

# The Micro-econometrics of Integrated Data

## Applications of Random Graph Models to Labor Markets

John M Abowd and Ian Schmutte

Cornell University

April 2008

# $p_1$ Random Graphs

- There are  $g$  actors in the network
- The data consist of adjacency matrices,  $x$ , that are realizations of a random matrix  $X$  ( $g \times g$ )
- Each dyad,  $D_{ij} = (X_{ij}, X_{ji})$ , is an independent bivariate random variable with possible values

$$\begin{aligned} D_{ij} &= (1, 1) \\ &= (1, 0) \text{ or } (0, 1) \\ &= (0, 0) \end{aligned}$$

- Model the probability,  $\Pr(D_{ij} = (k, h))$  for  $i, j = 1, \dots, g$  and  $k, h = 0, 1$
- Generalization of a Poisson random graph model in which each  $X_{ij}$  is an independent Bernoulli random variable
- Extension to multiple relations: dyad outcome is a  $2 \times R$  vector

# Formal Probability Model

- Let  $R = 1$
- Consider the random variable  $Y_{ijkh}$  such that

$$Y_{ijkh} = 1 \text{ when } D_{ij} = (k, h)$$

$$\pi_{ijkh} \equiv \Pr(D_{ij} = (k, h)) = \Pr(Y_{ijkh} = 1)$$

- This allows us to capture a very general model in which the probability of each dyad state depends on the actors' identities but also incorporates the dyad state itself.
- The Holland-Leinhardt  $p_1$ -class specifies a log-linear model in which the probability of each dyad state depends on actor identities and the possibility of reciprocity:

$$\begin{aligned} \log \Pr(D_{ij} = (k, h)) &= \log \Pr(Y_{ijkh} = 1) \\ &= \lambda_{ij} + k(\alpha_i + \beta_j + \theta) + h(\alpha_i + \beta_j + \theta) + kh\rho_{ij} \end{aligned}$$

# Formal Probability Model

- Impose the identifying restrictions

$$\begin{aligned}\rho_{ij} &= \rho \\ \sum_{i=1}^g \alpha_i &= \sum_{i=1}^g \beta_i = 0\end{aligned}$$

- The  $p_1$  model as specified above is fully identified
- Estimation of  $p_1$  requires software for a log-linear model for  $Y_{ijkh}$  with the particular set of two-way interactions
- Iterative Proportional Fitting or Maximum Likelihood on the  $g \times g \times 2 \times 2$  matrix  $Y = (Y_{ijkh})$

# $p^*$ Markov Random Graph Model

- Able to account for completely general specifications of dependencies between features of the graph.
- Anderson, Wasserman and Crouch note that the  $p^*$  model can be derived either by specifying an autologistic regression model that keeps track of the dependency structure, or alternatively, as a generalization of the theory of Markov random fields via a result called the Hammersley-Clifford theorem.
- The presentation here is based on the autologistic regression approach and borrows from the presentation in Anderson, Wasserman and Crouch (1999).

# $p^*$ Markov Random Graph Model

- Let  $g$  be the number of actors in the data, and  $N = \{1, \dots, g\}$ .
- We let  $\mathbf{x}$  be the observed data, which is a  $g \times g$  adjacency matrix, a realization of  $\mathbf{X}$  with elements  $X_{ij}$ .
- The function  $z(\mathbf{x}) = (z_1(\mathbf{x}), z_2(\mathbf{x}), \dots, z_r(\mathbf{x}))$  is a map from the space of all possible  $g \times g$  adjacency matrices to  $\mathbb{R}^r$ . These can be any functions of the data. Table 4 in Anderson, Wasserman and Crouch suggests some of the possible functions of the data that can be incorporated into  $z(\mathbf{x})$ . We assume that the probability over graphs is log-linear in the components of  $z$ .

$$\Pr(\mathbf{X} = \mathbf{x}) = \frac{\exp[\theta' z(\mathbf{x})]}{\kappa(\theta)}$$

where  $\theta \in \mathbb{R}^r$  is a parameter vector to be estimated.

# $p^*$ Markov Random Graph Model

- Formulate a logit model for the individual links.
- Unlike the standard i.i.d. case for these models, it is necessary to work with the conditional distribution of each link.
- To facilitate this, some special notation is required. Let  $\mathbf{X}_{ij}^c$  refer to the set of all random variables describing each edge, after removing the variable  $X_{ij}$ . This is the conditioning set for link  $X_{ij}$ .
- Let  $\mathbf{x}_{ij}^+$  be the matrix identical to  $\mathbf{x}$ , but with  $x_{ij} = 1$ . Define  $\mathbf{x}_{ij}^-$  as the data matrix identical to  $\mathbf{x}$  but with  $x_{ij} = 0$ .

