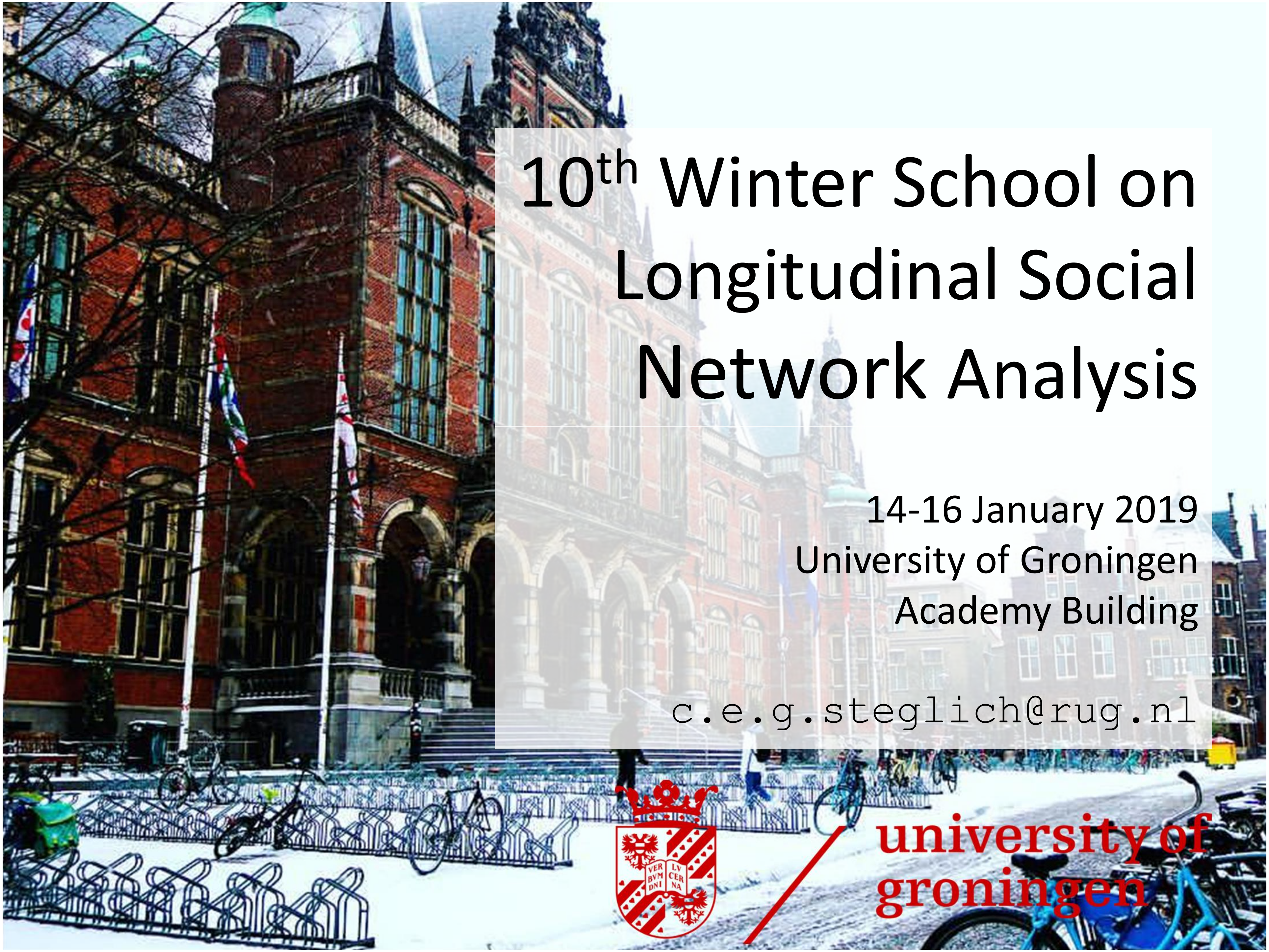

10th Winter School on Longitudinal Social Network Analysis
14-16 January 2019
University of Groningen
Academy Building
c.e.g.steglich@rug.nl
university of groningen

Co-evolution models for networks and behaviour

- Interdependence of networks and behaviour
- Extension of the stochastic actor-based modelling framework to "behaviour" dimensions
- The case of *homogeneity bias / network autocorrelation*
- An example:
  Co-evolution of academic performance & advice seeking
- Notes on the modelling of peer influence

Interdependence of networks and behaviour

LINKÖPING UNIVERSITY

university of groningen

Known:

Networks can depend on actor characteristics

Three main effect types in <u>directed</u> networks

- *“selective mixing” (one effect type, two signs:)*
  - assortative (homophily): interaction with similar others can be more rewarding than interaction with dissimilar others
  - disassortative (heterophily/exchange): selection of partners such that they complement own abilities and resources
- *“sociality” (two effect types:)*
  - popularity (receiver effect): some properties render actors more attractive as receivers of network ties
  - activity (sender effect): other properties may make actors send more network ties

New:
Actor characteristics can depend on network

Changeable individual characteristics can be affected by other individuals in the network: behaviour proper, but also opinions, attitudes, intentions, etc. – we use the term "*behaviour*" here.

