



university of
groningen

behavioural and
social sciences

sociology

Statistical Analysis of Social Networks

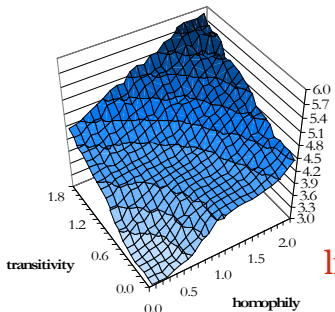
Stochastic Actor-based Modelling of Network Dynamics

Christian Steglich

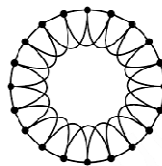
c.e.g.steglich@rug.nl



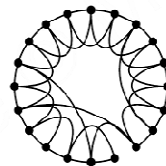
median geodesic distance between groups



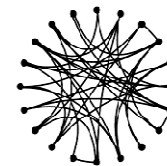
Regular



Small-world



Random



$$\ln\left(\frac{\Pr(x^c \rightarrow_i x^b)}{\Pr(x^c \rightarrow_i x^a)}\right) = \sum_{k=1}^K \beta_k (s_{ik}(x^b) - s_{ik}(x^a))$$





Outline: Methodological concerns & strategy, ...

The issues

- Why model network dynamics?
- And how? What must we pay attention to?

The stochastic actor-based framework

- What is the modelling strategy?
- What data sets can be analysed?
- Which model assumptions are made?



... statistical aspects & applications (w/ software).

Fitting models to data

- How are models estimated?
- How are hypotheses tested?

Examples

- Advice seeking among MBA students (Lomi, Torlò et al.)
- Sex segregation in secondary schools (Steglich & Knecht)
- Hierarchies in friendship networks
- Preferential trade agreements (Manger, Pickup & Snijders)

Notes on interpretation, goodness of fit, ...



Why model network dynamics?

- › Research often starts as a question about associations between network features and individual features...
 - *Do popular students smoke?*
 - *Are nations with high income levels more central in preferential trade agreement networks?*
- › ... or between network features and dyadic features.
 - *Do students ethnically segregate in school?*
 - *Do non-democratic countries trade with each other?*
- › Any such association begs the question of “Why?”...



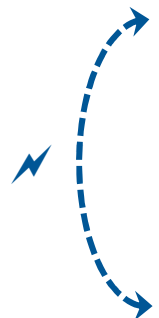
Explanatory mechanisms typically are dynamic

- › Competing explanations are the rule:
 - *Do students ethnically segregate...*
 - ... because they prefer their own ethnic group?*
 - ... or because they prefer ties to students living close by (& there is residential segregation)?*
 - ... or because inter-ethnic ties break up more quickly?*
- › These explanations invoke dynamic mechanisms:
 - ... same ethnicity precedes tie formation;*
 - ... geographic proximity precedes tie formation;*
 - ... same ethnicity precludes tie dissolution.*



How to model network dynamics?

› Ultimate criterion: *“Such that you can tell which of the competing mechanisms is consistent with the data.”*

- 
- **Statistical approach** needed to control different effects for each other
 - **Longitudinal approach** needed to link antecedents to consequences
 - **Complete network approach** needed because conceptually, selection can only be studied when the complete pool of candidates is known

⚡ **Statistics & non-independent data: specialised models!**



Actor-based models for network evolution

- › Modelling strategy
- › Data format considerations
- › Model assumptions
- › The network evolution algorithm
- › Model specification, selection of effects



Modelling strategy

Conceptually: Actor-driven model.

Actors are the locus of modelling, change is due to individual decisions. [Assumption: Luce's (1959) choice axioms.]

- actors control “their” network ties;
- two submodels:
 - When can actor i make a decision? (rate function)
 - Which decision does actor i make? (objective function)

Technically: Continuous time Markov process.

Assumption: *conditional independence of the future from the past, given the present network.*



Data requirements

Required are repeated measures of the same network:

- same group of actors
(some composition change is allowed)
- same relational variable. ****states, not events!****

Subsequent observations are assumed to be related through an unobserved continuous process of change.

Fully observed continuous-time data can in principle be analysed with standard statistical software – but this requires a lot of data organisation.



Example data: (Andrea Knecht, 2003/04)

Networks among first grade pupils at Dutch secondary schools (“bridge class”).

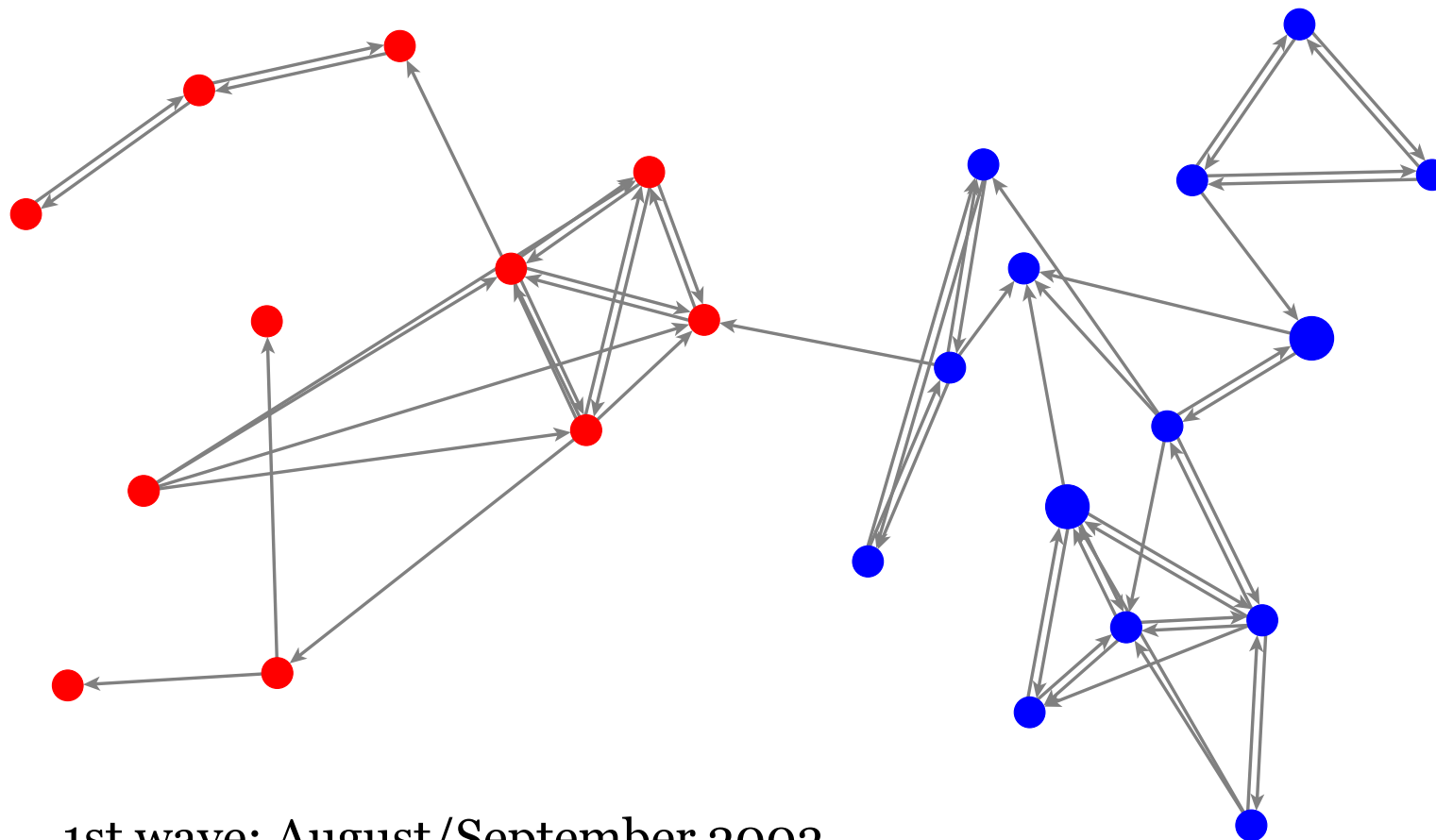
125 school classes

4 measurement points,

various network & individual measures.

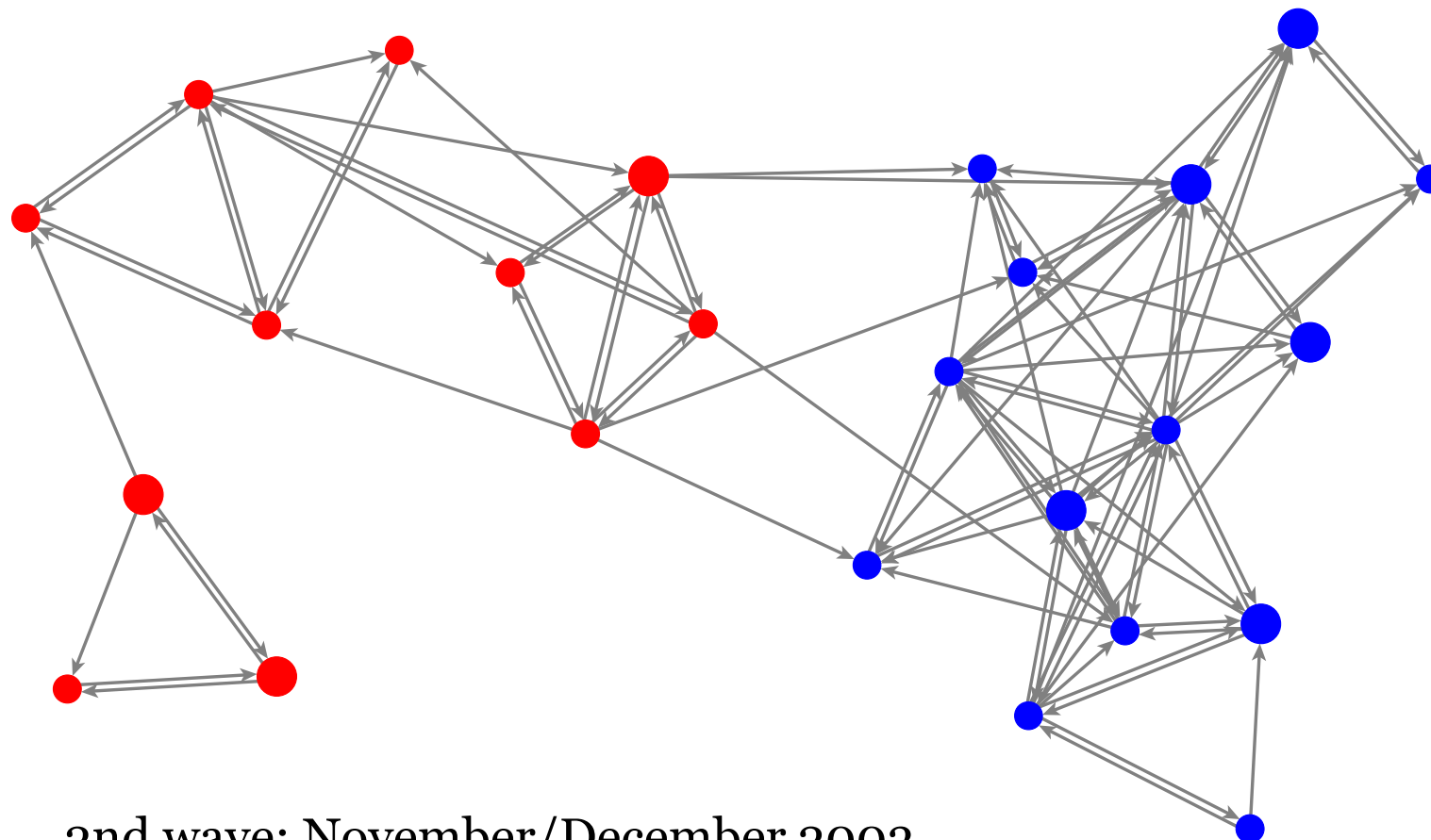
The following slides show the evolution of the friendship network in one classroom.

The graph layout is a bit messy for each observation alone, but optimal over time according to a stress minimisation algorithm.

