An optimal test for strategic interaction in network formation among heterogenous agents

European Summer Meeting of the Econometric Society

Barcelona, August 2023

Andrin Pelican & Bryan S. Graham

Strategic Network Formation

In an *economic* model of (directed) network formation agents *purposefully* direct links to one another in order to maximize utility.

A payoff function maps all possible network configurations into agent utilities.

Agents use this payoff function to weigh the benefits of directing any particular link against the costs of doing so.

A Nash Equilibrium (NE) network arises when all agents link choices are individually optimal given the choices made by other agents (Bala and Goyal, 2000).

Models and data

Bringing structural models of network formation to data is very challenging.

1. A directed network is the realization of a game with N players each making $N-1$ strategic decisions — extant tools from econometrics of games don't scale to problems of this size ($2^{N(N-1)}$; computation);

2. Agents are heterogeneous, in ways observed and unobserved (high dimensional parameterization);

3. Asymptotics of graphs $+$ many players games challenging (inference).

Our paper

1. Researcher specifies a (non-strategic) null model that allows for homophily on observed attributes as well as unobserved variation in the costs of sending and receiving links (degree heterogeneity);

2. Researcher also specifies an alternative *strategic* model (the network benefit function is freely specified);

3. We construct an *exact* test for the null that is *locally best* in the direction of the strategic alternative.

What can you learn from our test?

1. Under the maintained assumption that the true model is either in the null, or specified alternative model space, the researcher learns the sign of a strategic interaction parameter.

2. More generally, the test provides a valid assessment of null model adequacy.

3. Our test involves a new simulation algorithm which is also useful for exploratory data analysis in the networks setting.

Three technical challenges

1. *size control* (composite null: the null model includes K^2+2N parameters, a very high dimensional nuisance parameter);

2. finding the form of the *locally best* test (the model is incomplete under the alternative);

3. our test is *exact*, but in practice we need to *simulate* from the null distribution.

Basic Terms & Notation

- An **directed graph** $G(\mathcal{N}, \mathcal{A})$ consists of a set of **nodes** $\mathcal{N} = \{1, \ldots, N\}$ and a list of ordered pairs of nodes called **arcs/edges** $\mathcal{A} = \{\{i,j\}, \{k,l\}, \ldots\}$ for $i \neq j$, $k \neq l$ and $i, j, k, l \in \mathcal{N}$.

- A graph is conveniently represented by its **adjacency matrix** $\mathbf{D} = \left[D_{ij}\right]$ where

$$D_{ij} = \begin{cases} 1 & \text{if } \{i,j\} \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} . \qquad (1)$$

- No self-ties $\Rightarrow$ $\mathbf{D}$ is a binary matrix with a diagonal of so-called structural zeros.

Preferences

Let $\mathbf{d} \in \mathbb{D}$ be a feasible network.

The utility agent i gets from network $\mathbf{d}$ is

$$\nu_i\left(\mathbf{d}_i, \mathbf{d}_{-i}; \theta, \mathbf{U}_i\right) = \underbrace{\gamma_0 g_i\left(\mathbf{d}\right)}_{\text{Network Benefit}} \underbrace{- \sum_j d_{ij} c_{ij}\left(X_i, X_j; \delta, U_{ij}\right)}_{\text{Link Costs}}$$

where $g_i\left(\mathbf{d}\right)$ a known, but not necessarily closed-form, function of the network adjacency matrix (e.g., a measure of $i's$ centrality).

Agents weigh costs and benefits when deciding which links to send.

Network benefit example: “structural holes”

Inspired by formal model of Kleinberg et al (2008)...

Individuals that connect disparate groups gain “bridging” or intermediation benefits.

These benefits arise from lying on a (shortest) path connecting two agents not directly connected themselves.

If $d_{ki}d_{ij}\left(1 - d_{kj}\right) = 1$, then i serves as a “bridge" between k and j .

Network benefit example: "structural holes" (continued)

The summation $\sum_l d_{kl}d_{lj}\left(1 - d_{kj}\right)$ yields a count of the total number of bridging agents between k and j.

While agents benefit from serving as a bridge between two agents, these benefits decline in the number of other agents also serving as bridges for the same (directed) dyad.

This suggests a network payoff function of the form

$$g_i\left(\mathbf{d}\right) = \sum_j \sum_{k \neq j} \phi\left(d_{ki}d_{ij}\left(1 - d_{kj}\right), \sum_l d_{kl}d_{lj}\left(1 - d_{kj}\right)\right)$$

with $\phi\left(0, k\right) \equiv 0$ and $\phi\left(1, k\right) > \phi\left(1, k + 1\right) > 0$ for $k = 1, \ldots, N-2$.

Link Costs

The "cost" function captures the net costs and benefits for i of link ij that are invariant to other choices:

$$c_{ij}\left(X_i, X_j; \delta, U_{ij}\right) = -\left\{A_i + B_j + X_i' \Lambda_0 X_j - U_{ij}\right\}$$

1. A_i and B_j capture out- and in-degree heterogeneity;

2. X_i is a vector of K community membership dummies; $W_{ij}' \lambda_0 = X_i' \Lambda_0 X_j$: benefits of sending links may vary systematically with the sender and receiver types;

3. U_{ij} is an iid logistic random utility component.

Marginal Utility

Let $\nu_i(\mathbf{d}) \equiv \nu_i(\mathbf{d}_i, \mathbf{d}_{-i}; \theta, \mathrm{U}_i)$; the *marginal utility* of arc ij for agent i equals

$$MU_{ij}(\mathbf{d}) = \begin{cases} \nu_i(\mathbf{d}) - \nu_i(\mathbf{d} - ij) & \text{if } d_{ij} = 1 \\ \nu_i(\mathbf{d} + ij) - \nu_i(\mathbf{d}) & \text{if } d_{ij} = 0 \end{cases}$$

Marginal utility measures the utility gain (loss) to agent i from adding (subtracting) link ij holding the structure of all other links in the network constant (including any other links agent i directs).