Some examples:

- contagion / assimilation: innovations spreading in a professional community; adolescents adopting friends' attitudes; investment bankers copying behaviour of successful competitors
- differentiation: division of tasks in a connected work team
- effects of centrality or position: special portfolio of connections may lead to behaviour that other actors do not exhibit

"Natural matching" of effects in both directions

<u>Buying friends with sweets?</u>

- *Suppose there exists a mechanism such that the amount of candies a pupil brings to school attracts friendships.*
- *Over time, this mechanism will lead to a positive association between candies and (in)degree in the friendship network.*

Suppose further that in a cross-sectional data collection, we can measure this association. Is the mechanism proven?

- *No! The same association can also be explained by a mechanism in the other causal direction: a higher number of friends could make a student bring more candies.*

General point ("conjugate mechanisms")

Any *cross-sectional association* between network features and individual characteristics could come about by at least two *competing dynamic mechanisms*:

1. The network feature leads to adjustment of individual characteristics.
2. The individual characteristics lead to adjustment of the network feature.

Aims:

- *Construction of a model that allows a teasing-apart.*
- *Construction "as simple as possible" (close to existing stochastic, actor-based modelling).*

Extension of the stochastic actor-based framework

LINKÖPING UNIVERSITY

university of groningen

"Do as much in analogy as possible"

- Stochastic process (X,Z) on the (extended) space of all possible network-behaviour configurations (x,z).

  $2^{n(n-1)}$ states for network x (binary, directed case)

  r^n states for behaviour z (ordinal, finite range r)

  State space of (X,Z) has size $2^{n(n-1)} \times r^n$.

- Again, the first observation is not modelled but conditioned upon as the process' starting value.
- Discrete change is modelled as occurring in continuous time, but now there are two types of change.

Actor based approach now in two domains

Network actors drive the process (discrete choice model).

- two domains of decisions:
  - decisions about network neighbours,
  - decisions about own behaviour.
- per decision domain two model parts:
  - *When* can actor *i* make a decision? (rate functions λ^{net}, λ^{beh})
  - *Which* decision does actor *i* make? (objective functions f^{net}, f^{beh})

By again sampling waiting times and identifying the shortest one, it becomes clear *who* makes *which type* of change.

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 change and behaviour change therefore are *controlled for each other*'s parallel occurrence.

Monte Carlo simulation

Initialise system at 1st observation

Each actor draws one waiting time for his/her next network change and one for his/her next behaviour change

rate functions
$\lambda^{\text{net}}, \lambda^{\text{beh}}$

Actor and change type of the shortest waiting time are identified

Is time up?

no

yes

Terminate and return system state

objective functions
$f^{\text{net}}, f^{\text{beh}}$

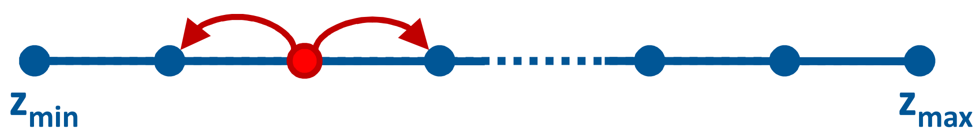
Actor implements change of given type

Time is updated

Mini steps assumed in behaviour change

Choice options:

(1) increase, (2) decrease, or (3) keep current score
on the ordinal behavioural variable, provided the range is not left



Choice probabilities:

Analogous to network part: conditionial logit model based on evaluations of options according to behavioural objective function.

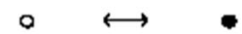
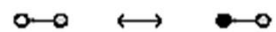
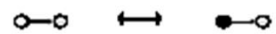
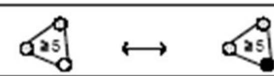
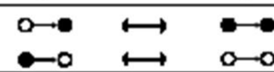
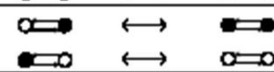
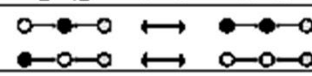
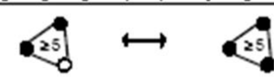
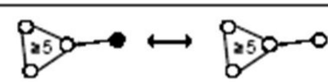
Explanatory model for behaviour change:

By inclusion of effect statistics in the objective function.

Also here, many effects can be formulated for inclusion in the objective function…

TABLE 3

SELECTION OF POSSIBLE EFFECTS FOR MODELING BEHAVIORAL EVOLUTION

effect	network statistic	effective transitions in network*	verbal description
1. tendency	z_i		main behavioral tendency
2. indegree × behavior	$z_i \sum_j x_{ji}$		effect of own popularity on behavior
3. outdegree × behavior	$z_i \sum_j x_{ij}$		effect of own activity on behavior
4. dense triads × behavior	$z_i \sum_{jh} \text{group}(ijh)$		effect of belonging to cohesive subgroups on behavior
5. peripheral × behavior	$z_i \sum_{jhk} \text{peripheral}(i;jhk)$		effect of being peripheral to cohesive subgroups on behavior
6. isolation × behavior	$z_i \text{ isolate}(i)$		effect of being isolated in the network on behavior
7. similarity	$\sum_j x_{ij} \text{ sim}_{ij}$		assimilation to friends (contagion / influence)
8. similarity × reciprocity	$\sum_j x_{ij} x_{ji} \text{ sim}_{ij}$		assimilation to reciprocating friends
9. similarity × pop. alter	$\sum_j x_{ij} \text{ sim}_{ij} \sum_k x_{kj}$		assimilation to popular friends
10. similarity × dense triads	$\sum_{jh} \text{group}(ijh)(\text{sim}_{ij} + \text{sim}_{ih})$		assimilation to the majority behavior in a cohesive subgroup
11. similarity × peripheral	$\sum_{jhk} \left(\text{peripheral}(i;jhk) \times (\text{sim}_{ij} + \text{sim}_{ih} + \text{sim}_{ik})\right)$		assimilation to those cohesive subgroups one unilaterally attaches to

