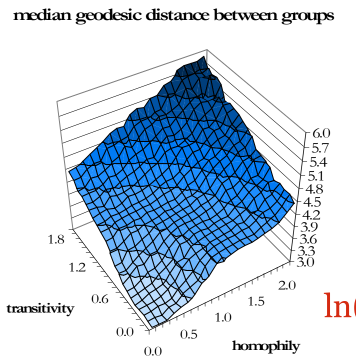
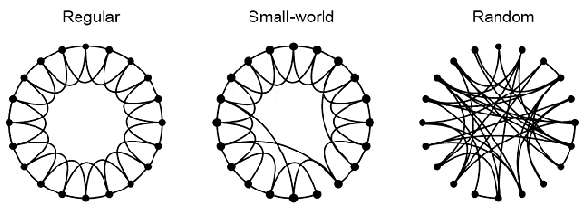
university of groningen / behavioural and social sciences / sociology

A short introduction to peer influence modelling with 'SIENA models'

SRA Biennial Meeting, 8-10 March 2012, Vancouver BC

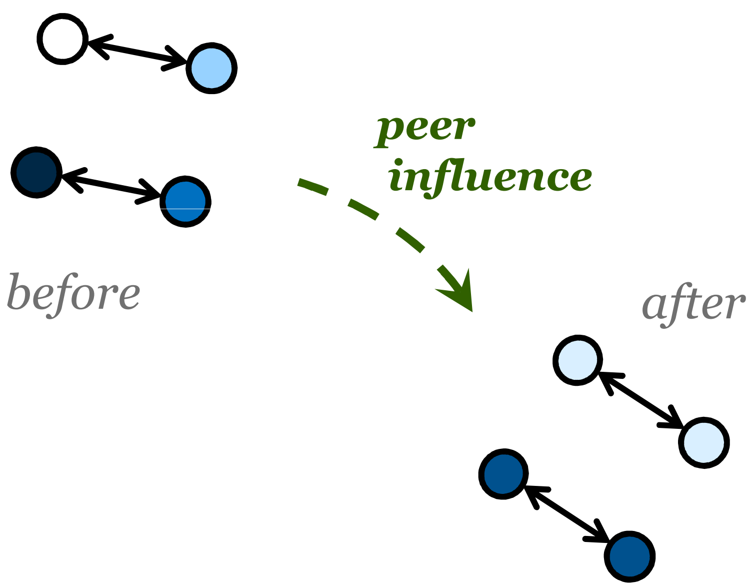
Christian Steglich c.e.g.steglich@rug.nl



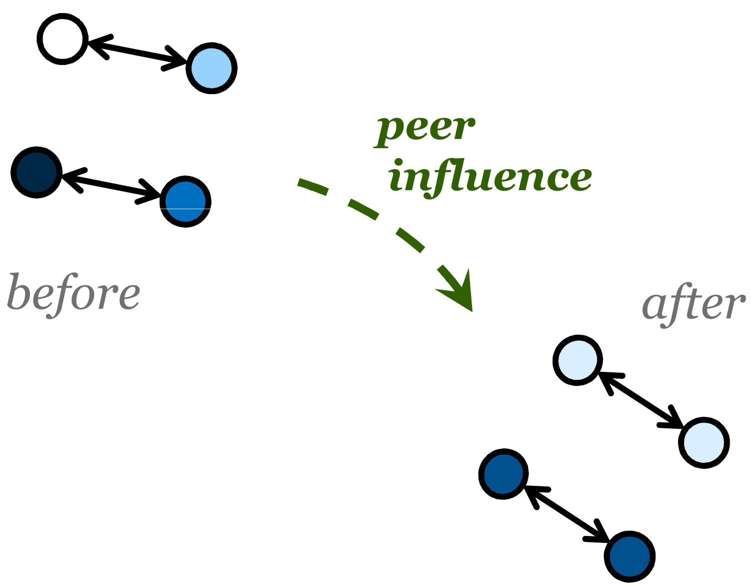


$$\ln\left(\frac{\Pr(x^c \to_i x^b)}{\Pr(x^c \to_i x^a)}\right) = \sum_{k=1}^{K} \beta_k \left(s_{ik}(x^b) - s_{ik}(x^a)\right)$$

Peer influence mechanism, basic longitudinal pattern



Peer influence mechanism, basic longitudinal pattern



<u>Example operationalisation:</u>

Influence = "getting more similar"

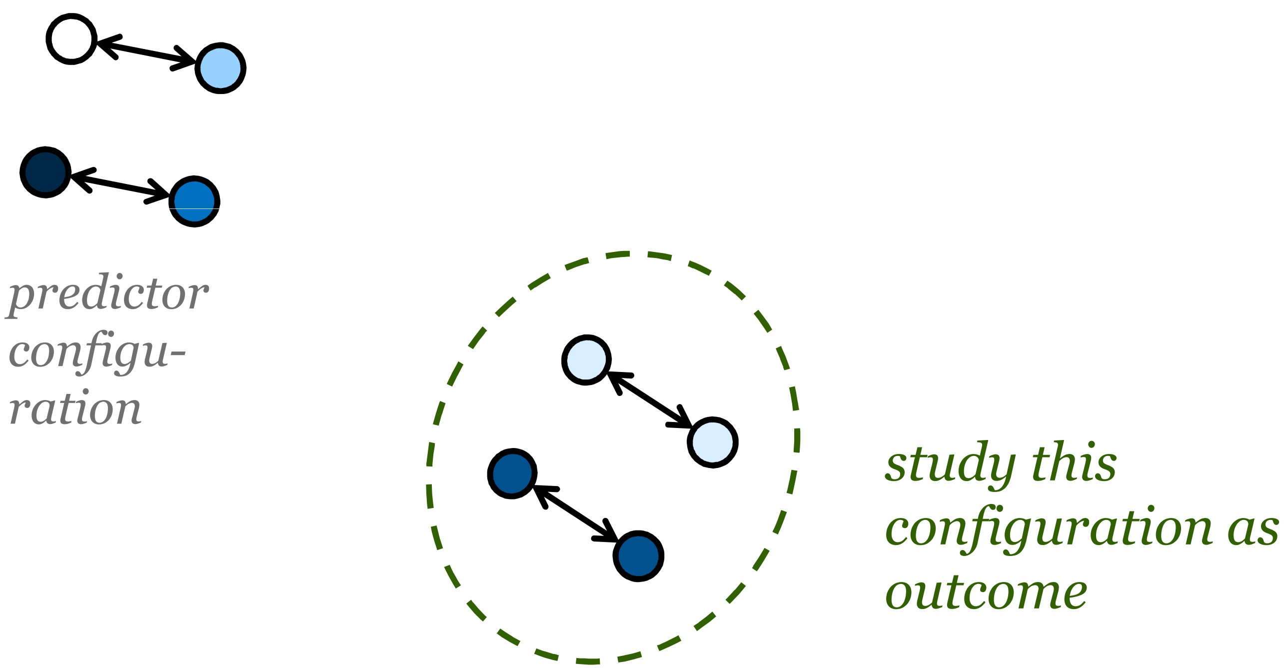
... is not always the best choice!