Marginal Utility (continued)

The component of marginal utility associated with the network benefit function $g_i(\mathbf{d})$ plays an important role in our analysis:

$$s_{ij}(\mathbf{d}) = \begin{cases} g_i(\mathbf{d}) - g_i(\mathbf{d} - ij) & \text{if } d_{ij} = 1 \\ g_i(\mathbf{d} + ij) - g_i(\mathbf{d}) & \text{if } d_{ij} = 0 \end{cases}$$

Putting things together yields an expression for marginal utility of

$$\boxed{MU_{ij}(\mathbf{d}) = A_i + B_j + W_{ij}'\lambda_0 + \gamma_0 s_{ij}(\mathbf{d}) - U_{ij}}$$

Equilibrium

We assume that the observed network $\mathbf{D}$ coincides with the equilibrium outcome of an N-player complete information game. Each agent

1. observes $\left\{(A_i, B_i, X_i')\right\}_{i=1}^{N}$ and $\left\{U_{ij}\right\}_{i \neq j}$ and then

2. decides which, out of the $N-1$ other agents, to send links to.

3. Agents may play mixed strategies.

Equilibrium (continued)

A mixed strategy profile σ^* is a NE when $\theta = \theta_0$ and $\mathbf{U} = \mathbf{u}$, if for all $i = 1, \ldots, N$,

$$\nu_i \left(\sigma_i^*, \sigma_{-i}^*; \theta_0, \mathbf{u}_i \right) \geq \nu_i \left(\mathbf{d}_i, \sigma_{-i}^*; \theta_0, \mathbf{u}_i \right)$$

for all possible pure strategy selections $\mathbf{d}_i$.

We assume that the *observed* network $\mathbf{D}$ is either a pure strategy NE or in the support of a mixed strategy NE.

Notation Redux

Out- and *in-degree sequences* equal

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{\text{out}} \\ \mathbf{S}_{\text{int}} \end{pmatrix}' = \begin{pmatrix} D_{1+}, \ldots, D_{N+} \\ D_{+1}, \ldots, D_{+N} \end{pmatrix}.$$

Here $D_{+i} = \sum_j D_{ji}$ and $D_{i+} = \sum_j D_{ij}$ equal the in- and out-degree of agents $i = 1, \ldots, N$.

The $K \times K$ *cross-link matrix* equals

$$\mathbf{M} = \sum_i \sum_j D_{ij} X_i X_j'$$

This matrix summarizes the inter-group link structure in the network (homophily).

Notation Redux (continued)

Let $\mathbf{S}, \mathbf{M}$ be a degree sequence and cross-link matrix.

We say $\mathbf{S}, \mathbf{M}$ is *graphical* if there exists at least one arc set $\mathcal{A}$ such that $G(\mathcal{V}, \mathcal{A})$ is a simple directed graph with degree sequence $\mathbf{S}$ and cross link matrix $\mathbf{M}$.

We call any such network a *realization* of $\mathbf{S}, \mathbf{M}$ (open problem).

The set of all possible realizations of $\mathbf{S}, \mathbf{M}$ is denoted by $\mathbb{G}_{\mathbf{S},\mathbf{M}}$ ($\mathbb{D}_{\mathbf{S},\mathbf{M}}$).

Equilibrium selection mechanism

Let $\mathcal{N}(\mathbf{d},\mathbf{u};\theta)$ be a function which assigns, for $\mathbf{U}=\mathbf{u}$, a probability weight to network $\mathbf{d}$:

$$\mathcal{N}(\mathbf{d},\mathbf{u};\theta):\mathbb{D}_N\times\mathbb{R}^n\rightarrow[0,1]$$

In order for $\mathcal{N}(\mathbf{d},\cdot;\theta)$ to be a valid NE selection function it must satisfy some conditions.

For $\mathbf{U}=\mathbf{u}$ the realized vector of logistic link preference shocks and θ_0 the payoff function parameter, let $\mathbf{d}^*(\mathbf{u};\theta_0)$ be a pure strategy NE (or a network contained in the support of a mixed strategy NE) and $\mathbb{D}_N^*(\mathbf{u};\theta_0)$ be the set of all such networks.

Equilibrium selection mechanism (continued)

$\mathcal{N}(\mathbf{d}, \cdot; \theta)$ is such that:

1. $\mathcal{N}(\mathbf{d}, \mathbf{u}; \theta_0) \geq 0$ for all $\mathbf{d} \in \mathbb{D}_N^*(\mathbf{u}; \theta_0)$;

2. $\sum_{\mathbf{d} \in \mathbb{D}_N^*(\mathbf{u}; \theta_0)} \mathcal{N}(\mathbf{d}, \mathbf{u}; \theta_0) = 1$;

3. $\mathcal{N}(\mathbf{d}, \mathbf{u}; \theta_0) = 0$ for all $\mathbf{d} \in \mathbb{D}_N \backslash \mathbb{D}_N^*(\mathbf{u}; \theta_0)$.

Model Parameters

$\theta = (\gamma, \delta')'$ with:

γ - parameter of interest (strategic interaction);

$\delta = (\lambda', \mathbf{A}', \mathbf{B}')'$ – homophily/heterogeneity;

we also have $\mathcal{N}$, the equilibrium selection rule;

δ and $\mathcal{N}$ are (high dimensional) nuisance parameters.