* 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. Actor i is the actor who changes color z_i in the transition indicated by the double arrows. Illustrations are not exhaustive.

Equation-based estimation of co-evolution models

Algorithm needs to be modified slightly because the *default equations for 'competing process explanations' are identical* and would imply an unsolvable, collinear system of equations.

Solution: work with *cross-lagged statistics* in the equations!

- Network change in response to prior behaviour,
- behaviour change in response to prior network.

(Likelihood-based estimation does not require such modification.)

Estimating equations

When *X*, *Z* are model-based simulated data and *x*, *z* the empirical data, the following statistics are used:

- For parameters in the network objective function:

$$\mathbf{S}_{\bullet}\left(\mathbf{X},\mathbf{Z}\right)=\sum_{\mathbf{k}}\sum_{\mathbf{i}}\mathbf{s}_{\mathbf{ih}}^{\mathbf{net}}\left(\mathbf{X}(\mathbf{t}_{\mathbf{k}+1}),\mathbf{z}(\mathbf{t}_{\mathbf{k}})\right)$$

- For parameters in the behaviour objective function:

$$\mathbf{S}_{\bullet}\left(\mathbf{X},\mathbf{Z}\right)=\sum_{\mathbf{k}}\sum_{\mathbf{i}}\mathbf{s}_{\mathbf{ih}}^{\mathbf{beh}}\left(\mathbf{x}(\mathbf{t}_{\mathbf{k}}),\mathbf{Z}(\mathbf{t}_{\mathbf{k}+1})\right)$$

The estimating equations are $\mathbf{E}(\mathbf{S}_{\bullet}(\mathbf{X},\mathbf{Z}))=\mathbf{S}_{\bullet}(\mathbf{x},\mathbf{z})$; everything else remains as in the case of the simple network evolution model.

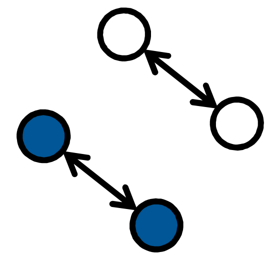
An important class of “social influence” applications

LINKÖPING UNIVERSITY

university of groningen

Explaining homogeneity bias

In networks connected actors are often behaviourally more similar than non-connected actors. Technically, this has been termed *homogeneity bias* or *network autocorrelation*.



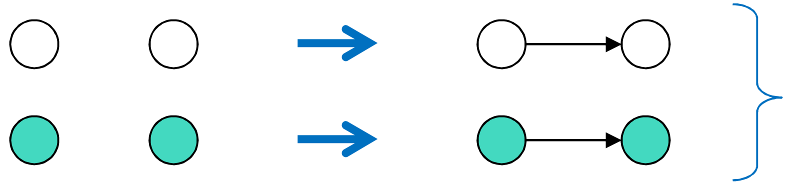
One measure (implemented in RSiena) is the *network similarity statistic* $\sum_j x_{ij}\,\mathrm{sim}_{ij}$, where sim_{ij} is a standardised measure of similarity of two actors based on their distance on a variable *z* :

$$\mathrm{sim}_{ij} = 1 - \left(| z_i - z_j | / \mathrm{range}_z \right)$$

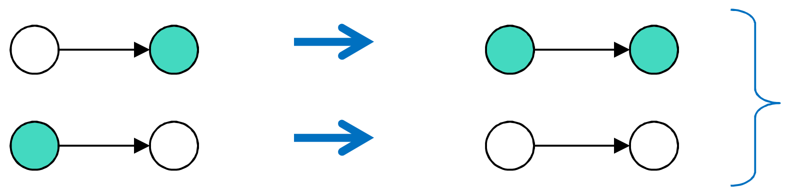
$\mathrm{sim}_{ij}=1$ means scores of *i* and *j* are identical; $\mathrm{sim}_{ij}=0$ means they are maximally apart (one maximal, the other minimal).