Think of ...

contagion / exposure / critical mass / initiation / 'spurring on' effects

these may actually imply *<u>increased dissimilarity</u>* in some developmental contexts!

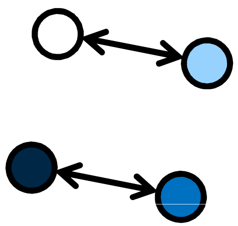
Peer influence mechanism, basic longitudinal pattern



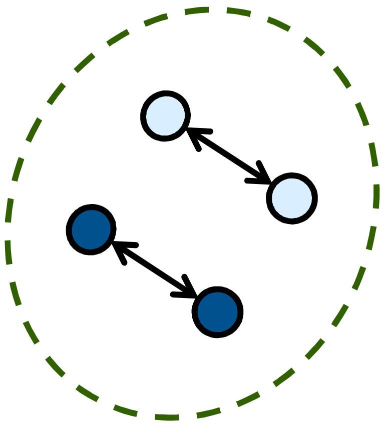
Peer influence mechanism, basic longitudinal pattern

Requires *longitudinal relational data*, emphasis is on what happens in *dyads that stay connected* only.

Conditional inference based on partial data is seldom a good idea ⇒ *analyse all dyads*, i.e., the network!

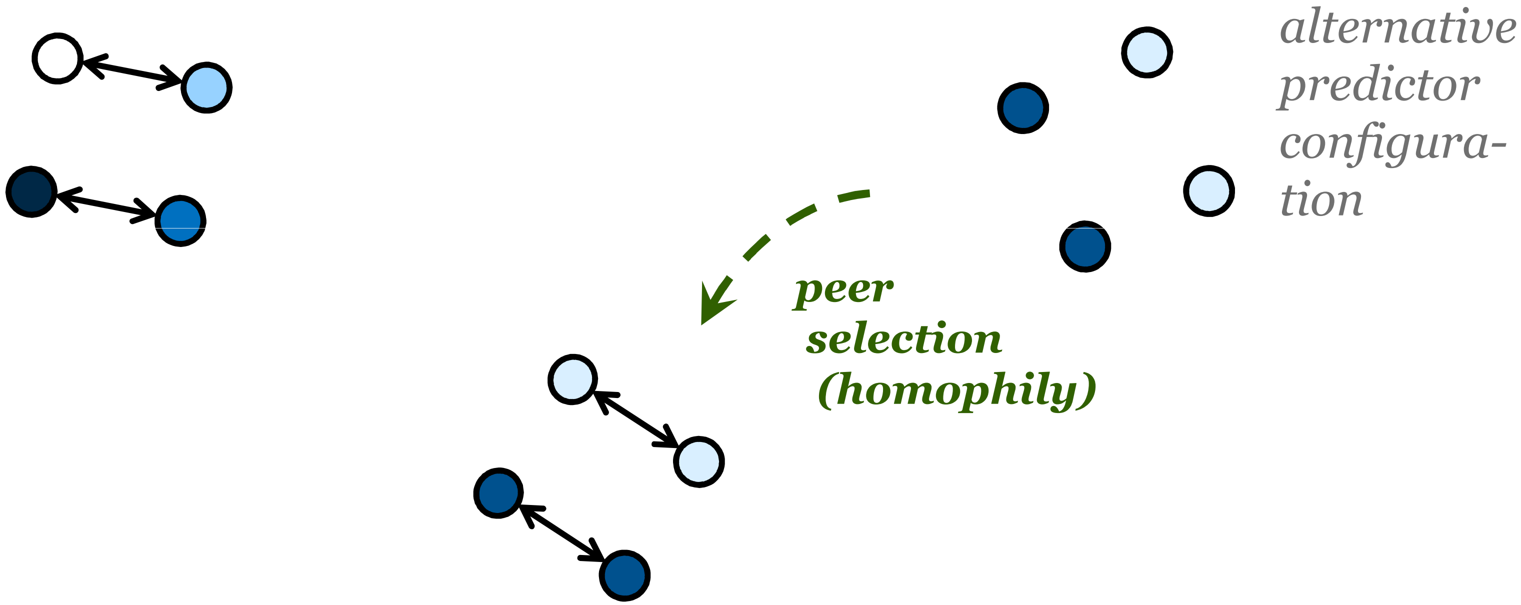


predictor configu-ration

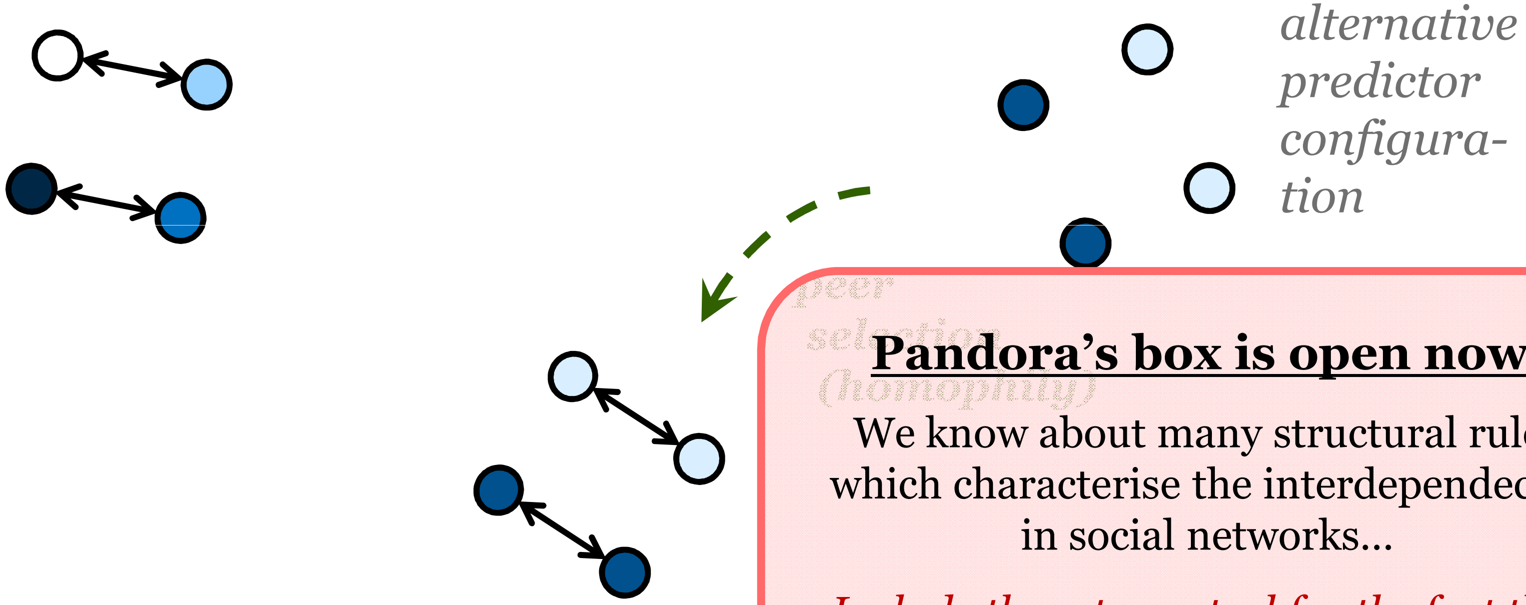


study this configuration as outcome

Peer selection *by homophily* as competing mechanism



Peer selection *by homophily* as competing mechanism

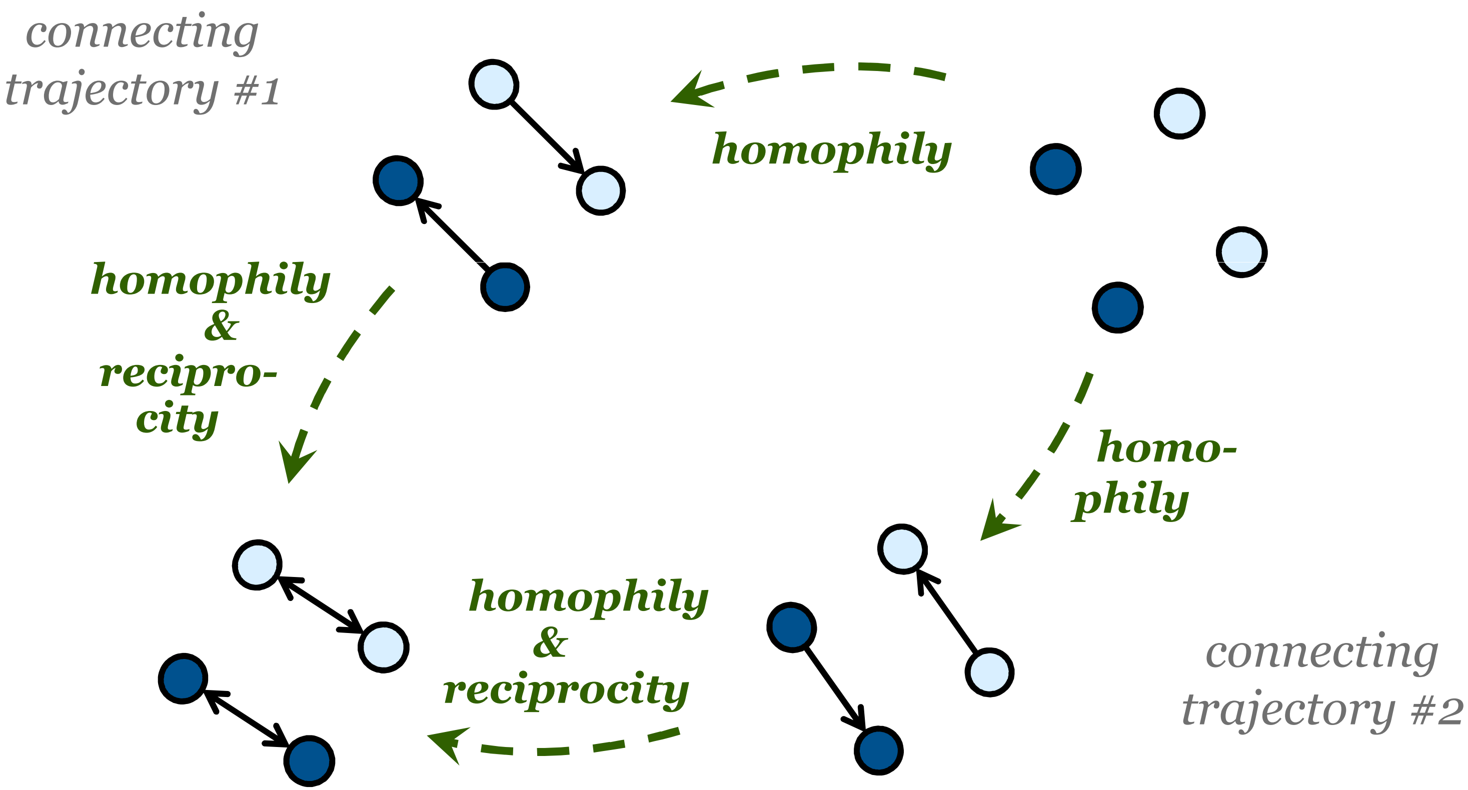


Pandora's box is open now.

