Testing for Externalities in Network Formation by Simulation

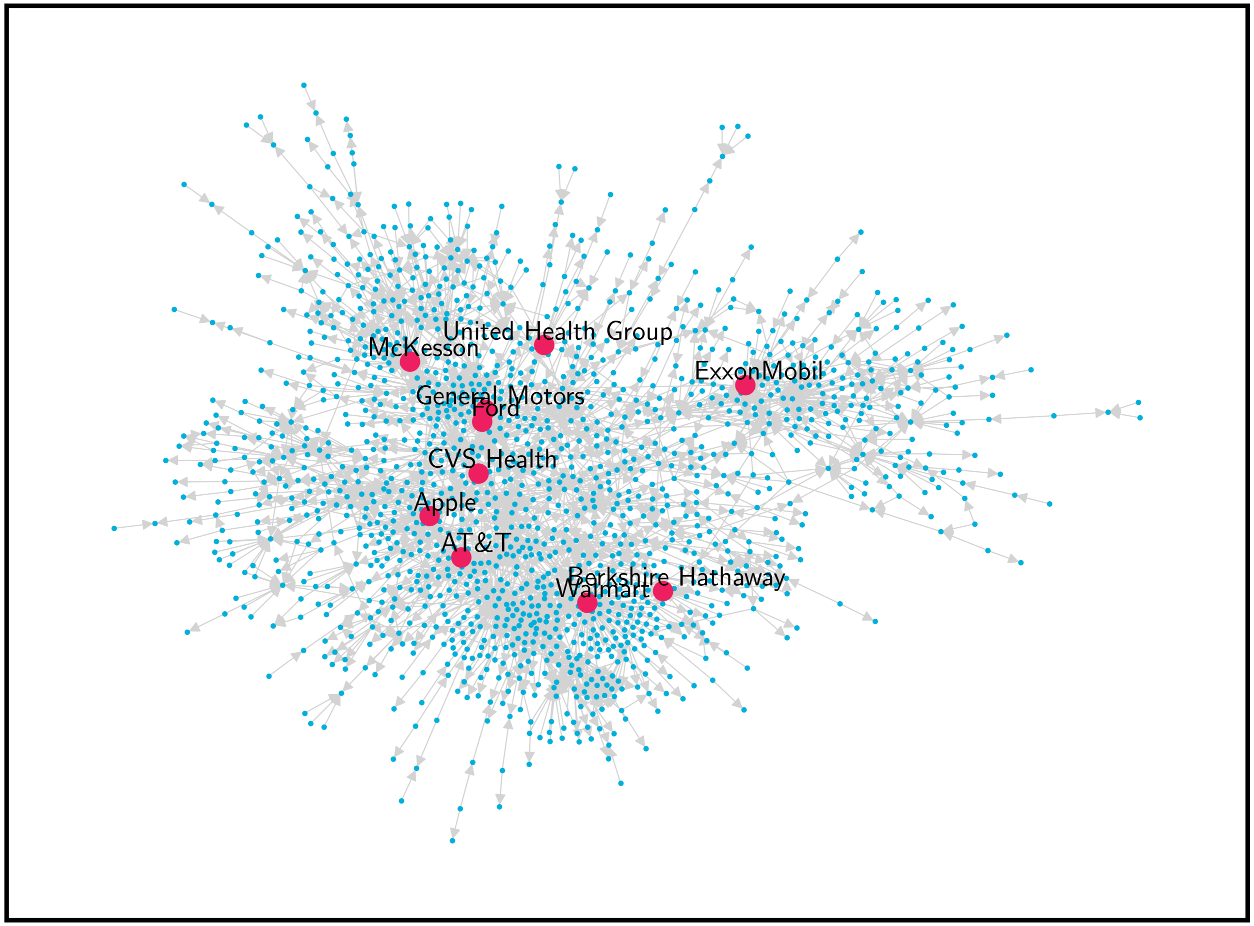
Econometric Methods for Social Spillovers and Networks