Competing explanatory stories

A story of network change: Actors base their social relations on similarity in individual features. As a *mini step*:

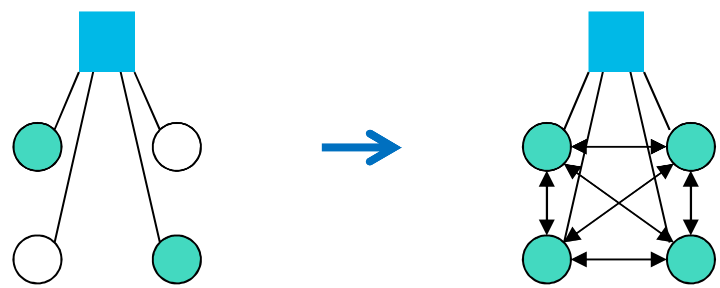


A story of behaviour change: Actors adjust their behaviour to the behaviour in their social environment. As a *mini step*:



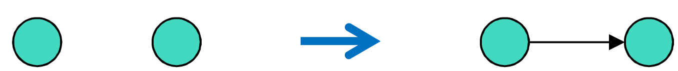
Never forget possibility of "confounders"

Notably 'shared context' can lead to both connectivity and individual change:



If connectivity happens faster, this looks like influence:

If individual change is faster, it looks like homophily:



Modelling selection and influence

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

…in the ***network objective function***, homophilous selection is modelled;

…in the ***behaviour objective function***, assimilation / social influence is modelled.

It can be of crucial importance to be able to control one effect for the occurrence of the other – e.g., in the design of "peer-led" social network interventions to reduce risk-taking behaviours at schools.

… & what to do about confounders?

Most relevant confounders are probably “shared social contexts”.

$\rightarrow$ Before the study, think about those & make inventory.

$\rightarrow$ In the study, measure them.

$\rightarrow$ In the analysis, control for them.

This procedure reduces danger of missing important unobserved confounders.

