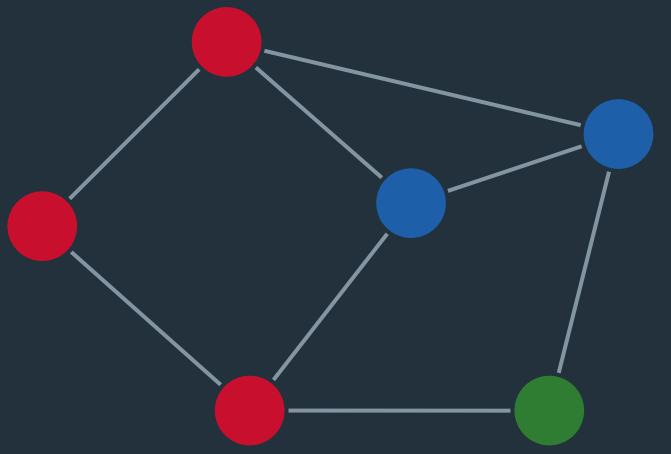
From one network to change between waves

Course C · Day 5, afternoon, option A: panel data

Chapters 14 and 25 · about three hours, with a break

Instructors are free to use, adapt and share these slides.



This afternoon: four blocks, one break

0:00	**1 Change between waves**	50 min
0:50	**2 Modeling the change**	25 min
1:15	Break	15 min
1:30	**3 An SAOM on the Glasgow waves**	70 min
2:40	**4 What the panel shows**	15 min

What the morning established

Read

An ERGM coefficient is a change in the conditional log-odds of one tie, given the rest of the network, read through a change statistic. The advice model: mutual 2.81, same office 2.76, partner is asked 1.49 as odds ratios.

Check

Converged is not a verdict. Read the effective sample sizes after every MCMLE fit, then compare simulated and observed networks on statistics the model does not contain.

Limit

An ERGM describes one snapshot. It cannot say how the network changed, or whether lawyers came to return each other's requests. A panel can.

This afternoon: how much did a network change between waves, and what do actors' choices look like in a model of that change?

Before we start

```
. which nwsaom
. nwwebuse glasgow, nwclear
. nwsummarize glasgow1
```

Check that every participant can run nwsaom. Every fit this afternoon takes seconds, and every one takes seed() so that results repeat.

Glasgow

50 girls; friendship nominations at three waves, about a year apart: glasgow1, glasgow2, glasgow3.

Attributes

smoke1 and sport1 at wave 1; alcohol1, alcohol2, alcohol3 at the three waves.

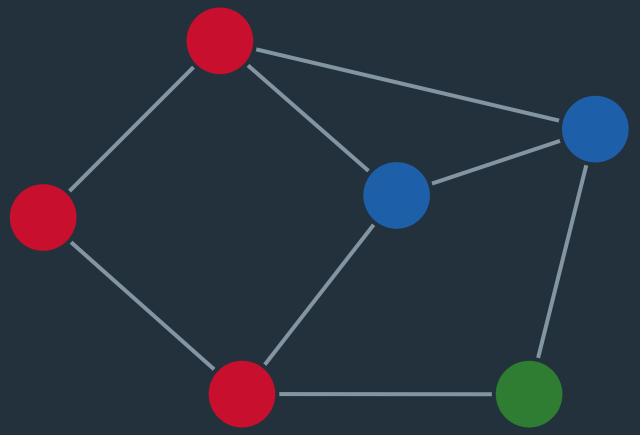
Open the do-files

The book's do-files for Chapters 14 and 25 hold every command.

1

Change between waves

50 min · Chapter 14: nwturnover, formed and dissolved ties



A network observed once is a snapshot of a process

Descriptive question

How much of the network persists, how much is new, how quickly is it changing? The tools of this block answer it.

Explanatory question

Why do ties form or end? That needs assumptions about who chooses whom and when: the models of Part IV.

A wave is one round of data collection; what makes three waves a panel is that they share the same actors. Anything that happens between the waves is invisible: a friendship formed and ended within the year leaves no trace, and a longer interval shows more change.

How much changed between waves 1 and 2

```
. nwwebuse glasgow, nwclear
. nwturnover glasgow1 glasgow2, generate(_turnover)
----------------------------------------
  Networks: glasgow1 -> glasgow2
  Stable: 57  Formed: 59  Dissolved: 56
  Jaccard index: .331
  Persistence: .504
```

0.33

Jaccard index: 57 stable ties of 172 present in either wave

0.50

persistence: 57 of the 113 ties of wave 1

Stable: present in both waves. Formed: absent, then present. Dissolved: present, then absent. The variable _turnover, which generate() requires, holds each girl's own Jaccard index.