Likelihood

If $\mathcal{N}(\mathbf{d}, \cdot; \theta)$ satisfies the conditions above, then the likelihood of observing network $\mathbf{D} = \mathbf{d}$ is

$$P(\mathbf{d}; \theta, \mathcal{N}) = \int_{\mathbf{u} \in \mathbb{R}^n} \mathcal{N}(\mathbf{d}, \mathbf{u}; \theta) f_{\mathbf{u}}(\mathbf{u}) \mathrm{d}\mathbf{u},$$

where $f_{\mathbf{u}}(\mathbf{u}) = \prod_{i \neq j} f_U(u_{ij})$ with

$$f_U(u) = e^u / [1 + e^u]^2$$

the logistic density.

For the likelihood to be well-defined we require that $\mathcal{N}(\mathbf{d}, \cdot; \theta)$ is measureable (Theorem 1).

Likelihood (Incompleteness)

Note that $s_{ij}(\mathbf{d})$ has finite range $\mathbb{S}$!

Example: $s_{ij}(\mathbf{d}) = d_{ji}$, such that agents prefer reciprocated links. Here $\mathbb{S} = \{0, 1\}$.

Can use $\mathbb{S}$ to partition the range of U_{ij} ($\mathbb{R}$) in *buckets*:

$$\left(-\infty, \mu_{ij}\right] \cup \left(\mu_{ij}, \mu_{ij} + \gamma\right] \cup \left(\mu_{ij} + \gamma, \infty\right)$$

with $\mu_{ij} = A_i + B_j + W_{ij}'\lambda_0$ the systematic "non-strategic" utility generated by arc ij.

Comment: when γ_0 is small the probability that U_{ij} falls into the inner bucket is low.

Likelihood (Incompleteness)

Three types of U_{ij} realizations:

1. If U_{ij} falls into the first *(outer) bucket*, then agent i *always* directs a link to j (irrespective of whether j reciprocates; strongly dominant strategy).

2. If U_{ij} falls into the *inner bucket,* then i sends a link only if j reciprocates ($\left(D_{ij}, D_{ji}\right) = (0,0)$ and/or $(1,1)$ depending on U_{ji}).

3. If U_{ij} falls into the last *(outer) bucket*, then agent i *never* directs a link to j.

Testing for Strategic Interaction

Let Δ denote a subset of the $K^2 + 2N$ dimensional Euclidean space in which δ_0 is, a priori, known to lie, and

$$\Theta_0 = \left\{ \left(\gamma, \delta' \right) : \gamma = 0, \delta \in \Delta \right\}.$$

Our null hypothesis is the *composite* one

$$H_0 : \theta \in \Theta_0 \tag{2}$$

since δ may range freely over $\Delta \subset \mathbb{R}^{K^2+2N}$ under the null.

Null Model

Null model is a variant of that studied by Charbonneau (2017), Graham (2017), Jochmans (2018) and others.

Links are conditionally independent with $P_0(\mathbf{d};\delta) \overset{def}{\equiv} P\left(\mathbf{d};(0,\delta')',\mathcal{N}_0\right)$ equal to

$$P_0(\mathbf{d};\delta) = \prod_{i=1}^{N}\prod_{j\neq i}\left[\frac{\exp\left(W_{ij}'\lambda + R_i'\mathbf{A} + R_j'\mathbf{B}\right)}{1+\exp\left(W_{ij}'\lambda + R_i'\mathbf{A} + R_j'\mathbf{B}\right)}\right]^{d_{ij}} \times \left[\frac{1}{1+\exp\left(W_{ij}'\lambda + R_i'\mathbf{A} + R_j'\mathbf{B}\right)}\right]^{1-d_{ij}}$$

with R_i an $N \times 1$ vector with 1 as its i^{th} element and zeros elsewhere.

Null Model (continued)

Note that $P_0(\mathbf{d};\delta)$ equals

$$P_0(\mathbf{d};\delta) = \int_{\mathbf{u}\in\mathbb{R}^n} \mathcal{N}_0(\mathbf{d},\mathbf{u};\theta)\, f_{\mathbf{u}}(\mathbf{u})\, \mathrm{d}\mathbf{u}$$

with

$$\begin{aligned}\mathcal{N}_0(\mathbf{d},\mathbf{u};\theta) = & \prod_i \prod_j \mathbf{1}\left(A_i + B_j + W_{ij}'\lambda \geq u_{ij}\right)^{d_{ij}} \\ & \times \mathbf{1}\left(A_i + B_j + W_{ij}'\lambda < u_{ij}\right)^{1-d_{ij}}.\end{aligned}$$

Things are more involved under the alternative.

Null Model: Exponential Family

The null model belongs to the exponential family:

$$P_0(\mathbf{d};\delta) = c(\delta)\,\text{exp}\left(\mathbf{t}'\delta\right)$$

with a (minimally) sufficient statistic for δ of

$$\mathbf{t} = \left(\text{vec}\left(\mathbf{m}'\right)', \mathbf{s}'_{\text{out}}, \mathbf{s}'_{\text{in}}\right)'.$$

In words, the $K^2 + N + N$ sufficient statistics are (i) the cross link matrix, (ii) the out-degree sequence and (iii) the in-degree sequence.

Null Model: Conditional Likelihood

Under H_0 the conditional likelihood of $\mathbf{D} = \mathbf{d}$ is

$$P_0\left(\mathbf{d} \middle| \mathbf{T} = \mathbf{t}\right) = \frac{1}{|\mathbb{D}_{\mathbf{s},\mathbf{m}}|}.$$

To simulate the distribution of a statistic under H_0 we need to be able to draw adjacency matrices (i.e., networks) uniformly at random from the set $\mathbb{D}_{\mathbf{s},\mathbf{m}}$.