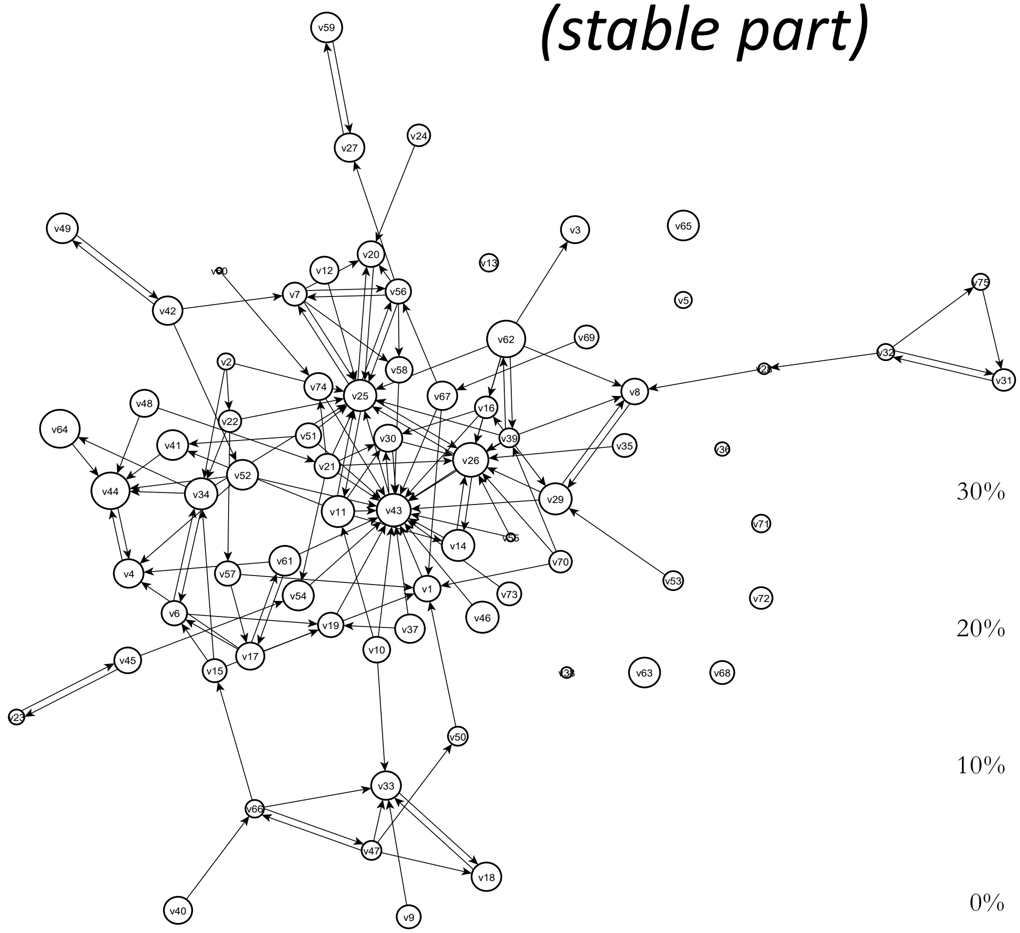
An example

Consider this MBA student data set

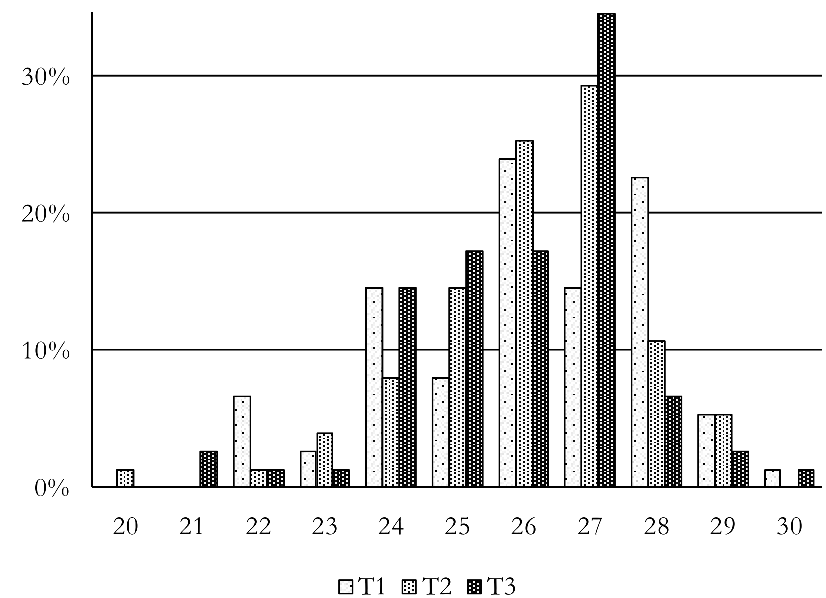
- **75 students** enrolled in an MBA program;
- **4 network variables**: advice-seeking, communication, friendship, acknowledge-contribution-to-learning;
- **co-evolving behavioural dimension**: performance in examinations;
- **several other actor variables**: gender, age, experience, nationality;
-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

We now focus on the co-evolution of students' *performance* and their *advice seeking* network.

Advice seeking network (stable part)



Performance distribution



We expect the following mechanisms

Expectation 1: High performers ask less for advice

- Association (negative) between outdegree & performance
- Conjugate process: high outdegree reduces performance

Expectation 2: High performers are asked more for advice

- Association (positive) between indegree & performance
- Conjugate process: high indegree increases performance

Expectation 3: Processes of homogeneity bias (see earlier)

Performance part of the estimates

Effects	Advice (restricted)	Advice (full)	
Rate period 1	4.123 (0.900)	4.127 (1.018)	
Rate period 2	2.647 (0.632)	2.685 (0.583)	
Linear shape	–0.685* (0.293)	–0.679* (0.307)	
Quadratic shape	–0.035 (0.039)	–0.037 (0.048)	
Average similarity	10.077* (4.491)	9.841* (4.778)	Exp.3 infl.
Indegree (popularity)	0.044* (0.022)	0.044* (0.021)	Exp.2 conj.
Outdegree (activity)	0.015 (0.033)	0.019 (0.034)	Exp.1 conj.
Gender	–	0.015 (0.165)	
Ability	–	0.009 (0.016)	
Age	–	–0.009 (0.037)	
Work experience	–	–0.001 (0.005)	
Nationality	–	0.115 (0.288)	
Time since graduation	–	0.002 (0.004)	

Advice seeking part of the estimates (I)

Effects	Advice (restricted)	Advice (FULL)
Rate Period 1	9.212 (0.888)	9.202 (0.881)
Rate Period 2	6.423 (0.537)	6.474 (0.586)
Endogenous network effects		
Outdegree (density)	-2.473^{***} (0.125)	-3.301^{***} (0.196)
Reciprocity	1.065^{***} (0.127)	0.996^{***} (0.133)
Transitive triplets	0.261^{***} (0.031)	0.249^{***} (0.030)
3-cycles	−0.030 (0.066)	−0.018 (0.064)
Betweenness	-0.035^{*} (0.019)	-0.044^{**} (0.018)
Indegree – Popularity	0.019^{***} (0.003)	0.258^{***} (0.051)
Exogenous network effects		
Friendship	0.972^{***}(0.091)	0.945^{***} (0.091)
Advice	–	–
Control covariates effects		
Gender (M) alter	–	−0.069 (0.096)
Gender (M) ego	–	-0.239^{**} (0.094)
Same Gender (M)	–	0.094 (0.088)

LINKÖPING UNIVERSITY

university of groningen

Advice seeking part of the estimates (II)

Effects	Advice (restricted)	Advice (FULL)	
Ability similarity	–	0.249 (0.175)	
Age alter	–	−0.024 (0.018)	
Age ego	–	−0.011 (0.019)	
Age similarity	–	0.113 (0.299)	
Same academic background	–	0.262** (0.079)	
Work experience alter	–	0.002 (0.003)	
Work experience ego	–	0.001 (0.003)	
Work experience similarity	–	−0.249 (0.468)	
Same nationality	–	0.444*** (0.122)	
Time since graduation alter	–	0.002 (0.003)	
Time since graduation ego	–	0.004* (0.002)	
Time since graduation similarity	–	0.303 (0.286)	
Performance feedback effects			
→ Performance alter	–	0.156** (0.055)	Exp.2
→ Performance ego	–	−0.090^{+} (0.055)	Exp.1
→ Performance similarity	–	1.424* (0.657)	Exp.3 sel.

LINKÖPING UNIVERSITY

university of groningen

Two more remarks on stochastic actor-based influence modelling

LINKÖPING UNIVERSITY

university of groningen

(1) Always consider distribution of behaviour!