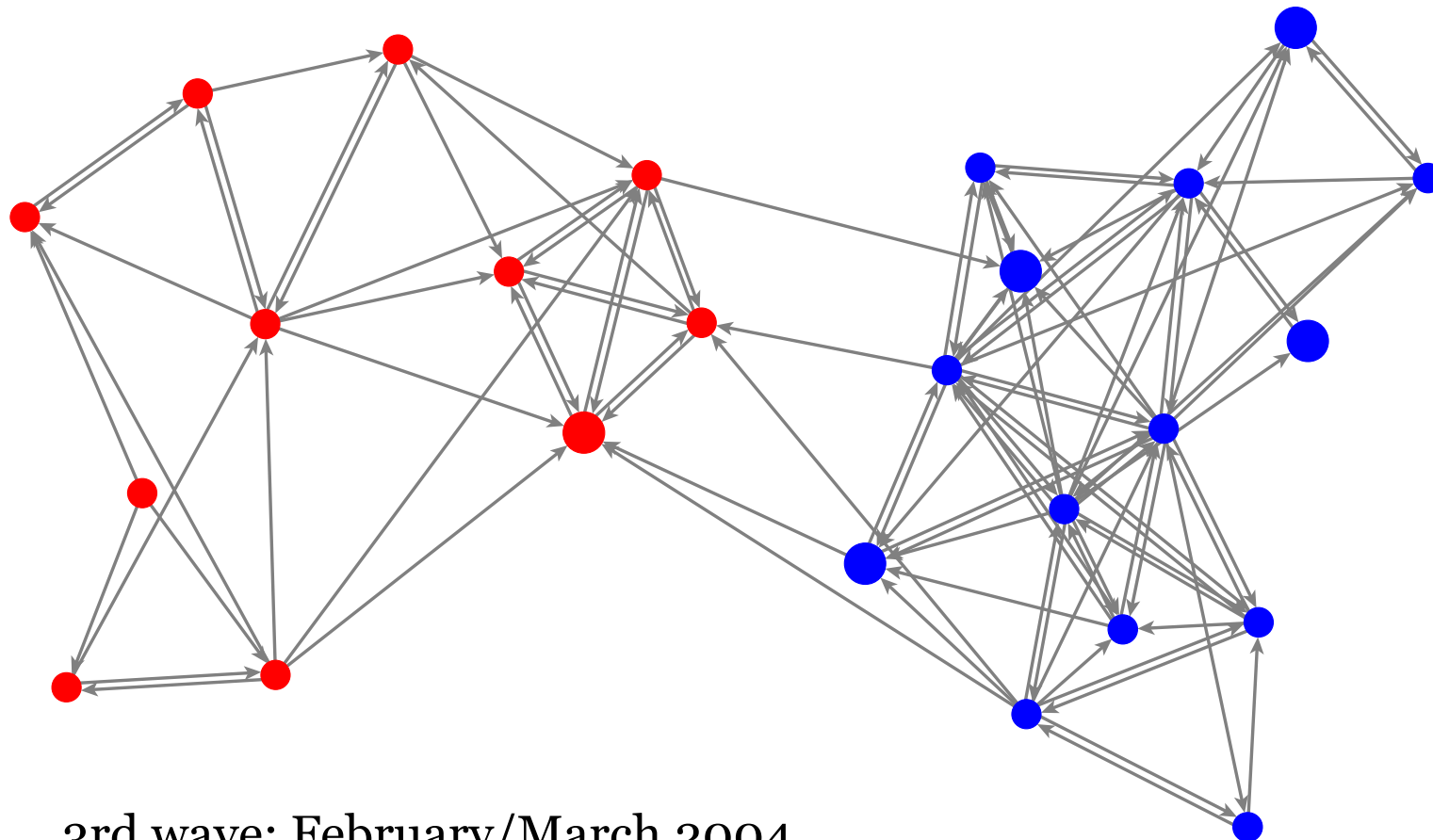


1st wave: August/September 2003

Node size indicates strength of delinquency...

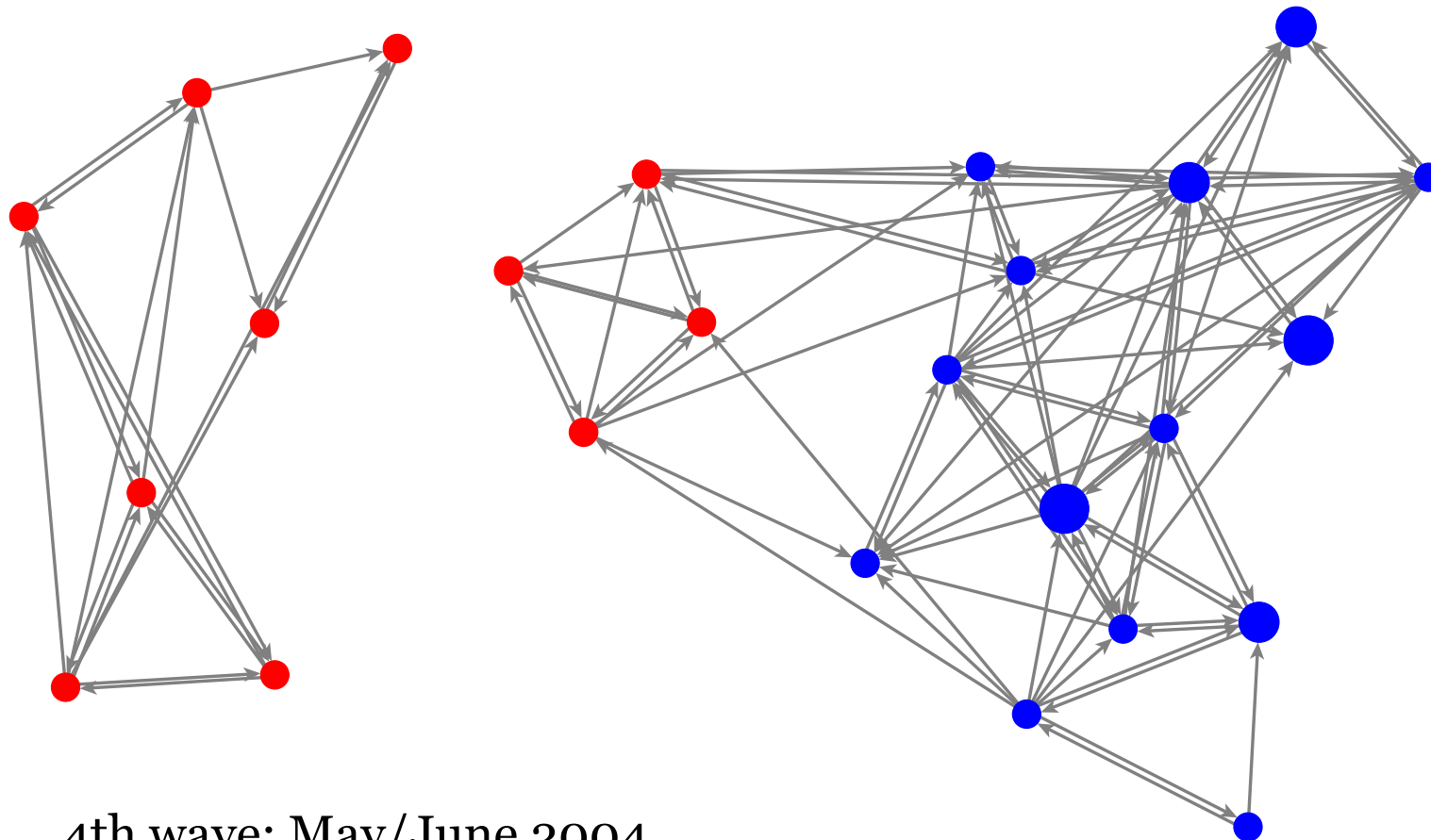


2nd wave: November/December 2003
... and node colour indicates sex.



3rd wave: February/March 2004

So-called *anchored layout* can be used to...



4th wave: May/June 2004

... animate the data in a movie (also with visone).



Points to consider before trying actor-based modelling

states ↔ **events**

NOT snapshots of e-mail traffic, BUT reliable measures of a social relation.

Event networks could be aggregated over time to obtain state networks!

change ↔ **stability**

The networks should change ‘slowly’, contain a stable part.

Rules for structural change typically are about individual ties changing in response to surrounding ties (which remain stable, for that moment).



Data format issues to consider

binary	\leftrightarrow signed \leftrightarrow valued
directed	\leftrightarrow undirected
tie loss possible	\leftrightarrow growth only networks
1-mode	\leftrightarrow bipartite
single dependent	\leftrightarrow multiplex

- › The standard model is developed for a single dependent, binary, directed, 1-mode network that can both grow and shrink over time.

Everything else is a non-standard model extension, and not necessarily supported by the software implementation.



Modelling principles for such data sets

Random walk: Network evolution proceeds as a stochastic process on the space of all possible networks;

No contamination by the past: The first observation is not modelled but conditioned upon as the process' starting value.

Continuous-time model: Change is modelled as occurring in continuous time.

Micro steps: Big change from one observation to the next is assumed to accrue from a sequence of smallest possible changes.

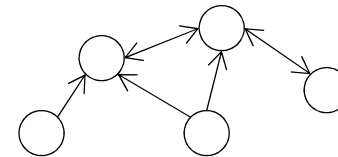
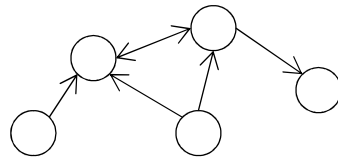
This assumption of *temporal decomposability / separability* is quite a strong one, but crucial for statistical power!



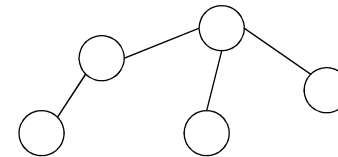
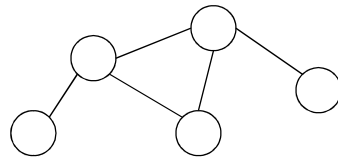
What are smallest possible changes?

- › Changes between two networks that differ by just one tie variable, while all others are identical.

- Example directed network:



- Example undirected network:



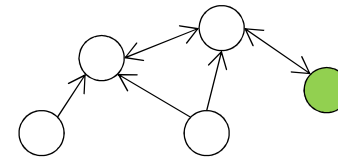
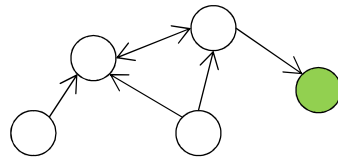
Terminology: these networks *differ by a micro step*.



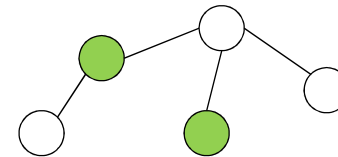
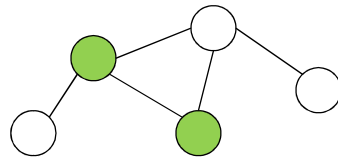
Micro steps and locus of control

- › A micro step involves uniquely identified actors – these are assumed to control & decide about the tie variable:

- Directed network: ONE actor

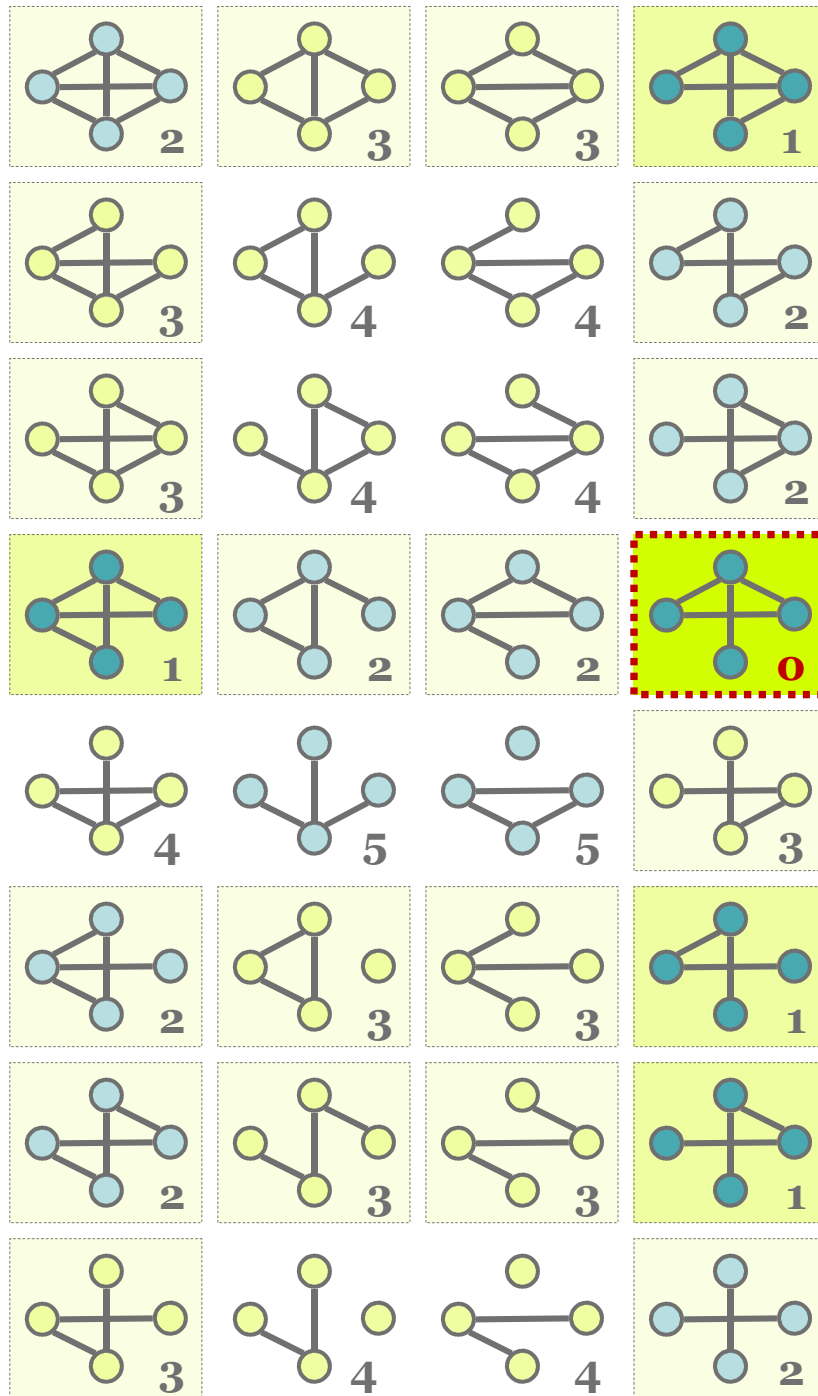


- Undirected network: TWO actors



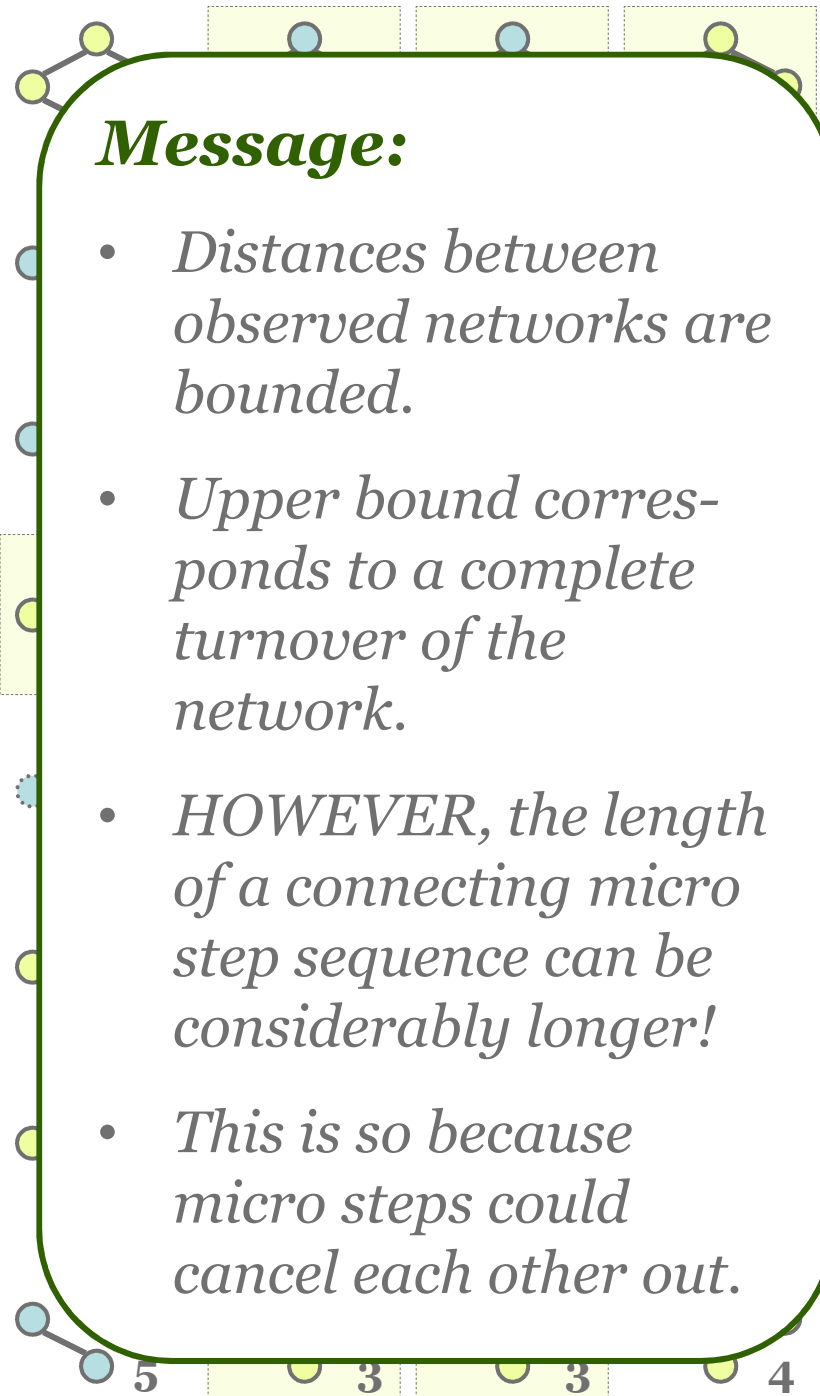
The directed case is therefore simpler to model, in an actor-based way.

Example: Distances from **o**-network in 'micro steps'



Message:

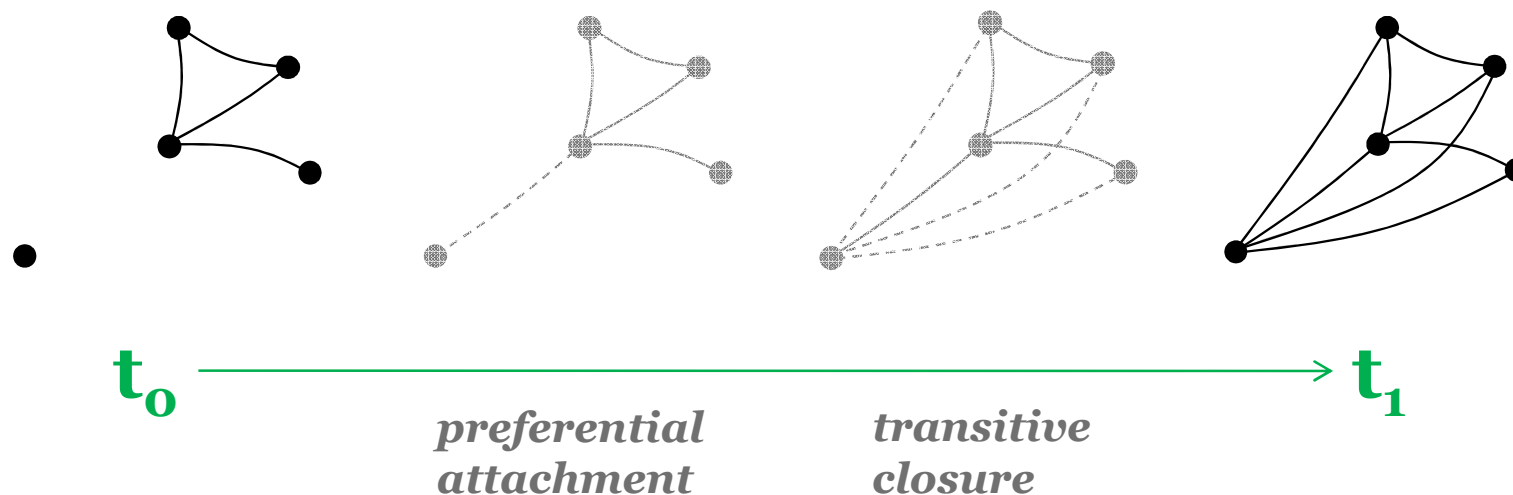
- Distances between observed networks are bounded.
- Upper bound corresponds to a complete turnover of the network.
- *HOWEVER*, the length of a connecting micro step sequence can be considerably longer!
- This is so because micro steps could cancel each other out.





An advantage of continuous-time modelling

Complex patterns emerging from simple(r) mechanisms



Some new ties may be realisation-contingent on other new ties.
Discrete time models cannot easily model their compound emergence.



The network evolution algorithm

Network evolution in observation period $t_0 \rightarrow t_1$ takes place as in this *straight simulation* algorithm:

1. Model time is set to $t = t_0$, and simulation starts out at the network observed at this time point.
2. For all actors, a *waiting time* is sampled according to the *rate function*.
3. The actor with the shortest waiting time τ is identified.
4. If $t + \tau > t_1$, the simulation terminates.
5. Otherwise, the identified actor gets the opportunity to set a micro step. This is determined by his *objective function*.
6. Model time is updated and simulation proceeds at step 2.

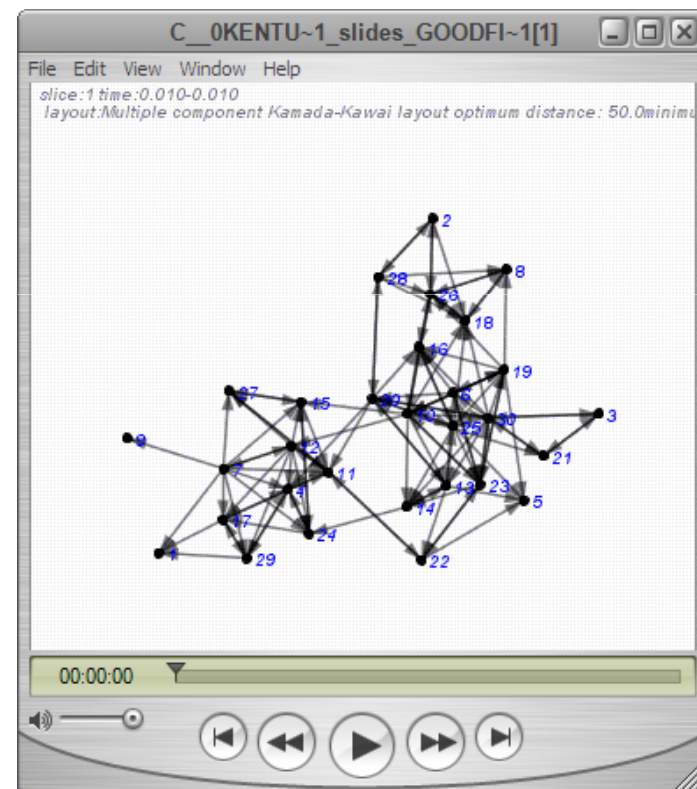


Visualisation with SONIA

SIENA-based imputation of the unobserved trajectory of changes between two consecutive observations.

The movie shows but *one* instantiation of the model.

*Classroom friendship data,
Andrea Knecht, 2003/04.*





The rate function

$$\lambda_i(\mathbf{x}) = \sum_k \rho_k r_{ik}(\mathbf{x})$$

- › Models *speed* differences between actors **i**.
- › Statistics **r_{ik}** of **i**'s neighbourhood in **x** are weighted by model parameters **ρ_k** .
- › These weights express whether the feature expressed in the statistic is related to more frequent (**$\rho_k > 0$**) or less frequent (**$\rho_k < 0$**) network changes by the actors.
- › They are estimated from the data.

Technically, λ_i is parameter of an exponential distribution of waiting times – as in Poisson regression.

Typically, it is good to start an analysis under the assumption of a periodwise constant rate function.



The objective function $f_i(\mathbf{x}) = \sum_k \beta_k s_{ik}(\mathbf{x})$

- › Models attractiveness of network states \mathbf{x} to actor \mathbf{i} .
- › Statistics s_{ik} of \mathbf{i} 's neighbourhood in \mathbf{x} are weighted by model parameters β_k .
- › These weights express whether the feature expressed in the statistic is desired ($\beta_k > 0$) or averted ($\beta_k < 0$).
- › Also they are estimated from the data.

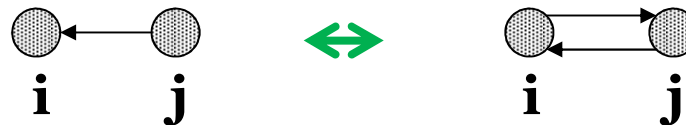
Technically, $f_i(\mathbf{x})$ is parameter of a multinomial logit model for discrete, probabilistic choice.

The objective function is the main part of modelling. Here, hypotheses typically are operationalised.



Some effect statistics

reciprocity effect $S_{i \text{ recip.}} = \sum_j X_{ij} X_{ji}$



transitivity effect $S_{i \text{ tr.trip.}} = \sum_{jk} X_{ij} X_{jk} X_{ik}$

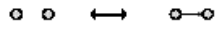
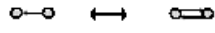

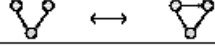
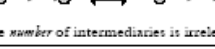
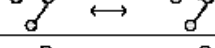
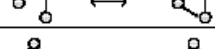
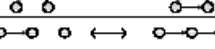
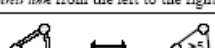
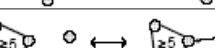
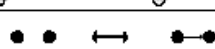
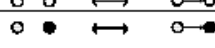
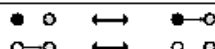
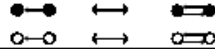
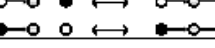
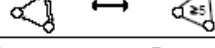
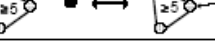
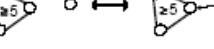
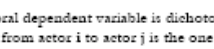





*Very often, effect
statistics are motif
(subgraph) counts.*

Effects measure attractiveness difference between right and left configuration, for the focal actor **i**.

Many other effects are possible to include in the objective function...

