# Beyond Homophily

Tom A.B. Snijders and Alessandro Lomi
using data collected by Vanina Torló

University of Groningen
University of Oxford


February, 2023

Domain of this approach:

$\Rightarrow$ directed networks $X$, where ties $i \rightarrow j$ may be regarded as a 'positive' choice by sender ego of receiver alter;

$\Rightarrow$ actor variables $V$ with range $[V^-, V^+]$ , assumed to have interval scale properties;

$\Rightarrow$ statistical model with a linear predictor for creation / existence / maintenance of ties $x_{ij}$ (for SAOM, the objective function); relevant component of linear predictor depending on $V$ is

$$x_{ij}\, a(v_j \mid v_i) \ .$$

$a(v_j \mid v_i)$ is called the *attraction function*.

Tom A.B. Snijders and Alessandro Lomi (2019). 'Beyond Homophily: Incorporating Actor Variables in Statistical Network Models'. *Network Science*, 7, 1-19.

A basic mechanism for how the existence of ties may be connected to actors' values of $V$ is *homophily*.

But there are more!

A basic mechanism for how the existence of ties
may be connected to actors' values of $V$ is *homophily*.

But there are more!

Four 'mechanisms' by which $V$ might affect ties are considered here:

1. homophily
   (depending on combination of ego's and alter's values $v_i$, $v_j$)
2. aspiration (attraction toward high, or to low, alter's values $v_j$)
3. conformity (attraction toward 'normal' alter's values $v_j$)
4. sociability (depending on ego's value $v_i$).

# Modeling attraction in SAOMs: four parameters

Homophily is usually expressed by attraction function

$$a_1(v_j \mid v_i) = \theta \, | \, v_j - v_i \, |$$

with $\theta < 0$. But this is very restricted!

The four mechanisms jointly can be expressed by

$$\begin{aligned} a_4(v_j \mid v_i) &= \theta_1 \, (v_j - v_i)^2 + \theta_2 \, v_j^2 + \theta_3 \, v_j + \theta_4 \, v_i \\ &\sim \theta_1 \, (v_j - v_i)^2 + \theta_2 \left( v_j + \frac{\theta_3}{2\theta_2} \right)^2 + \theta_4 \, v_i \; . \end{aligned}$$

`shortName:` `diffSqX, altSqX, altX, egoX`

Weight $-\theta_1$ for homophily, $-\theta_2$ for conformity to normative value

$$V^{\text{norm}} = -\frac{\theta_3}{2\,\theta_2} \; ,$$

parameter $\theta_4$ for varying sociability.

# Location of the optimum.

If $\theta_1 + \theta_2 < 0$ this function is unimodal,
with maximum for given $v_i$ attained at $v_j$ given by

$$v_i^{\max}(\theta) = \frac{\theta_1 v_i - \theta_3/2}{\theta_1 + \theta_2} = \frac{\theta_1 v_i + \theta_2 V^{\text{norm}}}{\theta_1 + \theta_2},$$

if $\theta_1 < 0$ and $\theta_2 < 0$ this is a weighted mean of $v_i$ and $V^{\text{norm}}$.

If $V^- \leq V^{\text{norm}} \leq V^+$ we may call the second term conformity,
attraction toward the social norm $V^{\text{norm}}$.

# Aspiration

What is aspiration? Three definitions, from weak to strong:

1. Norm $V^{\text{norm}}$ higher than mean $\bar{V}$.
   For centered $V$ (i.e., $\bar{V} = 0$) with $\theta_2 < 0$, equivalent to $\theta_3 > 0$.
2. Normative contribution
   $$\theta_2 \left( v_j + \frac{\theta_3}{2\theta_2} \right)^2$$
   increasing in $v_j$ throughout $V^- \leq v_j \leq V^+$ .
   If $\theta_2 < 0$, equivalent to $V^{\text{norm}} \geq V^+$.
   If $\theta_2 > 0$, equivalent to $-\theta_3/(2\,\theta_2) \leq V^-$.
3. Aspiration trumps homophily for everybody, i.e.,
   $a(v_j \mid v_i)$ increasing in $v_j$ for all $v_i$. If $\theta_1 < 0,\ \theta_2 < 0$,
   equivalent to $v_i^{\max}(\theta)\Big[v_i = V^-\Big] \geq V^+$,
   and to $V^{\text{norm}} \geq V^+ + \theta_1 (V^+ - V^-)/\theta_2$ .