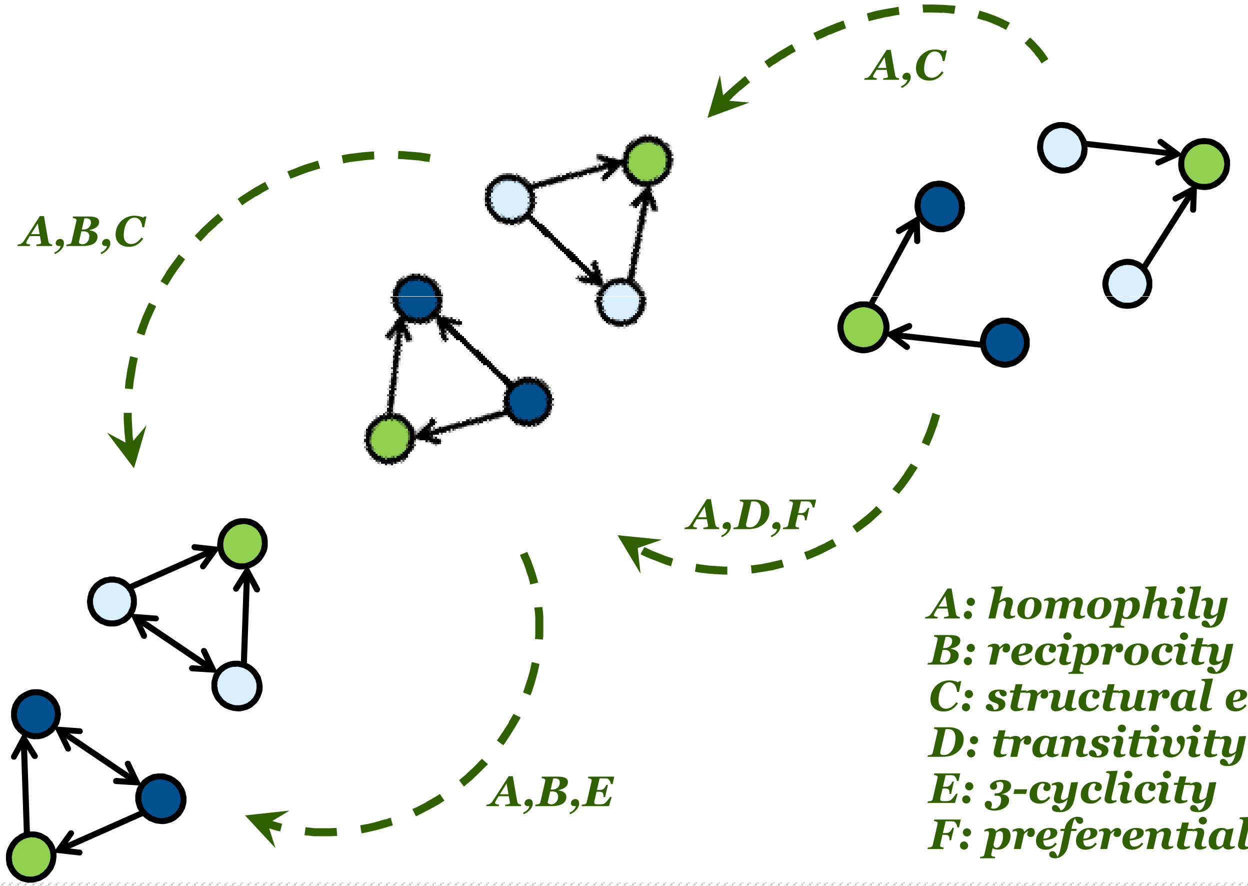
We know about many structural rules which characterise the interdependecies in social networks...

Include these to control for the fact that you handle non-independent data!

Peer selection governed by other structural rules



Peer selection affected by higher-order structure



A: homophily
B: reciprocity
C: structural equivalence
D: transitivity
E: 3-cyclicity
F: preferential attachment

Peer selection affected by higher-order structure

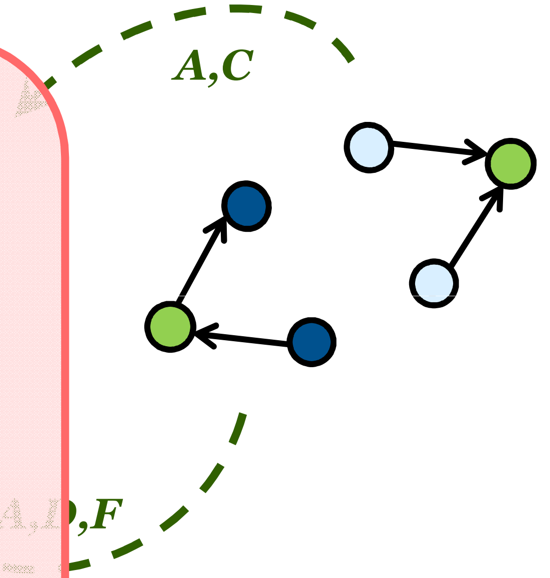
<u>Dynamic network approach forces you to think…</u>

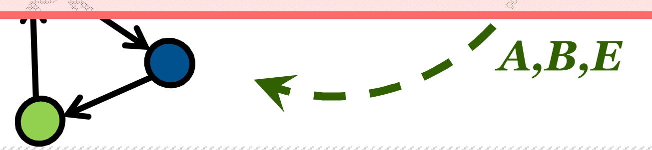
… less in terms of *observed change*

… but more in terms of *connecting trajectories* (i.e., sequences of small changes),

… and the *social mechanisms / dependencies* that these small changes instantiate.

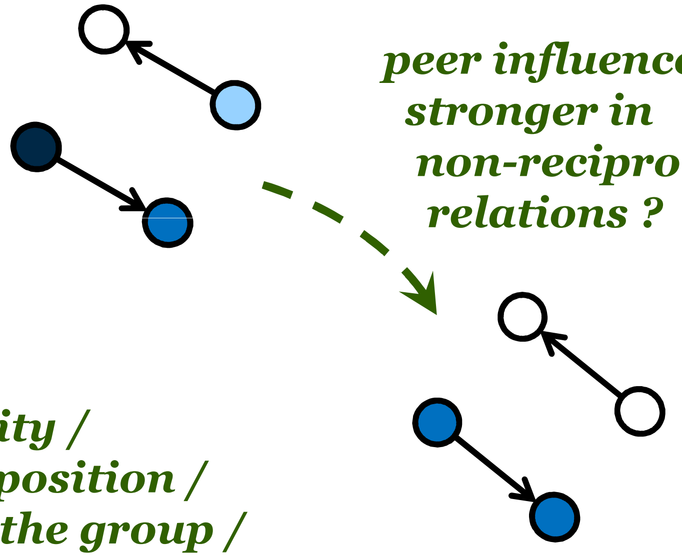
Also with respect to attitude & behaviour change in the network!