Universität St.Gallen, October 7 to 11, 2024

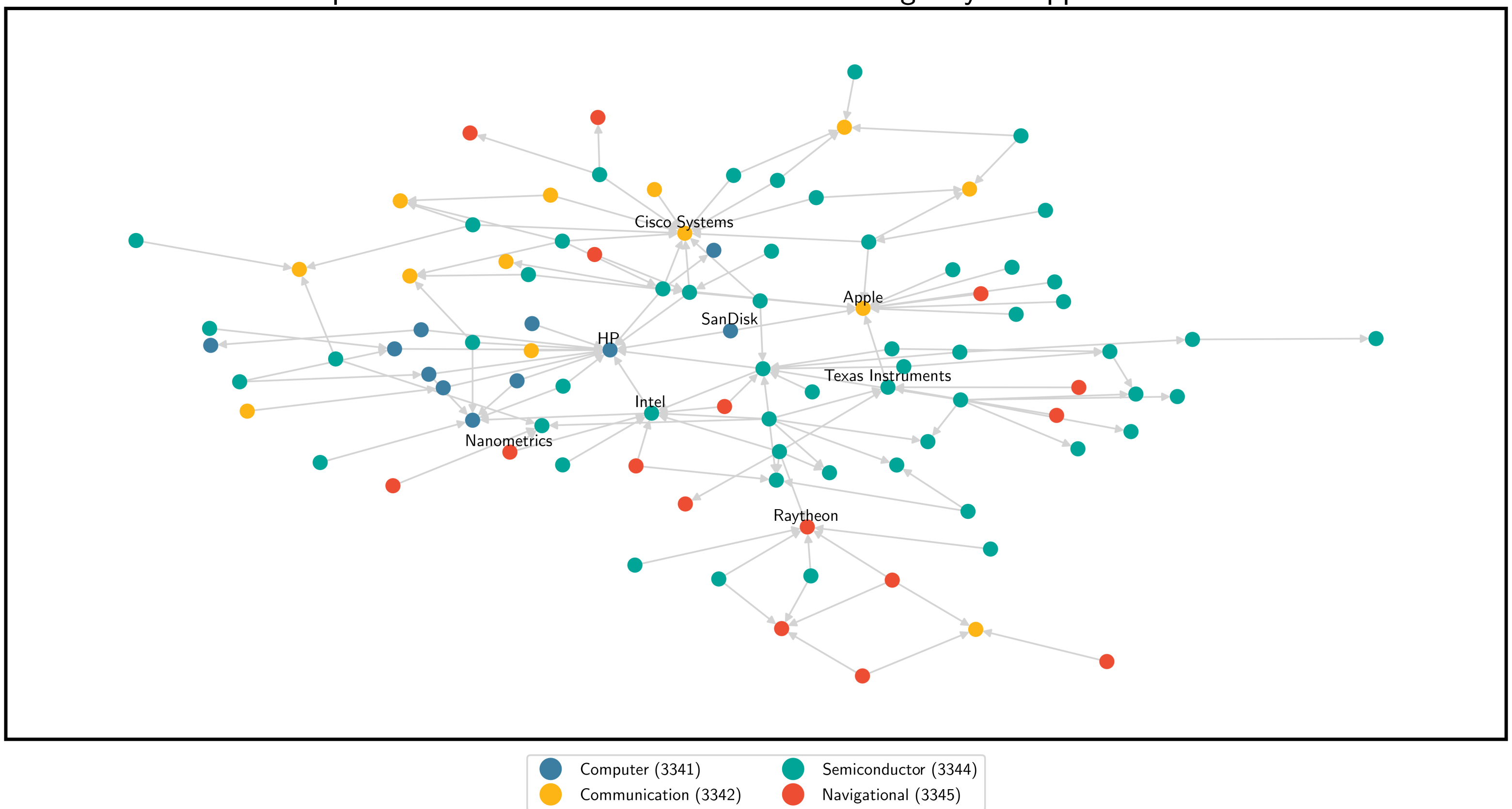
Bryan S. Graham

University of California - Berkeley

US Buyer-Supplier Network, 2015



Computer and Electronic Product Manufacturing Buyer-Supplier Network



Two classes of network formation models

Null model: the utility i generates by linking with j depends upon ego (i) and alter (j) attributes alone (attributes may be observed or unobserved).

- Stochastic Block Model

- β-Model

Alternative model: the utility generated by an i to j link *additionally varies* with the presence or absence of *other links* in the network.

- *Strategic* models

Research question

Can we determine whether the network in hand was generated according to null or alternative model?

Very little prior work in this space.

Why I care and you should (might?) too

With strategic behavior:

1. There may be multiple equilibrium network configurations.

2. The observed configuration may not maximize welfare.

3. Vertex removal (and/or local re-wirings) can trigger a process of link revision that is global in scope.

The effect of policies on the form of a network are very different under the null vs. the alternative.

Utility

Random utility framework *a la* McFadden (1973).