# Sociability

Tendency toward sociability for an actor $i$ as depending on $v_i$ can be expressed by maximum of attraction function

$$a^{\max}(v_i) = \max_{v_j} a(v_j \mid v_i) \ .$$

When this is increasing in $v_i$,
$V$ may be said to have a positive sociability dimension.

# Full quadratic model

The model may be extended to

$$\theta_1 (v_j - v_i)^2 + \theta_2 \left( v_j + \frac{\theta_3}{2\theta_2} \right)^2 + \theta_4 v_i + \theta_5 v_i^2 \ .$$

`shortName:  diffSqX, altSqX, altX, egoX, egoSqX`

This treats senders and receivers of ties similarly.

Include $\theta_5$ if there are good reasons for it;
these may be theoretical or empirical (fit, significance).

# Summary: four confounded mechanisms / dimensions

$$\theta_1 (v_j - v_i)^2 + \theta_2 \left( v_j + \frac{\theta_3}{2\theta_2} \right)^2 + \theta_4 v_i \qquad \left( + \theta_5 v_i^2 \right)$$

`diffSqX, altSqX, altX, egoX, ( egoSqX)`

1. Test homophily by $\theta_1$ (negative).
2. Test aspiration – weak definition by $\theta_3$ for centered $V$.
3. Test conformity by $\theta_2$ (negative).
4. Express aspiration by checking the three definitions involving $\theta_3$, $\theta_2$, and the distribution of $V$.
   Note that aspiration is a special case of conformity: all agree that high $v_j$ values are desirable.
5. Express sociability by looking at the function $a^{\max}(v_i)$, to which $\theta_4$ and $\theta_5$ have important contributions.

# Recommendation

For interval-scale actor variables $V$ that are expected
to have important effects on the network:

Unless you are a priori quite sure that only one of the mechanisms
homophily / aspiration / conformity / sociability plays a role,
try out the five-parameter model,
drop $\theta_5$ (egoSqX) if not significant and uninteresting;

at the very least, keep this model in the back of your mind
in case results are unexpected.

# Example: Vanina Torló's students

International MBA program in Italy;

75 students; 3 waves distributed over one year.

Dependent network:

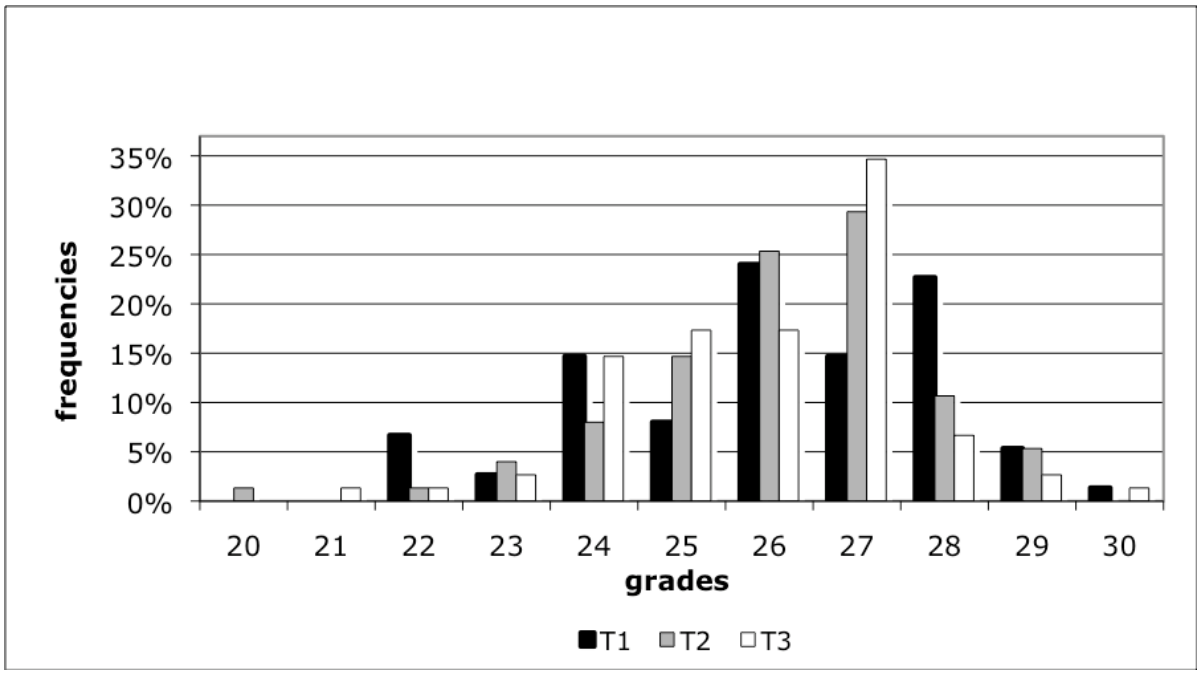
1. *Advice*: whom do you regularly consult for help and support on program-related tasks

