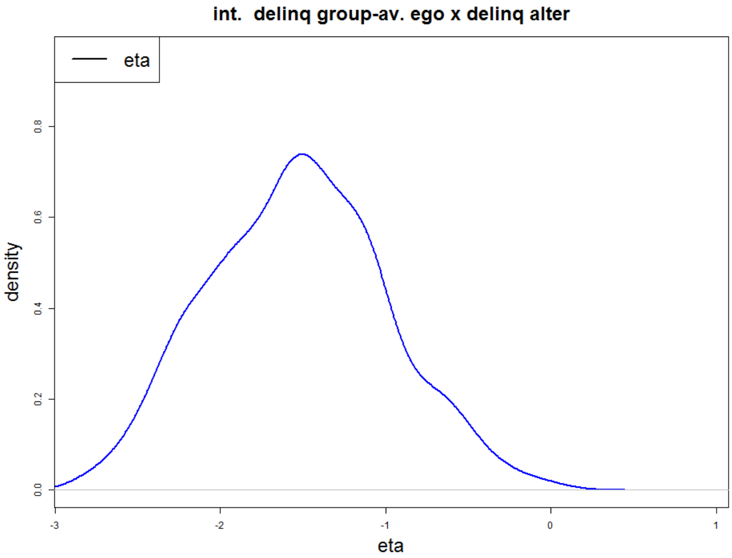
Density plot for del. ego × del. alter



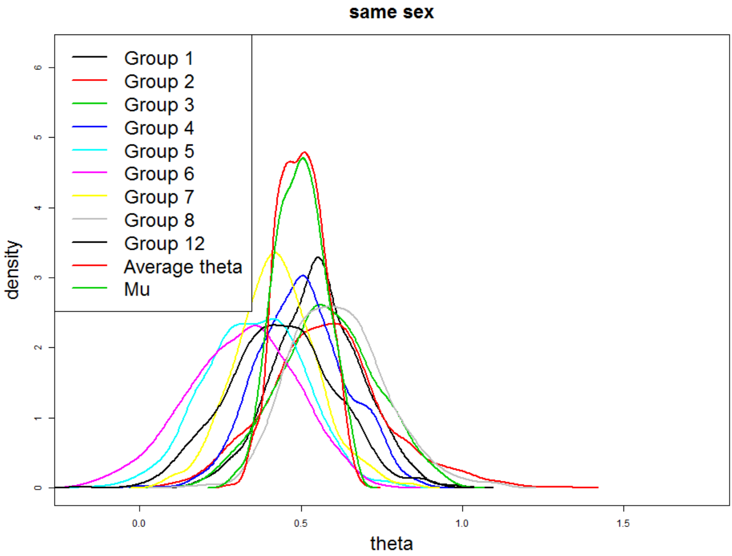
Density plot for delinquency group-average ego



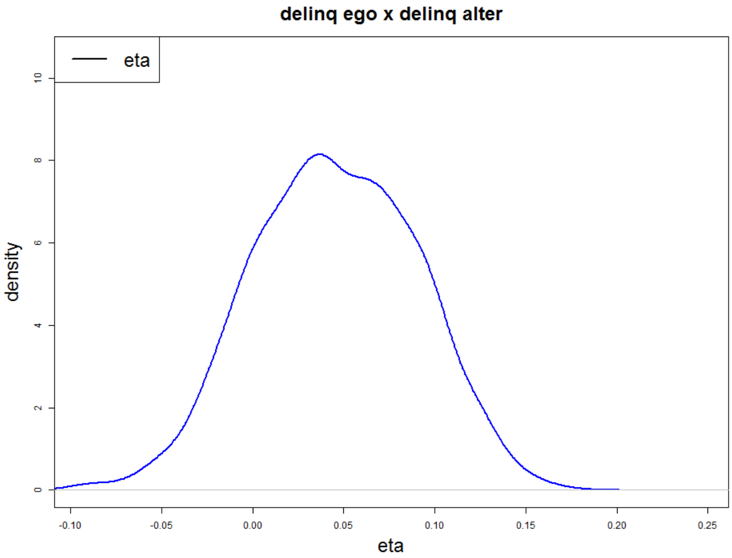
Density plot for delinquency group-average ego × delinquency alter



