Modelling Variations in Social Connections

Mark Tranmer
CCSR
School of Social Sciences
University of Manchester.
(and Honorary at The University of Stirling)

Is Britain Pulling Apart?
The Royal Statistical Society, London, May 13th, 2013.

Multilevel Models

- Multilevel Models have been used to investigate variations in a response of interest for various levels or classifications of a population structure.
- Pupils in schools: variations in exam scores between pupils, between classes & between schools.
- Individuals in areas: e.g. variations in health.
- Consequences of ignoring a level (Tranmer and Steel, 2001).

Multilevel Models for Social Network Analysis

- If ego-nets, where an individual nominates her friends (alters) have minimal overlap, can regard ego-net as two-level structure in multilevel model.
- Here unit of analysis is **tie** between ego and alter, for which a value can be attached.
- e.g. a tie to a person with a particular characteristic; number of people for which ties exist between two occupational groups.
- Snijders et al. (1995) used a multilevel model to investigate the nature of valued ties between ego and alter for cocaine users in Rotterdam.
- de Miguel Luken and Tranmer (2010) fitted multilevel logistic regression models to data on support networks of recent immigrants to Spain.

Ego-net as a Two-Level Structure

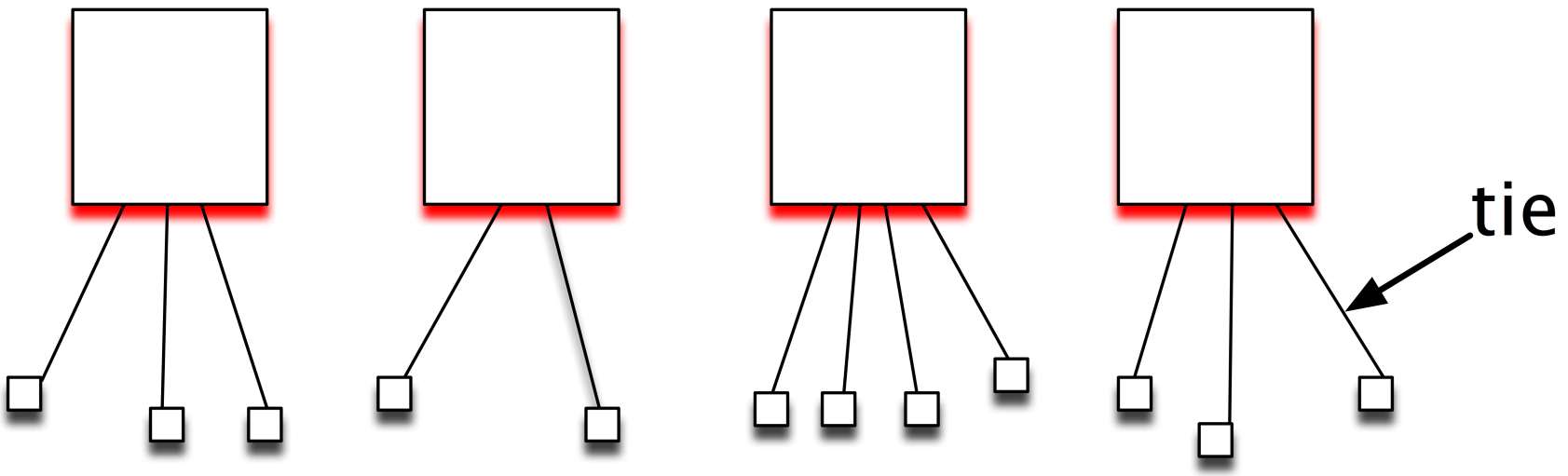


Figure: ego-net as a two-level structure

Ego-net as a Two-Level Structure

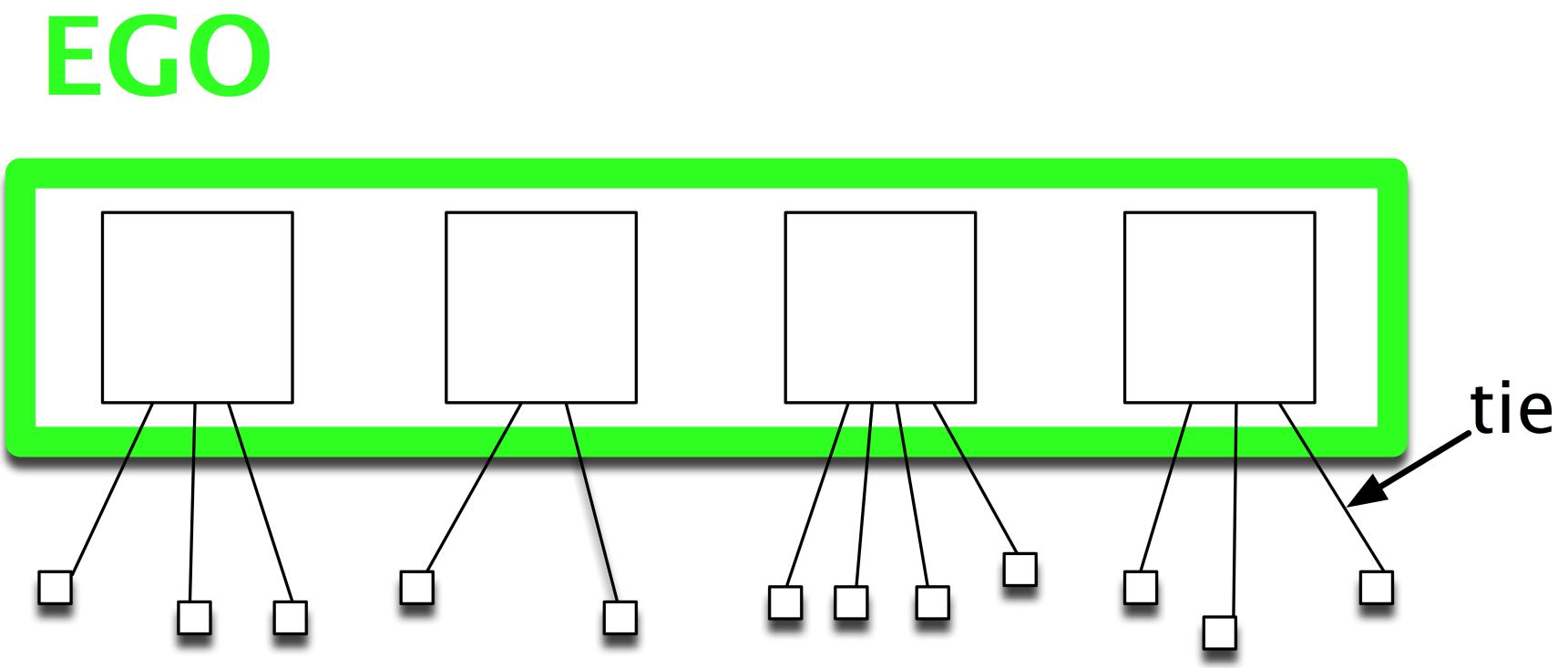
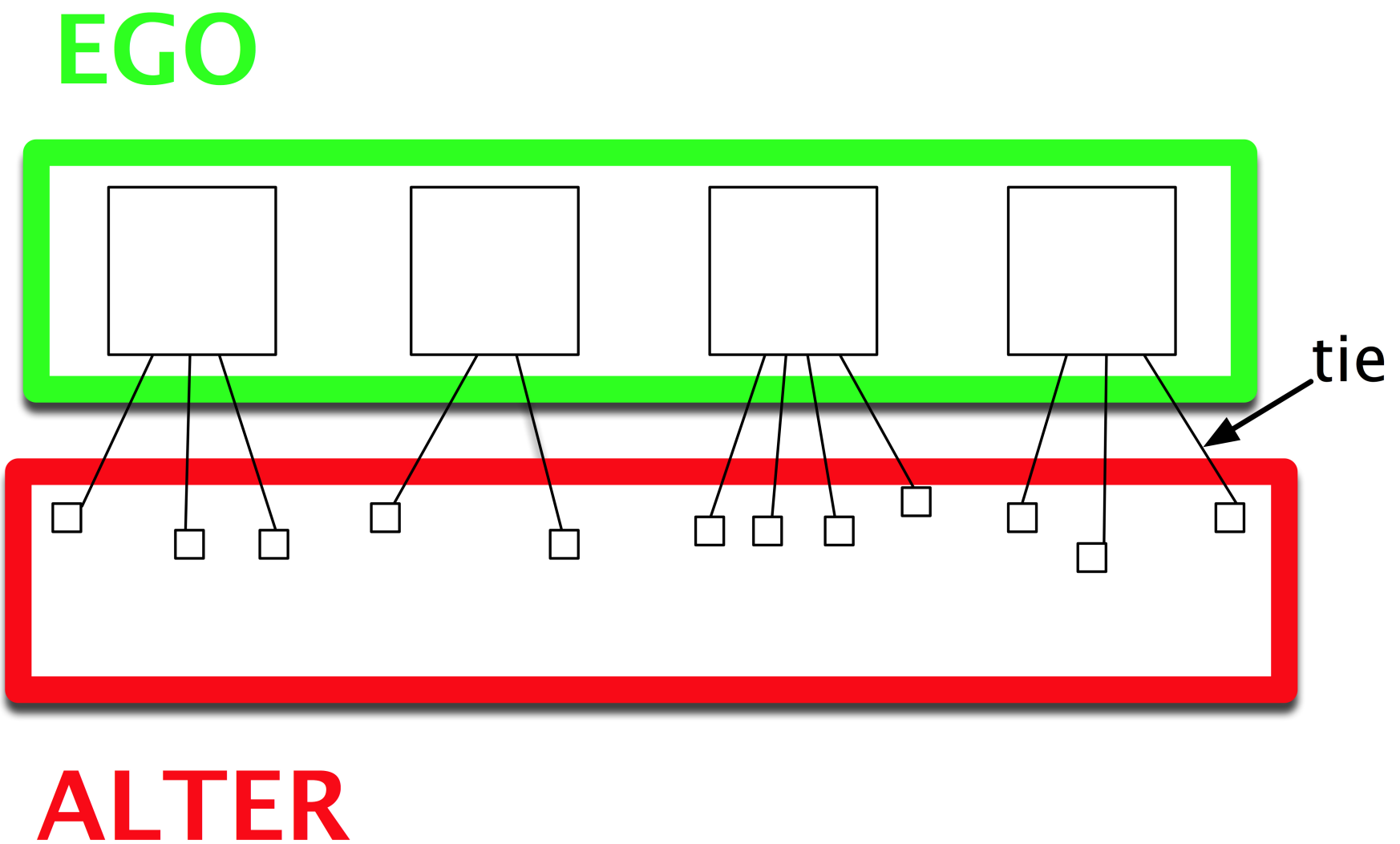
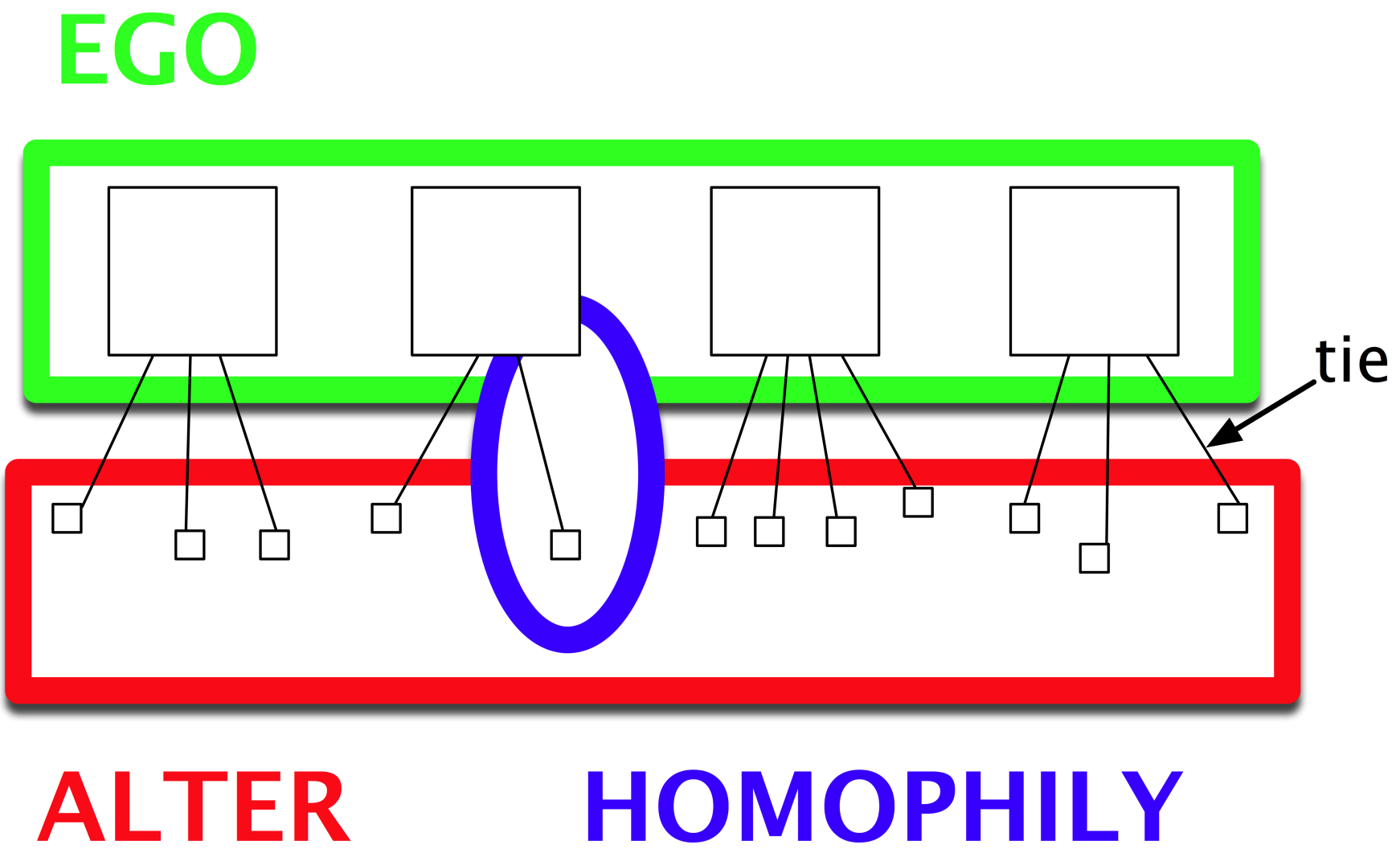