The second year looks somewhat more stable

```
. nwturnover glasgow2 glasgow3, generate(_turnover) ///
>     replace
  Networks: glasgow2 -> glasgow3
  Stable: 66  Formed: 56  Dissolved: 50
  Jaccard index: .384
  Persistence: .569
```

Waves	Stable	Formed	Dissolved	Jaccard	Persist.
1 to 2	57	59	56	0.331	0.504
2 to 3	66	56	50	0.384	0.569

Compared with the tie counts

The waves have 113, 116 and 122 ties. Formed minus dissolved is only 3 and 6, yet about half of the ties of one wave are not present in the next: near-constant density hides turnover. Two intervals from one cohort cannot show that the second year is more stable than chance.

What a Jaccard index of 0.33 says

It says
About a third of all friendship nominations present at either wave were present at both.

It does not say
That two thirds of the friendships were lost, or that the data are unreliable. Some turnover is measurement error and some is real change, and the statistics cannot tell which. Nor is it the speed of change: ties formed and dissolved within the interval are never seen.

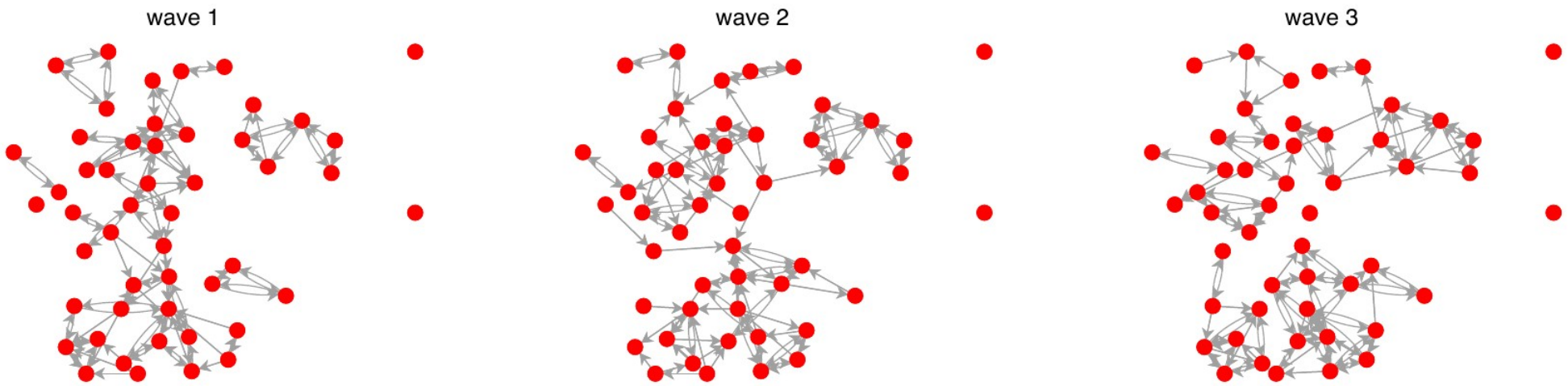
Intervals differ

A Jaccard index after one month and one after two years are not comparable.

Instruments differ

A different ceiling on how many friends may be named produces apparent turnover that is measurement.

Looking at change: every girl at the same position



Layout computed once on the union of the waves: every change from panel to panel is a real tie change (Chapter 14, Figure 15).

```
. nwplot glasgowall, generate(x y)
. nwplot glasgow1, nodexy(x y) title("wave 1") name(g1, replace)
```

Which formed ties return a tie or close a two-path?

```
. nwturnover glasgow1 glasgow2, generate(_turnover) ///
>     replace formation
  Formed ties, compared with all pairs without a tie in glasgow1:
                            Formed    No tie
    Pairs                       59      2337
    Return a tie             0.186     0.015
    Close a two-path         0.153     0.053
```

Share of	1 to 2		2 to 3	
	formed	no tie	formed	no tie
Return a tie	0.186	0.015	0.268	0.020
Close a two-path	0.153	0.053	0.446	0.057

Two mechanisms, descriptively

A formed tie from i to j returns a tie if j had named i in the first wave, and closes a two-path if i had named someone who named j. The comparison group is all pairs without a tie in the first wave: 2,337 and 2,334 pairs. Between waves 1 and 2, 11 of 59 formed ties returned a tie and 9 closed a two-path.