Peer influence doesn't necessarily mean *"connected people becoming / staying more similar over time"*

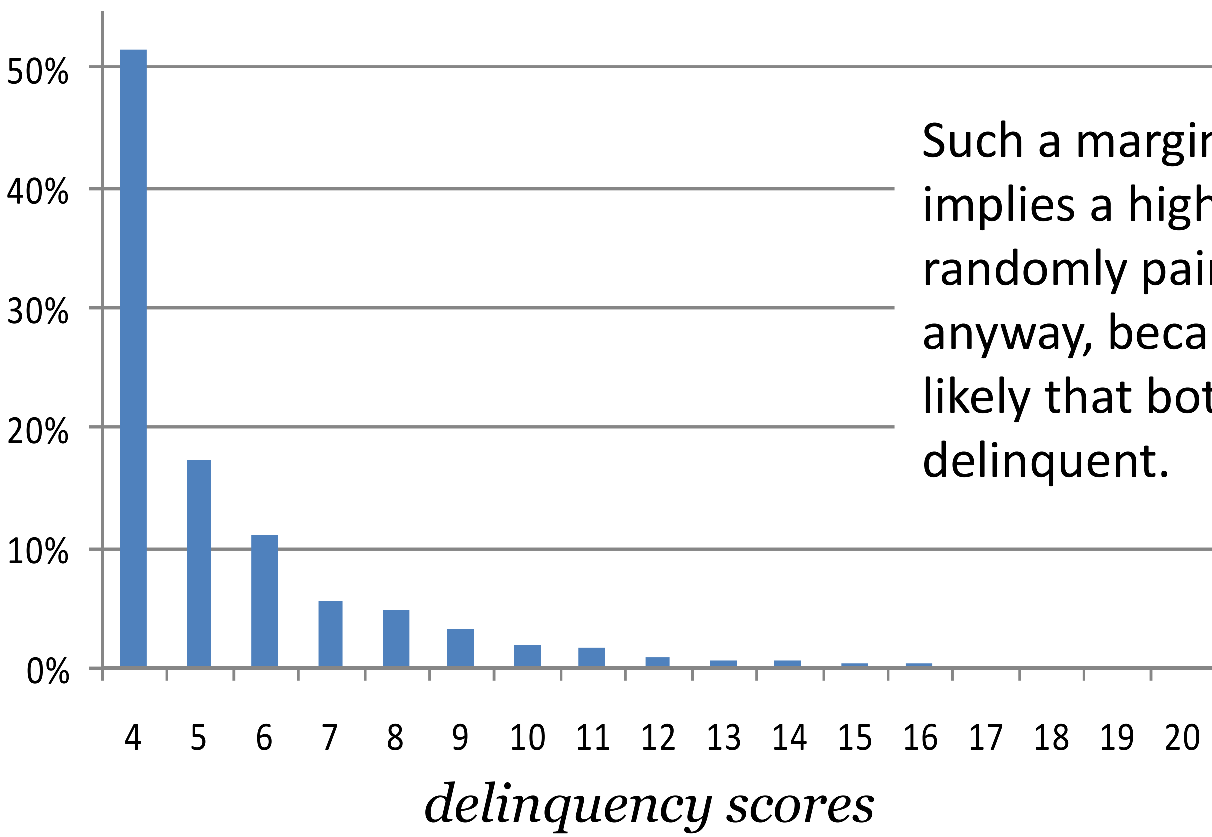
- For strongly skewed variables, peer influence may even coincide with connected people becoming less similar.

  Example: When entering secondary school, students initially are all non-delinquent, i.e., perfectly similar. Any subsequent movement implies a reduction of similarity.

- In such cases, the *similarity based* measures can be *wrong specifications* of peer influence!

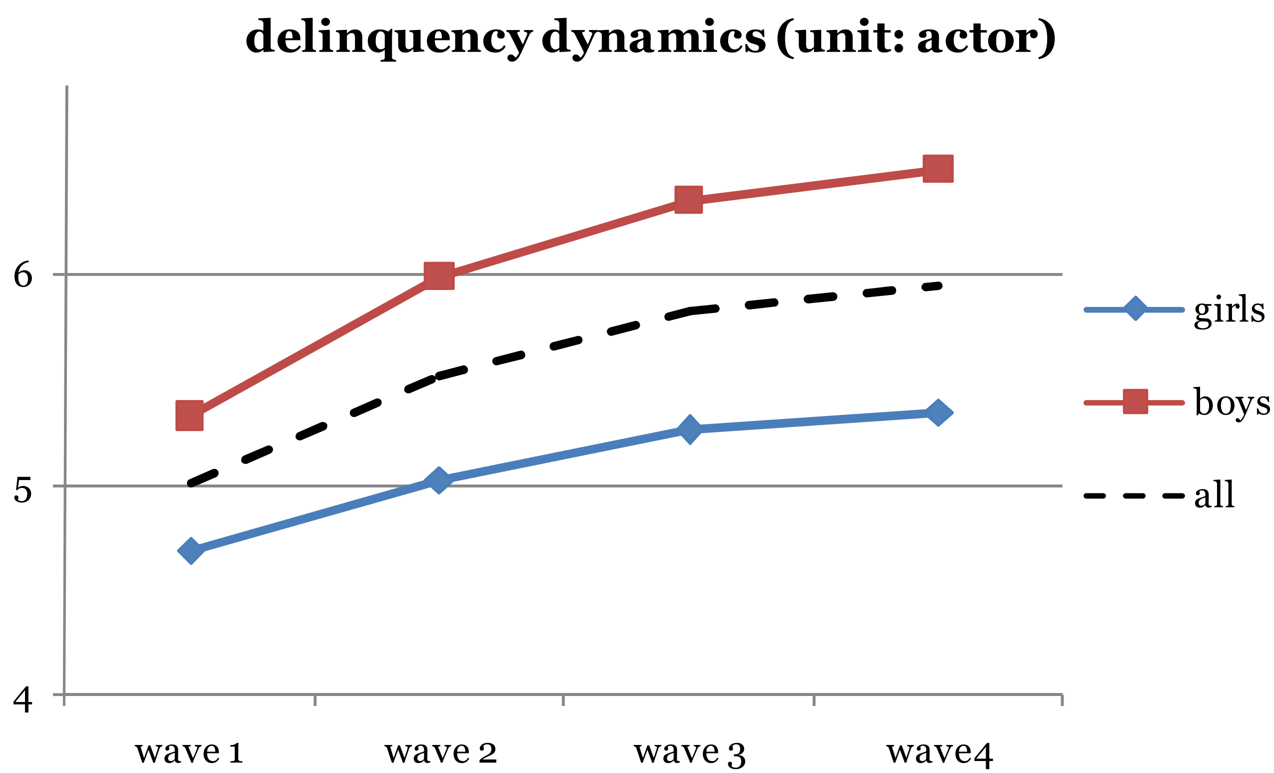
  Correlational measures may be the better choice here; see Knecht et al. (Social Development, 2010) & following slides.

