Multilevel analysis of network dynamics

Tom A.B. Snijders

University of Groningen
University of Oxford

September, 2019
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 can be achieved, 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 gave an overview of empirical work until then involving multiple networks.
Sample from Population of Networks

Suppose we have a sample indexed by $j = 1, \ldots, N$ from a population of networks on disjoint node sets, 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 on disjoint node sets, 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 \( \theta_1, \ldots, \theta_j, \ldots, \theta_N \).

Each disjoint network is called a group.
We assume we have network panel data for each group, and wish to analyze these by fitting Stochastic Actor-Oriented Models (‘SAOMs’, RSiena).
Several approaches are possible for combining such data:
Several approaches are possible for combining such data:

1. Multi-group analysis:
   assume all parameters are identical.
Several approaches are possible for combining such data:

1. Multi-group analysis:
   assume all parameters are identical.

2. Integrated hierarchical approach:
   Assumption: population of networks, normal distribution.
Several approaches are possible for combining such data:

1. Multi-group analysis: assume all parameters are identical.
2. Integrated hierarchical approach: Assumption: population of networks, normal distribution.
Several approaches are possible for combining such data:

1. Multi-group analysis:
   assume all parameters are identical.

2. Integrated hierarchical approach:
   Assumption: population of networks, normal distribution.

3. Meta analysis:
   Assumption: population of networks, no distributional assumptions:
   \textit{two-stage meta analysis}.

4. Meta analysis: no population assumption:
   \textit{Fisher combination of independent tests}.
Such data sets have multilevel structure; all caveats and considerations from usual multilevel analysis apply!

Multilevel designs, generalization, and the risks of having few groups

For some basic intuitive understanding of the consequences of multilevel designs, consider the simplest case: estimation of the population mean $\mu$ of some variable.

Left one-sided testing.

We have $N = 5$ randomly drawn groups each of size $n = 20$. Data range from $-2.5$ to $+3.5$, as in the histogram (next page). Is there evidence that the population mean is positive?

Test $H_0 : \mu = 0$ against $H_1 : \mu > 0$. 
mean = 0.36, s.d. = 1.18, $t_{99}=3.40$, $p = 0.0003$

*linear model (ols): yes*
multilevel analysis...

Frequency histogram of five groups

multilevel $t_4 = 1.49, p = 0.12$

multilevel model: no

but by group $p = 0.11, 0.84, 0.034, 0.021, 0.00002$
multilevel analysis...

multilevel \( t_4 = 1.49, p = 0.12 \)

*multilevel model: no*

but by group \( p = 0.11, 0.84, 0.034, 0.021, 0.00002 \)
An even more extreme example:

\[
\text{ols } t_{99} = 7.9, \ p \approx 0; \text{ multilevel } t_{4} = 1.50, \ p = 0.12
\]

by group \( p = 0.64, 0.002, 0.000012, < 0.0001, < 0.0001 \)
The conclusion for both data sets is not only, that the multilevel model tells us $\mu$ is not significantly positive.

There is also significant heterogeneity between the groups: $\mu$ is not zero in all groups (this can be concluded already from an ols analysis).

And one of the groups has such a small $p$, that the significance survives multiple testing with Bonferroni.

*If we wish to generalize to the population of all groups, there is no evidence of a positive mean; but there is evidence that at least one of these groups has a positive mean, and that the mean of these five groups is positive.*
After this sensitization, on with the four approaches:

1. Multi-group analysis: assume all parameters are identical.
2. Integrated hierarchical approach: Assumption: population of networks, normal distribution.

We treat how this can be handled with SAOMs, using the package RSiena.
Multi-group analysis

Multi-group analysis (sienaGroupCreate, siena07) assumes all parameters are the same. The groups are treated just like subsequent waves, strung in a sequence, where (of course) transitions between groups are not analyzed.

Example for 3 groups, with 3 / 2 / 3 waves; the analyzed 2+1+2 periods are indicated as bold lines.