Figure: ego-net as a two-level structure

Ego-net as a Two-Level Structure



Ego-net as a Two-Level Structure



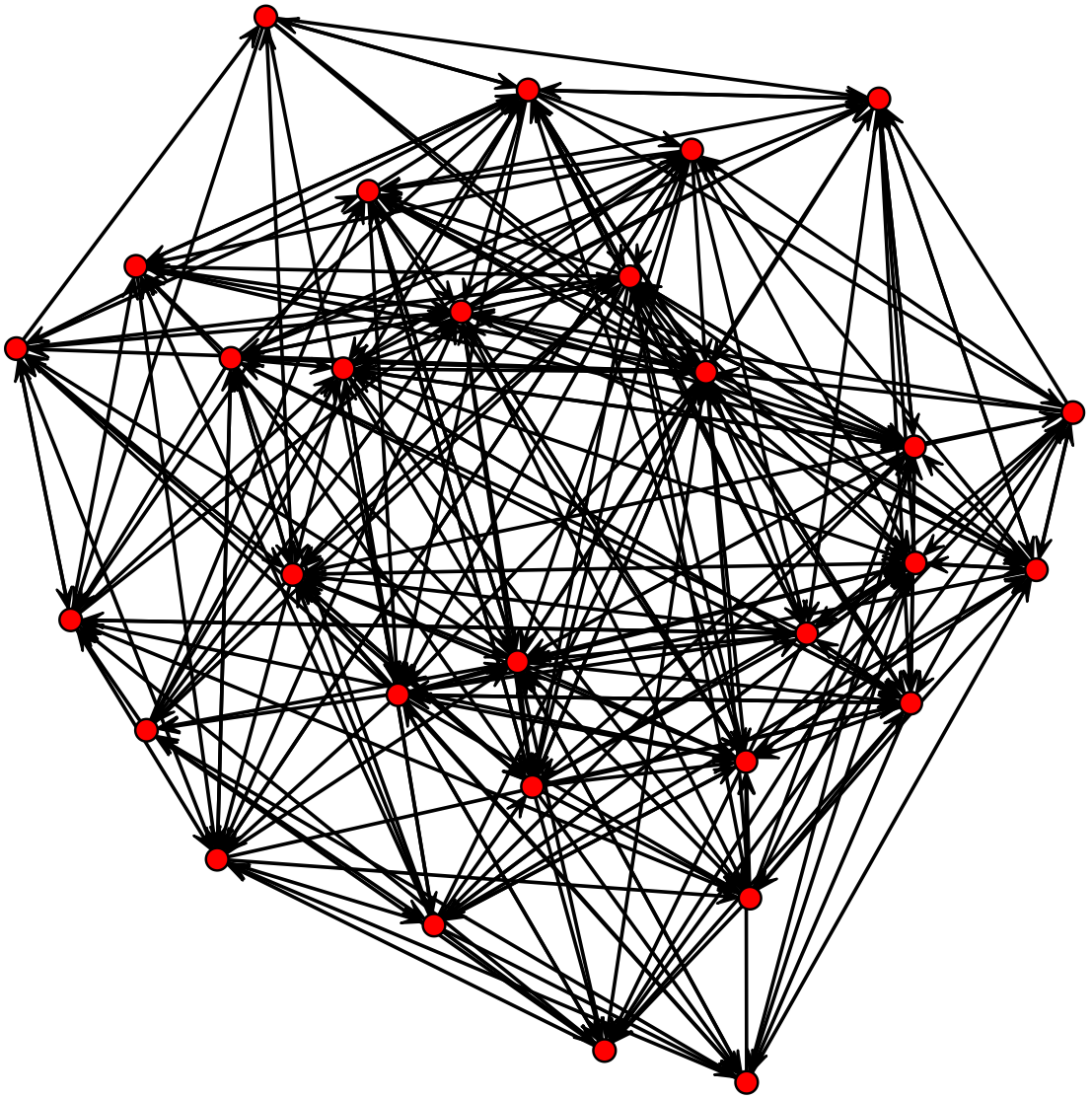
Multilevel Models for Social Network Analysis

- Some people have looked at social network dependencies in the context of other groups.
- Koskinen and Stenberg (2012) devised some methods for friendship dependencies and applied these to data from Swedish schools.
- Tranmer, Steel & Browne (forthcoming JRSS(A) 2014) looked at variation in exam scores at the individual, network, school, and area levels/classification using Multiple Membership Multiple Classification (MMMC) Models.
- Such models **do** allow for overlapping networks: individuals are members of one or more ego-nets: the y variable here is the exam score for a pupil (not the tie between pupils; though peer effects are taken into account).

Modelling Variations in Social Connections

- I think the next challenge is to develop and apply a hybrid approach of modelling **ties** for overlapping ego-nets using **multiple** membership models to investigate variations in social connections.
- Rather than making the nodes in the network people, these could be occupations, etc.
- The response could be the value or number of people tied between two occupations.
- There will be overlap here: doctors have partners that are also doctors, are nurses, are dentists; dentists have partners who are dentists, nurses, doctors.
- Where is the variation in the distribution of social ties - within occupational groups, or between them?

Simulated Directed Network with 30 Nodes, 264 ties.