These shares are descriptions, not yet effects

It says

New ties are many times likelier than a random pair without a tie to reciprocate a received tie or to close a two-path, in both intervals.

It does not say

That reciprocity or transitivity drives the formation of friendships. Girls in one class nominate each other and share friends, and the two shares overlap. With 59 and 56 formed ties in one cohort the difference between the intervals may be chance.

What a model adds

Separating the tendencies from one another, holding attributes and the other tendencies fixed, and attaching a rate to the opportunities: the models of Part IV.

Hands-on (5 min)

Ch. 14, Ex. 2

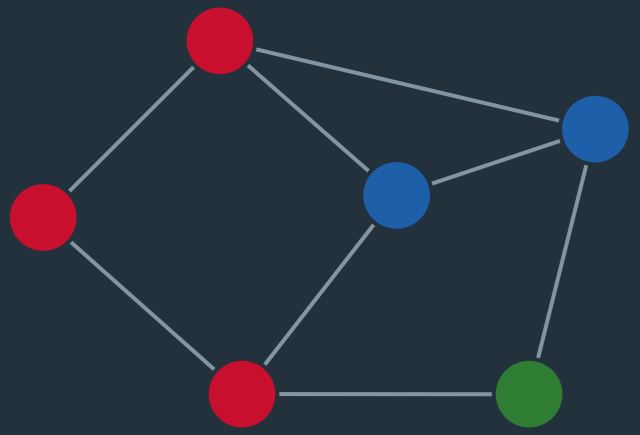
Check by hand the Jaccard index (0.384) and the persistence rate (0.569) for waves 2 to 3 from the counts of 66 stable, 56 formed and 50 dissolved ties.

Hint The Jaccard index is stable divided by stable + formed + dissolved, the ties present in either wave. Persistence is stable divided by stable + dissolved, the ties of wave 2.

2

Modeling the change

25 min · Chapter 25: ministeps, rate, objective function



Four assumptions turn the unseen process into a model

Small steps in continuous time

At any moment at most one tie changes; a larger difference between waves is the accumulation of many ministeps.

Markov

The next ministep depends on the current network and attributes, not on how the network got there.

Actors control outgoing ties

Each ministep is a choice by one actor about one of its own ties: create one, dissolve one, or leave everything as it is.

Myopic choice

An actor compares the networks its options would produce, with an objective function, without planning ahead.

Rate: how often. Objective function: which change

Rate function

Each actor has a clock that ticks at random moments; the rate is the expected number of opportunities to change a tie between two waves. A larger rate means more opportunities, not more consequential ones.

Objective function

What an activated actor does with the opportunity: it evaluates each alternative (add a tie, drop one, change nothing) by a weighted sum of statistics of its own ties.

```
f_i(x) = θ1 s_i1(x) + θ2 s_i2(x) + ...          objective function of actor i
P(actor i chooses x') is proportional to exp{ f_i(x') }          multinomial logit
```

A coefficient is read as a log-odds in the actor's choice: of two alternatives differing by one unit of a statistic, the one with the higher f_i is more likely, by a log-ratio equal to the coefficient.

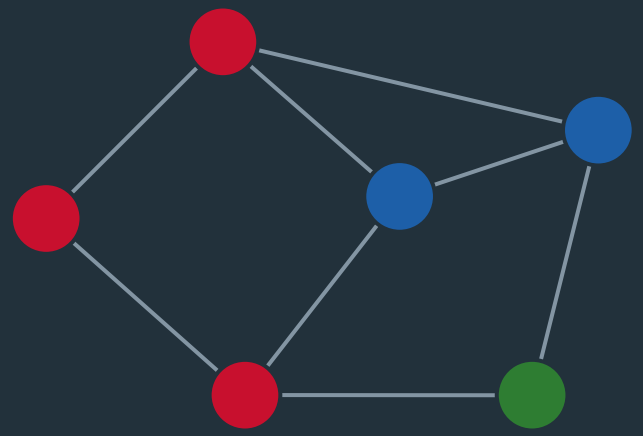
How the SAOM differs from the ERGM

	ERGM (this morning)	SAOM (this afternoon)
Outcome	the whole network, one joint distribution	the change of one actor's outgoing ties, a step at a time
Data	one observed network (standard form)	panel waves of the same actors (standard form)
Statistics	counted across the whole graph	nested in the actor's own ties
Estimation	Monte Carlo maximum likelihood	Method of Moments
Convergence	effective sample size, Geweke	t-ratios of the deviations

Effects with the same name are different statistics: adding the tie i to j raises an ERGM's count of transitive triads for every triad it completes; in a SAOM, i counts only the triads in which i sends two of the ties.

