An econometric model of link formation with degree heterogeneity

Bryan S. Graham, UC - Berkeley

Goals in empirical network analysis

Goal #1: Use observed configuration of links across agents, as well as agent characteristics, to infer the structure of their preferences.

Goal #2: Predict how policies will change the structure of a network (forecasting).

- planner may have direct preferences over the form of the network
- ...or her interest may be instrumental (cf., Goldsmith-Pinkham & Imbens, 2013)
- ...strong analogies with classic discrete choice problems

Challenges in empirical network analysis

Heterogeneity: A link may reflect (i) high returns generated by observed agent characteristics or (ii) by unobserved agent characteristics

1. some agents may have attributes which generate high levels of link surplus (*degree heterogeneity*)

2. agents similar on an unobserved dimension may generate more surplus (*homophily on unobservables*, *assortative matching*)

Challenges in empirical network analysis

Interdependency: Link surplus may vary with the presence of absence of links *elsewhere* in the network

1. coherency & completeness (Bresnahan and Reiss, 1991; Sheng, 2014; de Paula et al., 2014).

2. heterogeneity vs. interdependency (Graham, 2013; 2016).

Outline

1. Set-up, notation and model

2. Degree heterogeneity bias

3. Likelihood

   (a) conditional maximum likelihood

   (b) joint maximum likelihood (likely skip)

4. Monte Carlo

5. Ongoing work

   (a) Graham (2013, 2016) – homophily & transitivity

   (b) Bickel & Graham (in progress) – graphlet counts

Setup & Notation

Agents (actors, nodes): Let $i = 1, \ldots, N$ index a random sample of (potentially connected) agents

Dyad: a pair of agents

Links (ties, edges): Let $D_{ij} = 1$ if agent i is linked to agent j and zero otherwise

Links are undirected: $D_{ij} = D_{ji}$

- No self-links: $D_{ii} = 0$

- $\mathbf{D}$ denotes the $N \times N$ *adjacency matrix* with elements D_{ij}

The goal is to formulate an econometric model for $\mathbf{D}$

Setup (continued)

For each of the N agents we observe the vector of attributes X_i

Let W_{ij} be a $K \times 1$ vector of dyad-specific attributes (symmetrically) formed using X_i and X_j

- $W_{ij} = X_i X_j$
- $W_{ij} = \left| X_i - X_j \right|$, etc.

W_{ij} may also include measures intrinsically defined at the dyad level, e.g., coefficient of relationship (Wright, 1922)

Link formation

Let i and j form a link if the net surplus from doing so is positive

$$D_{ij} = \mathbf{1}\left(W_{ij}'\beta + A_i + A_j - U_{ij} \geq 0\right)$$

with:

- A_i (unrestricted) agent specific degree heterogeneity

- U_{ij} i.i.d. across dyads and logistic

Utility is transferable (across directly linked agents)

No externalities or "network interdependencies"

Econometrician observes $\left(D_{ij}, W_{ij}'\right)'$ for $i = 1, \ldots, N$ and $j < i$

Degree heterogeneity bias

Let $D_{i+} = \sum_{j \neq i} D_{ij}$ denote the *degree* of agent i

The *degree sequence* of the network is

$$D_{1+}, \ldots, D_{N+}$$

Degree sequences are often well-described by power laws (Albert and Barabasi, 2002).

Many agents with few links, a small number of agents with <u>many</u> links ("hubs").

Hubs link both within <u>*and*</u> across groups.

High degree agents tend to attenuate measured homophily.

Degree heterogeneity bias in basic model

A common empirical specification is

$$D_{ij} = \mathbf{1}\left(\alpha + \left(X_i + X_j\right)'\gamma + \left|X_i - X_j\right|'\beta - U_{ij} \geq 0\right).$$

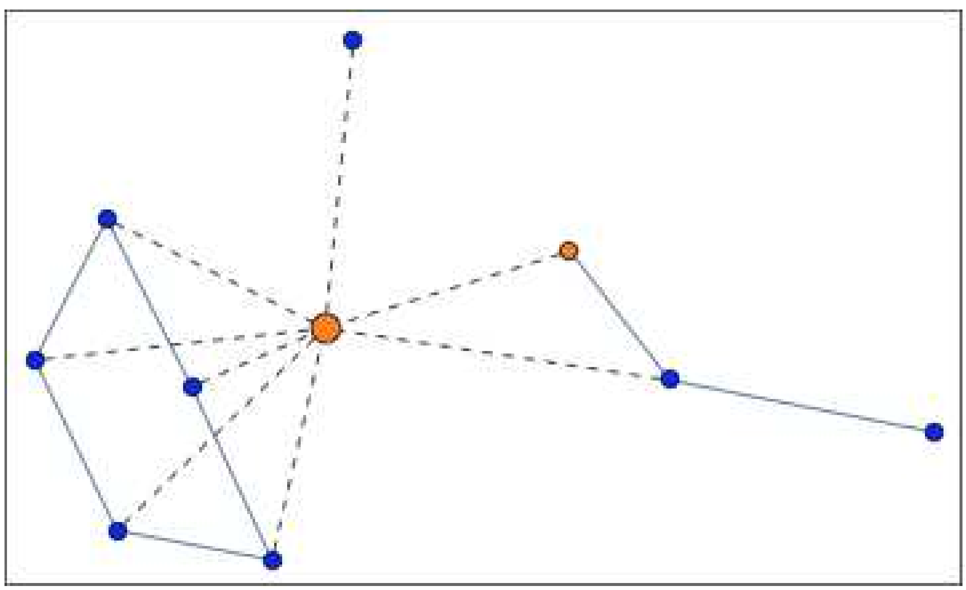
Examples: Lai and Reiter (2000, *JCR*), Attanasio et al. (2012, *AEJ*), Apicella et al. (2012, *Nature*).