This is a non-trivial problem. See Blitzstein & Diaconis (2010) and Tao (2016).

Test Formulation

In our setting, a test $\phi(\mathbf{D})$, will have size α if its null rejection probability (NRP) is less than or equal to α for *all* values of the nuisance parameter:

$$\sup_{\theta\in\Theta_0} \mathbb{E}_\theta\left[\phi(\mathbf{D})\right] = \sup_{\gamma=0,\delta\in\triangle} \mathbb{E}_\theta\left[\phi(\mathbf{D})\right] = \alpha.$$

Since δ is high dimensional, size control is non-trivial.

Intuition: transitivity/clustering example.

This motivates proceeding conditionally on $\mathbf{T}$ vs. using a single critical value.

Let $\mathbb{T} = \{(\mathbf{s},\mathbf{m}) : \mathbf{s},\mathbf{m}$ is graphical$\}$ be the set of possible $\mathbf{T}$.

Test Formulation (continued)

For each $\mathbf{t} \in \mathbb{T}$ we form a test with the property that, for all $\theta \in \Theta_0$,

$$\mathbb{E}_\theta \left[\phi(\mathbf{D}) | \mathbf{T} = \mathbf{t} \right] = \alpha.$$

Such an approach ensures *similarity* of our test since, by iterated expectations

$$\mathbb{E}_\theta \left[\phi(\mathbf{D}) \right] = \mathbb{E}_\theta \left[\mathbb{E}_\theta \left[\phi(\mathbf{D}) | \mathbf{T} \right] \right] = \alpha$$

for any $\theta \in \Theta_0$ (cf. Ferguson, 1967).

By proceeding conditionally we ensure the NRP is unaffected by the value of δ.

Test Formulation (continued)

By Ferguson (1967, Lemma 1, Section 3.6) $\mathbf{T}$ is a boundedly complete sufficient statistic for θ under the null.

By Ferguson (1967, Theorem 2, Section 5.4) *every* similar test will therefore take the form

$$\mathbb{E}_{\theta}\left[\phi\left(\mathbf{D}\right)\middle|\,\mathbf{T}=\mathbf{t}\right]=\alpha$$

for $\mathbf{t}\in\mathbb{T}$.

Therefore, if we desire similarity we can/must take the conditional approach.

A Conditional Test: Heuristic Approach

Let $R(\mathbf{D})$ be some statistics of the adjacency matrix, for example, the reciprocity index.

$$R(\mathbf{D}) = \frac{2\widehat{P}(\leftrightarrow)}{2\widehat{P}(\leftrightarrow) + \widehat{P}(\rightarrow)}. \tag{3}$$

A conditional test based upon $R(\mathbf{d})$ will have the critical function:

$$\phi(\mathbf{d}) = \begin{cases} 1 & R(\mathbf{d}) > c_\alpha(\mathbf{t}) \\ g_\alpha(\mathbf{t}) & R(\mathbf{d}) = c_\alpha(\mathbf{t}) \\ 0 & R(\mathbf{d}) < c_\alpha(\mathbf{t}) \end{cases}$$

where $c_\alpha(\mathbf{t})$ and $g_\alpha(\mathbf{t})$ are chosen to ensure correct size.

The null distribution of $R(\mathbf{D})$ corresponds to the one induced by a discrete uniform distribution on $\mathbb{D}_{\mathbf{s},\mathbf{m}}$.

A Conditional Test: Heuristic Approach

Two remaining challenges:

Its possible that the test based upon $R(\mathbf{d})$ will have good power to detect violations of the null in the direction of the alternative of interest, but there are no guarantees.

The cardinality of $\mathbb{D}_{\mathbf{s},\mathbf{m}}$ is generally intractably large – need a method for constructing uniform random draws from this set in order to approximate null distribution.

Locally Best Test

Under the alternative of strategic interaction the conditional likelihood is

$$P\left(\mathbf{d}\right|\mathbf{T}=\mathbf{t};\theta,\mathcal{N})=\frac{P\left(\mathbf{d};\theta,\mathcal{N}\right)}{\sum_{\mathbf{v}\in\mathbb{D}_{\mathbf{s},\mathbf{m}}}P\left(\mathbf{v};\theta,\mathcal{N}\right)}.$$

This likelihood is complicated and (logically) cannot be evaluated without specifying an explicit equilibrium selection mechanism.

Even then, it is not typically feasible to evaluate (see Graham and Pelican, 2022).

Locally Best Test

For each $\mathbf{t} \in \mathbb{T}$, we choose the critical function, $\phi(\mathbf{D})$ to maximize the *derivative* of the (conditional) power function

$$\beta(\gamma, \mathbf{t}) = \mathbb{E}\left[\phi(\mathbf{D}) \middle| \mathbf{T} = \mathbf{t}\right]$$

evaluated at $\gamma = 0$ subject to the (conditional) size constraint

$$\mathbb{E}_\theta\left[\phi(\mathbf{D}) \middle| \mathbf{T} = \mathbf{t}\right] = \alpha. \tag{4}$$

Such a $\phi(\mathbf{D})$ is *locally best* (Ferguson, 1967, Section 5.5).