Illustration: a very skewed distribution…



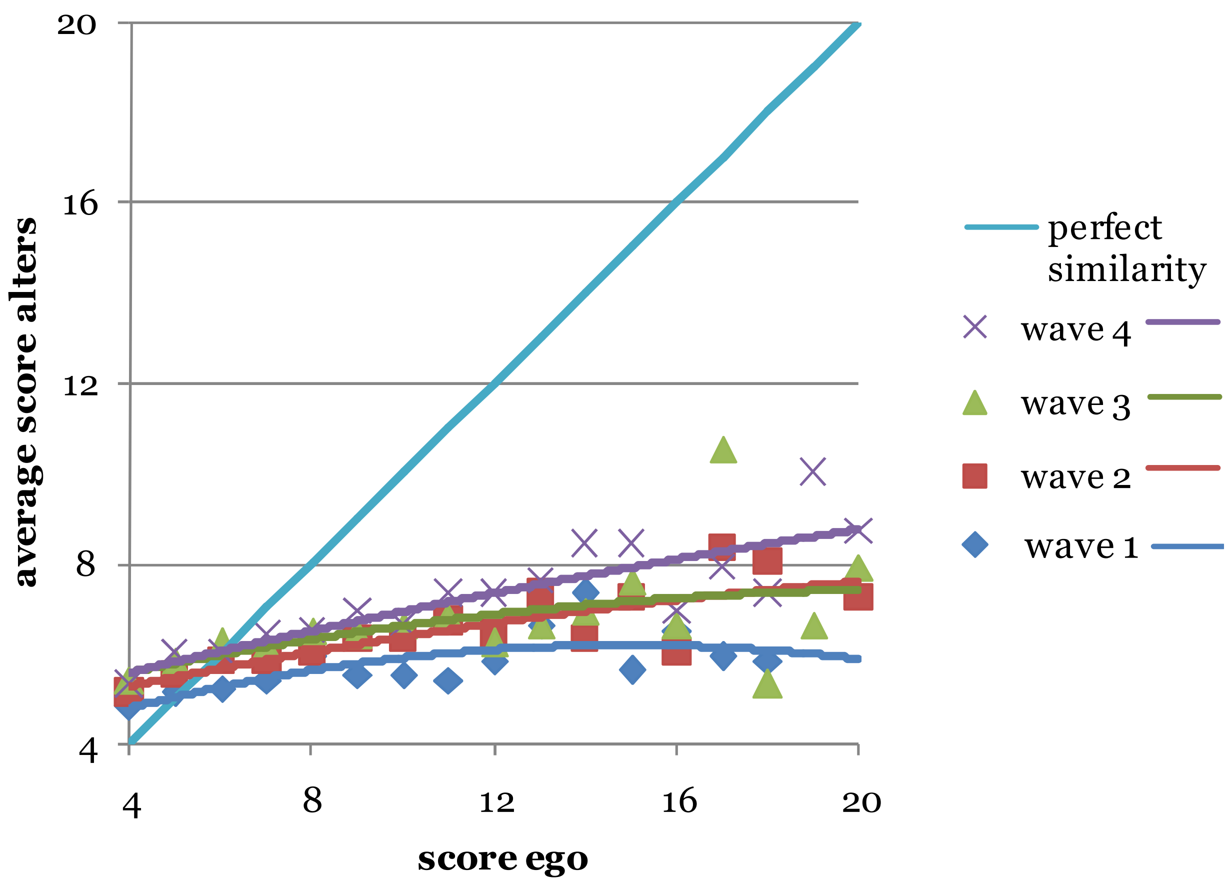
Such a marginal distribution implies a high similarity of randomly paired actors anyway, because it is very likely that both are non-delinquent.