A test of $H_0 : \beta_k \geq 0$ vs. $H_1 : \beta_k < 0$ is a test of homophily on X_{ki}.

Degree heterogeneity in this model is proxied by $\left(X_i + X_j\right)'\gamma$ term.

Unobserved degree heterogeneity in this model will bias estimates of β.

Degree heterogeneity bias in basic model



Model with degree heterogeneity

The conditional probability of an (i, j) link is:

$$\Pr\left(D_{ij} = 1 \middle| \mathbf{X}, \mathbf{A}\right) = \frac{\exp\left(W_{ij}'\beta + A_i + A_j\right)}{1 + \exp\left(W_{ij}'\beta + A_i + A_j\right)}.$$

High A_i agents generate more link surplus and form more links (degree heterogeneity, hubs).

β parameterizes which configurations of dyad attributes generate the most surplus *holding degree heterogeneity fixed*.

Model will fit the observed degree sequence of a network perfectly.

Likelihood

Links are conditionally independent with

$$\Pr\left(\mathbf{D} = \mathbf{d} | \mathbf{X}, \mathbf{A}\right)$$

equal to

$$\prod_{i=1}^{N} \prod_{j<i} \left[\frac{\exp\left(W_{ij}'\beta + T_{ij}'\mathbf{A}\right)}{1 + \exp\left(W_{ij}'\beta + T_{ij}'\mathbf{A}\right)} \right]^{d_{ij}} \times \left[\frac{1}{1 + \exp\left(W_{ij}'\beta + T_{ij}'\mathbf{A}\right)} \right]^{1-d_{ij}}$$

with T_{ij} an N-vector with ones in rows i and j and zeros elsewhere.

If $\mathbf{A}$ were observed (in addition to $\mathbf{W}$ and $\mathbf{D}$) estimation of, and inference on, β would be standard.

Unobserved Heterogeneity

In the observed network X_i (and hence W_{ij}) may covary with A_i.

Random effects: model distribution of $\mathbf{A}$ given $\mathbf{X}$ (cf., van Duijn, Snijders and Zijlstra, 2004; Goldsmith-Pinkham & Imbens, 2013).

Fixed effects: leave the joint distribution of $(\mathbf{X}, \mathbf{A})$ unrestricted.

Two approaches:

1. conditional (fixed effects) MLE
   - cf., large-N, fixed-T panel data (e.g., Chamberlain, 1980)
2. joint (fixed effects) MLE
   - cf., large-N, large-T panel data (e.g., Hahn & Newey, 2004)

Conditional MLE

Manipulating the likelihood yields Pr $\left(\mathbf{D}=\mathbf{d}\right|\mathbf{X},\mathbf{A})$ equal to

$$\frac{\exp\left(\sum_{i=1}^{N}\sum_{j<i}d_{ij}W_{ij}'\beta+\sum_{i=1}^{N}\sum_{j<i}d_{ij}T_{ij}'\mathbf{A}\right)}{\prod_{i=1}^{N}\prod_{j<i}\left(1+\exp\left(W_{ij}'\beta+T_{ij}'\mathbf{A}\right)\right)}$$

with $\sum_{i=1}^{N}\sum_{j<i}d_{ij}T_{ij}'=\left(D_{1+},\ldots,D_{N+}\right)'=\mathbf{D}_{+}'$.

The network degree sequence is a sufficient statistic for $\mathbf{A}$.

Variation in W_{ij} drives link formation conditional on $\mathbf{D}_{+}$.

Conditional MLE

Let $\mathbb{D}^s$ denote the set of networks with degree distributions coinciding with the observed one.

The conditional likelihood, which is constant in $\mathbf{A}$, is given by

$$\Pr\left(\mathbf{D} = \mathbf{d} \middle| \mathbf{X}, \mathbf{D}_+\right) = \frac{\exp\left(\sum_{i=1}^{N}\sum_{j<i} d_{ij} W_{ij}' \beta\right)}{\sum_{\mathbf{v}\in\mathbb{D}^s} \exp\left(\sum_{i=1}^{N}\sum_{j<i} v_{ij} W_{ij}' \beta\right)}$$

Conceptually straightforward, but evaluation of denominator is difficult.

Cf. Importance sampling algorithm of Blitzstein and Diaconis (2011)

Cf. Andersen (1973), Chamberlain (1980), Charbonneau (2014)

Conditional Inference Using Subgraphs: Tetrad Logit

Consider the sub-graph formed by agents i, j, k and l (*tetrad*).