Relevant covariate: *Achievement*: average exam grades.

See various publications by Vanina Torló, Alessandro Lomi, and co-authors.

# Descriptives

| | Advice | | |
|---|---|---|---|
| | $T_1$ | $T_2$ | $T_3$ |
| Av. degree | 4.1 | 4.9 | 4.5 |
| Reciprocity | 0.29 | 0.33 | 0.33 |
| Transitivity | 0.24 | 0.24 | 0.26 |



# Results

| Effect | par. | (s.e.) |
|---|---|---|
| outdegree | –2.181*** | (0.208) |
| reciprocity | 1.606*** | (0.197) |
| transitivity gwesp | 1.307*** | (0.121) |
| reciprocity × transitivity gwesp | –0.314 | (0.250) |
| indegree – popularity | 0.0253** | (0.0089) |
| outdegree – popularity | –0.101** | (0.033) |
| outdegree – activity | –0.0072 | (0.0092) |
| gender alter (M) | 0.027 | (0.098) |
| gender ego (M) | –0.239* | (0.100) |
| same gender | 0.130 | (0.092) |
| same nationality | 0.405*** | (0.122) |

(... to be continued ...)

# Results (continued)

| Effect | par. | (s.e.) |
|---|---|---|
| $\hat{\theta}^{\mathrm{g}}_1$ (grades ego minus alter) squared | –0.0288$^{***}$ | (0.0073) |
| $\hat{\theta}^{\mathrm{g}}_2$ grades squared alter | –0.003 | (0.012) |
| $\hat{\theta}^{\mathrm{g}}_3$ grades alter | 0.044 | (0.032) |
| $\hat{\theta}^{\mathrm{g}}_4$ grades ego | –0.095$^{**}$ | (0.031) |
| $\hat{\theta}^{\mathrm{g}}_5$ grades squared ego | 0.026$^{\dagger}$ | (0.010) |
| $\hat{\theta}^{\mathrm{a}}_1$ (age ego minus alter) squared | –0.0014 | (0.0023) |
| $\hat{\theta}^{\mathrm{a}}_2$ age squared alter | –0.0070 | (0.0045) |
| $\hat{\theta}^{\mathrm{a}}_3$ age alter | 0.039$^{*}$ | (0.019) |
| $\hat{\theta}^{\mathrm{a}}_4$ age ego | 0.038$^{*}$ | (0.018) |
| $\hat{\theta}^{\mathrm{a}}_5$ age squared ego | –0.0071$^{\dagger}$ | (0.0041) |