TABLE 2
SELECTION OF POSSIBLE EFFECTS FOR MODELING NETWORK EVOLUTION

effect	network statistic	effective transitions in network*	verbal description
1. outdegree	x_{ij}		preference for ties to arbitrary others
2. reciprocity	$x_{ij}x_{ji}$		preference for reciprocated ties
3. transitive triplets	$x_{ij} \sum_h x_{ih} x_{hj}$		preference for being friend of the friends' friends
4. balance	$x_{ij} \text{strsim}_{ij}$		preference for ties to structurally similar others
5. actors at distance two	$\begin{cases} 1 & \text{if between}(h;ij) = 1 \text{ for some } h \\ 0 & \text{else} \end{cases}$	 (the number of intermediaries is irrelevant)	preference for keeping others at social distance two
6. popularity alter	$x_{ij} \sum_h x_{hi}$		preference for attaching to popular others, i.e., others who are often named as friend ('preferential attachment')
7. activity alter	$x_{ij} \sum_h x_{jh}$		preference for attaching to active others, i.e., others who name many friends
8. 3-cycles	$x_{ij} \sum_h x_{jh} x_{hi}$		preference for forming relationship cycles (negative indicator for hierarchical relations)
9. betweenness	$\sum_h \text{between}(i;jh)$	 (no direct link from the left to the right actor)	preference for being in an intermediary position between unrelated others
10. dense triads	$\sum_h \text{group}(ijh)$		preference for being part of cohesive subgroups
11. peripheral	$\sum_{hk} \text{peripheral}(i;jhk)$		preference for unilaterally attaching to cohesive subgroups
12. similarity	$x_{ij} \text{sim}_{ij}$		preference for ties to similar others (selection)
13. behavior alter	$x_{ij} z_j$		main effect of alter's behavior on tie preference
14. behavior ego	$x_{ij} z_i$		main effect of ego's behavior on tie preference
15. similarity × reciprocity	$x_{ij} x_{ji} \text{sim}_{ij}$		preference for reciprocated ties to similar others
16. between dissimilar alters	$\sum_h (1 - \text{sim}_{jh}) \text{between}(i;jh)$		preference for being in an intermediary position between unrelated, dissimilar others (brokerage potential)
17. similarity × dense triads	$\sum_h \text{group}(ijh) (\text{sim}_{ij} + \text{sim}_{ih})$		preference for being part of behaviorally similar cohesive subgroups
18. behavior × peripheral	$z_i \sum_{hk} \text{peripheral}(i;jhk)$		behavior-specific preference for unilaterally attaching to cohesive subgroups
19. similarity × peripheral	$\sum_{hk} (\text{peripheral}(i;jhk) \times (\text{sim}_{ij} + \text{sim}_{ih} + \text{sim}_{jh}))$		preference for unilaterally attaching to behaviorally similar cohesive subgroups

* In the effective transitions illustrations, it is assumed that the behavioral dependent variable is dichotomous and centered at zero; the color coding is  = low score (negative),  = high score (positive),  = arbitrary score. The tie x_{ij} from actor i to actor j is the one that changes in the transition indicated by the double arrow. Illustrations are not exhaustive.



Choice probabilities $\Pr(x \rightarrow_i x') \propto \exp(f_i(x'))$

- › Choice probabilities for micro steps are proportional to the exponential function of the objective function.
- › Valid options are all possible micro steps, plus the option not to change the status quo.
- › This probability distribution can be interpreted as optimisation of a *random utility function*, namely the objective function f_i plus a Gumbel-distributed error term.
- › Note that the probabilities only depend on x' and not on past states, not even x . This can be relaxed (keyword: *endowment function*).



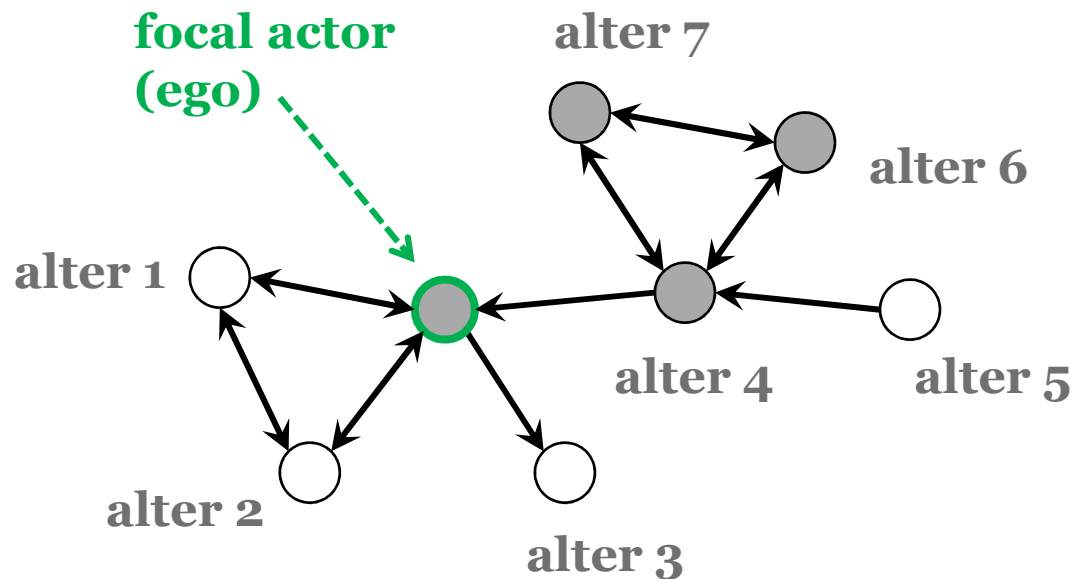
Illustration of “how a micro step works”

Assume that a model specification with the following objective function parameters was estimated on a classroom friendship network:

- outdegree $\beta_{\text{outdg.}} = -2.6$ *friendship is rare*
- reciprocity $\beta_{\text{recip.}} = 1.8$ *friendship is reciprocal*
- transitivity $\beta_{\text{tr.trip.}} = 0.4$ *friendship shows clustering*
- three-cycles $\beta_{\text{3-cycl.}} = -0.7$ *friendship shows hierarchy*
- same gender $\beta_{\text{same}} = 0.8$ *friendship is sex segregated*



Example of an actor's decision

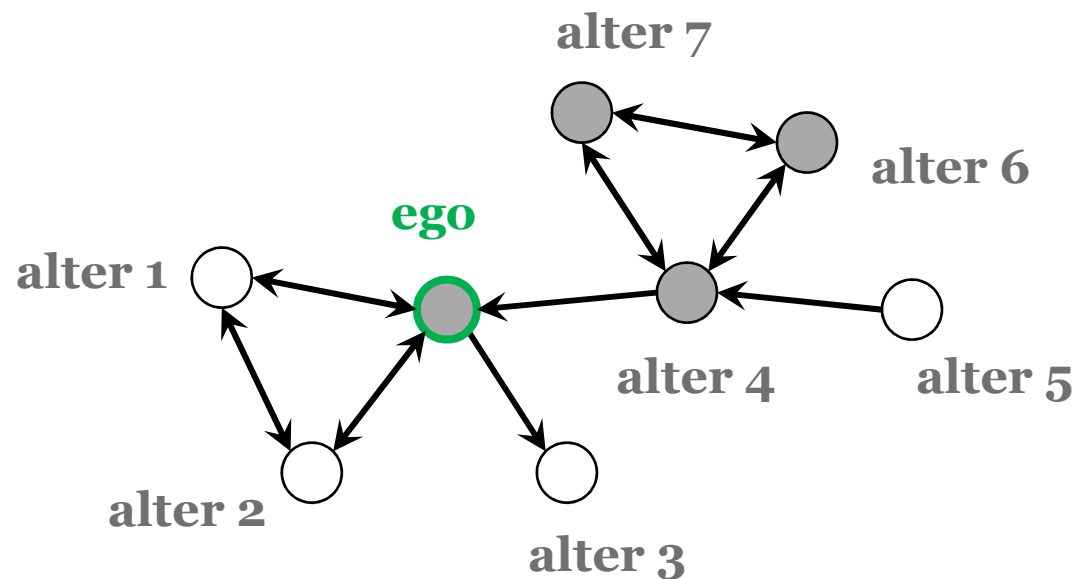


Options:

- drop tie to alter 1
- drop tie to alter 2
- drop tie to alter 3
- create tie to alter 4
- create tie to alter 5
- create tie to alter 6
- create tie to alter 7
- keep status quo



Count model-relevant motifs for all options

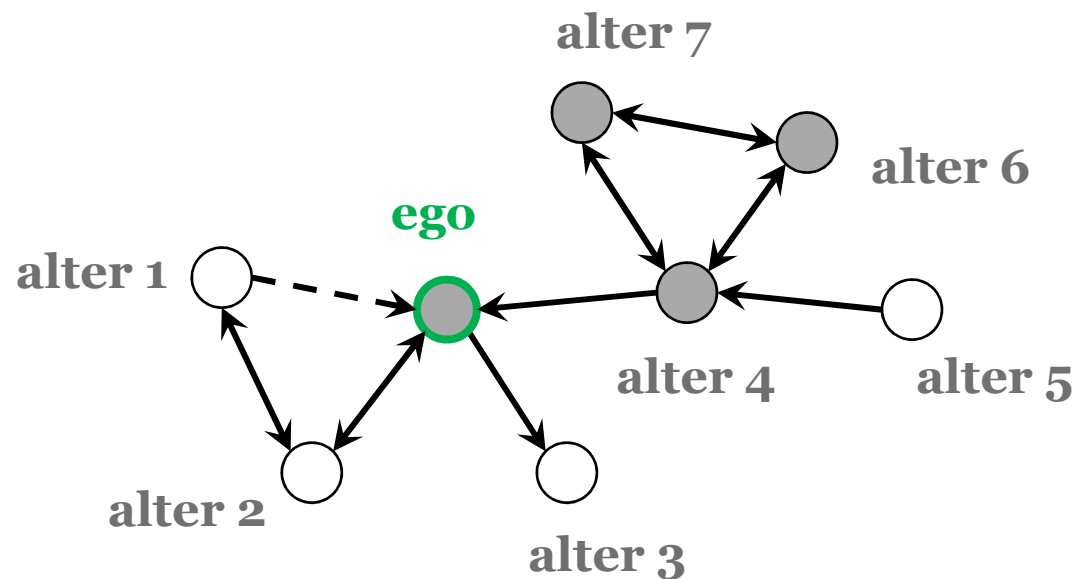


Status quo (ego):

- 3 outgoing ties
- 2 reciprocated ties
- 2 transitive triplets
- 2 three-cycles
- 0 same colour ties



Count model-relevant motifs for all options

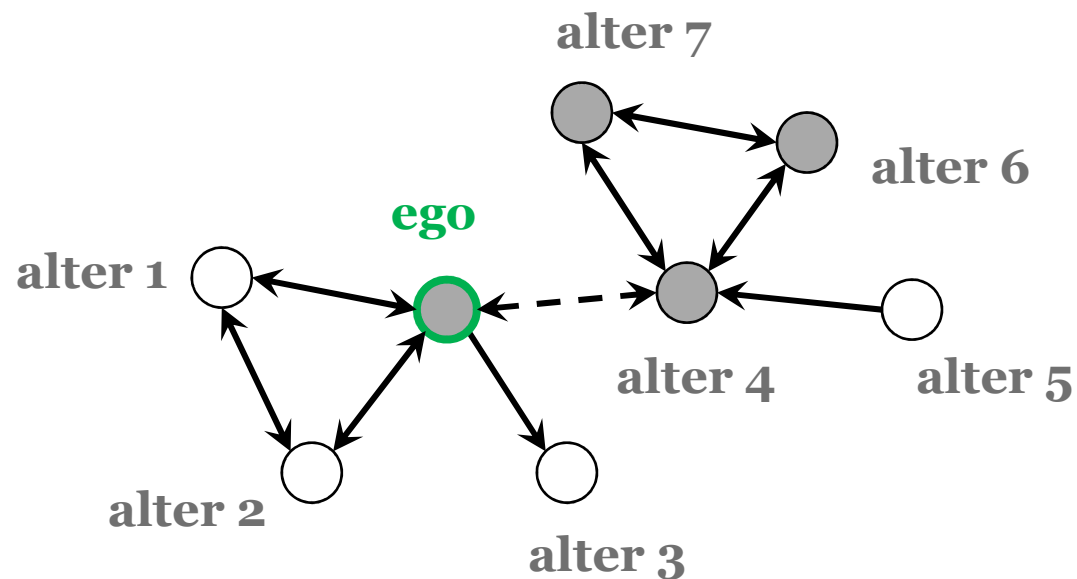


Drop tie to alter 1:

- 2 outgoing ties
- 1 reciprocated tie
- 0 transitive triplets
- 1 three-cycles
- 0 same colour ties



Count model-relevant motifs for all options



Create tie to alter 4:

- 4 outgoing ties
- 3 reciprocated ties
- 2 transitive triplets
- 2 three-cycles
- 1 same colour tie

*...these calculations
are done for all the
eligible options.*



Option	# out-ties	# recip. ties	# tr.triplets	# 3-cycles	# same col.
drop tie to alter 1	2	1	0	1	0
drop tie to alter 2	2	1	0	1	0
drop tie to alter 3	2	2	2	2	0
create tie to alter 4	4	3	2	2	1
create tie to alter 5	4	2	2	3	0
create tie to alter 6	4	2	2	3	1
create tie to alter 7	4	2	2	3	1
keep status quo	3	2	2	2	0