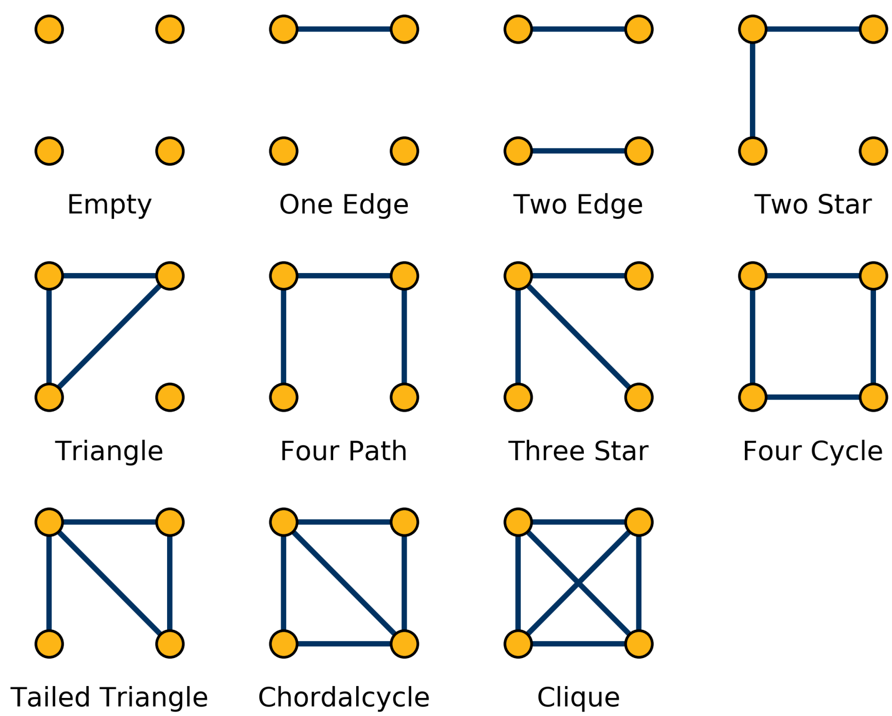
There are $2^6 = 64$ possible tetrad configurations (11 isomorphisms).

46 of these configurations are completely determined by their degree sequence.

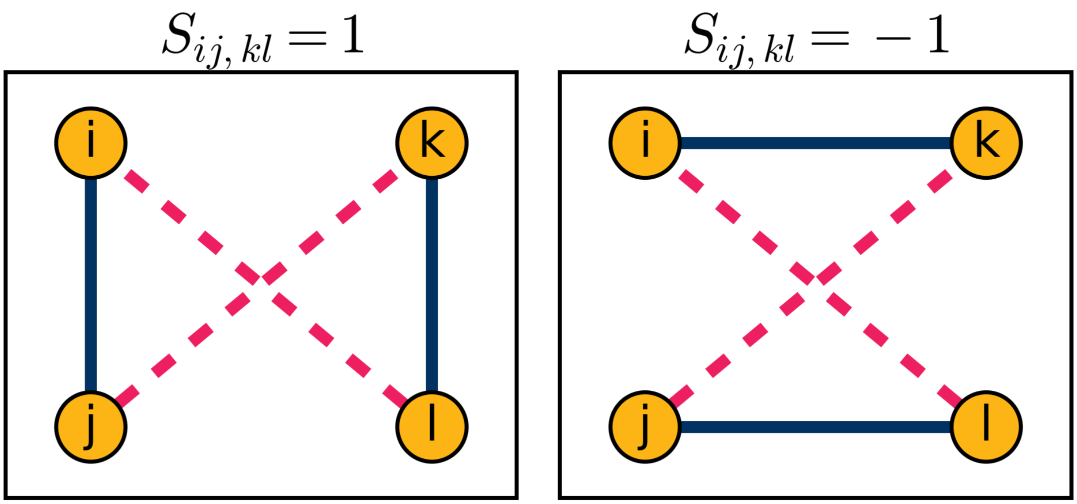
18 configurations share their degree sequence with at least one other configuration.

We can move from one subgraph to another with the same degree sequence by a sequence of *edge swaps*.

Tetrad Isomorphisms



Tetrad Logit



Both subgraphs have the same degree sequence (e.g., (2,2,2,2) or (2,1,2,1)).

Relative frequency of first configuration compared to second configuration does not depend on (A_i, A_j, A_k, A_l).

Tetrad Logit (continued)

Let

$$\begin{aligned} S_{ij,kl} &= D_{ij} D_{kl} \left(1 - D_{ik}\right) \left(1 - D_{jl}\right) \\ &\quad - \left(1 - D_{ij}\right) \left(1 - D_{kl}\right) D_{ik} D_{jl}. \end{aligned}$$

Some algebra gives

$$\Pr\left(S_{ij,kl} = 1 \middle| \mathbf{X}, \mathbf{A}, S_{ij,kl} \in \{-1, 1\}\right)$$

equal to:

$$\frac{e^{\left[W_{ij} + W_{kl} - \left(W_{ik} + W_{jl}\right)\right]' \beta_0}}{1 + e^{\left[W_{ij} + W_{kl} - \left(W_{ik} + W_{jl}\right)\right]' \beta_0}}.$$

Tetrad Logit (continued)

$\left[W_{ij} + W_{kl} - \left(W_{ik} + W_{jl}\right)\right]' \beta_0$ is a measure of *complementarity* (increasing differences).

If positive, then the net surplus from the (i, j) and (k, l) link configuration exceeds that from the (i, l) and (j, k) configuration

β is identified by homophily on observables "holding the degree distribution fixed".