# Attraction function for grades: 'selection table'

The effect of grades on advice is important:

joint Wald test: $\chi^2_5 = 23.3,\ p < 0.001.$

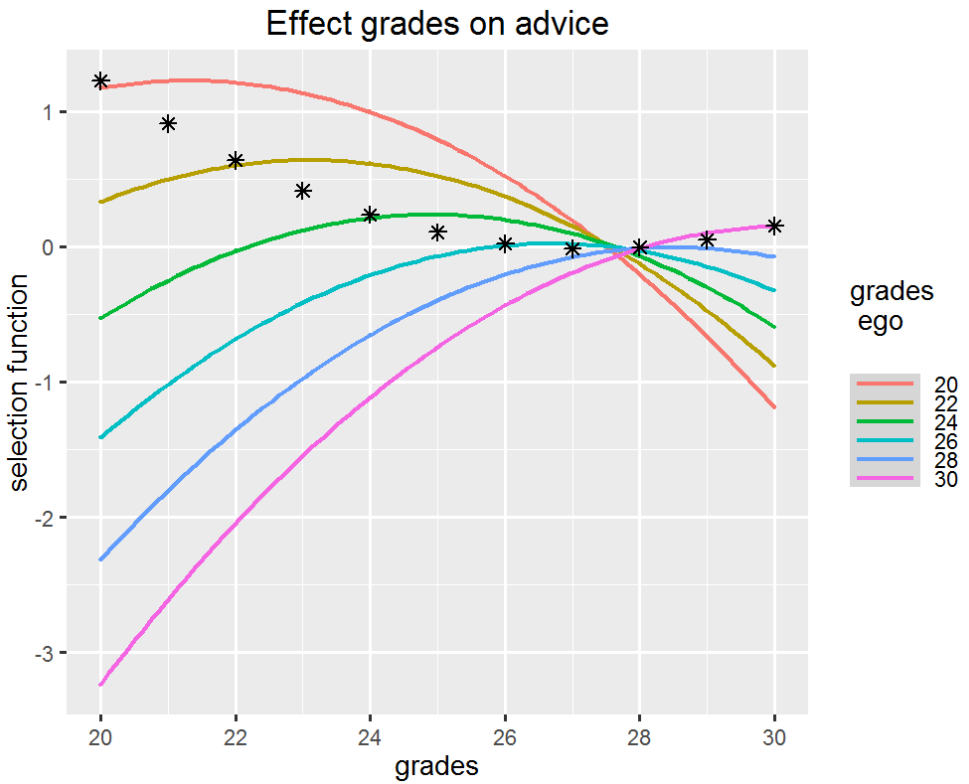
The effect of age on friendship also is significant, but less strongly:

joint Wald test: $\chi^2_5 = 11.9,\ p < 0.05.$

Attraction function of advice based on grade:

$$-0.0288\,(v_i - v_j)^2 - 0.003\,v_j^2 + 0.044\,v_j - 0.095\,v_i + 0.026\,v_i^2$$

where $v = \text{grade} - 26.1$.



# Interpretation of effects grades on advice

Mechanisms: homophily $\hat{\theta}_1 < 0$, conformity $\hat{\theta}_2 < 0$ (n.s.),
aspiration (definition 1) $\hat{\theta}_3 > 0$ (n.s.).

As a whole, the 4-parameter model fits better
than the homophily-only model.
Interpretation can only be done with the plot combining the 5 parameters.

