# Some effects for categorical covariates

Tom A.B. Snijders

University of Oxford
University of Groningen

June 2025

# Overview

1. Effects for categorical covariates

# 1. Effects for categorical covariates

This is part of my presentation at the 2025 Sunbelt in Paris,
modified to be independently understood.

Tom A.B. Snijders and Beata Łopaciuk-Gonczaryk (2025).
Double agency and co-evolution for two-mode networks, with an application to corporate interlocks and firms' environmental performance. *Social Networks*, 83, 92-104.

# Expressions reflecting heterogeneity between nodes

Especially in two-mode networks — but also more generally, there can be important differences between ‘kinds’ of nodes.

*‘Senders differ in the kind of receivers they choose.’*

*‘Receivers differ in the kind of senders choosing them.’*

*‘Different kinds of receivers are chosen by different kinds of senders.’*

Such expectations can be implemented in models when ‘kinds’ are operationalized by categorical variables.

Consider categorical nodal covariates $U,\ V$,
with values $0, 1, 2, \ldots$ ,
which divide the node set into several groups.

*especially helpful for more than two groups*

Suppose that $U$ is for the first node set (senders),
and $V$ is for the second node set (receivers).

*Node sets may be the same: one-mode.*

For two-mode networks especially, such category systems can be important.

The following pages give formulas for effects for actor $i$ in network $X$.

# *‘senders differ in the kinds of receivers they choose.’*

*[outAct]* restricted to ties to nodes of the same $V$ :

*outdegree activity to homogeneous covariate*

(internal effect parameter $p = 3$)

*[homXOutAct2]* defined by the sum of $i$’s outgoing ties
weighted by the proportion of $i$’s outgoing ties
to nodes with the same covariate value (as receiver $j$),

$$s_i(x) \;=\; \sum_j x_{ij} \, \frac{\sum_h x_{ih} \; I\{v_h = v_j > 0\}}{\sum_h \; I\{v_h = v_j > 0\}} \;.$$

(The restriction $v_j > 0$ is to give a possibility for handling missing data, or irrelevant categories.)

$\beta_k$ > 0: actors specialize in specific $V$-categories;
$\beta_k$ < 0: actors do not want many ties to same $V$-category.

## *'receivers differ in the kinds of senders choosing them.'*

*[inPop]* restricted to ties from actors of the same *U* :

*indegree popularity from same covariate*
(internal effect parameter $p = 3$)

*[sameXInPop]* defined by the sum over $j$ of outgoing ties $i \to j$
weighted by the proportion of the incoming ties of $j$
from actors with the same covariate value (as sender $i$),

$$s_i(x) = \sum_j x_{ij} \frac{\sum_h x_{hj} \; I\{u_h = u_i\}}{\sum_h I\{u_h = u_i\}} .$$

$\beta_k > 0$: actors follow others' choices in same *U*-category;
$\beta_k < 0$: actors repelled by others' choices in same *U*-category.

*'Different kinds of receivers chosen by different kinds of senders.'*

*indegree popularity from same cov1 to same cov2*, *[(sameXVInPop)]*
(name 'popularity' is not very appropriate....)
(internal effect parameter $p = 3$)
defined by the sum over $j$ of ties $i \to j$
weighted by the density of the $(u_i,\ v_j)$ block,

$$s_i(x) = \sum_j x_{ij} \frac{\sum_{h,k} x_{hk}\ I\{u_h = u_i > 0,\ v_k = v_j > 0\}}{\sum_{h,k} I\{u_h = u_i > 0,\ v_k = v_j > 0\}}\ .$$

This is similar to a pre-specified stochastic block model.
It is an elementary effect.
(The restriction $u_i > 0$, $v_j > 0$ is to give a possibility for handling missing data, or irrelevant categories.)

$\beta_k$ > 0: actors adapt to the current block structure;
$\beta_k$ < 0: actors equalize the current block structure.

Since *[(sameXVInPop)]* is an elementary effect,
it really should be indicated as

$$e^{\mathrm{el}}_{kij}(x) = \frac{\sum_{h,k} x_{hk}\, I\{u_h = u_i > 0,\ v_k = v_j > 0\}}{\sum_{h,k} I\{u_h = u_i > 0,\ v_k = v_j > 0\}} \ldots$$

It might be regarded as a contextual effect,
the context being given by the current $(U, V)$ block structure.

The motivation to have these effects
is the ambiguity that sometimes occurs in the delineation
of the second mode of a two-mode network,
which may lead to quite heterogeneous node sets.

But they may be useful quite generally.