Groups can have different sizes, must have same variables and a common model specification.
The multi-group analysis assumes all parameters are identical.
The multi-group analysis assumes all parameters are identical.

*possible but risky*

Gives a global impression, risk of underestimated standard-errors.

Try to find relatively homogeneous sets of groups for doing this.
The multi-group analysis assumes all parameters are identical.

*possible but risky*

Gives a global impression, risk of underestimated standard-errors.

Try to find relatively homogeneous sets of groups for doing this.

`sienaTimeTest()` can be used to test homogeneity assumption.
Random effects: hierarchical multilevel structure

1. On the tie level there is a dynamic process governed by the SAOM.

2. On the network level there is a Stochastic Actor-Oriented Model (SAOM) with parameter vector \((\theta^{(1)}_j, \eta)^t\) for group \(j\). Here \(\theta^{(1)}_j \sim \mathcal{N}(\mu, \Sigma)\).

3. On the global level there is a population of networks with either a multivariate normal distribution \(\mathcal{N}(\mu, \Sigma)\) for the randomly varying parameters \(\theta^{(1)}_j\), and a common parameter \(\eta\) for the rest; or without distribution assumption.
Hierarchical multilevel analysis

Assumption: population of networks;

$\theta_j \sim$ multivariate normal distribution, perhaps conditionally on network-level covariates.

1. mixed effects:
some parameters varying, others constant across groups: 

   restricted integrated hierarchical approach.

2. random effects:
all parameters varying across groups: 

   integrated hierarchical approach

but it is unusual to specify all parameters as randomly varying.

RSiena: multilevel option rjagsBayes(); (still experimental)
Integrated procedure:
Estimate the distribution of $\theta_j$ and consider the ‘posterior’ distribution of $\theta_j$ given the data.
**Integrated procedure:**
Estimate the distribution of $\theta_j$ and consider the ‘posterior’ distribution of $\theta_j$ 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 rather small networks.*

The method implemented in *sienaBayes* is in the paradigm of *Bayesian* as opposed to *frequentist* statistics. Some explanation follows later; frequentist methods would also be possible.
Hierarchical model: restricted variation

Terminology here:

*fixed parameters* have the same value across groups;
*random parameters* have different values across groups, with a priori a normal distribution.

From regular hierarchical linear models (*HLMs*) we know that ‘random slopes’ require a lot from the data.

In practice, HLMs have only a few random slopes.
For SienaBayes, however, a somewhat large number of random parameters seems not to be a problem.

But note the important distinction:

For fixed parameters ($\eta$), there is much more information than for the means and variances of random parameters ($\mu, \Sigma$).

Therefore, the posterior distribution will be more concentrated (smaller variance) for the $\eta$ parameters than for $\mu$. 
The Bayesian MCMC procedure produces, when/if there is convergence (i.e., hopefully, after a burn-in period), a sample from the posterior distribution of all the parameters, both the $\theta_j$ referring to the individual sampled networks, and $\eta$, $\mu$ and $\Sigma$ referring to the population of networks.

The inference is based on the sampled posterior distributions. Two kinds of plot are important:

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.
Prior distributions

For a smallish number of groups, the prior is consequential. It is given to SienaBayes as priorMu, priorSigma, priorDf.

For rate parameters, SienaBayes uses a data-dependent prior.

For evaluation function do not use the defaults (mean=0, var.=1) but use your prior experience; in particular, prior between-group variances can usually be much smaller than 1, because the Siena effects are calibrated to be usually less than 1 in absolute value.

For parameters that are the focus of research questions, use prior mean = 0.
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, 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 is not really known.

2. Other group-level variables may also be relevant.

However, the number of level-2 variables should not be too large! The same considerations apply as for the number of covariates, given sample size, in linear regression models.
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 $\Rightarrow$ dyadic.

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

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; $\log(n)$;
for delinquency $V$:
$V(\text{ego})$, $V(\text{alter})$; $V^2(\text{ego})$; $V(\text{ego}) \times V(\text{alter})$;
$V(\text{group mean ego})$; $V(\text{group mean ego}) \times V(\text{alter})$;

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 $V(\text{group mean ego})$. 
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.