Calculation of objective function:

$f_{\text{drop-1}}$

$f_{\text{drop-2}}$

$f_{\text{drop-3}}$

$f_{\text{create-4}}$

$f_{\text{create-5}}$

$f_{\text{create-6}}$

$f_{\text{create-7}}$

$f_{\text{stat. quo}}$

Option	# out-ties	# recip. ties	# tr.triplets	# 3-cycles	# same col.
drop tie to alter 1	2	1	0	1	0
drop tie to alter 2	2	1	0	1	0
drop tie to alter 3	2	2	2	2	0
create tie to alter 1	3	2	2	2	1
create tie to alter 2	4	2	2	3	0
create tie to alter 3	4	2	2	3	1
create tie to alter 4	4	2	2	3	1
keep status quo	3	2	2	2	0

Matrix S_{ego}

$\beta_{\text{outdg.}}$

$\beta_{\text{recip.}}$

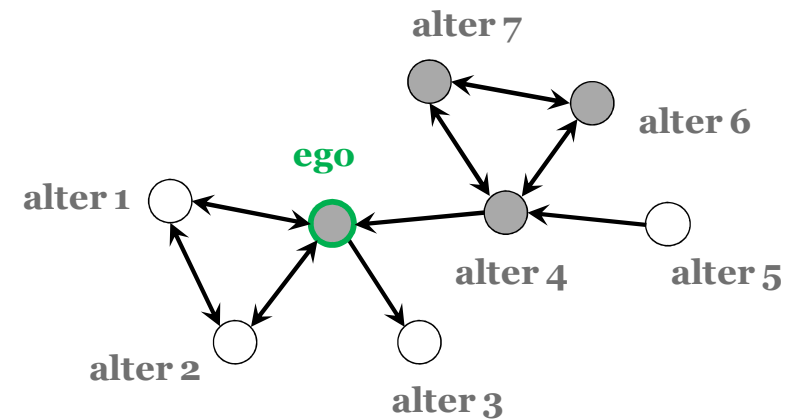
$\beta_{\text{tr.trip.}}$

$\beta_{\text{3-cycles}}$

β_{same}



Option	objective function	exponential transform	probability
drop tie to alter 1	-4.1	0.017	10%
drop tie to alter 2	-4.1	0.017	10%
drop tie to alter 3	-2.2	0.111	68%
create tie to alter 4	-4.8	0.008	5%
create tie to alter 5	-8.1	0.000	0%
create tie to alter 6	-7.3	0.001	0%
create tie to alter 7	-7.3	0.001	0%
keep status quo	-4.8	0.008	5%



***Dropping the tie to alter 3
clearly dominates this
decision situation.***

*Note: SIENA internally centers
many variables – this does not
affect the choice probabilities.*



Homogeneity assumptions

Unless otherwise specified (by including interaction terms with nuancing variables), the model assumes...

Actor homogeneity:

All actors follow the same behavioural rules in their networking activities.

Time homogeneity:

These behavioural rules do not change over time.

This can be problematic whenever actors or time periods are heterogeneous but there are no predictors for differences in the data. ***So: check this in your models!***



Significance testing of parameters

- › The RSiena software estimates parameters β_k and their standard errors $\text{st.err.}(\beta_k)$.
- › By calculating the *t-ratio* of those, parameter significance can be tested:

$$t = \beta_k / \text{st.err.}(\beta_k)$$

- is approximately normally distributed*
- under the assumption (null hypothesis) that *actual network evolution* follows a model in which the parameter is constrained to zero ($\mathbf{H}_0: \beta_k = \mathbf{0}$).

* Thus far, this claim largely rests on extensive simulation studies.



Model specification / effect selection

When investigating social network dynamics, researchers ususally do not come empty-handed but have theories or (at least) hypotheses about the mechanisms that might operate.

- › *These mechanisms [hopefully] can be expressed in terms of SIENA parameters, and the hypotheses can be restated in terms of the corresponding model parameters.*
- › *By estimating the parameters and calculating significance tests for them, the theories / hypotheses can be tested empirically.*

But... how do parameters & hypotheses relate to each other?



Local characterisation of choice probabilities

- › For two networks that could be obtained in competing micro steps from the same network of origin, the ratio of choice probabilities is this (“odds”):

$$\frac{\Pr(x^c \rightarrow_i x^a)}{\Pr(x^c \rightarrow_i x^b)} = \exp \left(\sum_{k=1}^K \beta_k \left(s_{ik}(x^a) - s_{ik}(x^b) \right) \right)$$

compared are two moves (‘micro steps’) made by actor *i* from a network x^c to two “neighbouring networks” x^a and x^b

model
parameters

difference in model statistics of actor *i* between the two compared moves



The main part of the formula in detail:

The sum $\sum_{k=1}^K \beta_k (s_{ik}(x^a) - s_{ik}(x^b))$ determines whether x^a or x^b is more likely to succeed x^c in the network evolution process.

β_k positive: states with higher scores s_{ik} are more likely than states with lower scores;

β_k negative: states with lower scores s_{ik} are more likely than states with higher scores.

This way, parameter values β_k express dynamic tendencies of network evolution: “**actors are moving towards a high [low] score on the corresponding network statistic $s_{.k}$** ”



Example (Torlò, Steglich, Lomi & Snijders, 2007)

- › **75 students** enrolled in an MBA program;
- › **4 network variables:** advice-seeking, communication, friendship, acknowledge-contribution-to-learning;
- › **co-evolving behavioural dimension:** performance in examinations;
- › **several other actor variables:** gender, age, experience, nationality;
- › **3 waves** in yearly intervals.

We focus here on the analysis of the evolution of the *advice network* only.

Which hypotheses are investigated? *[just 3 of them...]*



1. *You seek advice from your friends.*

Mechanism: presence of a friendship tie between two actors increases the likelihood that an advice tie is present between the same actors.

If x_{ij} stands for i seeking advice from j and w_{ij} stands for i naming j as a friend, then the effect

$$s_{i \text{ friend}}(x) = \sum_j x_{ij} w_{ij}$$

operationalises the above mechanism, and the corresponding parameter β_{friend} can be used to test it.



- › The effect statistic $\mathbf{s}_{i \text{ friend}}$ counts the degree to which advice seeking and friendship ‘overlap’.
- › The parameter β_{friend} expresses whether by changing the advice network, such an overlap is sought or avoided, i.e., whether friendship enhances or weakens advice seeking:
 - β_{friend} positive: advice seeking is more likely when it coincides with friendship; ☒
 - β_{friend} negative: advice seeking is less likely when it coincides with friendship. ☐
- › In SIENA, the effect can be included as main effect of a dyadic covariate (friendship) on network evolution.

Hypothesis 1: $\beta_{\text{friend}} > 0$; test the null hypothesis $\beta_{\text{friend}} = 0$.



2. *The lower your performance, the more advice you need [and the more you will seek].*

Mechanism: actors with low performance scores are likely to have more outgoing advice ties than actors with high performance scores.

If z_i stands for performance of actor i , then the effect

$$s_{i \text{ own-performance}}(x) = z_i \sum_j x_{ij}$$

operationalises the above mechanism, and the parameter $\beta_{\text{own-performance}}$ can be used to test it.



- › The effect statistic $S_{i \text{ own-performance}}$ counts the degree to which active advice seeking and performance coincide.
- › The parameter $\beta_{\text{own-performance}}$ expresses whether by changing the advice network, such a coincidence is sought or avoided, i.e., whether own performance enhances or weakens advice seeking:
 - $\beta_{\text{own-performance}}$ positive: high performers seek more advice than low performers; ☐
 - $\beta_{\text{own-performance}}$ negative: high performers seek less advice than low performers. ☒
- › In SIENA, the effect can be included as an ego-effect of an actor variable (performance) on network evolution.

Hypothesis 2: $\beta_{\text{own-p.}} < 0$; test the null hypothesis $\beta_{\text{own-p.}} = 0$.



3. *The higher your performance, the better the advice you can give [and the more you will be asked for advice].*

Mechanism: actors with high performance scores are likely to attract more incoming advice ties than actors with low performance scores.

Let z_j now stand for performance of actor j , then effect

$$s_{i \text{ others-performance}}(x) = \sum_j z_j x_{ij}$$

operationalises the above mechanism, and the parameter $\beta_{\text{others-performance}}$ can be used to test it.



- › The effect statistic $\mathbf{S}_{i \text{ others-performance}}$ counts the degree to which passive advice seeking ('being asked') and performance coincide.
- › The parameter $\beta_{\text{others-performance}}$ expresses whether by changing the advice network, such a coincidence is sought or avoided, i.e., whether others' performance makes them more or less attractive as sources of advice:
 - $\beta_{\text{others-perf.}}$ positive: high performers are more often asked for advice than low p'fs. ☒
 - $\beta_{\text{others-perf.}}$ negative: high performers are less often asked for advice than low p'fs. ☐
- › In SIENA, this is the alter-effect of an actor variable.

Hypothesis 3: $\beta_{\text{oth-p.}} > 0$; test the null hypothesis $\beta_{\text{oth-p.}} = 0$.



Results on these particular hypotheses

Dependent network: Advice	Estimate	St.Error	t-stat	pred.?
1. eval outdegree (density)	-2.6541	(0.0890)	0.0131	
2. eval reciprocity	0.9973	(0.1231)	-0.0307	
3. eval transitive triplets	0.2781	(0.0291)	-0.0246	
4. eval 3-cycles	-0.1199	(0.0534)	-0.0362	
5. eval indegree - popularity	0.0410	(0.0055)	-0.0368	
6. eval friendship	1.0230	(0.0823)	-0.0324	✓
7. eval same background	0.1661	(0.0726)	-0.0311	
8. eval same experience	0.1174	(0.0757)	-0.0704	
9. eval performance alter	0.1035	(0.0272)	-0.0643	✓
10. eval performance ego	-0.0840	(0.0256)	-0.0564	✓
11. eval performance similarity	0.8371	(0.3044)	-0.0358	



More specifically: tests & p-values

Hypothesis 1:

“You seek advice from your friends”

$$t = 1.0230 / 0.0823 = 12.4 ; \quad p < 0.001$$

Hypothesis 2:

“The lower your performance, the more advice you seek”

$$t = -0.0840 / 0.0256 = -3.28 ; \quad p = 0.001$$

Hypothesis 3:

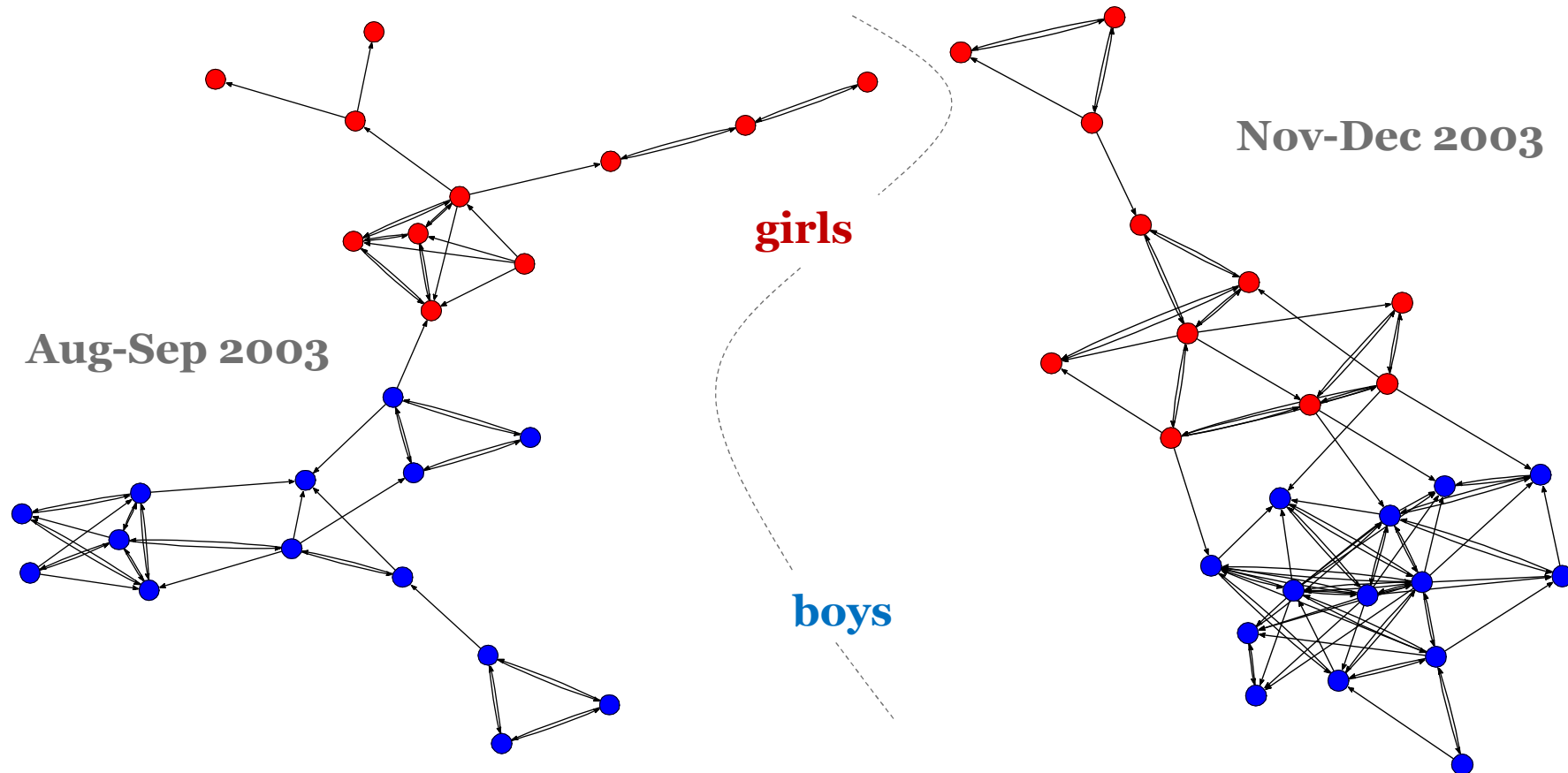
“The higher your performance, the more others ask you for advice”

$$t = 0.1035 / 0.0272 = 3.81 ; \quad p < 0.001$$

All three hypotheses are confirmed!