Locally Best Test (continued)

Differentiating the power function we get

$$\left.\frac{\partial \beta(\gamma, \mathbf{t})}{\partial \gamma}\right|_{\gamma=0} = \mathbb{E}\left[\phi(\mathbf{D})\, \mathbb{S}_\gamma(\mathbf{D}|\, \mathbf{T}; \theta)|\, \mathbf{T} = \mathbf{t}\right] \qquad (5)$$

with $\mathbb{S}_\gamma(\mathbf{d}|\, \mathbf{t}; \theta)$ the conditional score function

$$\begin{aligned}\mathbb{S}_\gamma(\mathbf{d}|\, \mathbf{t}; \theta) &= \frac{1}{P_0(\mathbf{d}; \delta)} \left.\frac{\partial P(\mathbf{d}; \theta)}{\partial \gamma}\right|_{\gamma=0} - \sum_{\mathbf{v} \in \mathbb{D}_{\mathbf{s},\mathbf{m}}} \left.\frac{\partial P(\mathbf{v}; \theta)}{\partial \gamma}\right|_{\gamma=0} \\ &= \frac{1}{P_0(\mathbf{d}; \delta)} \left.\frac{\partial P(\mathbf{d}; \theta)}{\partial \gamma}\right|_{\gamma=0} + k(\mathbf{t})\end{aligned}$$

and $k(\mathbf{t})$ only depending on the data through $\mathbf{T} = \mathbf{t}$.

Locally Best Test (continued)

By the Neyman-Pearson lemma the test with critical function

$$\phi(\mathbf{d}) = \begin{cases} 1 & \frac{1}{P_0(\mathbf{d};\delta)} \frac{\partial P(\mathbf{d};\theta)}{\partial \gamma}\Big|_{\gamma=0} > c_\alpha(\mathbf{t}) \\ g_\alpha(\mathbf{t}) & \frac{1}{P_0(\mathbf{d};\delta)} \frac{\partial P(\mathbf{d};\theta)}{\partial \gamma}\Big|_{\gamma=0} = c_\alpha(\mathbf{t}) \\ 0 & \frac{1}{P_0(\mathbf{d};\delta)} \frac{\partial P(\mathbf{d};\theta)}{\partial \gamma}\Big|_{\gamma=0} < c_\alpha(\mathbf{t}) \end{cases}$$

where the values of $c_\alpha(\mathbf{t})$ and $g_\alpha(\mathbf{t}) \in [0, 1]$ are chosen to satisfy (4), will be locally best.

Locally Best Test (continued)

Several (serious) implementation challenges:

1. Form of the likelihood gradient $\left.\frac{\partial P(\mathbf{d};\theta)}{\partial \gamma}\right|_{\gamma=0}$ (incompleteness is an issue)?

2. Locally best test statistic may depend on nuisance parameters δ (manageable) and $\mathcal{N}$ (problematic)?

3. To find $c_\alpha(\mathbf{t})$ and $g_\alpha(\mathbf{t})$ we need to be able to simulate the (null) distribution of $\frac{1}{P_0(\mathbf{D};\delta)} \left.\frac{\partial P(\mathbf{D};\theta)}{\partial \gamma}\right|_{\gamma=0}$ conditional on $\mathbf{T} = \mathbf{t}$.

Derivative Calculation: Buckets

Recall that $\mathbb{S} = \{\underline{s}, s_1, \ldots, s_M, \overline{s}\}$ equals the set of possible values for the network benefit part of marginal utility, $s_{ij}(\mathbf{d})$, ordered from smallest to largest.

$\mathbb{S}$ induces a partition of $\mathbb{R}$. We call each element $b \in \mathbb{B}$ of this partition a *bucket*, buckets are naturally ordered:

$$\mathbb{R} = \left(-\infty, \mu_{ij} + \gamma\underline{s}\right] \cup \left(\mu_{ij} + \gamma\underline{s}, \mu_{ij} + \gamma s_1\right] \cup \cdots \\ \cup \left(\mu_{ij} + \gamma s_M, \mu_{ij} + \gamma\overline{s}\right] \cup \left(\mu_{ij} + \gamma\overline{s}, \infty\right).$$

All buckets, with the exception of the first and the last, we call *inner buckets*.

For any draw of the utility shifter we have $U_{ij} \in b$, $b \in \mathbb{B}$.

Derivative Calculation: Buckets (continued)

If a realization of U_{ij} is in bucket b, we say U_{ij} falls in (or is in) b.

We suppress the dependence of the partition on ij in the notation.

Observe that for $\gamma \approx 0$, the probability that U_{ij} falls into an inner bucket is close to zero.

Derivative Calculation: Buckets (continued)

Let the boldface subscripts $\mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots$ index the $n = N(N-1)$ directed dyads in arbitrary order (e.g., $\mathbf{i}$ maps to some ij and vice-versa).

Let $\mathbf{b} \in \mathbb{B}^n = \mathbb{B} \times \cdots \times \mathbb{B}$ and $\mathbf{U} = (U_{\mathbf{1}}, \ldots, U_{\mathbf{n}})'$.

We have that $\mathbf{U} \in \mathbf{b}$ for $\mathbf{b} \in \mathbb{B}^n$ so that each element of the n-vector of utility shifters $\mathbf{U}$ falls into a bucket.

Derivative Calculation: Likelihood (continued)

Using our bucket notation we can re-write the likelihood as:

$$P(\mathbf{d}; \theta, \mathcal{N}) = \sum_{\mathbf{b} \in \mathbb{B}^n} \int_{\mathbf{u} \in \mathbf{b}} \mathcal{N}(\mathbf{d}, \mathbf{u}; \theta) f_{\mathbf{U}}(\mathbf{u}) \, \mathrm{d}\mathbf{u} \tag{6}$$

For a given bucket combination $\mathbf{b} \in \mathbb{B}^n$, $\int_{\mathbf{u} \in \mathbf{b}} \mathcal{N}(\mathbf{d}, \mathbf{u}; \theta) f_{\mathbf{u}}(\mathbf{u}) \, \mathrm{d}\mathbf{u}$ gives the associated contribution to the likelihood of observing $\mathbf{D} = \mathbf{d}$.

Summation over all possible bucket combinations gives the overall likelihood of observing $\mathbf{D} = \mathbf{d}$.