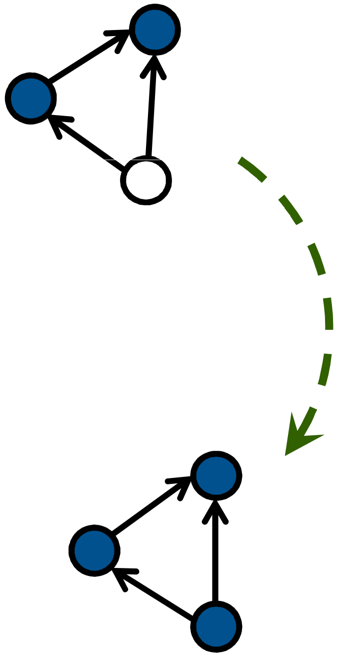


A: homophily
B: reciprocity
C: structural equivalence
D: transitivity
E: 3-cyclicity
F: preferential attachment

Peer influence probably also affected by structure



peer influence stronger in non-reciprocal relations ?



peer influence depending on others' popularity / own hierarchy position / homogeneity of the group / etc.

How to apply SIENA: Data requirements

Required are *repeated measures of the same network*:

- same group of actors

  (some composition change is allowed)
- same *slowly changing* relational variable.

Subsequent measures are assumed to be related through a continuous process of change.

In principle, continuous-time data should be easier to analyse this way – but the methods are not (yet) accessible to general public; see Brandes, Lerner & Snijders (2009); Stadtfeld (2012); others.

Example data: *(Andrea Knecht, 2003/04)*

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

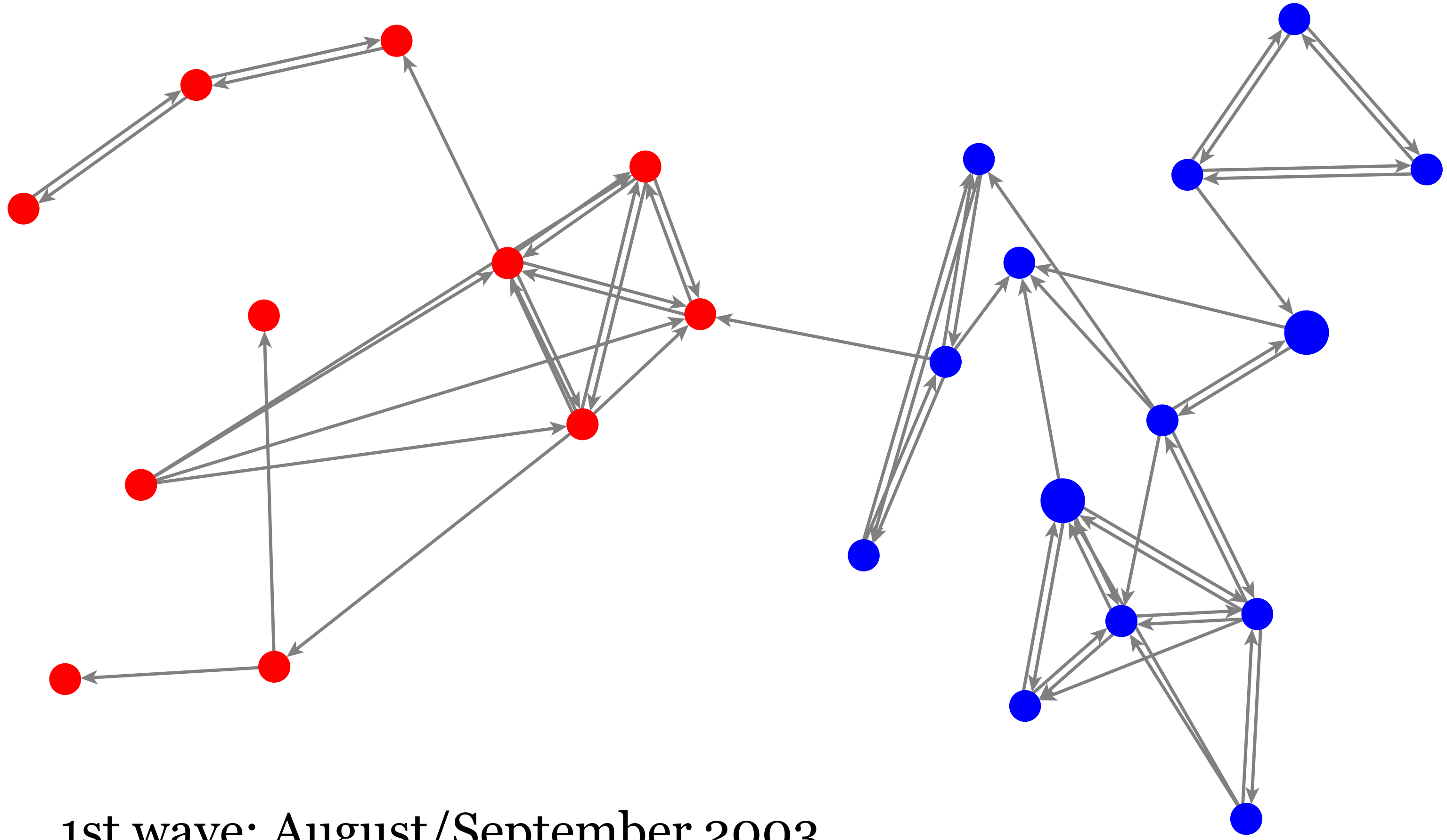
125 school classes

4 measurement points,

various network & individual measures.