# $p^*$ Markov Random Graph Model

- Specify the conditional odds ratio:

$$\begin{aligned}\exp\{\omega_{ij}\} &= \frac{\Pr(X_{ij} = 1 | \mathbf{X}_{ij}^c)}{\Pr(X_{ij} = 0 | \mathbf{X}_{ij}^c)} \\ &= \frac{\Pr(\mathbf{X} = \mathbf{x}_{ij}^+)}{\Pr(\mathbf{X} = \mathbf{x}_{ij}^-)} \\ &= \frac{\exp[\theta' z(\mathbf{x}_{ij}^+)]}{\exp[\theta' z(\mathbf{x}_{ij}^-)]} = \exp(\theta'(z(\mathbf{x}_{ij}^+) - z(\mathbf{x}_{ij}^-))) \equiv \exp \theta' d_{ij}\end{aligned}$$

where  $d_{ij} = z(\mathbf{X}_{ij}^+) - z(\mathbf{X}_{ij}^-)$ . So we end up with the system of equations

$$\omega_{ij} = \theta' d_{ij}$$

- Note the similarity between this formulation and a conditional logit analysis as it is usually formulated in the discrete choice literature.
- This model can be estimated using standard conditional logit techniques.



Table 4  
Some parameters and graph statistics for  $p^*$  models

Type Label	Parameter	Graph statistic $z(x)$
<i>Dyadic</i>		
Choice	$\phi$	$L = \sum_{i,j} X_{ij} = X_{++}$
Mutuality	$\rho$	$M = \sum_{i < j} X_{ij} X_{ji}$
<i>Triadic</i>		
Transitivity	$\tau_T$	$T_T = \sum_{i,j,k} X_{ij} X_{jk} X_{ik}$
Intransitivity	$\tau_I$	$T_I = \sum_{i,j,k} X_{ij} X_{jk} (1 - X_{ik})$
Cyclicity	$\tau_C$	$T_C = \sum_{i,j,k} X_{ij} X_{jk} X_{ki}$
2-in-stars	$\sigma_I$	$S_I = \sum_{i,j,k} X_{ji} X_{ki}$
2-out-stars	$\sigma_O$	$S_O = \sum_{i,j,k} X_{ij} X_{ik}$
2-mixed-stars	$\sigma_M$	$S_M = \sum_{i,j,k} X_{ji} X_{ik}$
<i>Subgroup effects</i>	$\phi^{rs}$	$B^{rs} = \sum_{i,j} X_{ij} \delta_{i,j,rs}$
<i>Individual level</i>		
Differential expansiveness	$\alpha_i$	$X_{i+} = \text{outdegree (degree centrality)}$
Differential attractiveness	$\beta_i$	$X_{+i} = \text{indegree (degree prestige)}$

The indicator quantity  $\delta_{i,j,rs} = 1$  if  $i$  is in the  $r$ th subgroup and  $j$  is in the  $s$ th, and 0 otherwise.

# Snijders (2001) Model for Longitudinal Social Network Data

Snijders (2001) considers the

- Dynamic evolution of social interactions between a set of  $n$  actors
- Directed social network at any point in time,  $t$ , modeled as an adjacency matrix,  $\mathbf{x}(t)$  ( $n \times n$ )
- Let  $\mathfrak{X}$  be the space of all such matrices
- The network evolution is modeled as a continuous time Markov chain whose states are the adjacency matrices in the set  $\mathfrak{X}$
- Transition probabilities depend on actor attributes, edge attributes, and various network statistics
- Actors in the model alter the network by changing their outgoing links
- Estimation of the parameters in the transition probabilities is based on matching simulations of the model to the observed data

# The Formal Model

- Each actor in the model has an objective function,  $f_i(\beta, x)$  which represents preferences over the various network configurations,  $x \in \mathfrak{N}$ , and which vary with the parameter  $\beta$ .
- At any point in time, at most one actor is selected to make a change to his outgoing links.
- We will adopt the following notation: let  $\mathbf{x}_{ij}^c(t)$  be the adjacency matrix that is identical to  $\mathbf{x}(t)$  except that the link from  $i$  to  $j$  has changed. Therefore,  $x_{ij}(t) = 1$  if and only if  $x_{ij}^c(t) = 0$ .
- The agent's objective function can depend upon his own characteristics, the characteristics of other agents, and on characteristics of their relationship.
- It can also depend on arbitrary characteristics of the complete network.
- Thus, the calculation by the agent of the benefit of extending or retracting any particular link can be based on any features of the network one cares to specify, as in the  $p^*$  model.