Tetrad Logit (continued)

Let $\tilde{W}_{ij,kl} = W_{ij} + W_{kl} - \left(W_{ik} + W_{jl}\right)$ and define a Bernoulli log-likelihood contribution $l_{ij,kl}(\beta_0)$ equal to

$$\left|S_{ij,kl}\right| \times \left\{S_{ij,kl}\tilde{W}_{ij,kl}'\beta_0 - \ln\left[1 + \exp\left(S_{ij,kl}\tilde{W}_{ij,kl}'\beta_0\right)\right]\right\}.$$

To impose symmetry in the index set I take the average over all permutations:

$$\begin{aligned} g_{ijkl}(\beta) &= \frac{1}{4!}\sum_{\pi\in\Pi_4} l_{\pi_1\pi_2,\pi_3\pi_4}(\beta) \\ &= \frac{1}{3}\left[l_{ij,kl}(\beta) + l_{ij,lk}(\beta) + l_{ik,lj}(\beta)\right]. \end{aligned}$$

Tetrad Logit (continued)

The *Tetrad Logit* estimator chooses $\widehat{\beta}$ to maximize:

$$L_N(\beta) = \binom{N}{4}^{-1} \sum_{i<j<k<l} g_{ijkl}(\beta).$$

Only Tetrad's with $T_{ijkl} = 1$ contribute where

$$T_{ijkl} = \begin{cases} 1, & S_{ij,kl} \in \{-1,1\} \vee S_{ij,lk} \in \{-1,1\} \\ & \vee S_{ik,lj} \in \{-1,1\} \\ 0, & \text{otherwise} \end{cases}$$

(i.e., tetrads with degree sequences which do not uniquely determine the subgraph).

Criterion function is a summation over a random set of indices (cf., Chamberlain, 1980).

Similar to a (4th-order) U-Process minimizer (cf., Honore and Powell, 1994).

Sparse & Dense Graph Sequences

Average *density* is

$$\begin{aligned}\rho_N &= \Pr\left(D_{ij} = 1\right) \\ &= \mathbb{E}\left[\Pr\left(D_{ij} = 1 \middle| \mathbf{X}, \mathbf{A}_0\right)\right].\end{aligned}$$

Average degree is

$$\lambda_N = (N-1)\,\rho_N.$$

The heterogeneity sequence $\{A_{0i}\}_{i=1}^{N}$ may induce *dense* ($\lambda_N = O(N)$) or *sparse* ($\lambda_N = O(1)$) networks as N grows large.

I assume that heterogeneity sequence is such that $N\lambda_N \to \infty$ as $N \to \infty$.

Also assume that $\lambda_N = \Omega(1)$ (i.e., $\lambda_N \geq \lambda_0 > 0$ for large N).

Tetrad Logit: Consistency (continued)

Let

$$\alpha_{q,N} = \Pr\left(T_{i_1 i_2 i_3 i_4} = 1, T_{j_1 j_2 j_3 j_4} = 1\right) \qquad (1)$$

be the probability that tetrads $\{i_1, i_2, i_3, i_4\}$ and $\{j_1, j_2, j_3, j_4\}$ *both* contribute when they share $q = 0, 1, 2, 3, 4$ agents in common.

For networks with $\rho_N \to 0$, the probability that a random tetrad contributes to $L_N(\beta)$,

$$\alpha_{4,N} = \Pr\left(T_{ijkl} = 1\right),$$

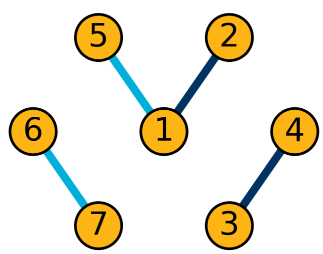
is of order ρ_N^2.

Consistency therefore requires that $\binom{N}{4}\alpha_{4,N} \to \infty$ or, equivalently, that $N\lambda_N \to \infty$.

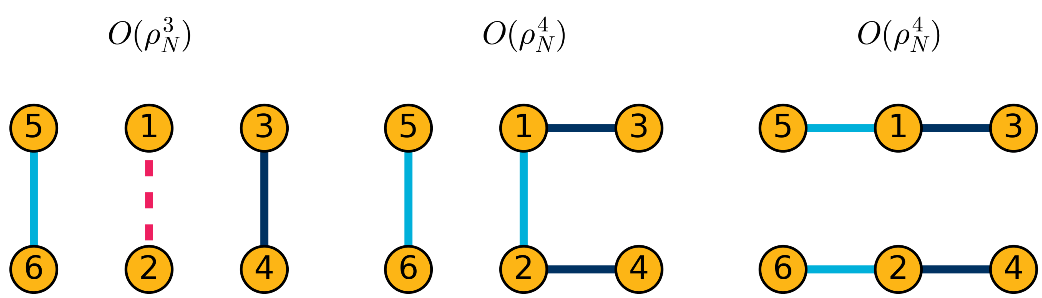
Determining the order of $\alpha_{q,N}$ for $q = 1, 2, 3$ is more complicated.

Tetrad Stitchings ($\alpha_{q,N}$, $q = 1$)

$O(\rho_N^4)$

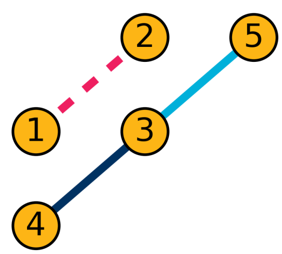


Tetrad Stitchings ($\alpha_{q,N}$, $q = 2$)



Tetrad Stitchings ($\alpha_{q,N}$, $q = 3$)

$O(\rho_N^3)$



Tetrad Logit: Consistency (continued)

Let $q_{ij,kl}(\beta) = \left[1 + \exp\left(-\tilde{W}'_{ij,kl}\beta\right)\right]^{-1}$, $q_{ij,kl} = q(\beta_0)$ and

$$Q(\beta) = -\mathbb{E}\left[D_{\mathsf{KL}}\left(q_{ij,kl} \middle\| q_{ij,kl}(\beta)\right) + \mathbf{S}\left(q_{ij,kl}\right) \middle| S_{ij,kl} \in \{-1,1\}\right].$$