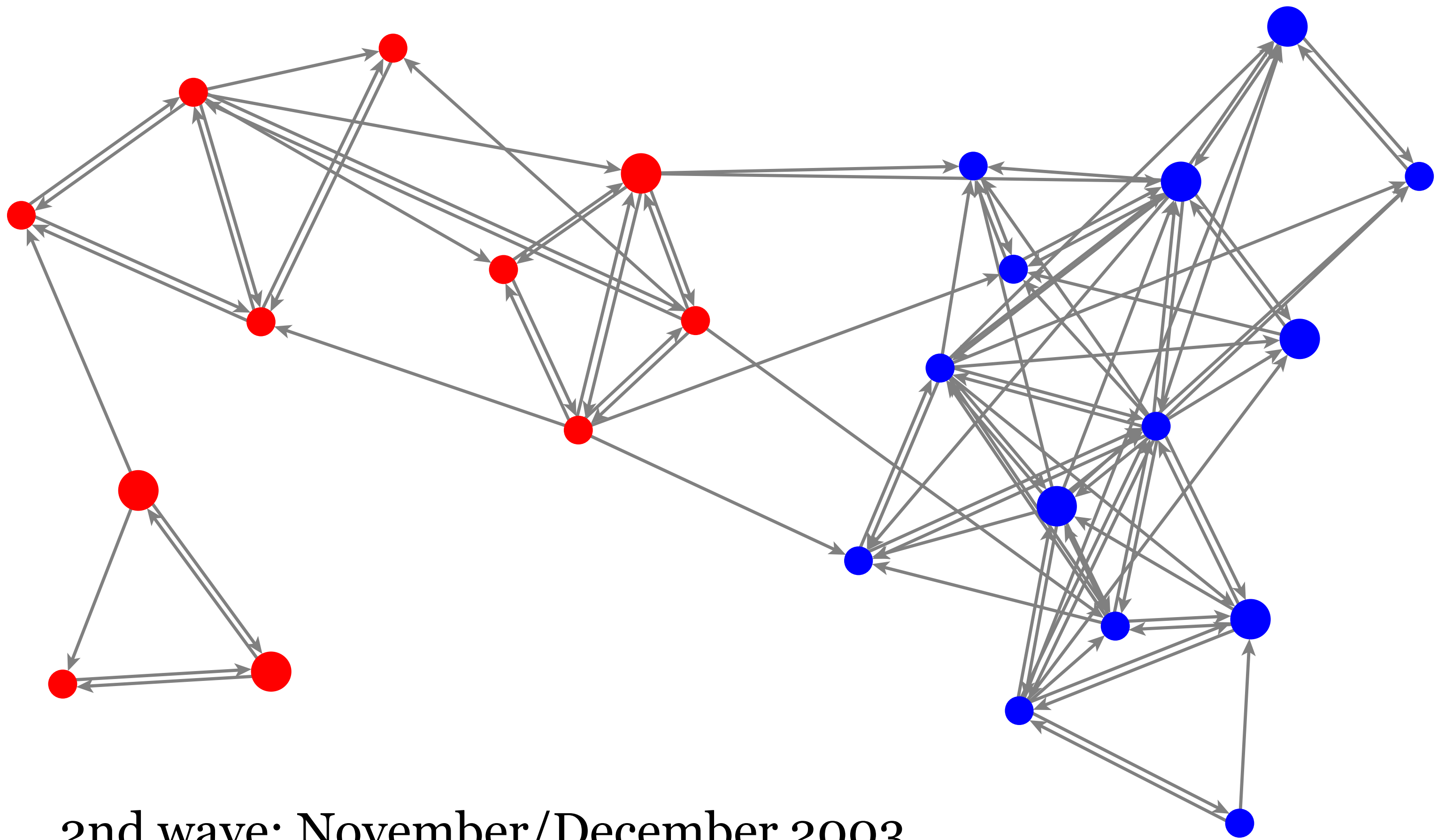
The following slides show the evolution of the friendship network in *one* of the school classes.

The graph layout is a bit messy for each observation alone, but optimal over time according to aggregated layout by visone.