# Transition Probabilities

- The various configurations of social network can be thought of as states in a stochastic process,  $\mathbf{X}(t)$  that evolves continuously over time.
- Transition probabilities in the continuous time Markov process are given by the “transition intensity” defined as

$$q_{ij}(\mathbf{X}) = \lim_{dt \downarrow 0} \frac{1}{dt} \Pr\{\mathbf{X}(t + dt) = \mathbf{x}_{ij}^c | \mathbf{X}(t) = \mathbf{x}\}$$

This is the probability of changing from state  $\mathbf{x}$  to state  $\mathbf{x}_{ij}^c$  at any particular moment in time.

- The matrix  $[q_{ij}(\mathbf{x})]$  is the continuous time analogue to the usual transition matrix. In Snijders setup, the transition intensity is

$$q_{ij}(\mathbf{x}) = \lambda_i(\mathbf{x}) p_{ij}(\mathbf{x})$$

where  $\lambda_i(\mathbf{x})$  is the waiting time between each change made by  $i$ , and  $p_{ij}(\mathbf{x})$  determines whether  $i$  changes his link to  $j$  conditional on the fact that he actually makes a change at  $t$ .

# Waiting Times between Changes

- $\lambda_i(\mathbf{x})$  can either be specified as coming from a particular distribution, or modeled as a function of node and edge covariates.
- What matters is that the distribution of waiting times is such that the probability that any two agents make a change at the same time is zero.
- The idea is that the model evolves through “ministeps” of which we observe only a few.
- This assumption is more crucial in the limited panel data available in most social network research, where only a few timepoints are observed
- Snijders (2001) assumes that the rate function is identical for all agents, given by  $\lambda_k$ , which means that for any time point,  $t \in (t_k, t_{k+1})$  the waiting time until the next change made by any actor has negative exponential distribution with parameter  $N\lambda_k$
- When an event occurs, the probability that it is made by any particular actor is  $1/N$

# Objective Functions

- That agent chooses whether to alter its links according to the objective function:

$$f_i(x) = \sum_l \beta_l s_{il}(\mathbf{x})$$

- Here,  $s_{il}(\mathbf{x})$  are arbitrary statistics of the network that can either capture individual attributes, match attributes, or more general aspects of the network's current configuration
- Given that  $i$  makes a change, he chooses to change the link to  $j$  that maximizes  $f_i(\mathbf{x}_{ij}^c) + U_i(j, t, \mathbf{x})$ , where  $U_i()$  is unobserved heterogeneity that is assumed to be independent of the deterministic term,  $f_i$ , and that depends on  $j, x$ , and  $t$ .
- If  $U_i$  has the Gumbel (extreme value) distribution, then the probability that  $i$  changes his link to  $j$  is given by the multinomial logit distribution

$$p_{ij}(x) = \frac{\exp(f_i(\mathbf{x}_{ij}^c))}{\sum_{h \neq i} \exp(f_i(\mathbf{x}_{ih}^c))}$$

# Labor Markets as Bipartite Graphs

- Imagine a set of  $I$  individuals,  $A(t)$ , and a set of  $J$  employers,  $F(t)$  arranged in a bipartite graph
- There is a link between  $i \in A(t)$  and  $j \in F(t)$  if and only if  $i$  is employed by  $j$  at date  $t$
- The totality of these links can be represented by the  $I \times J$  adjacency matrix  $B(t)$
- This graph is changing over time, so it makes sense to refer to  $B(t)$  as the adjacency matrix of the bipartite graph representing the individual-employer matches at time  $t$
- Since the employment relations between firms and workers change at any time, it is reasonable to think of  $t$  as a continuous variable

# Employer Degree Distribution

- We distinguish primary employment from other forms of employment
- This assumption puts constraints on the row degree distribution in  $B(t)$
- Specifically, assume that  $j = 0$  refers to the non-employment state.
- Including the column  $j = 0$  ensures that every individual in the population at date  $t$  has exactly one “employer”
- Hence,  $B(t) e_j = e_j$ , where  $e_j$  is the  $J \times 1$  column vector of 1s.
- Given this setup the column degree distribution,  $e_j' B(t)$ , is what is known in labor economics as the size distribution of employers (technically only the columns 1 to  $J$  are included in this distribution).
- We note that the (very hard) problem of entry and exit of employers can be included in this formalism by including columns in  $F$  for potential and defunct employers. For the moment, we are not going to worry about this complication.