Break

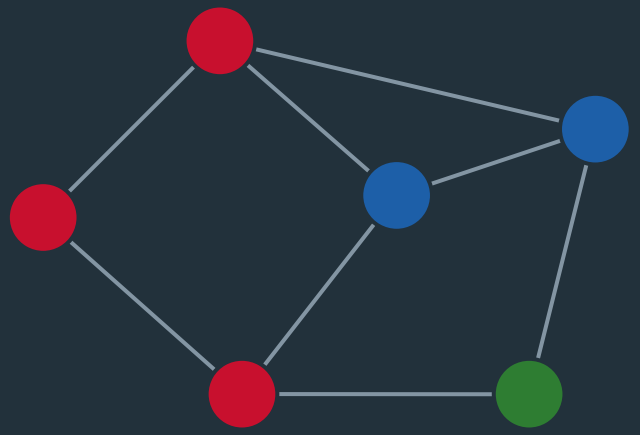
15 minutes



3

An SAOM on the Glasgow waves

70 min · Chapter 25: nwsaom describe, nwsaom, nwsaom refit, estat gof



First, how much change is there to model?

```
. nwsaom describe, waves(glasgow1 glasgow2 glasgow3)
       Wave |       Ties     Density   Avg. degree
          1 |        113      0.0461         2.260
          2 |        116      0.0473         2.320
          3 |        122      0.0498         2.440

     Period |    Kept   Created  Dissolved   Jaccard   Hamming
       1->2 |      57        59         56     0.331       115
       2->3 |      66        56         50     0.384       106
```

Glasgow's two periods are on the lower, but acceptable, side. Too little change leaves nothing to explain; too much makes one continuous process doubtful.

Rule of thumb

A Jaccard index of at least about 0.3. Below that, few ties are stable and estimation tends to be unstable.

Hamming distance

Ties created plus ties dissolved: 59 + 56 = 115 and 56 + 50 = 106.

A first model: outdegree and reciprocity

```
. nwsaom, wave1(glasgow1) wave2(glasgow2) outdegree reciprocity seed(1)
SAOM (Method of Moments, conditional), waves: glasgow1 -> glasgow2
Actors: 50                              Estimated rate:  5.552 (0.856)

        SAOM | Coefficient  Std. err.      z    P>|z|
-------------+----------------------------------------
   outdegree |   -2.26413    .127996   -17.69   0.000
 reciprocity |    2.50883   .2328995    10.77   0.000

Overall maximum convergence ratio: 0.178 (criterion: below 0.25; met)
Convergence t-ratios (criterion: all below 0.1 in absolute value): met
```

5.55

estimated rate: opportunities for change per pupil between the two waves

12

odds in a choice: exp(2.51) for a reciprocated tie

Outdegree (−2.26) works like an intercept: an additional friendship is evaluated unfavorably unless other effects make up for it, which is how the model reproduces a sparse network. Reciprocity (2.51) is strongly positive.

Two different kinds of number below the table

z and P>|z|

An ordinary Wald test of each effect against zero, as in any Stata table.

Convergence lines

Two criteria for the Method-of-Moments search: the overall ratio below 0.25 and every t-ratio below 0.1. Here 0.178: met.

It says

Networks simulated at the estimates reproduce the observed target statistics closely, so the Method of Moments has found a stable solution.

It does not say

That the effects are significant (that is the z column), or that the model fits the data. It says only that the algorithm has settled.

Building up: closure and homophily

```
. nwsaom, wave1(glasgow1) wave2(glasgow2) outdegree ///
>     reciprocity transtrip nodematch(smoke1) seed(12345)
SAOM (Method of Moments, conditional), waves: glasgow1 -> glasgow2
Actors: 50                              Estimated rate:  6.047 (1.003)

            SAOM | Coefficient  Std. err.      z    P>|z|
-----------------+------------------------------------------
       outdegree |  -2.710452   .2222379   -12.20   0.000
     reciprocity |   2.143375   .2537855     8.45   0.000
nodematch_smoke1 |   .2934813   .2079884     1.41   0.158
       transtrip |   .5790748   .1130786     5.12   0.000
```

0.58

transitive triplets: pupils befriend friends of their friends

0.29

same smoking at wave 1: positive, but p = 0.16

Both convergence criteria are met (overall ratio 0.115). transtrip counts closed triads within the pupil's own ties: the actor-oriented counterpart of gwesp, not the same statistic.