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 $\mu$: outdegree $-2$, same sex $0.4$, others $0$; other prior variances for $\mu$: $0.01$; prior covariances: all $0$; prior variances for $\eta$: $2$, but for group-level variables $1$; prior $\kappa$: $0.01$.

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 micro-steps varies between 100–600 depending on distance between observed networks;

2. ⇒ 500 iterations sampling $\theta_j^{(1)}, \eta, \mu, \Sigma$ for warmup

⇒ 1000 iterations sampling $\theta_j^{(1)}, \mu, \Sigma$

and 3000 for sampling $\eta$ 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 ($\mu$) of structural network effects

- outdegree (density)
- reciprocity
- transitive triplets
- indegree-popularity
- outdegree-activity
- rec.degree(1/1)-activity
Trace plots for fixed parameters ($\eta$) of structural network effects
Trace plots for means ($\mu$) of covariate network effects

- **delinquency ego**
- **same sex**
Trace plots for fixed parameters ($\eta$) of covariate network effects
Trace plots for means ($\mu$) of behavior effects

- delinq linear shape
- delinq quadratic shape
Trace plots for fixed parameters ($\eta$) 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).
### Friendship dynamics

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

Post. means, standard dev.s, and 95% credibility intervals for $\mu, \eta$; varying between classrooms: + = "yes", – = "no".
<table>
<thead>
<tr>
<th>Effect</th>
<th>Posterior mean (s.d.)</th>
<th>Interval from to</th>
<th>Varying</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delinquency dynamics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear shape</td>
<td>−0.068 (0.059)</td>
<td>−0.178 to 0.047</td>
<td>+</td>
</tr>
<tr>
<td>Quadratic shape</td>
<td>−0.264 (0.054)</td>
<td>−0.368 to −0.155</td>
<td>+</td>
</tr>
<tr>
<td>Average alter</td>
<td>0.268 (0.147)</td>
<td>−0.030 to 0.546</td>
<td>−</td>
</tr>
<tr>
<td>Effect from sex</td>
<td>0.212 (0.104)</td>
<td>0.007 to 0.421</td>
<td>−</td>
</tr>
</tbody>
</table>

Posterior means, standard deviations, and 95% credibility intervals for $\mu, \eta$; 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 $\mu$ and that of average $\theta_j$;

the average $\theta_j$ has a bearing on this sample only; for $\mu$, 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

outdegree (density)

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6
- Group 7
- Group 8
- Group 12
- Average theta
- Mu

Density

Theta

© Tom A.B. Snijders
Multilevel Networks
Density plot for network rate

constant friends rate (period 1)

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6
- Group 7
- Group 8
- Group 12
- Average theta
- Mu

theta

density
Density plot for reciprocity

The graph shows the distribution of reciprocity for different groups, with each group represented by a different line color. The x-axis represents the value of theta, and the y-axis represents the density of reciprocity. The graph also includes a line for the average theta and another for Mu.
Density plot for transitive triplets
Density plot for transitive reciprocated triplets

transitive recipr. triplets

eta

density

eta

© Tom A.B. Snijders
Multilevel Networks 46 / 84
Density plot for indegree-popularity

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6
- Group 7
- Group 8
- Group 12
- Average theta
- Mu
Density plot for outdegree-activity
Density plot for reciprocal degree-activity

rec.degree^{(1/1)} - activity

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6
- Group 7
- Group 8
- Group 12
- Average theta
- Mu
Density plot for old ties
Density plot for delinquency alter

![Density plot for delinquency alter](image)

- Integrated hierarchical random effects models
- Andrea Knecht’s data
- Density plot for delinquency alter
- Delinquency alter
- Density
- Eta
- Eta
- Multilevel Networks
Density plot for delinquency ego
Density plot for delinquency squared ego

delinq squared ego

-1.5 -1.0 -0.5 0.0 0.5 1.0 1.5
-1.5 -1.0 -0.5 0.0 0.5 1.0
0 0.5 1.0 1.5
-0.15 -0.10 -0.05 0.00 0.05 0.10 0.15