Example: classroom friendship network





Analyse this network the lab

- › ... making use of the following effects:
 - outdegree (density),
 - reciprocity,
 - transitive triplets,
 - gender effects of sender and receiver,
 - a gender homophily effect.
- › (see exercise on segregation & homophily).



Results of classroom data lab exercise

	Model 1	Model 2	Model 3
<i>Rate function friendship</i>			
Rate of change $t_1 \rightarrow t_2$	7,54 (0,97)	8,81 (1,31)	10,87 (2,63)
Rate of change $t_2 \rightarrow t_3$	2,73 (0,45)	2,92 (0,50)	3,04 (0,52)
Rate of change $t_3 \rightarrow t_4$	3,29 (0,49)	3,56 (0,54)	3,80 (0,65)
<i>Objective function friendship</i>			
Outdegree	-1,92 (0,17) ***	-2,03 (0,16) ***	-2,19 (0,16) ***
Reciprocity	—	1,09 (0,16) ***	0,84 (0,17) ***
Transitive triplets	—	—	0,18 (0,03) ***
primary school friendship	0,54 (0,21) *	0,30 (0,21)	0,40 (0,20) *
Male alter	0,30 (0,18)	0,28 (0,18)	0,05 (0,17)
Male ego	0,11 (0,19)	0,07 (0,19)	-0,17 (0,18)
Same sex	1,70 (0,18) ***	1,39 (0,18) ***	0,93 (0,18) ***

*'significantly
 biased'*



Model estimation by estimating equations

For each model parameter θ

whether part of the rate or of the objective function doesn't matter

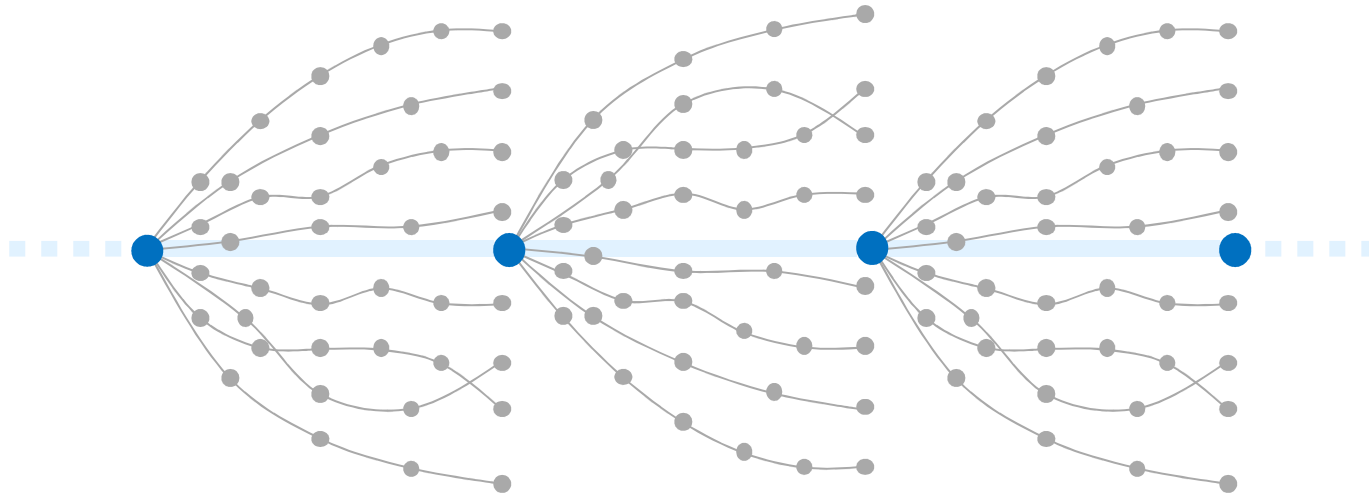
- › a statistic **S**. is identified that can be evaluated on a data set **y** (e.g. the *observed data x* or draws from the probability model **X** of these data, i.e., *simulated by straight simulation*)
- › this statistic defines an *estimating equation* for its parameter:

“Expected value over simulations must equal the observed value.”

The vector θ of parameter estimates is obtained as the joint solution to the corresponding system of equations.



Simulations under Estimating Equations algorithm



- › Observed data are met in expectation over simulations, on a vector of k statistics corresponding to the k model parameters (Snijders, 1996, *Journal of Mathematical Sociology*).
- › Trajectories are sampled by straight simulations.



Estimating statistics used

- › Basic rate parameter period m : $\theta_g = \lambda_m$

Estimating statistic: $\mathbf{S}_g(\mathbf{y}) = \sum_{ij} |y_{ij}(\mathbf{t}_{m+1}) - \mathbf{x}_{ij}(\mathbf{t}_m)|$

- › Objective function parameters: $\theta_g = \beta_h$

Estimating statistic: $\mathbf{S}_g(\mathbf{y}) = \sum_k \sum_i \mathbf{s}_{ih}(\mathbf{y}(\mathbf{t}_{k+1}))$

Estimating equations (for all parameters):

$$\mathbf{E}(\mathbf{S}_g(\mathbf{X})) = \mathbf{S}_g(\mathbf{x})$$



Conditional estimation with estimating equations

- › It can be useful to not estimate the rate parameter (modelling the observed amount of network change):
 - Can improve model convergence
 - Focus often is not on rate of change anyway
- › Then, the straight simulation algorithm needs to be slightly modified:
 - **Stopping rule** is not any more “model time exceeds period end time”...
 - ...but becomes *“Hamming distance in simulations reaches observed Hamming distance”*



Parameter updating based on simulations

- › Parameters are iteratively updated according to the rule

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - a_{k+1} \mathbf{D}_0^{-1} (\mathbf{S}_k^{\text{sim}} - \mathbf{S}^{\text{obs}})$$

where...

- \mathbf{D}_0 is the approximation of the derivative matrix of statistics \mathbf{S} by parameters $\boldsymbol{\theta}$, evaluated at the parameter's starting value $\boldsymbol{\theta}_0$
- \mathbf{a}_k is a sequence of numbers that approach zero at rate k^{-c}
- c is chosen ($0.5 < c < 1$) so as to obtain good convergence properties.

- › The final parameter estimate $\hat{\boldsymbol{\theta}}$ is the tail average $\frac{1}{N} \sum_{r=1}^N \hat{\boldsymbol{\theta}}_{k+r}$



Estimation of covariance structure

- › The approximative covariance matrix of the estimator function, evaluated at the estimate, is given by

$$\text{cov}_{\hat{\theta}}(\tilde{\theta}) = \mathbf{D}_{\hat{\theta}}^{-1} \Sigma_{\hat{\theta}}(\mathbf{S}) \mathbf{D}_{\hat{\theta}}$$

where...

- $\mathbf{D}_{\hat{\theta}}$ again is an approximation of the derivative matrix of statistics \mathbf{S} by parameters θ , now evaluated at the estimate $\hat{\theta}$,
 - $\Sigma_{\hat{\theta}}(\mathbf{S})$ is the matrix of simulated covariance of the vector \mathbf{S} of estimation statistics, also evaluated at the estimate.
- › *Standard errors* of the estimates are calculated as the square roots of the diagonal elements of this matrix.



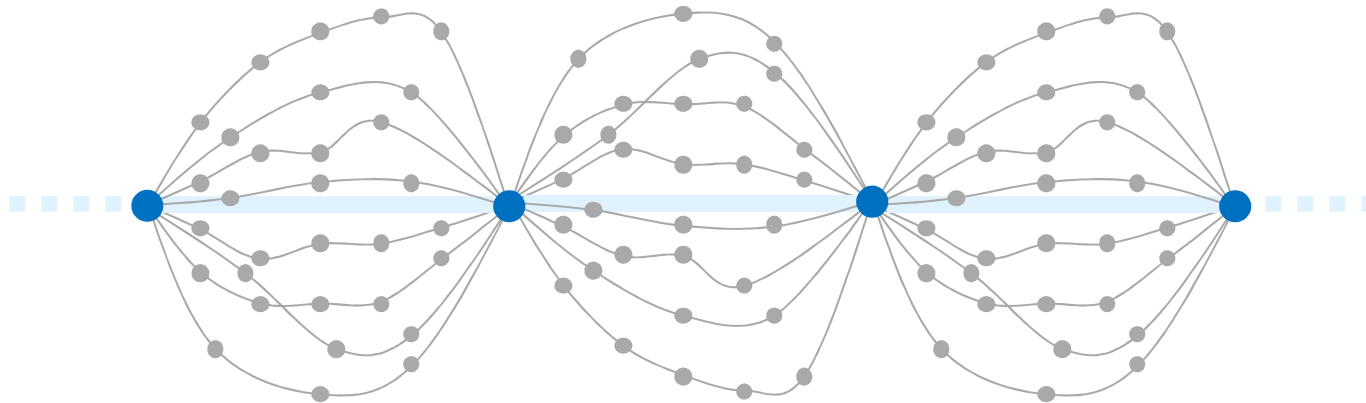
Model estimation by MCMC maximum likelihood

ML estimation requires approximation of the likelihood of the data, therefore...

- › Straight simulation inappropriate (typically doesn't end up in the observed data set)
- › Needed: construction of model-consistent distribution of simulated *network evolution trajectories that connect observed data points*
- › Is achieved by **MCMC techniques** (Snijders, Koskinen & Schweinberger, 2010, *Annals of Applied Statistics*)



Simulations under Likelihood-based algorithms



- › Observed data are met exactly.
- › Connecting trajectories are sampled from the model by MCMC technique (Snijders, Koskinen & Schweinberger, 2010, *Annals of Applied Statistics*).



When to use ML estimation?

Pro:

- › Makes more efficient use of available information
- › Therefore has higher statistical power
- › Relevant for datasets that have low information content already (many missings, little change, small size,...)

Contra:

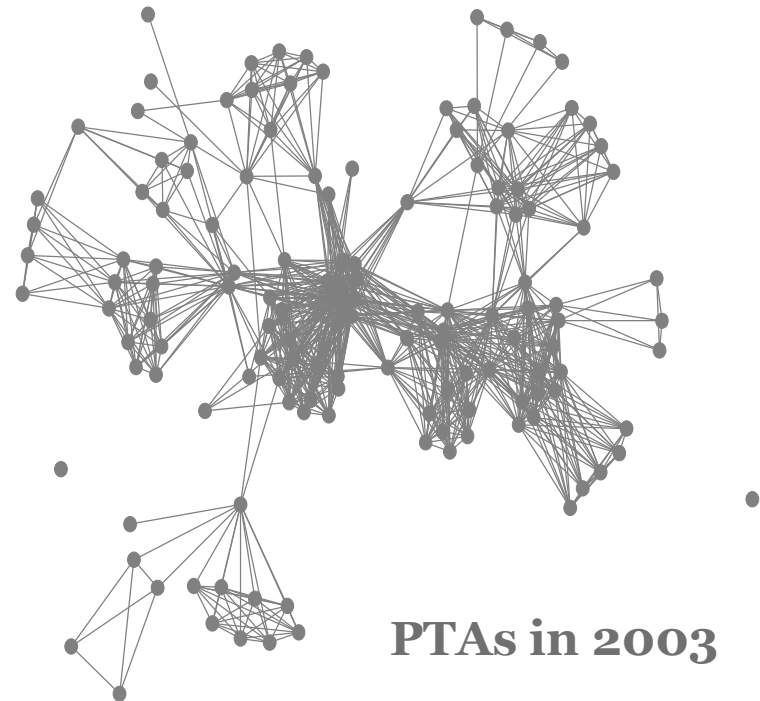
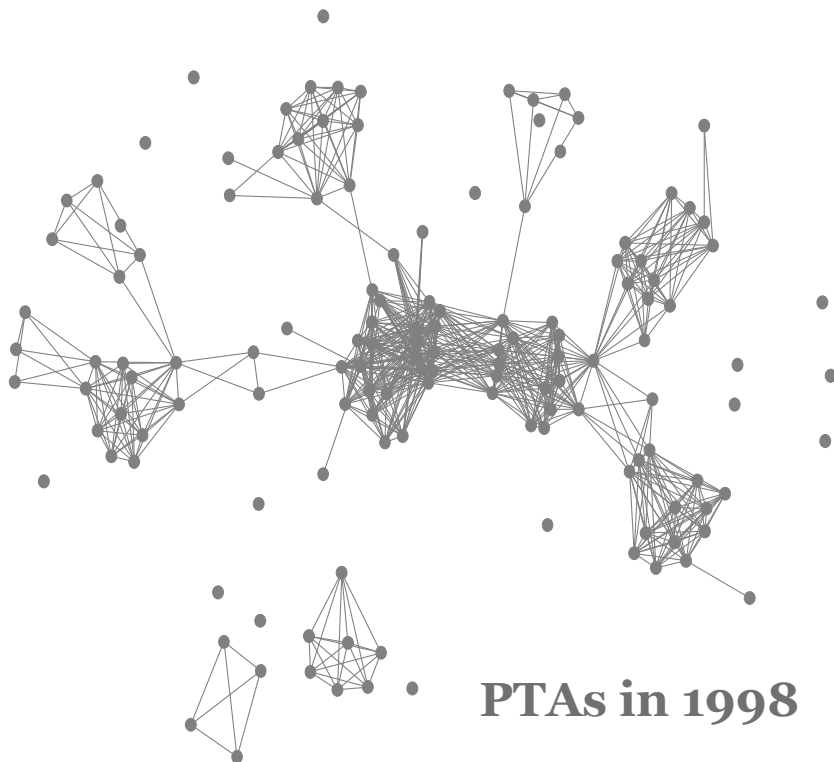
- › Takes considerably more time