# The Evolution of the Labor Market

- The existing data are snapshots of the labor market at points in time,  $B(t_1), \dots, B(t_T)$ , where  $T$  is the total number of available time periods.
- These adjacency matrices describe describe outcomes sampled at discrete points in time from the  $I \times (J + 1)$  potential outcomes at each moment of time.
- The objective is to use these snapshots of the labor market to test various assumptions about how the labor market evolves over time.

# Employment Transitions

- Individual has the objective function,  $u_i(B(t))$  assigns a value to every possible network configuration
- At every feasible point in time, the individual can opt to alter his link in the graph by changing employers
- Let  $b_i$  be the row vector of the adjacency matrix corresponding to the  $i^{\text{th}}$  worker
- Assume a worker has at most one employer, so  $\sum_j b_{ij} = 1$ , since non-employment is included in the columns

# Employment Transitions II

- The notation  $B(i \rightarrow j)$  means that  $i$  changes his link with  $j$ .
- If  $j$  is  $i$ 's current employer, and  $i$  takes a job with  $j'$ , we denote this by  $B(i \rightarrow j')$
- If  $i$  leaves his job with  $j$  without taking a new job, we denote this by  $B(i \rightarrow 0)$ .

[This notation, which was introduced by Snijders, represents the assumption that at any instant at most one link in the graph can change. The resulting change dissolves one employment relation and initiates another, including in either case the possibility of non-employment.]

# The Match Function

- The objective function for the individual-employer graph can be represented by a match function that depends upon
  - the graph  $B(t)$ ,
  - characteristics of the individual  $X(t)$ , an  $I \times k$  matrix,
  - characteristics of the employers  $Z(t)$ , a  $((J+1) \times q)$  matrix,
  - and characteristics of the match  $W(t)$ , a  $(I(J+1) \times p)$  matrix
  - Note that while  $B$ ,  $X$  and  $Z$  are observable, at least in principle,  $W$  is (mostly) latent since it contains data for all potential matches.
- Express the match function at a point in time as  $F(B(t), X(t), Z(t), W(t))$ , an  $I \times (J+1)$  matrix function with elements  $f_{ij}$ .

# The Match Function II

- The match function can, in principle, incorporate many features of the graph
  - any statistics about the graph structure that are relevant,
  - covariates associated with the nodes,  $X(t)$  and  $Z(t)$ ,
  - and covariates associated with the edges,  $W(t)$ .
- Theories of labor market equilibrium are descriptions of specializations of the match function  $f$ .

# The Transition Probabilities

- For modeling purposes, we now adopt Snijders' formalism of assuming that the state of the labor market adjacency matrix can only change a single match at a time.
- The configurations of employer-employee matches can be thought of as states in a stochastic process that evolves continuously over time.
- The states are characterized by relevant adjacency matrices,  $B(t)$ .
- The "transition intensity" is defined as

$$q_{\ell m}(B) = \lim_{dt \downarrow 0} \frac{1}{dt} \Pr\{B(t+dt) = B(i \rightarrow j) | B(t) = B\}$$

where  $\ell, m = 1, \dots, I(J+1)$ .

# The Transition Probabilities II

- The intensity  $Q(t)$  matrix is the continuous analogue of a transition probability matrix.
- In this application  $Q(t)$  is  $(I(J+1) \times I(J+1))$  and its dependence on  $B(t)$  means that only certain rows and columns, which depend upon the current state of the labor market, have non-zero transition rates.
- Snijders' simplification, adapted to the bipartite graph case, restricts the the transition intensity intensity as follows

$$q_{\ell m}(B, X, Z, W) = \lambda_{ij}(B, X, Z, W) p_{\ell m}(B, X, Z, W)$$

where  $\lambda_{ij}(B, X, Z, W)$  is the waiting time between each change made by  $ij$ , and  $p_{\ell m}(B, X, Z, W)$  determines whether the pair represented by row  $\ell$  changes to the pair represented by column  $m$ .