Both periods: shared coefficients, one rate each

```
. nwsaom, waves(glasgow1 glasgow2 glasgow3) outdegree ///
>     reciprocity transtrip nodematch(smoke1) seed(12345)
       outdegree |  -2.761655   .1525161   -18.11   0.000
     reciprocity |   2.443945   .1847386    13.23   0.000
nodematch_smoke1 |   .1341893   .1517608     0.88   0.377
       transtrip |   .6221807   .0796816     7.81   0.000
```

Rate parameters	period 1	period 2
rate	6.4343	5.2125
se	1.1104	0.8909

Overall maximum convergence ratio 0.092: met.

The smoking estimate falls

from 0.29 to 0.13 (p from 0.16 to 0.38): more precise, and smaller. The rates differ (6.43, 5.21), but with standard errors near 1 the data do not show that opportunity slowed. Equal coefficients across periods is an assumption of the pooled model.

Restart from the estimates, then test the fit

```
. nwsaom, waves(glasgow1 glasgow2 glasgow3) outdegree ///
>     reciprocity transtrip seed(12345)
. nwsaom refit, seed(12346)
nwsaom refit: 1 run; largest |t| 0.076, overall ratio 0.173 (converged)
. estat gof, nsim(100) seed(1)
           Statistic |   Mahalanobis  P-value      Fit
---------------------+------------------------------
           Outdegree |         9.922    0.380       ok
            Indegree |         8.373    0.480       ok
            Geodesic |        28.119    0.000 rejected
```

A converged first run is checked by a second. Rerun estat gof with another seed() when a p-value is near 0.05.

nwsaom refit

restarts from the last estimates with a new seed: here no estimate moves by more than 0.01.

estat gof

simulates from the observed first wave of each period; the p-value is the share of simulated distributions at least as far from the center.

Reciprocity and closure help, but not enough

Model	Outdegree	Indegree	Geodesic
outdegree only	5.408 (p 0.56)	6.728 (p 0.56)	219.780 (p 0.000)
+ reciprocity, transtrip	9.922 (p 0.38)	8.373 (p 0.48)	28.119 (p 0.000)

The model is useful and incomplete

It reproduces the degrees, yet the geodesic distribution still fails at any conventional level: its reciprocity and closure estimates describe tendencies in a network whose overall shape it does not capture.

What to do

Goodness of fit does not say what to add, but where to look: effects for the longer-range structure, tested with estat score (Chapter 26). Report the limit if the distances still fail.

Hands-on (25 min)

Ch. 25, Ex. 3

Refit the two-wave model (wave1(glasgow1) wave2(glasgow2) outdegree reciprocity transtrip nodematch(smoke1) seed(12345)) with nodematch(sport1) in place of nodematch(smoke1). sport1 records whether a pupil did sport regularly at wave 1. Report the coefficient and p-value of nodematch_sport1 next to those of smoking (0.29, p = 0.16), and say whether either kind of homophily is distinguishable from zero once reciprocity and transitivity are in the model. Then run estat gof, nsim(100) seed(1) on the sport model, and again with seed(2). Does the geodesic distribution fail as clearly as it did for the three-wave model, and what do the two runs together tell you?

Hint Only the variable in nodematch() changes. tabulate sport1 shows the two groups (37 of the 50 pupils do sport regularly); check the two convergence lines before reading the table. The two estat gof runs differ only in the seed, so any difference between them is simulation noise: ask whether it is large enough to change your reading.

Selection or influence? Letting drinking evolve too

```
. nwsaom, waves(glasgow1 glasgow2 glasgow3) outdegree ///
>     reciprocity transtrip behsim ///
>     behavior(alcohol1 alcohol2 alcohol3) ///
>     linear quadratic avalt seed(12345)

       behsim |   1.460717   .6218014     2.35   0.019
   beh_linear |    .3948453   .2135712     1.85   0.064
beh_quadratic |   -.6002813    .320045    -1.88   0.061
    beh_avalt |   1.305018   .7780128     1.68   0.093
```

Network effects omitted; both convergence criteria met (overall ratio 0.138).

A behavior must change often enough: alcohol (five levels) supports the model, smoking (three levels, rarely changing) gives very imprecise estimates.