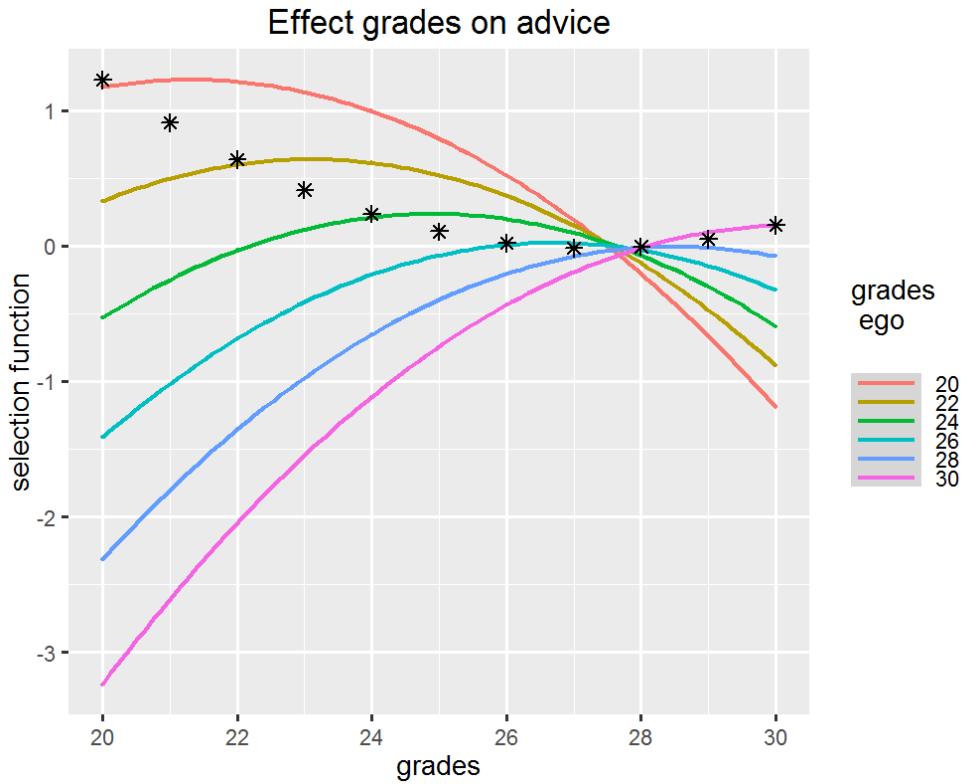
The estimated social norm is

$$V^{\text{norm}} = -\frac{\hat{\theta}_3}{2\hat{\theta}_2} = 6.9 \ ,$$

higher than the maximum of $V$.
So the second definition of aspiration is satisfied;
but without significance.

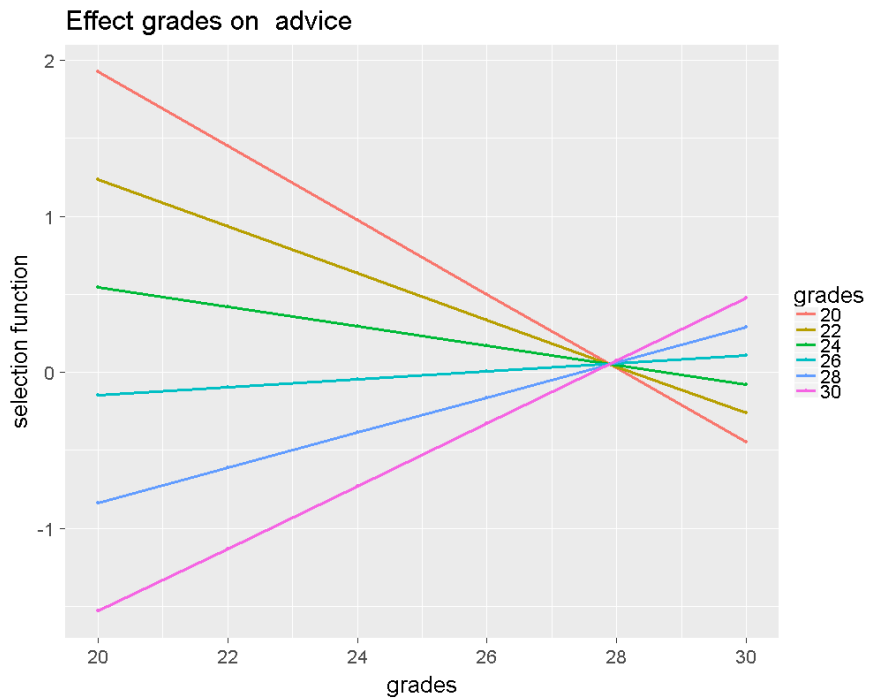
The weight for conformity is $\hat{\theta}_2/(\hat{\theta}_1 + \hat{\theta}_2) = 0.1$ and 0.9 for homophily:
homophily dominates conformity.