Derivative Calculation: Likelihood (continued)

Let $\tilde{\mathbb{B}}^n$ be the set of bucket configurations *with two or more inner buckets*. Define

$$\tilde{P}(\mathbf{d};\theta,\mathcal{N}) = \sum_{\mathbf{b}\in\mathbb{B}^n\setminus\tilde{\mathbb{B}}^n} \int_{\mathbf{u}\in\mathbf{b}} \mathcal{N}(\mathbf{d},\mathbf{u};\theta)\, f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u}$$

$$Q(\mathbf{d};\theta,\mathcal{N}) = \sum_{\mathbf{b}\in\tilde{\mathbb{B}}^n} \int_{\mathbf{u}\in\mathbf{b}} \mathcal{N}(\mathbf{d},\mathbf{u};\theta)\, f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u}.$$

Trivially we have the decomposition

$$P(\mathbf{d};\theta,\mathcal{N}) = \tilde{P}(\mathbf{d};\theta,\mathcal{N}) + Q(\mathbf{d};\theta,\mathcal{N}).$$

Derivative Calculation

To calculate $\partial P(\mathbf{d};\theta,\mathcal{N})/\partial\gamma$ we show that for $\gamma \to 0$

$$P(\mathbf{d};\theta,\mathcal{N}) = \tilde{P}(\mathbf{d};\theta,\mathcal{N}) + \mathcal{O}\left(\gamma^2\right).$$

Furthermore we show that

$$\left.\frac{\partial P(\mathbf{d};\theta,\mathcal{N})}{\partial\gamma}\right|_{\gamma=0} = \left.\frac{\partial \tilde{P}(\mathbf{d};\theta,\mathcal{N})}{\partial\gamma}\right|_{\gamma=0}. \qquad (7)$$

Hence to derive the form of $\left.\frac{\partial P(\mathbf{d};\theta,\mathcal{N})}{\partial\gamma}\right|_{\gamma=0}$ we need only calculate $\left.\frac{\partial \tilde{P}(\mathbf{d};\theta,\mathcal{N})}{\partial\gamma}\right|_{\gamma=0}$.

This calculation is non-trivial, but doable (i.e., it is tedious).

Derivative Calculation

Only need to worry about cases where (i) no draws of U_{ij} are in inner buckets or (ii) just one draw (out of n) is.

In the first case every player has a strictly dominating strategy profile.

Strong preferences: regardless of other players' action it is either optimal, or not, to form specific links.

Network is uniquely defined: $\mathcal{N}(\mathbf{d}, \mathbf{u}; \theta)$ is either zero or one.

Derivative Calculation

Second case: if all but one component of $\mathbf{U}$ falls into the first or last bucket, then the resulting network is uniquely defined except for the presence or absence of one edge, say, ij.

For any such draw of $\mathbf{U}$, since all other links are formed according to a strictly dominating strategy, player i will either benefit from forming the link ij or not.

Hence $\mathcal{N}(\mathbf{d}, \mathbf{u}; \theta)$ is also either zero or one in this case as well.

Derivative Calculation

For small values of γ the derivative is driven by summands where the precise details of the (unspecified) equilibrium selection mechanism are *not* relevant.

Those summands where the form of $\mathcal{N}(\mathbf{d}, \mathbf{u}; \theta)$ *is* germane contribute very little to the derivative when γ is small.

We are able to differentiate the likelihood with respect to the strategic interaction parameter and evaluate that derivative for small γ (specifically for $\gamma = 0$).

Derivative Calculation: Likelihood (continued)

Lemma: $P(\mathbf{d};\theta)$ is twice differentiable with respect to γ at $\gamma = 0$. Its first derivative at $\gamma = 0$ is

$$\left.\frac{\partial P(\mathbf{d};\theta)}{\partial\gamma}\right|_{\gamma=0} = P_0(\mathbf{d};\delta) \times \left[\sum_{i\neq j} s_{ij}(\mathbf{d})\left\{d_{ij}\frac{f_U\left(\mu_{ij}\right)}{\int_{-\infty}^{\mu_{ij}} f_U(u)\,\mathrm{d}u} - \left(1-d_{ij}\right)\frac{f_U\left(\mu_{ij}\right)}{\int_{\mu_{ij}}^{\infty} f_U(u)\,\mathrm{d}u}\right\}\right]$$

With a little manipulation we can simplify:

$$\boxed{\frac{1}{P_0(\mathbf{d};\delta)}\left.\frac{\partial P(\mathbf{d};\theta)}{\partial\gamma}\right|_{\gamma=0} = \sum_{i\neq j}\left[d_{ij} - F_U\left(\mu_{ij}\right)\right] s_{ij}(\mathbf{d})}$$

where $F_U(u) = e^u/[1+e^u]$ is the logistic CDF.

Operational Details

Locally best test statistic is large when links which have low probability under the null, tend to form precisely where their "strategic utility" is high.

Controlling for heterogeneity appears to be important for power.

Lots of triangles vs. "surprising" triangles.

Operational Details

Although the form of the locally optimal statistic does not depend on $\mathcal{N}$ (equilibrium selection; phew!) it does depend on δ (heterogeneity).

Plugging in any $\delta \in \Delta$ results in an admissible test.

We take a "best guess" approach, replacing $\mu_{ij} = A_i + B_j + W_{ij}'\lambda$ with its JMLE $\widehat{\mu}_{ij}$ (cf., Graham, 2017; Dzemski, 2018; Yan et al., 2018).

This is ad hoc, but appears to work well in practice.