© Tom A.B. Snijders
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

- Group 1
- Group 2
- Group 3
- Group 4
- Group 5
- Group 6
- Group 7
- Group 8
- Group 12
- Average theta
- Mu
Density plot for quadratic shape

![Graph showing density plots for different groups and group averages.](image-url)
Density plot for average alter

density

eta

delinquency average alter

© Tom A.B. Snijders  Multilevel Networks
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.
Meta-Analysis ~ Random Effects Model

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

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

\[
H_0^{\text{total}} : \quad \text{all } \theta_{kj} = 0;
\]
\[
H_0^{\text{mean}} : \quad E\{ \theta_{kj} \} = 0;
\]
\[
H_0^{\text{spread}} : \quad \text{var}\{ \theta_{kj} \} = 0.
\]
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 $\mu$ 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 \( \mu \) 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 \( \sigma^2 \),
and \( \text{var}(E_j) = \text{s.e.}_j^2 \) (‘\( V \)-known’).

Meta-Analysis ~ Random Effects Model (contd.)

This has been applied in quite many studies. Examples:

Lubbers (2003):
homophily in 57 classrooms with 1466 students
(also with random coefficient $p^*$ approach);

Baerveldt, van Duijn, Vermeij, van Hemert (2004):
ethnic homophily in 20 schools, 1317 students;

friendship & obesity, 17 classrooms with 617 students;

Mercken, Snijders, Steglich, & de Vries (2009):
fr. & smoking, 7704 adolescents, 70 schools, 6 countries;
also other studies by Liesbeth Mercken.
There are a variety of R packages for meta-analysis. For example: metafor, mvmeta.

RSiena also includes the function siena08 for meta-analysis. This carries out meta-analysis according to S & B (2003), but with an iterated weighted least squares (‘WLS’) method. metafor calls this WLS analysis of the random effects model. Viechtbauer (2005): REML is somewhat better than WLS.

Fixed effects: conclusion only about the average of these \( N \) groups (usually weighted inversely by variances);

random effects: generalization to population of groups.

Therefore: the advice is to use metafor, which also has a lot more options than siena08.

siena08 can be used for what is not provided by metafor: overall test of \( H_0^{\text{total}} \); Fisher combination (see below).
There are a variety of R packages for meta-analysis. For example: metafor, mvmeta.

**RSiena** also includes the function siena08 for meta-analysis. This carries out meta-analysis according to S & B (2003), but with an iterated weighted least squares (‘WLS’) method.

Therefore: the advice is to use metafor, which also has a lot more options than siena08.

siena08 can be used for what is not provided by metafor: overall test of $H_0^{\text{total}}$; Fisher combination (see below).

Some terminology:

The method carried out by siena08 is, in the language of metafor, the \textit{WLS analysis} of the \textit{random effects} model.

\textit{Fixed effects}: conclusion only about the average of these $N$ groups (usually weighted inversely by variances);

\textit{random effects}: generalization to population of groups.

Viechtbauer (2005): REML is somewhat better than WLS.
Example: Chris Baerveldt’s data

Houtzager and Baerveldt (1999): *Dutch Social Behavior study*.  
19 classrooms in secondary schools, 
evolution of friendship network and delinquent behavior.  
Also used in Snijders & Baerveldt (*J. Math. Soc.*, 2003) and  
3 classrooms had difficulties for convergence and were dropped; in 2 further classrooms,  
one parameter had to be fixed to 0 for stable convergence.  
Number of pupils per group varied 33–91.  
Average degrees per wave 1.0–3.2 (low!)  
Jaccards 0.24–0.38.
The following pages show a table that was produced as a hybrid from metafor and siena08 results.

It has results for the global parameters, $\mu$ as well as $\sigma$.