The normalized criterion function has expectation

$$\mathbb{E}\left[\alpha_{4,N}^{-1} L_N(\beta)\right] = Q(\beta) \times \Pr\left(S_{ij,kl} = \{-1,1\} \middle| T_{ijkl} = 1\right)$$

and variance (Hoeffding decomposition)

$$\mathbb{V}\left(\frac{L_N(\beta)}{\alpha_{4,N}}\right) = O\left(\frac{1}{N}\right) + O\left(\frac{1}{N\lambda_N}\right) + O\left(\frac{1}{N^2\lambda_N}\right) + O\left(\frac{1}{N^2\lambda_N^2}\right).$$

We therefore have

$$\alpha_{4,N}^{-1} L_N(\beta) \xrightarrow{p} Q(\beta) \Pr\left(S_{ij,kl} = \{-1,1\} \middle| T_{ijkl} = 1\right)$$

uniformly in $\beta \in \mathbb{B}$ if $N\lambda_N \to \infty$.

Tetrad Logit: Asymptotic Normality

(1) Mean value expansion.

(2) Convergence of Hessian term (Newey & McFadden, 1994, Lemma 2.9).

(3) Verify that, when appropriately scaled, the "score term"

$$U_N \stackrel{\text{def}}{=} \binom{N}{4}^{-1} \sum_{i<j<k<l} s_{ijkl}(\beta_0)$$

obeys a CLT (here $s_{ijkl}(\beta_0) = \nabla_\beta g_{ijkl}(\beta_0)$).

(4) U_N is similar to a 4th order U-Statistic.

Tetrad Logit: Asymptotic Normality (continued)

(5) (Hoeffding) variance calculation indicates "score term" is degenerate, with degeneracy of order 1 (rate of convergence is $1/n = \binom{N}{2}^{-1}$ in dense case).

(6) Hajek Projection argument to replace 4th order sum over *all tetrads* with a double sum over *all dyads*:

$$U_N^* = \frac{6}{n} \sum_{i<j} \bar{s}_{ij} (\beta_0),$$

for $\bar{s}_{ij} (\beta_0) \stackrel{\text{def}}{=} \mathbb{E} \left[s_{ijkl} (\beta_0) \middle| X_i, X_j, A_i, A_j, U_{ij} \right]$.

Tetrad Logit: Asymptotic Normality (continued)

(7) Elements of double sum are not independent from one another, but can be shown to have a martingale structure.

(8) Using Chatterjee (2006, *Annals of Probability*) I get a final result of (heuristically stated):

$$\sqrt{n\alpha_{2,N}^{-1}}\alpha_{4,N}\left(\widehat{\beta}_{TL}-\beta_0\right)\xrightarrow{D}\mathcal{N}\left(0,36\Gamma_0^{-1}\Omega_2\Gamma_0^{-1}\right).$$

Rate-of-Convergence

Observe that $\sqrt{n\alpha_{2,N}^{-1}\alpha_{4,N}} = O\left(\sqrt{n}\rho_N^{-3/2}\rho_N^{4/2}\right) = O\left(\sqrt{n\rho_N}\right) = O\left(\sqrt{N\lambda_N}\right)$, so that $\widehat{\beta}_{\mathsf{TL}} \xrightarrow{p} \beta_0$ at rate:

- $n^{-1/2}$ (or N^{-1}) if $\rho_N \to \rho_0 > 0$ as $N \to \infty$ (dense case);

- $n^{-1/4}$ (or $N^{-1/2}$) if $\lambda_N = (N-1)\rho_N \to \lambda_0 > 0$ as $N \to \infty$ (sparse case).

Under dense graph squares $\widehat{\beta}_{\mathsf{TL}}$ converges at the usual parametric rate.

When average density tends toward zero as the graph grows large, the rate of convergence slows.

"Sample size" is number of dyads!

Unconditional MLE

Treat $\mathbf{A}$ as parameters to be estimated along with β.

The dimension of $\mathbf{A}$ grows with N, the number of agents.

But the number of dyads, $n = \frac{1}{2}N(N-1)$, grows more quickly.

The ratio of the number of parameters to "observations" is $O\left(\frac{1}{N}\right)$.

Is there an incidental parameters problem? If so, how does it manifest itself?

Bounded link probabilities (Dense Graphs)

Let p_{ij} denote the probability of an (i,j) link at the population parameter; let $p_{i+} = \sum_{j \neq i} p_{ij}$ be agent i's *expected* degree.

I impose the restriction $p_{ij} \in [\kappa, 1-\kappa]$.

This implies that the support of the individual effects is bounded.

While this assumption can be weakened, we do require that the network is (fairly) dense.

The tetrad logit estimator does not require such an assumption – it can accommodate $\{A_i\}_{i=1}^{N}$ sequences which diverge (relatively rapidly) with N.

Computation: concentrated MLE

Let $\widehat{A}(\beta) = \arg\max_{\mathbf{A}} l_N(\beta, \mathbf{A})$

Define $\varphi(\mathbf{A};\beta) = (\varphi_1(\mathbf{A};\beta), \ldots, \varphi_N(\mathbf{A};\beta))$ with

$$\begin{aligned}\varphi_i(\mathbf{A};\beta) \;=\; & \ln D_{i+} - \\ & \ln\left(\sum_{j\neq i} \frac{\exp\left(W_{ij}'\beta\right)}{\exp\left(-A_j\right) + \exp\left(W_{ij}'\beta + A_i\right)}\right).\end{aligned}$$

The (unique) fixed point $\hat{\mathbf{A}}(\beta) = \varphi\left(\hat{\mathbf{A}}(\beta);\beta\right)$ coincides with the (concentrated) MLE of $\mathbf{A}$ given β (if it exists).