Simulation Study

- Simulated response based on network disturbance model (Leenders, 2002) with $\rho = 0.4$ and single β with value 1. **NET** is the network of connections.

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\beta + \epsilon \\ \epsilon &= \rho\mathbf{NET}\epsilon + \xi \\ \xi &\sim N(0, \sigma_\xi^2 \mathbf{I}_n) \end{aligned} \tag{1}$$

- Generated 264 values - the difference between the responses on each connected pair of the 30 nodes. these give valued ties, which we would expect to have a mean of zero.
- Model these using MMMC model (MCMC estimation).

Results

	Overlap		No overlap		Single level	
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
$\hat{\beta}_0$	-.065	.550	.109	.214	.009	.094
DIC	908		821		972	
	Estimate	%	Estimate	%	Estimate	%
ego var.	8.602		1.215		-	-
av. ego wt.	.129		1		-	-
adj. ego var.	1.113	39.8	1.215	50.1	-	0
alter variance	1.685	60.2	1.188	49.9	2.317	100

Mean weight in overlap model .129 : people tend to nominate around 8 or 9 friends.

Conclusion

- Multilevel models can be used to investigate variations in social connections.
- The model that has ties as the unit of analysis and allows for overlap through the multiple membership approach allows us to model counts of connections between paris of nodes, or sums or differences in scores for these pairs.
- One consideration: does the data allow us to get the full *potential* network, rather than model the *observed* network?
- If we can also model the zeroes (that are not structural zeroes) we can investigate combinations of social connection where ties do not occur to contrast with combinations where they do.

Thank You

- Thank you for listening!

- *Note: Some of the models I have discussed are defined algebraically below, and the references are given below.*

Two-level model for ego-nets [null model] - y variable is tie between ego-alter

$$y_{ij} = \beta_0 + u_j + e_{ij}$$

$$E[u_j] = E[e_{ij}] = 0$$

$$var(u_j) = \sigma_u^2; var(e_{ij}) = \sigma_e^2; cov(u_j, e_{ij}) = 0$$

$$\text{Hence } var(y_{ij}) = \sigma^2 = \sigma_u^2 + \sigma_e^2$$

i is the alter index

j is the ego index

$$\text{Intra-ego correlation } \sigma_u^2/(\sigma_u^2 + \sigma_e^2) = \sigma_u^2/\sigma^2 \quad (2)$$

MMMC model for ego nets and groups (e.g. schools). y variable is value at node e.g. exam score of pupil.

$$y_i = \mathbf{x}_i'\beta + u^{(2)}_{school(i)} + \sum_{j \in egonet(i)} w_{1i,j} u^{(1)}_j + e_i$$

$$i = 1, ..., n \quad egonet(i) \subset (i, ..., J_1)$$

$$u^{(2)}_{school(i)} \sim N(0, \sigma^2_{u^{(2)}}), \quad u^{(1)}_j \sim N(0, \sigma^2_{u^{(1)}}), \quad e_i \sim N(0, \sigma^2_e) \quad (3)$$

References : I

de Miguel Luken, V., Tranmer, M., Oct. 2010. Personal support networks of immigrants to Spain: A multilevel analysis. Social Networks 32 (4), 253–262.

Koskinen, J., Stenberg, S.-a., Aug. 2012. Bayesian Analysis of Multilevel Probit Models for Data With Friendship Dependencies. Journal of Educational and Behavioral Statistics 37 (2).
URL http://jeb.sagepub.com/cgi/doi/10.3102/1076998611402504

Leenders, R., 2002. Modeling social influence through network autocorrelation: constructing the weight matrix. Social Networks 24, 21–47.

References : II

Snijders, T., Spreen, M., Zwaagstra, R., 1995. The use of multilevel modelling for analysing personal networks: networks of cocaine users in an urban area. Journal of Quantitative Anthropology 5, 85–105.

Tranmer, M., Steel, D. G., 2001. Ignoring a level in a multilevel model: evidence from UK census data. Environment and Planning 33, 941–948.