Note: it is important for a meta-analysis to present results for mean tendency as well as between-group variability!!!

(Many researchers unfortunately present only the former.)

All results in the table are from metafor, except $p_T$, the $p$-value for $H_0^{\text{total}}$ from siena08.
### Random effects: two-stage procedure

Example: Chris Baerveldt’s data

<table>
<thead>
<tr>
<th>Friendship dynamics</th>
<th>N</th>
<th>$\hat{\mu}$</th>
<th>se($\hat{\mu}$)</th>
<th>$p_\mu$</th>
<th>$\hat{\sigma}$</th>
<th>$p_\sigma$</th>
<th>$p_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>outdegree (density)</td>
<td>16</td>
<td>-2.955</td>
<td>0.149</td>
<td>0.00</td>
<td>0.00</td>
<td>0.65</td>
<td>0.00</td>
</tr>
<tr>
<td>reciprocity</td>
<td>16</td>
<td>3.588</td>
<td>0.253</td>
<td>0.00</td>
<td>0.44</td>
<td>0.52</td>
<td>0.00</td>
</tr>
<tr>
<td>GWESP</td>
<td>16</td>
<td>2.377</td>
<td>0.120</td>
<td>0.00</td>
<td>0.18</td>
<td>0.27</td>
<td>0.00</td>
</tr>
<tr>
<td>indegree - popularity</td>
<td>16</td>
<td>-0.118</td>
<td>0.024</td>
<td>0.00</td>
<td>0.02</td>
<td>0.42</td>
<td>0.00</td>
</tr>
<tr>
<td>outdegree - activity</td>
<td>16</td>
<td>0.121</td>
<td>0.024</td>
<td>0.00</td>
<td>0.00</td>
<td>0.91</td>
<td>0.00</td>
</tr>
<tr>
<td>rec.degree - activity</td>
<td>16</td>
<td>-0.204</td>
<td>0.067</td>
<td>0.00</td>
<td>0.08</td>
<td>0.63</td>
<td>0.10</td>
</tr>
<tr>
<td>indegree - activity</td>
<td>16</td>
<td>-0.204</td>
<td>0.072</td>
<td>0.00</td>
<td>0.05</td>
<td>0.70</td>
<td>0.17</td>
</tr>
<tr>
<td>same ethnicity</td>
<td>16</td>
<td>-0.005</td>
<td>0.089</td>
<td>0.96</td>
<td>0.19</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>sex (M) alter</td>
<td>16</td>
<td>-0.329</td>
<td>0.086</td>
<td>0.00</td>
<td>0.00</td>
<td>0.64</td>
<td>0.03</td>
</tr>
<tr>
<td>sex (M) ego</td>
<td>16</td>
<td>-0.169</td>
<td>0.173</td>
<td>0.33</td>
<td>0.42</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>same sex</td>
<td>16</td>
<td>0.554</td>
<td>0.080</td>
<td>0.00</td>
<td>0.14</td>
<td>0.39</td>
<td>0.00</td>
</tr>
<tr>
<td>crimi alter</td>
<td>16</td>
<td>-0.053</td>
<td>0.046</td>
<td>0.24</td>
<td>0.00</td>
<td>0.83</td>
<td>0.74</td>
</tr>
<tr>
<td>crimi ego</td>
<td>16</td>
<td>0.074</td>
<td>0.080</td>
<td>0.35</td>
<td>0.16</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>crimi ego × alter</td>
<td>16</td>
<td>0.006</td>
<td>0.057</td>
<td>0.92</td>
<td>0.11</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>reciprocity × GWESP</td>
<td>14</td>
<td>-0.568</td>
<td>0.272</td>
<td>0.04</td>
<td>0.00</td>
<td>0.96</td>
<td>0.68</td>
</tr>
</tbody>
</table>

© Tom A.B. Snijders
Multilevel Networks
72 / 84
Random effects: two-stage procedure
Example: Chris Baerveldt’s data