$\widehat{A}_i(\beta)$ depends on *all* elements of $\mathbf{D}$ and $\mathbf{W}$, not just those associated with agent i.

We then compute $\widehat{\beta}$ as the maximizer of the *concentrated* log-likelihood $l_N\left(\beta, \hat{\mathbf{A}}(\beta)\right) = l_N^c(\beta)$.

Asymptotic Normality

Let $\mathcal{I}(\beta)$ be the probability limit of $-H^c_{N,\beta\beta}$ (Hessian of concentrated log-likelihood) divided by $n = \frac{1}{2}N(N-1)$.

A standard argument gives:

$$\begin{aligned}\sqrt{n}\left(\widehat{\beta}-\beta\right) &= \mathcal{I}_0^{-1}(\beta) \\ &\quad \times \frac{1}{\sqrt{n}} \sum_{i=1}^{N} \sum_{j<i} s_{\beta ij}\left(\beta_0, \hat{\mathbf{A}}(\beta_0)\right) \\ &\quad + o_p(1)\end{aligned}$$

However, since $\mathbb{E}\left[s_{\mathbf{A}ij}\left(\beta_0, \hat{\mathbf{A}}(\beta_0)\right)\right] \neq 0$, we cannot directly apply a CLT (e.g., Arellano and Hahn, 2007).

Asymptotic Normality

If we replace $s_{\mathbf{A}ij}\left(\beta_0, \hat{\mathbf{A}}\left(\beta_0\right)\right)$ with a second order Taylor series approximation we get the refined representation of $\sqrt{n}\left(\widehat{\beta}-\beta\right)$ equal to

$$\mathcal{I}_0^{-1}\left(\beta\right) \times \frac{1}{\sqrt{n}} \sum_{i=1}^{N} \sum_{j<i} \left\{ s_{\beta ij}\left(\beta_0, \mathbf{A}_0\right) \right.$$
$$\left. -H_{N,\beta\mathbf{A}} H_{N,\mathbf{AA}}^{-1} s_{\mathbf{A}ij}\left(\beta_0, \mathbf{A}_0\right) \right\}$$
$$+\mathcal{I}_0^{-1}\left(\beta\right) B_0 + o_p\left(1\right)$$

We can apply a CLT to the term in $\{\cdot\}$.

The second term implies the limit distribution is not mean zero.

Asymptotic Normality

Let

$$\begin{aligned} s^{o}_{\beta ij}(\beta_0, \mathbf{A}_0) &= s_{\beta ij}(\beta_0, \mathbf{A}_0) \\ &\quad - H_{N,\beta\mathbf{A}} H^{-1}_{N,\mathbf{AA}} s_{\mathbf{A}ij}(\beta_0, \mathbf{A}_0) \end{aligned}$$

be the score function associated with the concentrated log-likelihood.

We can show that (using martingale structure)

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{N} \sum_{j<i} s^{o}_{\beta ij}(\beta_0, \mathbf{A}_0) \xrightarrow{D} N\left(0, \mathcal{I}_0^{-1}(\beta)\right).$$

Asymptotic Normality

The presence of the bias term – $\mathcal{I}(\beta) B_0$ – means that the limiting distribution of $\sqrt{n}\left(\widehat{\beta}-\beta\right)$ is not centered at zero:

$$\sqrt{n}\left(\widehat{\beta}-\beta\right) \stackrel{D}{\to} \mathcal{N}\left(\mathcal{I}_0^{-1}(\beta) B_0, \mathcal{I}_0^{-1}(\beta)\right)$$

or

$$N\left(\widehat{\beta}-\beta\right) \stackrel{D}{\to} \mathcal{N}\left(\mathcal{I}(\beta) B_0 \sqrt{2}, 2\mathcal{I}_0^{-1}(\beta)\right)$$

Accurate inference requires bias correction.

Monte Carlo: Data Generating Process

Let $X_i \in \{-1, 1\}$ with $\Pr(X_i = 1) = 1/2$, $W_{ij} = X_i X_j$, and

$$A_i = \alpha_L \mathbf{1}(X_i = -1) + \alpha_H \mathbf{1}(X_i = 1) + V_i$$

where $\alpha_L \leq \alpha_H$ and

$$V_i | X_i \sim 2\left\{\text{Beta}(\lambda_0, \lambda_1) - \frac{\lambda_0}{\lambda_0 + \lambda_1}\right\}.$$

Lower values of α_L and α_H induce sparser networks.

Choices of λ_0 and λ_1 can be used to induce skewness in the degree distribution.

Results for $N = 200$, $n = \binom{N}{2} = 19,900$, $\binom{N}{4} = 64,684,950$

Monte Carlo: Bias & Standard Deviation

	Right-Skewed Correlated Heterogeneity			
Panel A	B.1	B.2	B.3	B.4
α_L	-1/3	-1	-5/3	-7/3
α_H	0	-2/3	-4/3	-2
λ_0	1/4	1/4	1/4	1/4
λ_1	3/4	3/4	3/4	3/4
Panel B				
Avg. Degree	86.5	41.7	15.0	4.5
Std. of Degree	15.3	12.7	6.7	2.8
Transitivity	0.51	0.31	0.14	0.05
Frac. Giant	1.0	1.0	0.99	0.96