Operational Details (continued)

For $s = 1, \ldots, S$ we draw (uniformly at random) $\mathbf{V}_s \in \mathbb{D}_{\mathbf{s},\mathbf{m}}$ and calculate $\frac{1}{P_0(\mathbf{V_s};\widehat{\delta})} \frac{\partial P(\mathbf{V_s};(\gamma,\widehat{\delta}'))}{\partial \gamma}\Big|_{\gamma=0}$.

If $\frac{1}{P_0(\mathbf{D};\widehat{\delta})} \frac{\partial P\left(\mathbf{D};(\gamma,\widehat{\delta}')'\right)}{\partial \gamma}\Bigg|_{\gamma=0}$, observed in the network in hand, is greater than 95 percent of our simulated statistics we reject the null (and, possibly, “accept” the alternative).

Simulation Algorithm

We begin with $\mathbf{D}$ and randomly rewire it, preserving the cross link structure and degree sequence at each step.

Our MCMC converges to the null distribution, generating a uniform random draw from $\mathbb{D}_{\mathbf{S},\mathbf{M}}$.

Key references: Rao et al. (1996) and Tao (2015).

Our contribution is to also account for the cross-link group structure.

Importance sampling approach not possible (cf., Blitzstein and Diaconis, 2010).

Alternating Walks

A: Alternating Walk

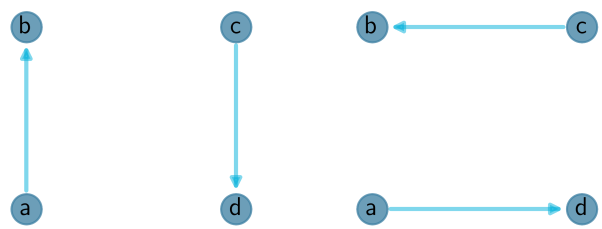
	a	b	c	d	e	f	g	h	i	j
a	0	1	0	0	1	0	0	0	0	0
b	0	0	0	0	0	0	0	0	0	0
c	0	0	0	1	0	0	0	1	0	0
d	0	0	0	0	0	0	0	0	0	0
e	0	0	1	0	0	0	0	0	0	0
f	0	0	0	0	0	0	0	1	1	0
g	0	0	0	0	0	0	0	0	0	0
h	0	0	1	0	0	0	1	0	0	0
i	0	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	1	0	1	0

B: Degree Sequence

	Indegree	Outdegree
a	0	2
b	1	0
c	2	2
d	1	0
e	1	1
f	0	2
g	2	0
h	2	2
i	2	0
j	0	2

Alternating Cycles

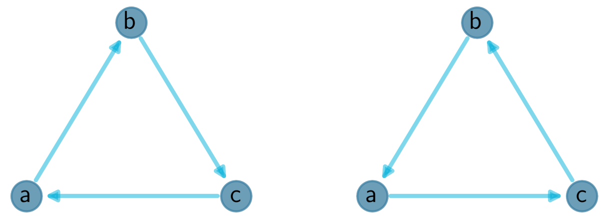
Alternating Rectangle



	a	b	c	d	a
i	1	2	3	4	5
sel. Pr	$\frac{1}{4}$	$\frac{1}{1}$	$\frac{1}{3}$	$\frac{1}{1}$	$\frac{1}{2}$

$\Pr(R_1) = \frac{1}{24}$

Compact Alternating Hexagon

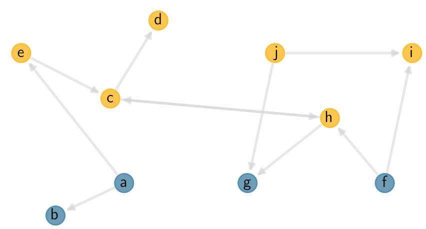


	a	b	c	a	b	c	a
i	1	2	3	4	5	6	7
sel. Pr	$\frac{1}{3}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{1}$

$\Pr(R_1) = \frac{1}{3}$

Schlaufen Sequences

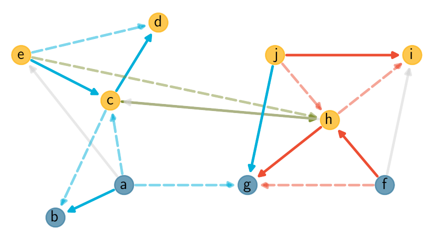
A: Network prior to edge swaps



D: Cross-link (PAM) matrix

	Blue	Gold
Blue	1	3
Gold	2	5

B: Network with
three schlaufen shown



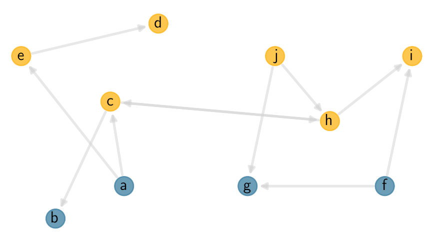
E: Three schlaufen

	j	g	a	b	c	d	e	c	a
i	1	2	3	4	5	6	7	8	9
sel. $\Pr$	$\frac{1}{10}$	$\frac{1}{2}$	$\frac{1}{7}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{1}$	$\frac{1}{8}$

	c	h	e
i	1	2	3
sel. $\Pr$	$\frac{1}{10}$	$\frac{1}{1}$	$\frac{1}{7}$

	f	h	j	i	h	g	f
i	1	2	3	4	5	6	7
sel. $\Pr$	$\frac{1}{10}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{1}$	$\frac{1}{7}$	$\frac{1}{1}$	$\frac{1}{6}$