- The waiting time matrix  $\lambda_{ij}(B, X, Z, W)$  can either be specified as coming from a particular distribution, or modeled as a function of node and edge covariates.
- In terms of the modelling, what matters is that the distribution of waiting times is such that the probability that any employer-employee pair make a change “two at a time.”
- The idea is that the model evolves through “ministeps,” of which we observe the cumulative effect after a single “period” of time has elapsed, from  $t_k$  to  $t_{k+1}$ .
- This assumption is crucial in the limited panel data available in most social network research, where only a relatively few timepoints are observed.



# The Likelihood Function

- Snijders (2001) has the rate identical for all agents, which in our application translates to all pairs  $\lambda_{ij}$ , hence for any time point,  $t \in (t_k, t_{k+1})$  the waiting time until the next change made by employer-employee pair has negative exponential distribution with parameter  $l(J+1)\lambda_{ij}$ .
- When an event occurs, the probability that it is made by any particular pair is  $1/l(J+1)$ . That pair chooses whether to alter its links according to the objective function:

$$f_{ij}(B, X, Z, W) = \sum_{k=1}^K \beta_k s_{ijk}(B(t), X(t), Z(t), W(t))$$

where  $s_{ijk}(B(t), X(t), Z(t), W(t))$  are specific characteristics of the network.

# The Likelihood Function II

- Given that  $ij$  make a change, the resulting link maximizes

$$f_{ij}(B, X, Z, W) + u_{ij}(i, j, t, B)$$

$u_{ij}()$  is some unobserved heterogeneity term that is independent of the deterministic term,  $f_{ij}$ , and that depends on  $i, j, t$ , and  $B$ .

- If  $u_{ij}$  has the Gumbel (extreme value) distribution, then the probability that  $ij$  changes links is given by the multinomial logit distribution

$$p_{ij}(\mathbf{x}) = \frac{\exp(f_{ij}(B, X, Z, W))}{\sum_{\ell m} \exp(f_{\ell m}(B, X, Z, W))}$$

# Estimation

- The parameters of the model are  $\theta' = (\beta', \lambda')$  where  $\lambda' = (\lambda_{11}, \dots, \lambda_{I(J+1)})$ .
- $\beta' = (\beta_1, \dots, \beta_K)$ , so  $\theta$  is  $I(J+1) + K \times 1$
- The idea is to use a  $I(J+1) + K$ -dimensional statistic  $M$  such that

$$E_{\theta}M = m$$

where  $m$  are the observed moments.

- Snijders suggests a stochastic iterative algorithm for estimation of  $\theta$ .
- The basic iteration step is

$$\hat{\theta}_{N+1} = \hat{\theta}_N - a_N I(M_N - m(B))$$

where  $a_N$  is some sequence converging to zero,  $M_N$  is the network statistic generated by simulating the model.

- Since the  $\lambda_{ij}$  are changing over time, they are estimated as

$$E_{\theta} [M_{ij}(B(t_{k-1}), B(t_k)) | B(t_{k-1})] = m_{ij}(B(t_{k-1}), B(t_k))$$

# Estimation II

- Since the  $\beta$  do not change over time, the relevant moment equation is

$$\sum_{k=1}^K E_{\theta} [M_{ij} (B(t_{k-1}), B(t_k)) | B(t_{k-1})] = \sum_{k=1}^K m_{ij} (B(t_{k-1}), B(t_k))$$

- For estimating such coordinates of  $M$ , it is necessary to simulate the model.
- So, taking the parameter estimate  $\hat{\theta}$ , for each  $k = 2, \dots, K$ , simulate the process starting with initial state  $B(t_{k-1})$  and let time run from  $k - 1$  to  $k$ .
- The network that results is  $B^{sim}(t_k)$ .
- Then, define

$$M = \sum_{r=1}^R m_r (B(t_{k-1}), B^{sim}(t_k))$$

- While the observed outcome is

$$m(B) = \sum_{r=1}^K m_r (B(t_{k-1}), B(t_k))$$

# Individual and Employer Effects in Wage Determination

- Recall the specification for the wage determination equation with individual and employer heterogeneity

$$y = X\beta + D\theta + F\psi + \epsilon$$

- And the fixed-effects moment equations

$$\begin{bmatrix} X'X & X'D & X'F \\ D'X & D'D & D'F \\ F'X & F'D & F'F \end{bmatrix} \begin{bmatrix} \beta \\ \theta \\ \psi \end{bmatrix} = \begin{bmatrix} X'y \\ D'y \\ F'y \end{bmatrix}$$