… plus a trend of over time can imply similarity decrease.



If the dynamic process starts with perfect similarity (“nobody delinquent”) it can only get <u>less</u> similar from there on…

Here, a <u>correlational</u> measure for social influence is better operationalisation than a <u>distance-based</u> one.



E.g., use the statistic

$$\sum_j x_{ij} z_i z_j$$

as alternative operationalisation of social influence into the model part expressing behaviour change.

Message of this:

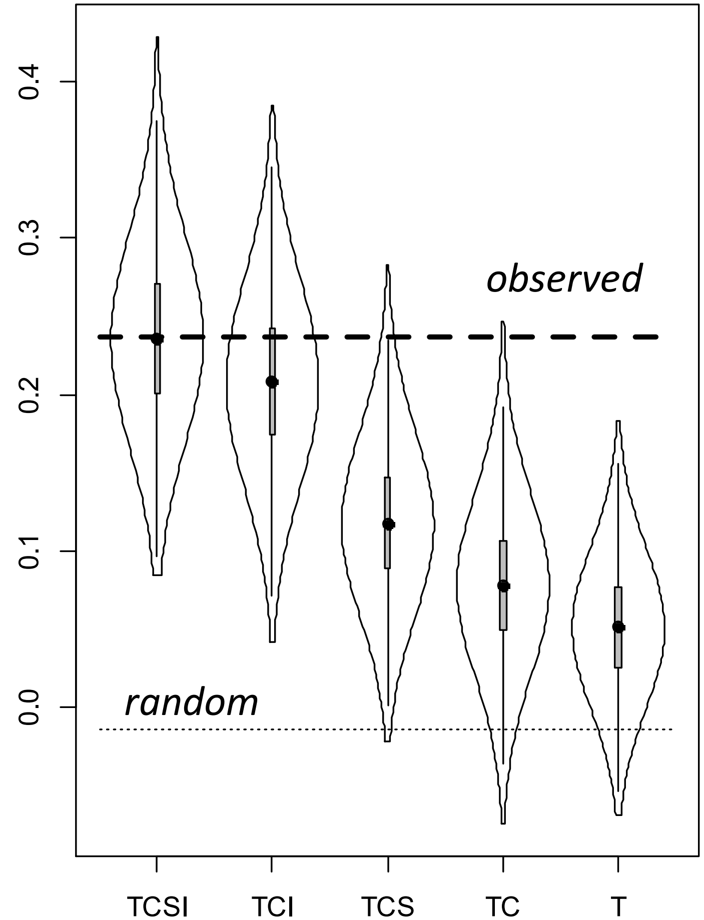
- “Influence” is <u>not</u> unequivocally tied to one specific operationalisation!
- It is <u>not always</u> about “similarity” – sometimes “alignment” / “association” is the better way to phrase it – and sometimes it is a “connectedness” issue.
- <u>Always</u> take a close look at your data set to find out what makes sense in your context.
- In the stochastic actor-based framework, goodness of fit tests (score type) facilitate the technical part of decision making – but doesn’t substitute thought!

(2) Comparison over decision domains?

- To what degree are performance of advice giver and advice recipient associated?
- Indicator Moran's autocorrelation:

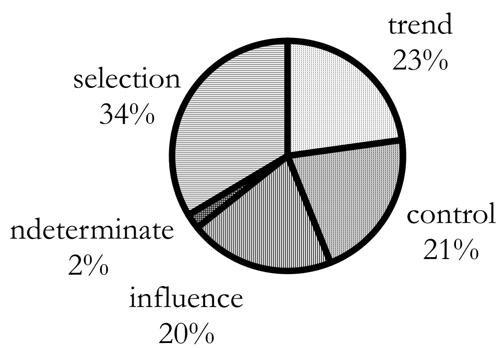
$$I = \frac{n\sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_i (z_i - \overline{z})^2\right)}$$

- Compared are (partially) nested models including these components:
  T rend (rewiring, perf. drift, etc.)
  C ontrol (sex, experience, etc.)
  S election (homophily, etc.)
  I nfluence (assimilation, etc.)

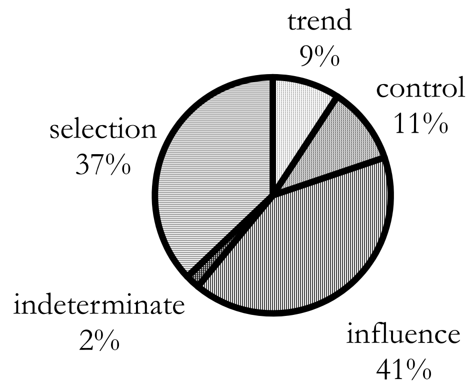


Pie chart diagrams based on these violin plots

The issue of <u>comparing the strength</u> of influence and selection requires a common metric for comparison, e.g. *network autocorrelation coefficients* (here: Moran's I).



SMOKING MODELS



ALCOHOL MODELS

Steglich, Snijders & Pearson, 2010. Dynamic Networks and Behavior: Separating Selection from Influence. *Sociological Methodology* 40: 329-393.

c.e.g.steglich@rug.nl

https://steglich.gmw.rug.nl

LINKÖPING UNIVERSITY

university of groningen