It says

Pupils prefer friends whose drinking resembles their own (selection, p = 0.019), and pupils whose friends drink more tend to drink more (influence, p = 0.093).

It does not say

That selection is stronger because 1.46 exceeds 1.31: the scales differ. A common cause, such as shared evenings out, could produce both.

What can go wrong

t-ratio or z?

Reading the convergence t-ratios as significance tests, or the table's z as a convergence check. Both sit in the same output.

Any behavior?

Expecting any behavior to identify a co-evolution model. The model runs and prints estimates whatever the behavior does; it must change often enough between waves.

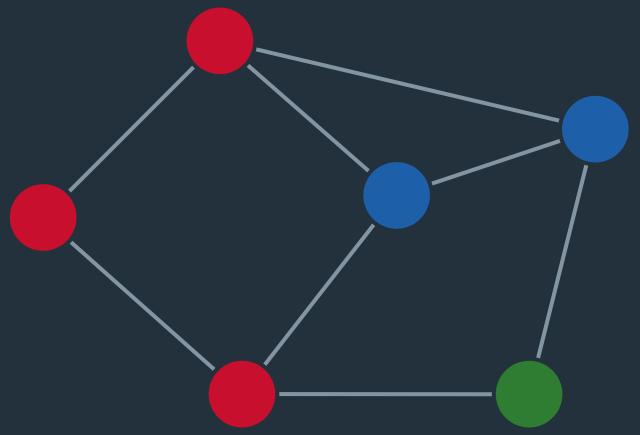
Not a panel

Fitting an SAOM to data that is not a panel. nwsaom needs two or more waves on the same actors, with joiners and leavers declared through present().

4

What the panel shows

15 min · Chapters 14 and 25, with a look at Chapters 26 to 29



What the Glasgow friendships show

0.33

Jaccard index of the first period (0.38 in the second): about two thirds of the ties present in either wave were not present in both

2.44

reciprocity in the pooled SAOM, with transtrip 0.62

0.019

selection p-value; influence is weaker (p = 0.093)

Formation and dissolution roughly cancel, so the number of ties barely moves while the network turns over. Reciprocity and transitivity appear as strong tendencies in how pupils choose and keep friends; the model is useful and incomplete, since the geodesic distances are not reproduced.

Not causal

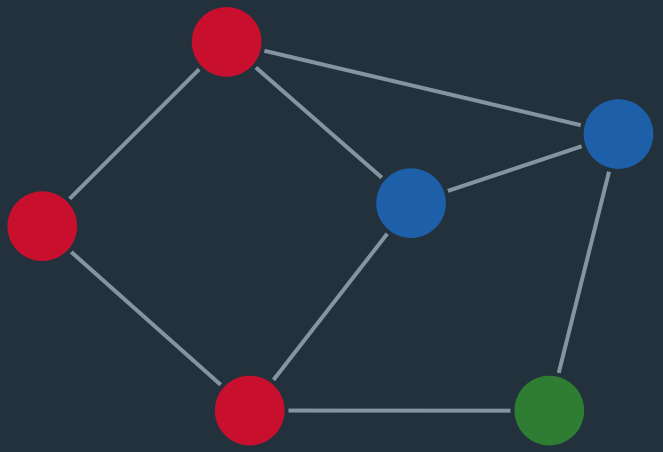
Selection and influence are estimates under the model: an unmodeled common cause could produce both.

The afternoon in five commands

Question	Command	Look at
How much of the network stayed, formed, dissolved?	`nwturnover`	Jaccard, persistence; formation
Is there enough change between the waves to model?	`nwsaom describe`	Jaccard, Hamming, behavior changes
Which tendencies go with tie change?	`nwsaom`	waves(), seed(); rate and coefficients
Has the search settled?	`nwsaom refit`	both convergence lines; stable estimates
Does the model reproduce what it did not fit?	`estat gof`	degrees and distances; nsim(), seed()

Take away. Describe the change before modeling it, check that the search has settled, then check the fit on statistics the model does not contain. An SAOM models actors' choices between waves; it does not show why the choices are what they are.

To read after the session



Chapter 14 · Networks over time · Chapter 25 · Stochastic actor-oriented models

Next: Chapter 26 (testing and extending an SAOM), Chapters 27 and 28 (events)

Exercises not done today: Ch. 14, Ex. 1 and 3; Ch. 25, Ex. 1, 2, 5 and 7

Do-files and teaching material: nwcommands.org/book

Instructors are free to use, adapt and share these slides.