# Restating in Terms of the Adjacency Matrix Sequence

- Note that if we change the sort order from  $(i, t)$  to  $(t, i)$  then

$$F = \begin{bmatrix} B(1) \\ B(2) \\ \vdots \\ B(T) \end{bmatrix}$$

where  $B(t)$  is the adjacency matrix from the bipartite labor market graph

- In particular  $E_{\theta} [M_{ij} (B(t_{k-1}), B(t_k)) | B(t_{k-1})]$  is the output from the dynamic random graph model above
- So a strategy for instrumenting  $F$  emerges from the model

# Restating in Terms of the Adjacency Matrix Sequence

- Define the instrument matrix

$$\tilde{F} = \begin{bmatrix} \hat{B}(1) \\ \hat{B}(2) \\ \vdots \\ \hat{B}(T) \end{bmatrix}$$

- Solve

$$\begin{bmatrix} \beta \\ \theta \\ \psi \end{bmatrix} = \arg \min \left\{ \begin{array}{l} (y - X\beta - D\theta - F\psi)' \begin{bmatrix} X & D & \tilde{F} \end{bmatrix} \\ \begin{bmatrix} X'X & X'D & X'\tilde{F} \\ D'X & D'D & D'\tilde{F} \\ \tilde{F}'X & \tilde{F}'D & \tilde{F}'\tilde{F} \end{bmatrix}^{-1} \\ \begin{bmatrix} X & D & \tilde{F} \end{bmatrix}' (y - X\beta - D\theta - F\psi) \end{array} \right\}$$

[CPVR06]



John M. Abowd, Patrick Corbel, and Francis Kramarz.

The entry and exit of workers and the growth of employment: An analysis of french establishments.

*Review of Economics and Statistics*, 81:170–187, 1999.



John M. Abowd, Robert H. Creecy, and Francis Kramarz.

Computing person and firm effects using linked longitudinal employer-employee data.

Technical Report TP-2002-06, LEHD, U.S. Census Bureau, 2002.



John M. Abowd, Francis Kramarz, and David N. Margolis.

High wage workers and high wage firms.

*Econometrica*, 67(2):251–333, March 1999.



John M. Abowd, Francis Kramarz, and Sébastien Roux.

Wages, mobility and firm performance: Advantages and insight from using matched worker-firm data.

*Economic Journal*, 116:F245–F285, 2006.



John M. Abowd, Paul Lengermann, and Kevin L. McKinney.

The measurement of human capital in the u.s. economy.





Carolyn J. Anderson, Stanley Wasserman, and Bradley Crouch.

A  $p^*$  primer: logit models for social networks.

*Social Networks*, 1999:37–66, 21.



Pierre Cahuc, Fabien Postel Vinay, and Jean-Marc Robin.

Wage Bargaining with On-the-job Search: Theory and Evidence.

*Econometrica*, 74(2):323364, March 2006.



Stephen E. Fienberg, Michael M. Meyer, and Stanley S. Wasserman.

Statistical analysis of multiple sociometric relationships.

*Journal of the American Statistical Association*, 80(389):51–67, 1985.



Fabien Postel-Vinay and Jean-Marc Robin.

Wage dispersion with worker and employer heterogeneity.

*Econometrica*, 70:2295–350, 2002.



O. Frank and D. Strauss.

Markov graphs.

*Journal of the American Statistical Association*, 81:832–842, 1986.



Paul W. Holland and Samuel Leinhardt.

An exponential family of probability distributions for directed graphs.  
*Journal of the American Statistical Association*, 76(373):33–50, 1981.



Paul W. Holland and Samuel Leinhardt.

An exponential family of probability distributions for directed graphs.  
*Journal of the American Statistical Association*, 76(373):33–50, 1981.



Robert Shimer.

The assignment of workers to jobs in an economy with coordination frictions.  
*Journal of Political Economy*, 113:996, 2005.



Tom A. B. Snijders.

The statistical evaluation of social network dynamics.  
*Sociological Methodology*, 30:361–395, 2001.



Simon Woodcock.

Wage differentials in the presence of unobserved worker, firm, and match heterogeneity.  
*Labour Economics*, forthcoming.