Monte Carlo: Bias & Standard Deviation

	Right-Skewed Correlated Heterogeneity			
	B.1	B.2	B.3	B.4
TL	0.9885 (0.020)	1.004 (0.027)	1.024 (0.049)	1.046 (0.095)
JML	1.012 (0.017)	1.010 (0.024)	1.006 (0.041)	-
BC	1.001 (0.017)	1.011 (0.024)	1.076 (0.048)	-
FT	0.1706	0.0760	0.0143	0.0015
#Cvg	1000	994	994	5

Monte Carlo: Size

	Right-Skewed Correlated Heterogeneity			
	B.1	B.2	B.3	B.4
$\alpha = 0.05$				
TL	0.826	0.892	0.868	0.930
JML	0.822	0.860	0.887	-
BC	0.906	0.850	0.411	-
$\alpha = 0.10$				
TL	0.893	0.944	0.931	0.965
JML	0.892	0.924	0.941	-
BC	0.945	0.917	0.511	-

Simple extension

With $T \geq 2$ we can consider a linking rule of

$$\begin{aligned} D_{ijt} &= \mathbf{1}\Bigg(\gamma D_{ijt-1} + \delta \sum_{k=1}^{N} D_{ikt-1} D_{jkt-1} \\ &\quad + W_{ijt}'\beta + A_i + A_j - U_{ijt} \geq 0\Bigg) \end{aligned}$$

Introduces state-dependence (γ) and a taste for transitivity (δ) in link formation.

For $t = 0, 1$ case set $D_{ij} = D_{ij1}$,

$$W_{ij}^{*} = \left(D_{ij0}, \sum_{k=1}^{N} D_{ik0} D_{jk0}, W_{ij1}'\right)',$$

and $\beta^{*} = (\gamma, \delta, \beta')'$ and proceed as in the cross-sectional case (cf., Nadler, 2015).

Related work

Antecedents:

Conditional: Andersen (1970), Chamberlain (1980), Blitzstein & Diaconis (2011), Charbonneau (2014)

Joint: Hahn & Newey (2004), Chatterjee, Diaconis & Sly (2011), Fernandez-Val & Weidner (2016)

Extensions:

Dzemski (2014), Nadler (2015), Jochmans (2016a,b), Yan, Jiang, Feinberg & Leng (2016), Candelaria (2016)

Ongoing Work

With $T \geq 4$ we can consider a linking rule of

$$D_{ijt} = \mathbf{1}\left(\beta D_{ijt-1} + \gamma \sum_{k=1}^{N} D_{ikt-1} D_{jkt-1} \right.$$
$$\left. A_{ij} - U_{ijt} \geq 0\right)$$

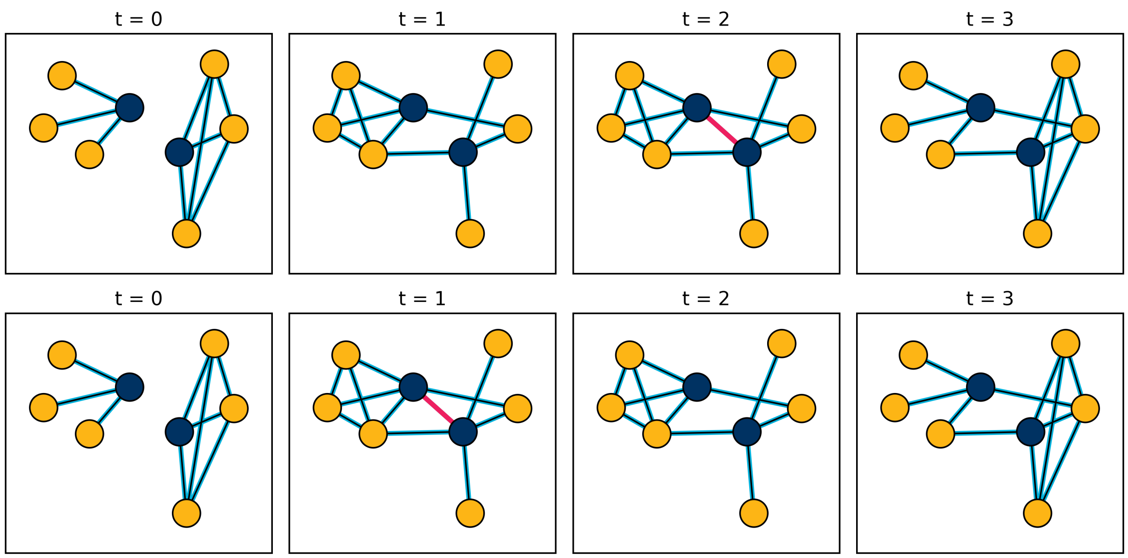
Allows for *very* rich unobserved heterogeneity structure *as well as* network interdependencies.

Graham (2013, 2016) shows "fixed effects" identification of β and γ.

Basic idea (high level): look at *relative* frequency of different link histories for dyads embedded in *stable* local network architectures.

Graham (2016) presents an estimator. Consistent and asymptotically normal under large (sparse) network asymptotics.

Ongoing Work: Stable Neighborhoods



Ongoing Work: Monte Carlo

Latent-space heterogeneity:

$\xi_x \sim \mathcal{U}\left[0, \sqrt{N}\right]$ and $\xi_y \sim \mathcal{U}\left[0, \sqrt{N}\right]$.

$$V_{ij} = \sqrt{\left(\xi_{xi} - \xi_{xj}\right)^2 + \sqrt{\left(\xi_{yi} - \xi_{yj}\right)^2}}$$

$A_{ij} = \alpha_0 \mathbf{1}\left(V_{ij} \leq r\right) + \alpha_1 \mathbf{1}\left(V_{ij} > r\right)$ with $\alpha_1 = -\infty$

Initial condition:

$D_{ij0} = \mathbf{1}\left(A_{ij} - U_{ij0} \geq 0\right)$ with U_{ijt} logistic

Average degree at $t = 0$ is $\approx \pi r^2 \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)} = 4$.