C: Network after
edge swaps



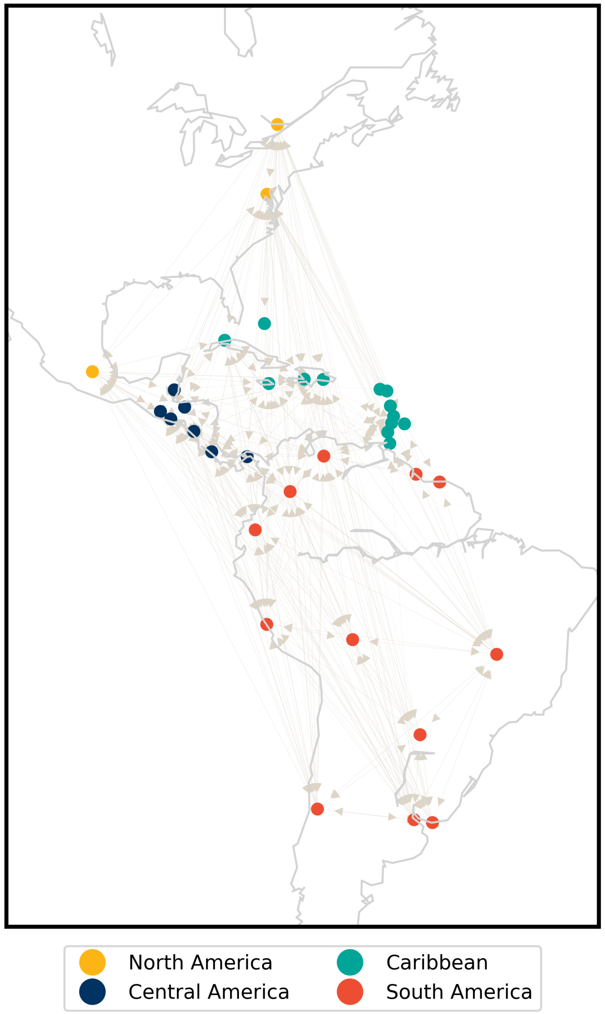
F: Violation matrices for
the three schlaufen

	Blue	Gold
Blue	-1	+1
Gold	+1	-1

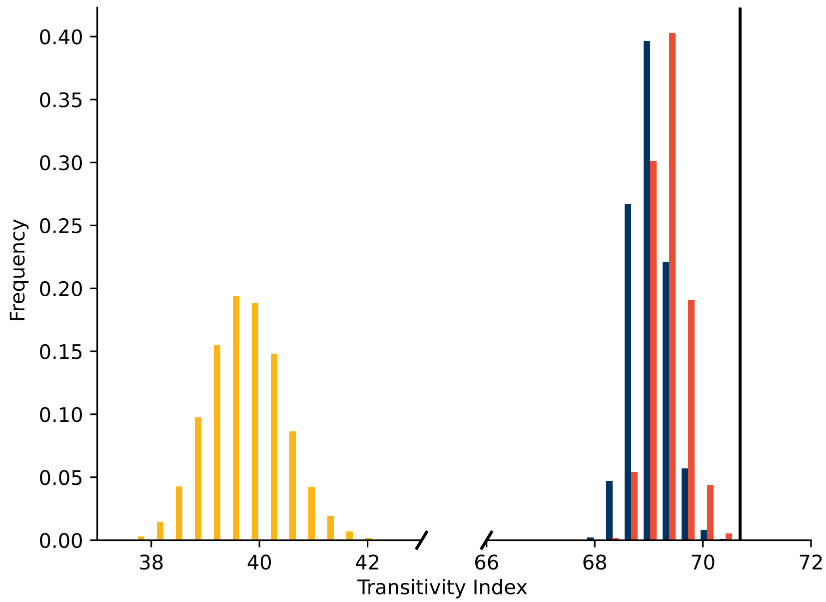
	Blue	Gold
Blue	0	0
Gold	0	0

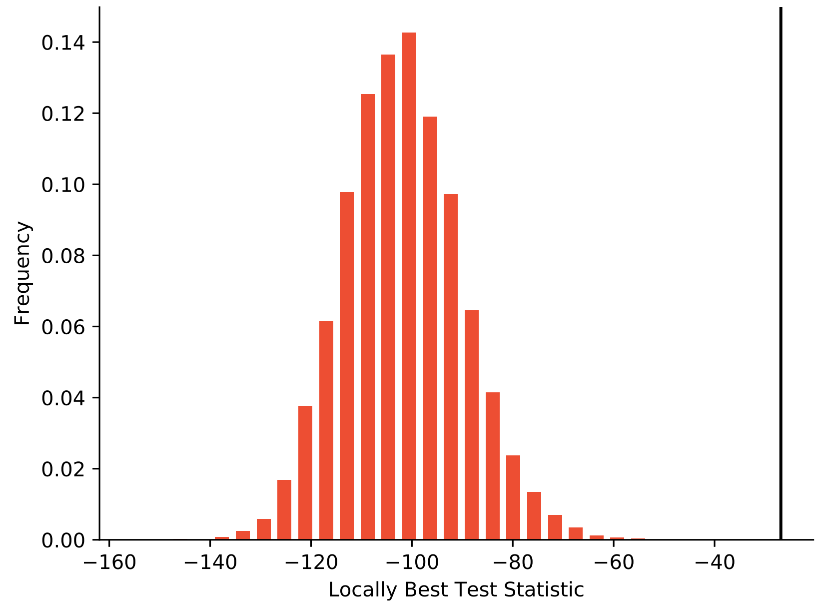
	Blue	Gold
Blue	+1	-1
Gold	-1	+1

Diplomacy Network of the Americas, 2005



Testing for strategic interaction in diplomatic relations





Wrapping-Up

We provide a method for assessing whether a specific form of strategic interaction characterizes the network in hand (under assumptions).

More generally we provide an exact test of model adequacy for the null model.

Our simulation algorithm also facilitates CMLE, exploratory data analysis etc.

Many open questions.