Recommendation:

- › Use it only when *estimating equations* gives problems



Example: preferential trade agreements (Manger, Pickup & Snijders)





Analyse the 1998-1999-2000 period in lab

- › Consider hypotheses related to PTA triads...

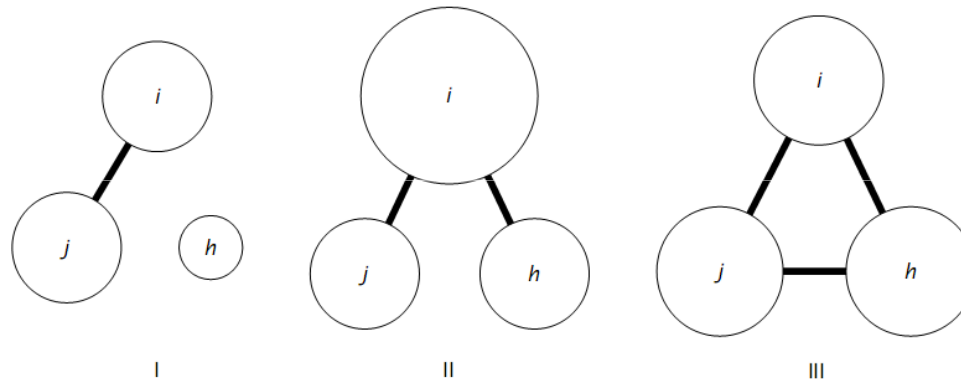


Figure 1: Distribution of Gains from Trade in Triads

- › ... next to actor- or dyad-level predictors
GDP, democratisation, bilateral trade, geogr. distance



Results of the PTA analysis in the lab

<i>Rate function PTA</i>	<i>est.</i>	<i>st.err.</i>	
parameter period 1	0.857	0.110	
parameter period 2	1.276	0.138	
<i>Objective function PTA</i>	<i>est.</i>	<i>st.err.</i>	<i>p-value</i>
degree	-0.968	0.270	<0.001
transitive triads	0.318	0.050	<0.001
Distance	-0.011	0.003	<0.001
Trade	0.070	0.021	0.001
binDemocracy	-0.013	0.124	0.916
same binDemocracy	0.455	0.158	0.004
binDemocracy x same binDemo	0.033	0.124	0.789
GDP_inv	1.161	0.190	<0.001
int. Trade x GDP_inv	0.117	0.045	0.010



More on interpretation of parameter estimates

Several types of interpretation:

1. At face value: parameter values and odds
2. Preference? Constraint? Artifact?
3. In relation to the data
4. As extrapolation into the distant future?

Beware: **model-based inference!**



1. *Parameter values...*

The linear shape of the objective function allows the direct comparison of different predictor variables' effects.

- *Parameters for two effects with same scale (e.g., “same gender” and “same ethnicity”) can be directly compared,*
- *otherwise, scaling needs to be taken into account (e.g., “reciprocity” and “transitive triplets”)*

Note that such comparisons take place on the objective function's scale – NOT on some tangible outcome measure!

[A predicament common to all logistic models.]



... and odds

The local characterisation of the model allows to calculate conditional odds and conditional odds ratios.

- The impact of a unit difference in statistic s_{ik} on the odds of choosing x^a vs. x^b is given by $\exp(\beta_k)$.
- Odds ratios $\exp(\beta_k)/\exp(\beta_m) = \exp(\beta_k - \beta_m)$ allow to compare different effects' sizes.
- From both, binary (or other) comparison probabilities can be calculated.

Note that while such comparisons take place on the probability scale, they refer to rather artificial choice situations!



2. Preference? Constraint? Artifact?

Typically, the parameter estimated for the outdegree statistic $s_{i \text{ outdg.}} = \sum_j x_{ij}$ is quite significantly negative.

Does this mean social actors prefer not to have social ties?

- Suppose $\beta_k = -2.6$.
- Then the odds of having another tie vs. not having it are $\exp(\beta_k) = \exp(-2.6) = 0.07$
- And the binary probability to have one vs. not to have one is $\exp(\beta_k) / (1 + \exp(\beta_k)) = 0.07 / 1.07 = 0.07 = 7\%$
- ***This reflects the overall density of the network!***



Zero objective function = density 50%

An objective function that does not discriminate between options implies model actors' indifference to everything – hence, all ties will be present (or absent) with equal probability. The density then will be 50%.

Because most networks commonly studied have a density way below 50% (and hence also most network evolution processes take place in a low-density region of the network space), the outdegree parameter is estimated as significantly negative.

Similar arguments can be made about other parameters, BUT beware of control effects in the model!



Zero reciprocity effect = “reciprocity index equals density”

An objective function that *controls for density* but does not discriminate between reciprocation of existing ties and creation of asymmetric ties has a reciprocity effect of zero. The probability of a reciprocated tie then is identical to the probability of any tie, which is the density.

Because many networks commonly studied have a reciprocity index way above the density (and hence also most network evolution takes place in such network regions), the reciprocity parameter is estimated as significantly positive.

The more effects are controlled for, the more difficult it gets to tie parameters to descriptive measures...



Beware of data collection artefacts!

As shown above, the outdegree parameter typically is estimated as significantly negative, reflecting a lower than 50% density of the network.

In many data collection designs, it is impossible to ever obtain a density of 50% (e.g., “Pick up to 12 best school friends, from your cohort of size 100+”).

Hence, the parameter’s significant departure from zero must not be treated as “result” of an analysis!

Its inclusion in a model must be viewed as the necessary control for density, without which other conclusions cannot be obtained.



Don't mis-diagnose constraint as preference!

Several parameters may not necessarily reflect the expression of actual preference in the actors' decisions, but features of the opportunity structure they face when making them.

Case in point: transitive closure.

"Friends of my friends are my friends"

... because I prefer to attain cognitive balance?

... or because I have a higher chance to interact with them?

Unique conclusion typically not possible without validation by additional (e.g., qualitative) data.

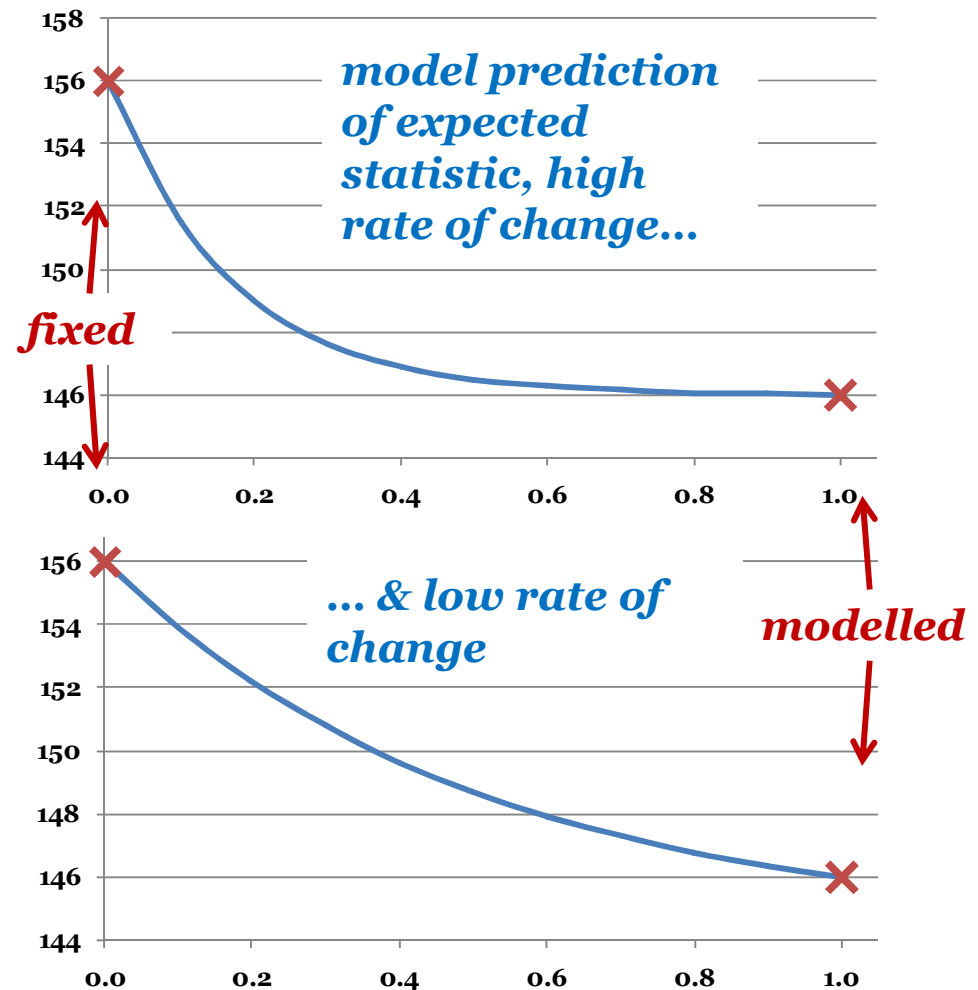


3. Interpolated region

Suppose a modelled network statistic s changes from **156** to **146** during an observation period.

The corresponding parameter is adjusted such that data point **146** is “hit” in expected value when starting out from **156**.

Steepness of the curve is co-determined by the total amount of change in the period (as modelled by rate parameters).





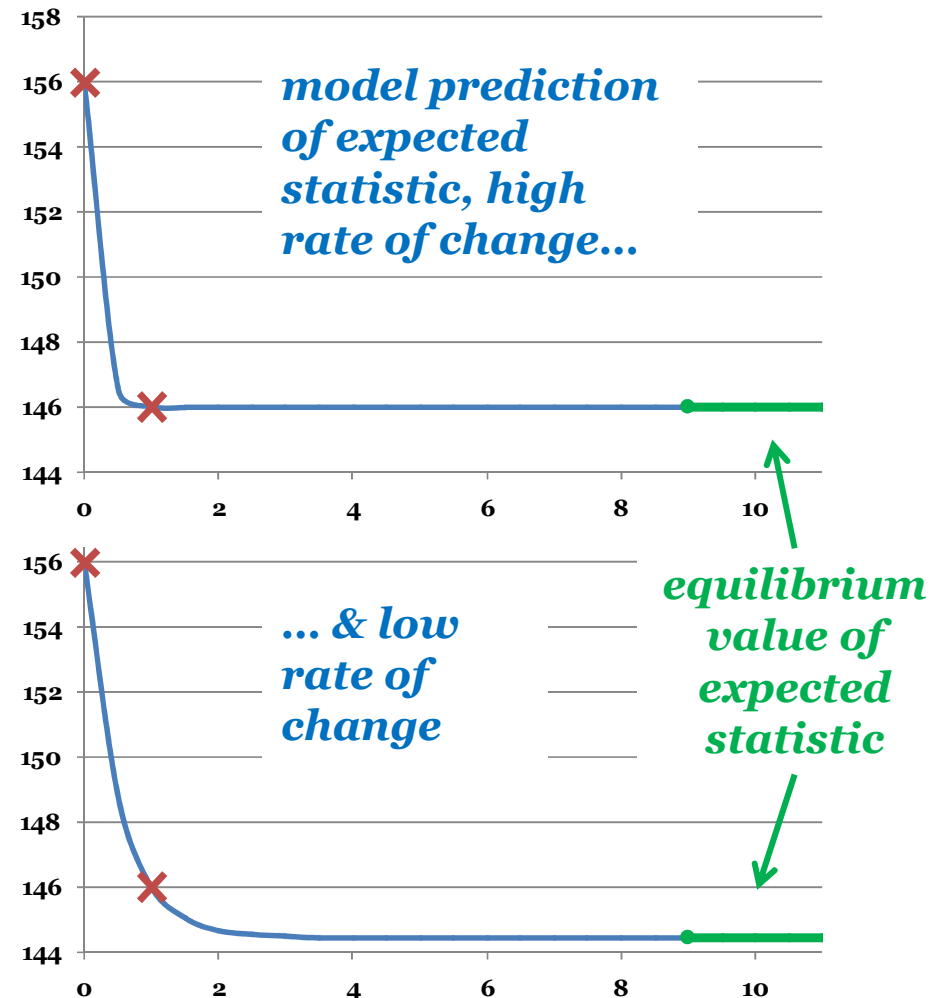
4. Projected equilibrium

Like all Markov processes, these models eventually lock in to an **equilibrium distribution** (here: on the network space).