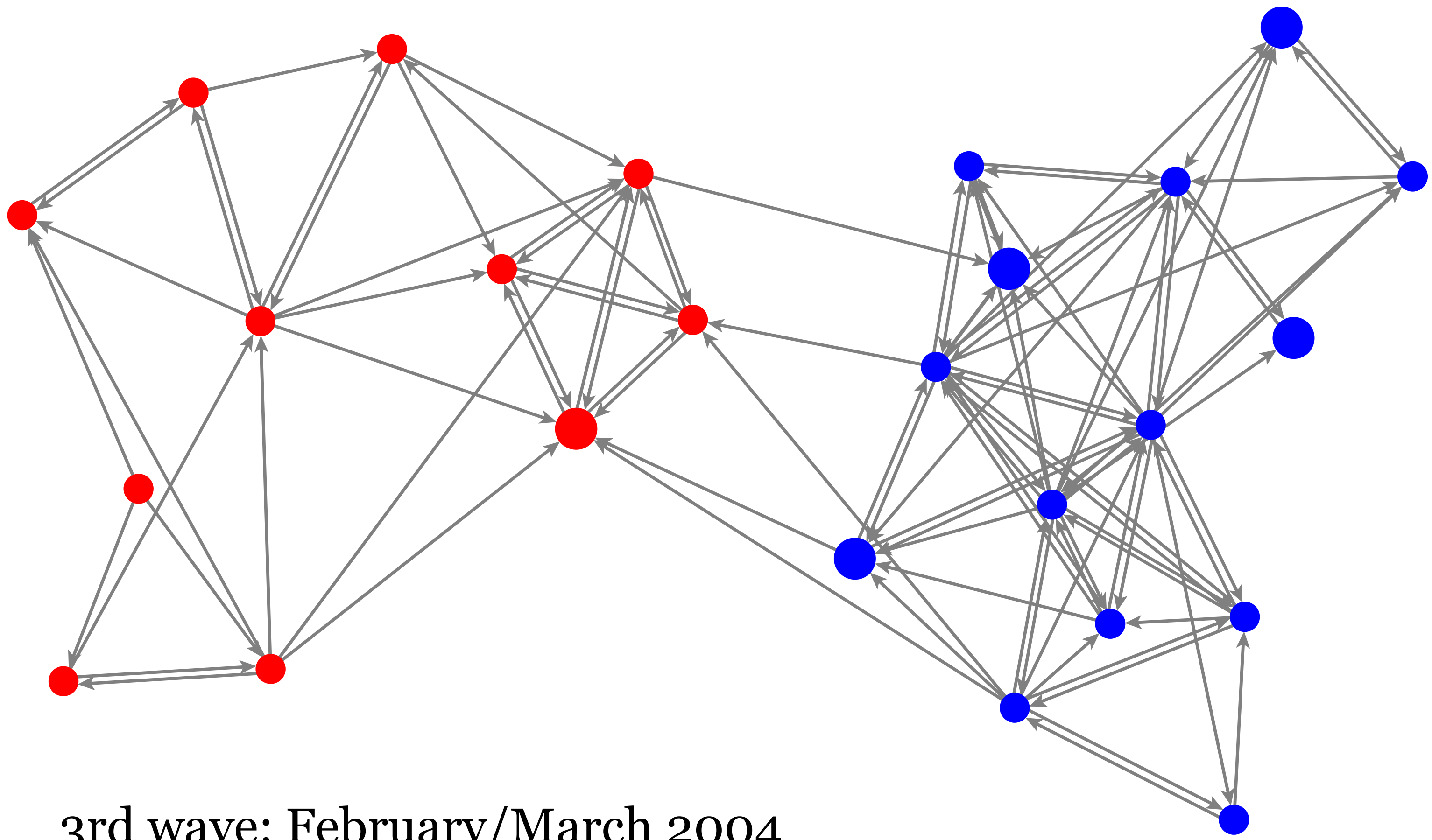
<u>1st wave:</u> August/September 2003

Node size indicates strength of delinquency…



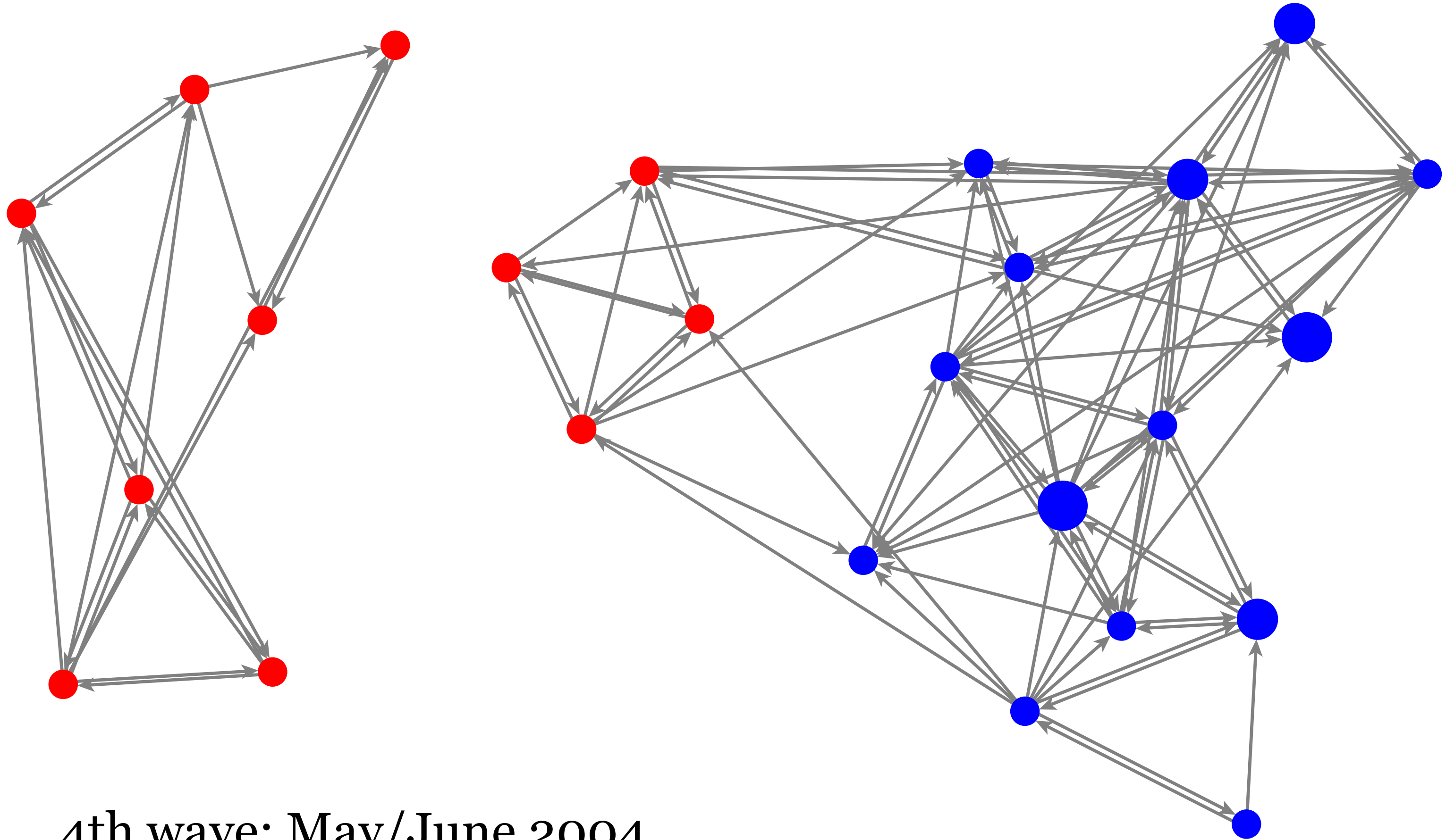
2nd wave: November/December 2003

... and node node colour indicates sex.



<u>3rd wave:</u> February/March 2004

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



<u>4th wave:</u> May/June 2004

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

How to apply SIENA: Modelling principles

Continuous-time model: Change is modelled as occurring in continuous time, in the unobserved period between observations.

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

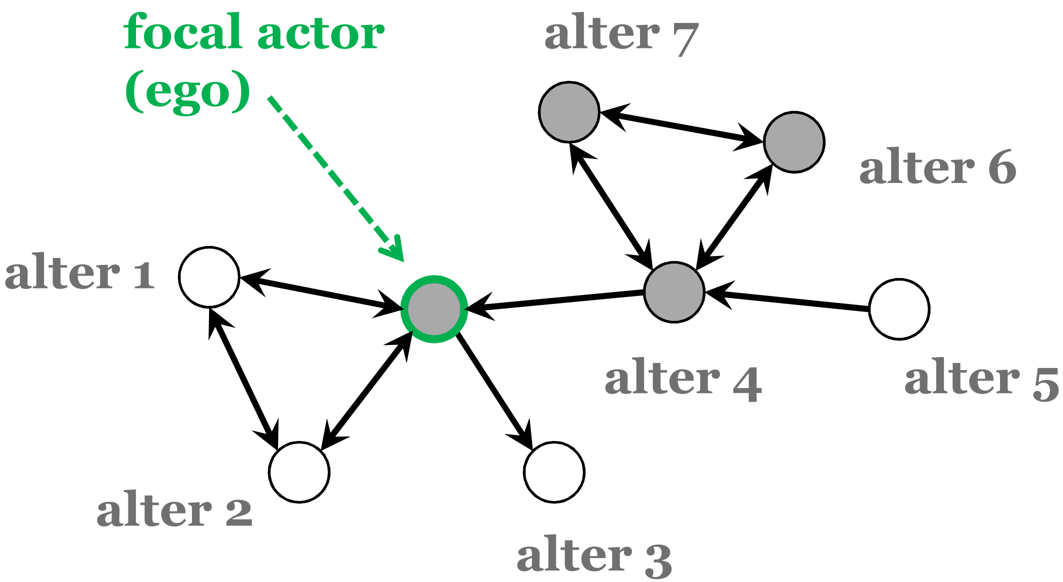
Actor-driven model: The network actors are the locus of modelling, change is outcome of individual decisions.