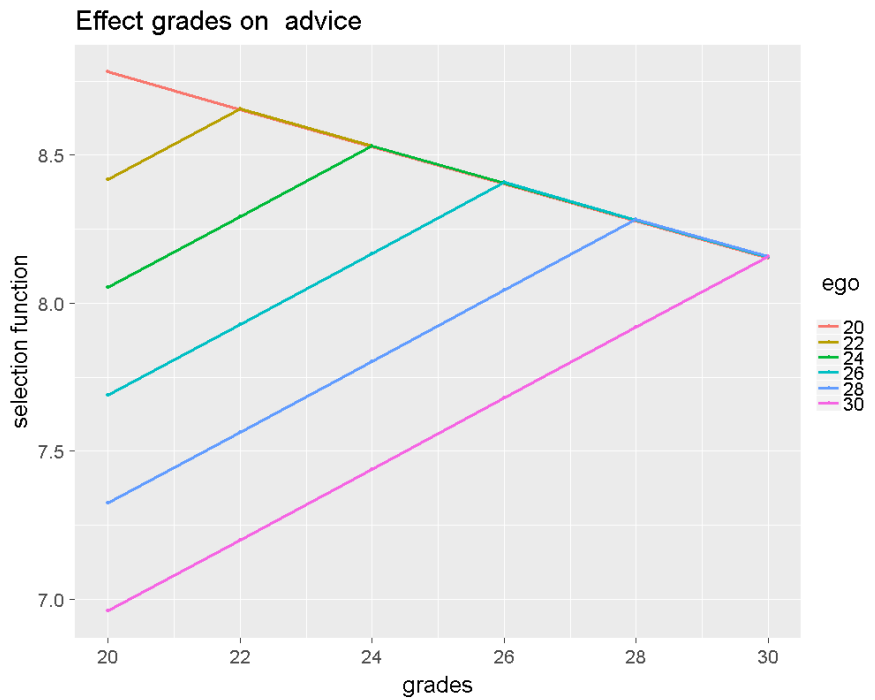
Sociability, represented by the optimum of the attraction function, is plotted by the dashed line in the figure. It is decreasing for the lower half of the range of grades, and approximately constant for the upper half. Although not decreasing uniformly, this nevertheless suggests a weakly negative sociability aspect for grades: higher grades $\Rightarrow$ less advice asking; — but weaker even than the weak definition.



# Comparison with other attraction functions

Comparisons show that the model with the usual specifications of the effects of grades on advice (main effects ego and alter, together with absolute difference ['similarity'] or ego $\times$ alter interaction) fit worse than the quadratic model.




Left: linear interaction specification; right: similarity specification.

# Discussion

- Tendencies toward homophily may and often will coexist
  with aspirations toward high (low) values
  and with tendencies toward a common (average) value.
- Specifications that have been in use in statistical network modeling
  for expressing homophily on numerical actor covariates
  are, in isolation, too rigid (absolute or squared difference)
  or no good representations of homophily (ego $\times$ alter interaction).
  Briefly, they are often misspecifications.
- A quadratic specification of the dependence on ego's and alter's value
  offers an important improvement.
  This four- or five-parameter model yields richer interpretation,
  and estimates each of the mechanisms while controllling for the other
  three.

# Discussion – continued

- Effects `egoSqX`, `altSqX`, `diffSqX` implemented in RSiena.
- Using 4 or 5 parameters seems like a lot, but is more meaningful for theoretical modeling than absolute difference or cross-product interaction.
- For actor variables with moderate effects on the network, the extra parameters (quadratic effects) make little difference, and simpler models may still be used.
- For actor variables with strong effects on the network, the extra parameters (quadratic effects) may make a large difference.

# Discussion – continued

- Homophily component can be tested by
  the ego-alter squared difference interaction effect
  in this four- or five-parameter model (‘three or four control effects’).
- Aspiration component can be tested by the alter effect
  if the variable is centered.
- This model can resolve interpretation problems in cases where
  ‘negative homophily’ (repulsion) seemed to be found
  when using an absolute difference or simple interaction model.
- The value for alter where the optimum is assumed
  (dependent on ego’s value) is an interesting descriptive feature.
- The attribute effects are net of further effects included in the model,
  and these may be correlated in complex ways with attribute effects.
  ‘Attraction function’ is not directly a kind of preference function!