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>( \hat{\mu} )</th>
<th>se(( \hat{\mu} ))</th>
<th>( p_\mu )</th>
<th>( \hat{\sigma} )</th>
<th>( p_\sigma )</th>
<th>( p_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delinquency dynamics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linear shape</td>
<td>16</td>
<td>0.015</td>
<td>0.108</td>
<td>0.89</td>
<td>0.00</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>quadratic shape</td>
<td>16</td>
<td>–0.295</td>
<td>0.048</td>
<td>0.00</td>
<td>0.05</td>
<td>0.70</td>
<td>0.00</td>
</tr>
<tr>
<td>outdegree</td>
<td>16</td>
<td>0.042</td>
<td>0.047</td>
<td>0.37</td>
<td>0.00</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>total alter</td>
<td>16</td>
<td>0.076</td>
<td>0.064</td>
<td>0.23</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>sex (M)</td>
<td>16</td>
<td>0.610</td>
<td>0.139</td>
<td>0.00</td>
<td>0.00</td>
<td>0.99</td>
<td>0.09</td>
</tr>
</tbody>
</table>

- \( N \): number of groups combined;
- \( \hat{\mu} \), se(\( \hat{\mu} \)), \( p_\mu \): inference for \( \mu \);
- \( \hat{\sigma} \), \( p_\sigma \): inference for \( \sigma \);
- \( p_T \): \( p \)-value for \( H_0^{\text{total}} \).

It may be noted that \( \sigma^2 \) never is significantly positive.

However, sienaTimeTest, which uses a more powerful test (multivariate, and using also information used for estimation) concludes that the parameters are significantly different.
Some example conclusions:

- reciprocity: strongly significant, but also quite variable across groups;
- reciprocal degree-activity: clearly significant, variability across groups might exist; if so, no or hardly any groups with positive parameter;
- same sex: strongly significant, variable across groups, will be positive in all groups;
- total alter: not significant; if non-zero, then more likely to be positive; note that $p_\sigma = p_T = 1$ but $p_\mu = 0.23$, which can be explained by tests not producing certainty, and the test for $\mu$ being more focused.

For ‘positive in all groups’, think of $\hat{\mu} \pm 2\hat{\sigma}$. 
The funnel plot produced by `metafor` shows groupwise estimates and standard errors. Important to get a feeling for what’s going on!
Note that for the reciprocated degree-activity effect, none of the individual estimates is significant, but jointly the effect is strongly significant.
Random effects: two-stage procedure

Example: Chris Baerveldt’s data
Meta-Analysis without a population assumption

\[ \theta_1, \ldots, \theta_j, \ldots, \theta_N \] are arbitrary values, no assumption about a population is made.
Meta-Analysis without a population assumption

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

two-stage procedure:
estimate each \( \theta_j \) separately, combine the results by Fisher’s procedure for combining independent tests:

‘is there any evidence for a hypothesized effect?’

Without a population assumption

Meta-Analysis \sim without population assumption (contd.):

For coordinate $k$ of the parameter, test null hypothesis

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

against alternative hypothesis

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

RSiena: first estimate for all groups, then use siena08(). Useful when there are few groups.

Mercken, Snijders, Steglich, & de Vries (2009) applied this in a study of smoking initiation: 7704 adolescents in 70 schools in 6 countries.
This principle here is applied in a double test for $\theta_{kj}$ (for each given effect $k$)

1. for detecting if any groups $j$ have $\theta_{kj} > 0$, the null hypothesis tested is
   $H_0^r$: For all groups $j$ we have $\theta_{kj} \leq 0$
   with the alternative right-sided hypothesis;
   $H_1^r$: For at least one group $j$, $\theta_{kj} > 0$;

2. for detecting if any groups $j$ have $\theta_{kj} < 0$, the null hypothesis tested is
   $H_0^\ell$: For all groups $j$ we have $\theta_{kj} \geq 0$
   with the alternative left-sided hypothesis;
   $H_1^\ell$: For at least one group $j$, $\theta_{kj} < 0$;

For each effect, siena08 gives the right-sided $p$-value $p_r$, and the left-sided $p$-value $p_\ell$. 
To account for the double testing, the significance level $\alpha$ should be divided by 2; corresponding to doubling the $p$-values.