This equilibrium...

- is uniquely identified by the parameter estimates,
- does not allow to draw conclusions about the observation period!

Similar issues as known for exp. random graph distributions.





Adding more nuance

- › Differentiating tie creation & tie maintenance
 - Endowment & creation effects
- › Goodness of fit checking
 - Score type tests (not in slides)
 - ‘Violin plots’
 - Time heterogeneity tests (not in slides)



Differentiating tie creation & tie maintenance

Remember the ‘objective function’: $f_i(\mathbf{x}) = \sum_k \beta_k s_{ik}(\mathbf{x})$

- › Models attractiveness of network states \mathbf{x} to actor \mathbf{i} .
- › Statistics \mathbf{s}_{ik} of \mathbf{i} ’s neighbourhood in \mathbf{x} are weighted by model parameters β_k .
- › These weights express whether the feature expressed in the statistic is desired ($\beta_k > 0$) or averted ($\beta_k < 0$).
- › Also they are estimated from the data.

Parameters at the same time express conditions under which new ties are established and existing ties are maintained.



What if you want to distinguish the 2 tendencies?

Earliest versions in **StOCNET**:

- › So-called ‘gratification effects’ for selected parameters.
- › To single out effects on creation vs. continuation of ties, one had to manually exploit the covariance matrix of estimates.

Later **StOCNET** versions, early RSiena:

- › Systematic availability of ‘endowment effects’ expressing continuation as an additional feature.
- › Same heavy manual data handling necessary to single out effects on creation vs. continuation of ties.



Since spring 2011:

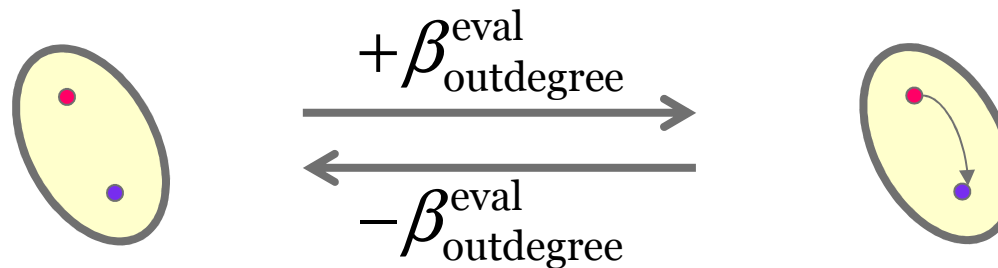
RSiena now features three separate parameters per effect:

- › The usual, non-discerning ‘evaluation effect’ expressing effects on the creation as well as maintenance of ties;
- › the ‘endowment effect’ only expressing continuation of existing ties;
- › the ‘creation effect’ only expressing establishment of new ties.

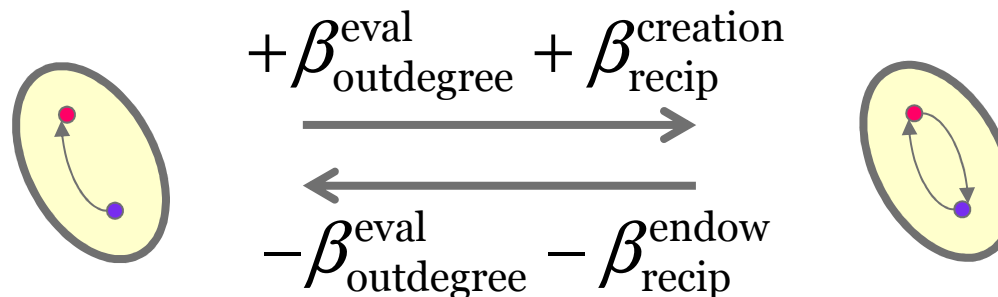
When replacing an ‘evaluation’ effect by the pair of corresponding ‘endowment’ and ‘creation’ effects, no clumsy number crunching is necessary anymore; interpretation is more transparent.



How to interpret network endowment and creation effects, vs. evaluation effects.



Diagrams show changes in the objective function for the **purple** (upper left) actor that are implied by the transitions indicated by the arrows between dyad states.





Example 1 (friendship, data courtesy of Gerhard van de Bunt)

outdegree = -1.55 , *recip.creation* = 0.98 ,
recip.endowment = 2.17

Unilateral link formation / dissolution:



Reciprocation / ending reciprocation:



Interpretation:

- formation of reciprocal ties is evaluated higher than formation of unilateral ties (upper arrows),
- dissolution of reciprocal ties is evaluated MUCH lower than dissolution of unilateral ties (lower arrows), EVEN lower than formation of reciprocal ties.



Example 2 (director provision, data courtesy of Olaf Rank)

$outdegree = -3.1$, $recip.creation = 2.9$,
 $recip.endowment = 0.7$

Unilateral link formation / dissolution:



Reciprocation / ending reciprocation:

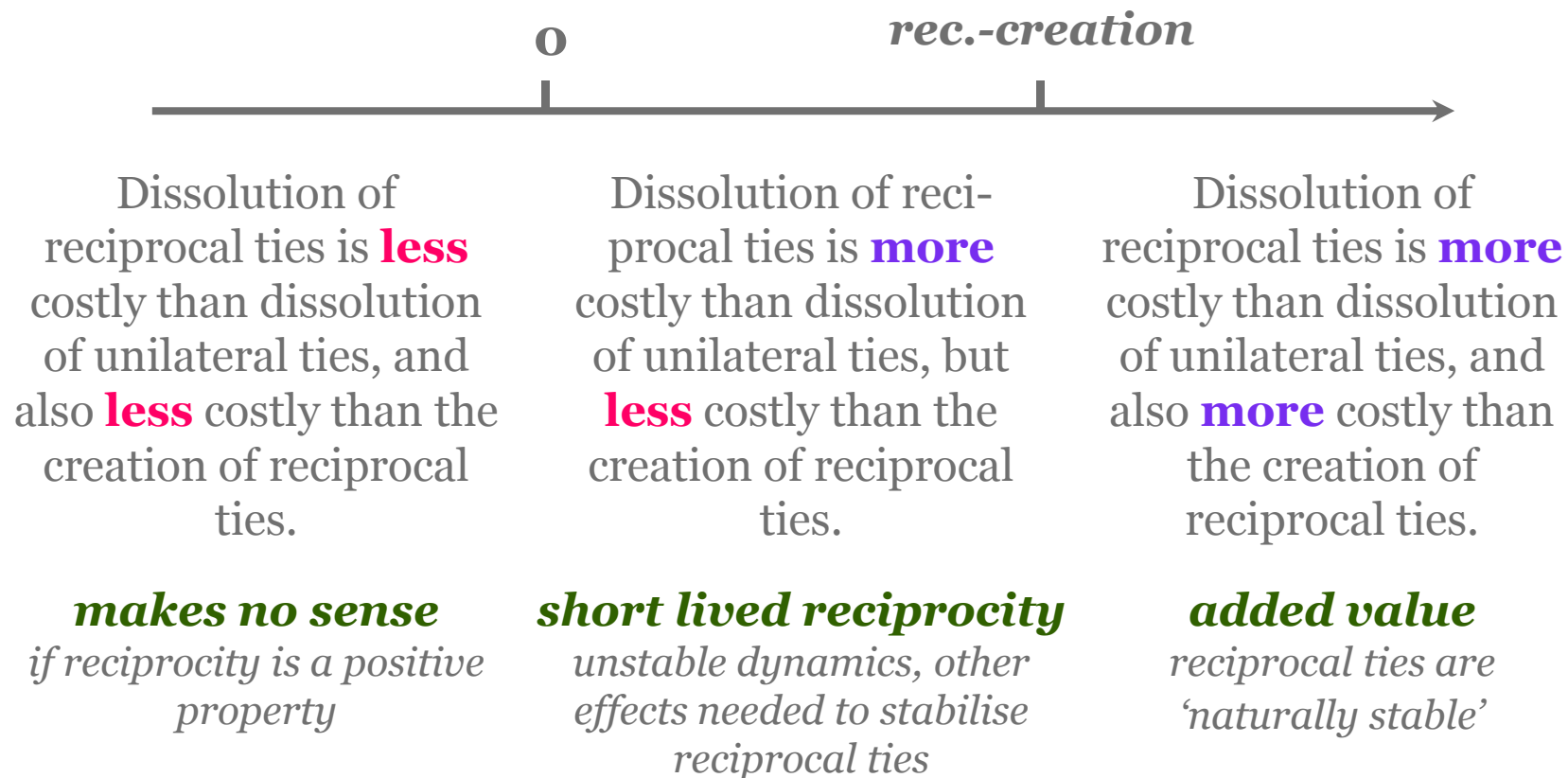


Interpretation:

- formation of reciprocal ties is evaluated higher than formation of unilateral ties (upper arrows),
- dissolution of reciprocal ties is evaluated lower than dissolution of unilateral ties (lower arrows), BUT HIGHER than the formation of reciprocal ties: *temporally limited reciprocation*.



Two 'reference points' for interpretation of the reciprocity-endowment parameter (assuming reciprocity > 0)





Extension of goodness of fit checking

What is “goodness of fit” for stochastic network models?

- › Simulate many networks from the estimated model, and see how well these simulated networks replicate features of the data that were not part of the model.

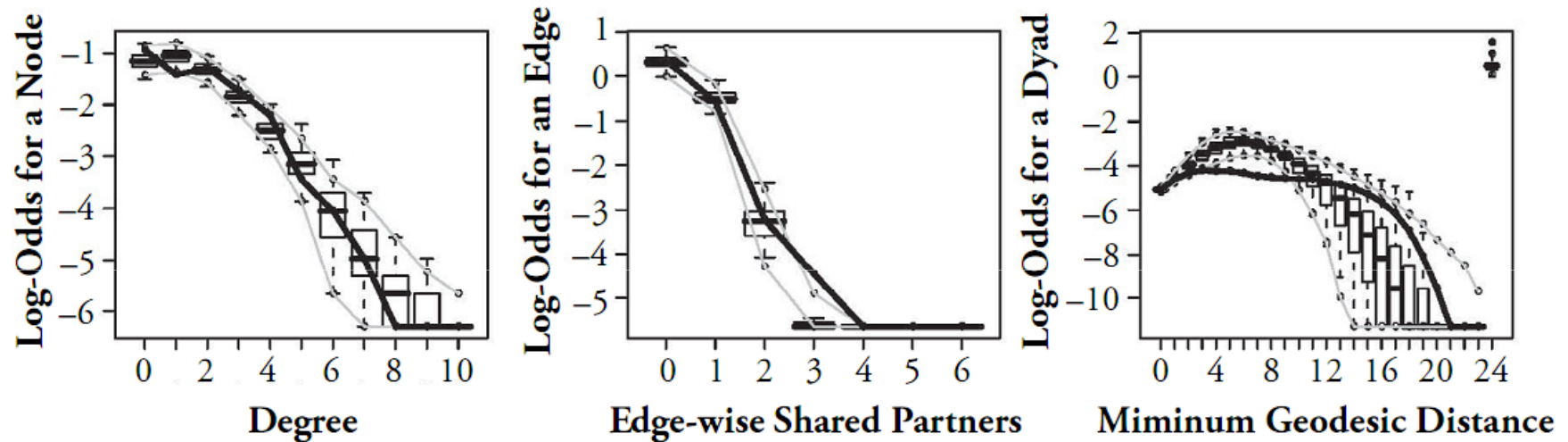
Since long (in **StOCNET** & early RSiena):

$$t = \frac{\text{E(simulated feature)} - \text{observed feature}}{\text{st.dev. (simulated feature)}}$$

- › these t-ratio indicators should not be too different from zero (ideally remain in the non-significant region)



Aim was for a while to mimic 'ergm'-package gof options



- › Black solid line shows observed values,
- › boxplots show distribution of simulated values.

This reveals a bit more about the goodness of fit than a collection of t-ratios does.



Now available in RSiena: 'violin plots'

- › Red solid line shows observed values,
- › boxplots & violins show distribution of simulated values,
- › p-value is based on a test of Mahalanobis' distance.

Examples will be elaborated in the workshop.