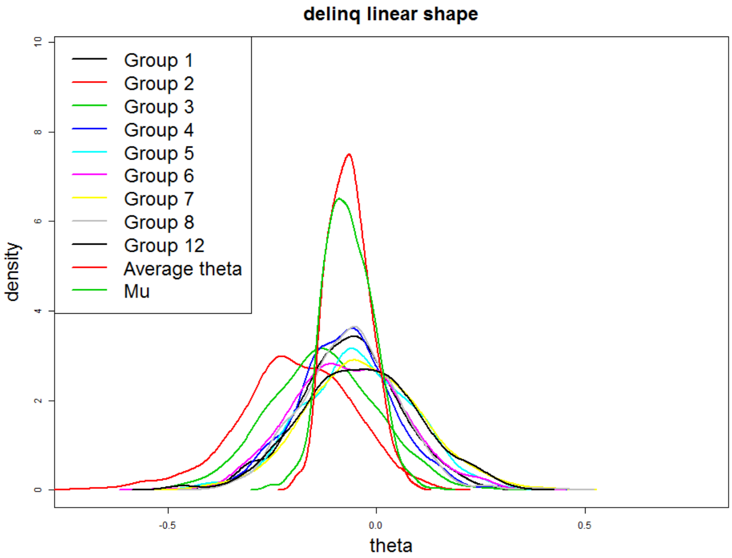
Density plot for same sex



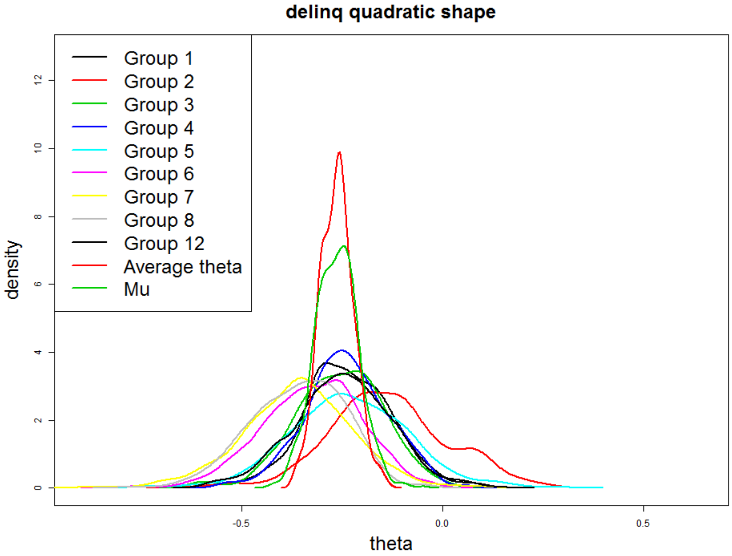
Density plot for log(n) ego



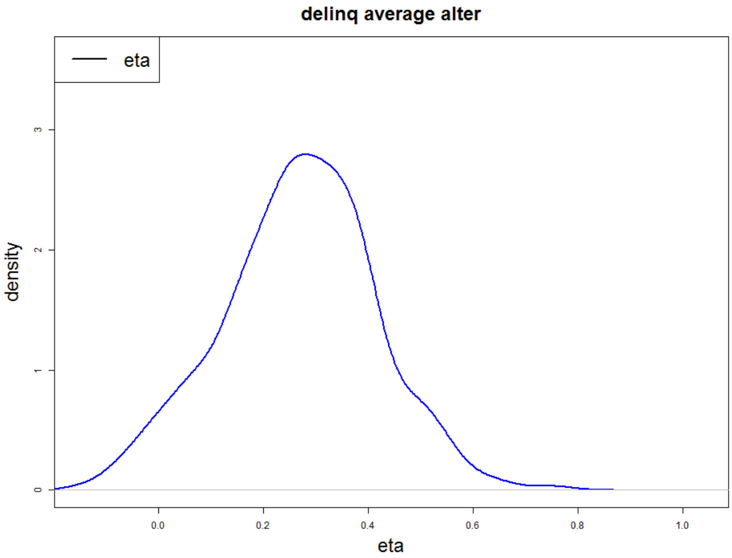
Density plot for linear shape



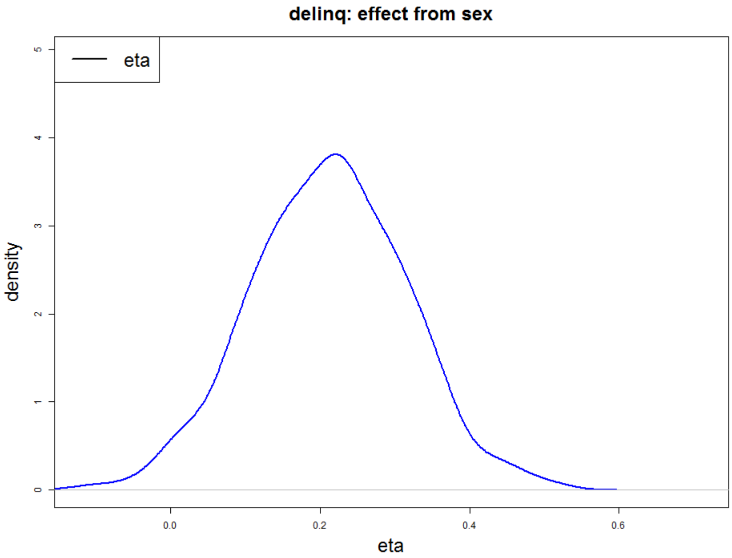
Density plot for quadratic shape



Density plot for average alter



Density plot for effect sex on delinquency



Conclusion

The method seems to work well.

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

Conclusion

The method seems to work well.

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

The posterior uncertainty about parameters is much larger when they are assumed to vary between groups.

Thus, prior assumptions do matter...

Making inference about a population of networks
is associated with much larger uncertainty
than making inference about a single network.

Conclusion

The method seems to work well.

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

The posterior uncertainty about parameters is much larger when they are assumed to vary between groups.

Thus, prior assumptions do matter...

Making inference about a population of networks
is associated with much larger uncertainty
than making inference about a single network.

Especially if there are not so many groups.