Lots of "non-structural" transitivity in links generated by A_{ij}

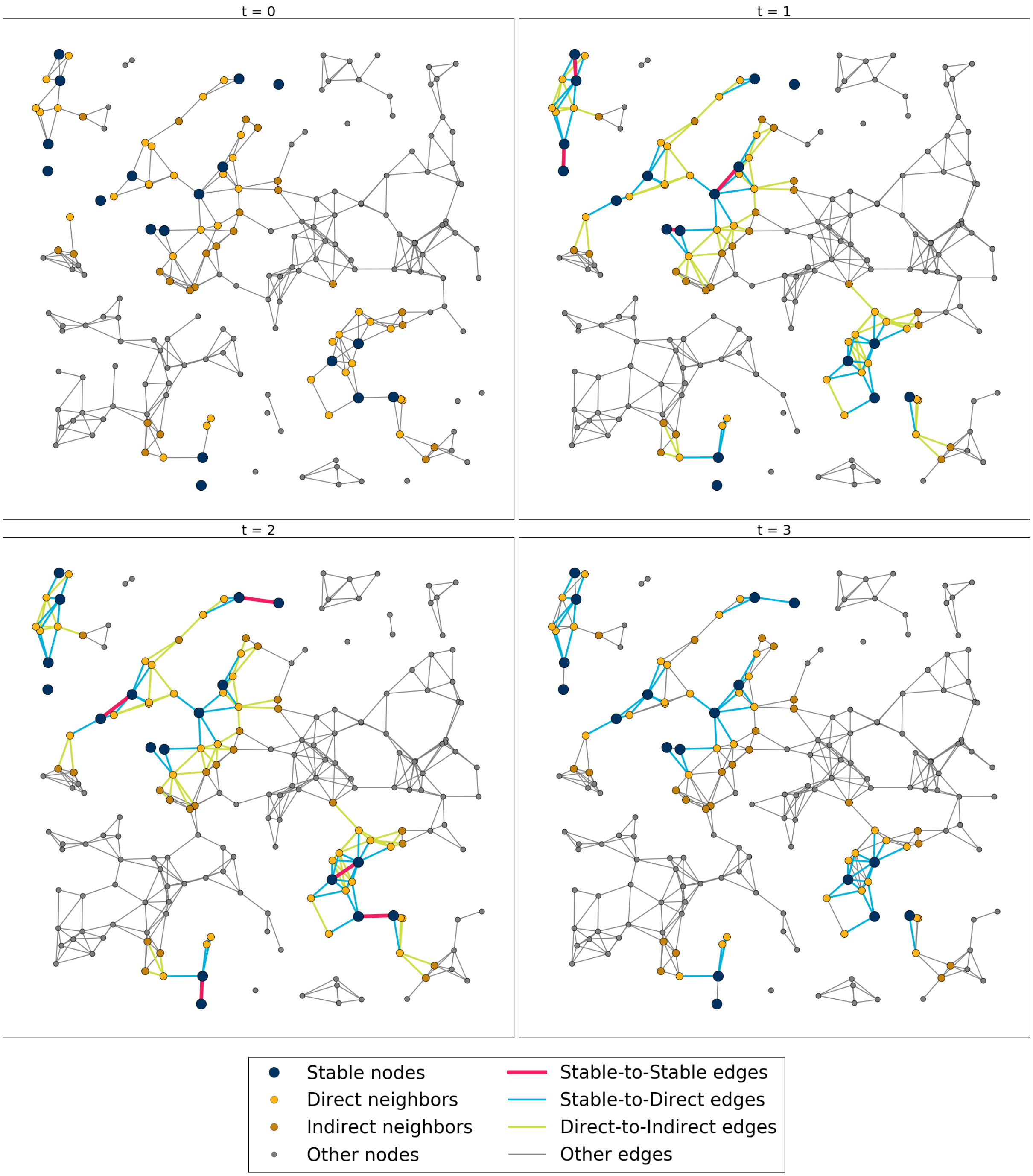
Ongoing Work: Monte Carlo (continued)

- $N = 5,000$

- $\beta = \gamma = 1$

- Average number of dyads embedded in stable neighborhoods across 1000 samples was 110.6

	β	γ
Average	1.044	1.046
Std. Dev.	0.458	0.298
Avg. Std. Err.	0.449	0.292
Coverage	0.962	0.965


t = 0
t = 1
t = 2
t = 3
Stable nodes
Direct neighbors
Indirect neighbors
Other nodes
Stable-to-Stable edges
Stable-to-Direct edges
Direct-to-Indirect edges
Other edges

Future Directions: Network Moments

- Older literature on subgraph counts & testing (e.g., triad census of Holland & Leinhardt, 1976).

- Bickel, Chen & Levina (2011) call subgraph frequencies *network moments*.

- In joint work with Peter Bickel I am studying the asymptotic sampling distribution of vectors of network moments in digraphs.

Future Directions

Many open areas:

- bi-partite graphs (one-to-one, many-to-one, many-to-many) (Nadler, 2015)

- large sample theory (Bickel, Chen & Levina, 2011; Leung, 2015)

- panels, dynamics (Graham 2013, 2016)

- 'large game' aspect (Menzel, 2015; de Paula et al., 2014)

- network endogeneity (Goldsmith-Pinkham & Imbens, 2013; Auerbach, 2016)

- empirical work!