- actors control *whom they nominate* & *how they behave*;
- two submodels for each domain:

  When can an actor make a decision? (*rate functions*)

  Which decision does the actor make? (*objective functions*)

Micro steps for network change (example)



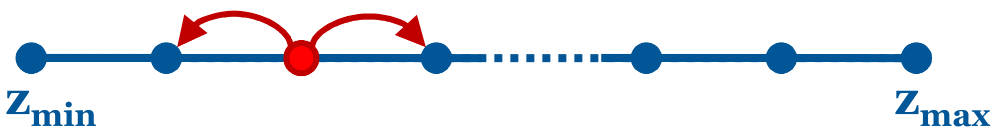
Possible micro steps:

- 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

Micro steps for behaviour change

Possible micro steps:

- increase current score or
- decrease current score on the ordinal behavioural variable, provided the range is not left



For continuous behaviour variables, the model is not yet developed.

Rate functions $\lambda_i^{dom}(x,z) = \sum_k \rho_k^{dom} r_{ik}^{dom}(x,z)$

- Models *speed* differences between actors **i** in domain dom .
- Statistics $\mathbf{r_{ik}}$ of **i**'s neighbourhood in **x,z** 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 ($\rho_k < 0$) changes in domain dom .
- 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.

Objective functions $f_i^{dom}(x,z) = \sum_k \beta_k^{dom} s_{ik}^{dom}(x,z)$

- Models attractiveness of state **x,z** to actor **i** in domain dom.
- Statistics $\mathbf{s_{ik}}$ of **i**'s neighbourhood in **x,z** are weighted by model parameters $\beta_\mathbf{k}$.
- These weights express whether the feature expressed in the statistic is desired ($\beta_\mathbf{k}$>0) or averted ($\beta_\mathbf{k}$<0).
- Also they are estimated from the data.

Technically, $f_i(x)$ ***is parameter of a multinomial logit model for discrete, probabilistic choice.***

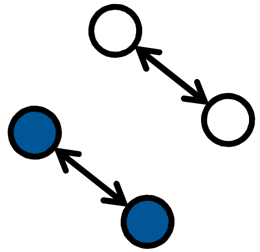
The objective function is the main part of modelling. Here, hypotheses typically are operationalised. Interpretation is very similar to logistic regression!

Schematic overview of model components

	Timing of decisions	**Decision rules**
Network evolution	Network rate function λ^{net}	Network objective function f^{net}
Behavioural evolution	Behaviour rate function λ^{beh}	Behaviour objective function f^{beh}

- *By simultaneously operating both processes on the same state space (conditionally independent, given the current state), feedback processes are instantiated.*
- *Network evolution model and behavioural evolution model therefore are controlling for each other!*

So, how is peer influence analysed?



A measure implemented in SIENA that measures prevalence of our *‘outcome configuration’* on the right is the ***network similarity statistic*** $\sum_j x_{ij} sim_{ij}$, where sim_{ij} is a standardised measure of similarity of two actors based on their distance on a variable z , $sim_{ij} = 1 - \left(|z_i - z_j| / range_z \right)$.

$sim_{ij}=1$ means scores of i and j are identical;

$sim_{ij}=0$ means they are maximally apart (one maximal, the other minimal).

Modelling selection and influence, technically

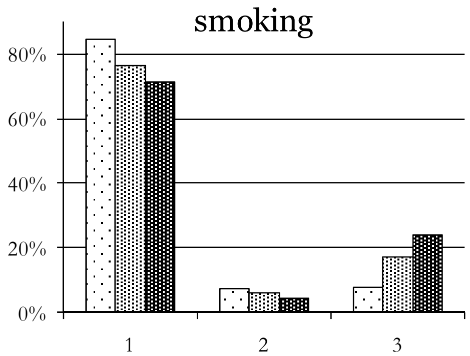
By including the network similarity statistic $\sum_j x_{ij} \mathrm{sim}_{ij}$

...in the <u>*network objective function*</u>, homophilous selection is modelled,

...in the <u>*behaviour objective function*</u>, assimilation / social influence is modelled.

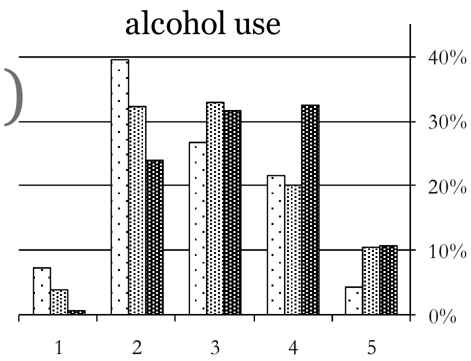
Other peer effects are handled in the same way: the 'outcome configuration' is quantified in a statistic, the peer effect is estimated as part of the behaviour change model, the selection confounder is part of the network change model.

Patterns in applications to the substance use field *(very briefly; NEW results to come!)*



Tobacco: (primarily by Mercken et al.)

- Selection important,
- influence not reliably found.
- Polarisation dynamics (addiction).



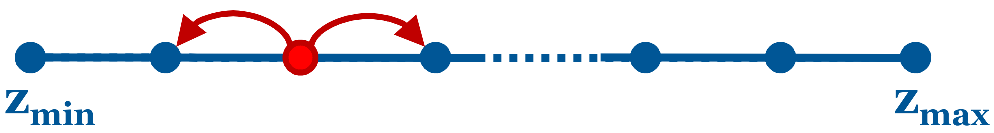
Alcohol: (Steglich et al.; Knecht et al.; others)

- Influence and selection very important.
- Normative dynamics (agreement).

Today's new work:

- Light et al. on *alcohol use* and *initiation*
- de la Haye et al. on *marijuana use* and *initiation*
- Green et al. on *smoking intiation*

Initiation is the methodologically interesting keyword here… it is in violation of this little diagram I presented:



It will require some modified approach…

Let's listen to the speakers to find out!