It is very well possible that $p_r$ as well as $p_\ell$ are significant (e.g., for a combined $\alpha = 0.05$, that $p_r$ as well as $p_\ell$ are $< 0$).

This means that some groups have positive, others negative parameter values; a perfectly possible conclusion.
To account for the double testing, the significance level $\alpha$ should be divided by 2; corresponding to doubling the $p$-values.

It is very well possible that $p_r$ as well as $p_\ell$ are significant (e.g., for a combined $\alpha = 0.05$, that $p_r$ as well as $p_\ell$ are $< 0$).

This means that some groups have positive, others negative parameter values; a perfectly possible conclusion.

The following pages present the two one-sided $p$-values for the Baerveldt data for comparison combined with the earlier results for $\mu$. 
<table>
<thead>
<tr>
<th>Friendship dynamics</th>
<th>$N$</th>
<th>$\hat{\mu}$</th>
<th>se($\hat{\mu}$)</th>
<th>$p_\mu$</th>
<th>$p_l$</th>
<th>$p_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>outdegree (density)</td>
<td>16</td>
<td>-2.955</td>
<td>0.149</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>reciprocity</td>
<td>16</td>
<td>3.588</td>
<td>0.253</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>GWESP</td>
<td>16</td>
<td>2.377</td>
<td>0.120</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>indegree - popularity</td>
<td>16</td>
<td>-0.118</td>
<td>0.024</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>outdegree - activity</td>
<td>16</td>
<td>0.121</td>
<td>0.024</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>rec. degree - activity</td>
<td>16</td>
<td>-0.204</td>
<td>0.067</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>indegree - activity</td>
<td>16</td>
<td>-0.204</td>
<td>0.072</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>same ethnicity</td>
<td>16</td>
<td>-0.005</td>
<td>0.089</td>
<td>0.96</td>
<td>0.36</td>
<td>0.42</td>
</tr>
<tr>
<td>sex (M) alter</td>
<td>16</td>
<td>-0.329</td>
<td>0.086</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>sex (M) ego</td>
<td>16</td>
<td>-0.169</td>
<td>0.173</td>
<td>0.33</td>
<td>0.66</td>
<td>0.04</td>
</tr>
<tr>
<td>same sex</td>
<td>16</td>
<td>0.554</td>
<td>0.080</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>crimi alter</td>
<td>16</td>
<td>-0.053</td>
<td>0.046</td>
<td>0.24</td>
<td>0.88</td>
<td>0.31</td>
</tr>
<tr>
<td>crimi ego</td>
<td>16</td>
<td>0.074</td>
<td>0.080</td>
<td>0.35</td>
<td>0.09</td>
<td>0.73</td>
</tr>
<tr>
<td>crimi ego $\times$ alter</td>
<td>16</td>
<td>0.006</td>
<td>0.057</td>
<td>0.92</td>
<td>0.25</td>
<td>0.52</td>
</tr>
<tr>
<td>reciprocity $\times$ GWESP</td>
<td>14</td>
<td>-0.568</td>
<td>0.272</td>
<td>0.04</td>
<td>0.98</td>
<td>0.13</td>
</tr>
</tbody>
</table>
In this case, the results of Fisher’s combination of tests are in line with the random-effects meta-analysis; there is a difference in ‘significance’ only for reciprocity × GWESP.

This is related to the small variability between the groups, and the fact that no parameters have ambiguous signs.

In other cases, it may well be different! Statistical reasoning gives no certainty...

See the manual for further explanation.
Literature

The example using Chris Baerveldt’s data is in the script
`RscriptMultipleGroups_meta.R` on the Siena website.

References

RSiena manual: Chapter 11.


Wolfgang Viechtbauer (2010). ‘Conducting meta-analyses in R with the metafor package’, *Journal of Statistical Software*. 