Let $\mathbf{d} \in \mathbb{D}$ be an *undirected* adjacency matrix. The utility agent i gets from some feasible network wiring $\mathbf{d}$ is

$$\nu_i\left(\mathbf{d}, \mathbf{U}; \gamma_0, \delta_0\right) = \sum_j d_{ij}\left[A_i + B_j + \gamma_0 s_{ij}\left(\mathbf{d}\right) - U_{ij}\right],$$

where:

1. A_i is a "extroversion effect";

2. B_j is a "popularity effect";

Utility (continued)

1. $s_{ij}(\mathbf{d}) = s_{ij}(\mathbf{d} - ij) = s_{ij}(\mathbf{d} + ij)$ is a network/strategic effect; can be used to model:

   (a) "rich-get-richer": $s_{ij}(\mathbf{d}) = \sum_k d_{jk}$;

   (b) transitivity: $s_{ij}(\mathbf{d}) = \sum_k d_{ik} d_{jk}$;

4. $\left\{U_{ij}\right\}_{i \neq i}$ idiosyncratic utility-shifter (i.i.d. logistic)

Pelican and Graham (2019) work with a much more general model.

Utility (continued)

The marginal utility for agent i associated with (possible) edge (i,j) is

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

where $\mathbf{D} - ij$ is the adjacency matrix associated with the network obtained after deleting edge (i,j)...

...and $\mathbf{D} + ij$ the one obtained via link addition.

Equilibrium

Network is undirected.

It is convenient to assume utility is transferable.

Use *pairwise stable with transfers* equilibrium concept from Bloch and Jackson (2006).

(Pairwise stability with Transfers) The network $G(\mathcal{V}, \mathcal{E})$ is pairwise stable with transfers if
(i) $\forall (i,j) \in \mathcal{E}(G)\,,\ MU_{ij}(\mathbf{D}) + MU_{ji}(\mathbf{D}) \geq 0$
(ii) $\forall (i,j) \notin \mathcal{E}(G)\,,\ MU_{ij}(\mathbf{D}) + MU_{ji}(\mathbf{D}) < 0$

Equilibrium (continued)

The marginal utility agent i gets from a link with j is

$$MU_{ij}(\mathbf{d}, \mathbf{U}; \gamma_0, \delta_0) = A_i + B_j + \gamma_0 s_{ij}(\mathbf{d}) - U_{ij}.$$

Pairwise stability then implies that, conditional on the realizations of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{U}$ and the value of externality parameter γ_0, the observed network must satisfy, for $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$,

$$D_{ij} = \mathbf{1}\left(\tilde{A}_i + \tilde{A}_j + \gamma_0 \tilde{s}_{ij}(\mathbf{D}) \geq \tilde{U}_{ij}\right) \tag{2}$$

with $\tilde{A}_i = A_i + B_i$, $\tilde{s}_{ij}(\mathbf{D}) = s_{ij}(\mathbf{D}) + s_{ji}(\mathbf{D})$ and $\tilde{U}_{ij} = U_{ij} + U_{ji}$.

Defines a system of $\binom{N}{2} = \frac{1}{2}N(N-1)$ *nonlinear* simultaneous equations

Equilibrium: Fixed-Point Representation

Consider, similar to Miyauchi (2016), the mapping $\varphi(\mathbf{D}) : \mathbb{D}_N \to \mathbb{I}_{\binom{N}{2}}$:

$$\varphi(\mathbf{d}) \equiv \begin{bmatrix} \mathbf{1}\left(\tilde{A}_1 + \tilde{A}_2 + \gamma_0 \tilde{s}_{12}(\mathbf{d}) \geq U_{12}\right) \\ \mathbf{1}\left(\tilde{A}_1 + \tilde{A}_3 + \gamma_0 \tilde{s}_{13}(\mathbf{d}) \geq U_{13}\right) \\ \vdots \\ \mathbf{1}\left(\tilde{A}_{N-1} + \tilde{A}_N + \gamma_0 \tilde{s}_{N-1N}(\mathbf{d}) \geq U_{N-1N}\right) \end{bmatrix}. \qquad (3)$$

The observed adjacency matrix corresponds to the fixed point

$$\mathbf{D} = \mathsf{vech}^{-1}\left[\varphi(\mathbf{D})\right].$$

There may be other $\mathbf{d} \in \mathbb{D}_N$ such that $\mathbf{d} = \mathsf{vech}^{-1}\left[\varphi(\mathbf{d})\right]$.

Existence can be shown using Tarski's fixed point theorem for some forms of $s_{ij}(\mathbf{d})$.

Testing goal: challenges

Goal is to construct a test of the no strategic interaction ($\gamma_0 = 0$) null.

Three key challenges:

1. null is composite – nuisance parameter $\delta = \mathbf{A}$ is high dimensional (worry: size distortion);

2. can't evaluate likelihood under the alternative (worry: how to maximize power?);

3. characterizing/simulating null distribution (worry: feasibility).

Testing goal: solutions

1. Apply exponential family theory (Ferguson, 1967; Lehmann & Romano, 2005).

2. Find *locally* best test:

   (a) derivative of likelihood w.r.t to γ difficult to compute (incompleteness);

   (b) exploit insights from the econometrics of games (e.g., Tamer, 2003; Bajari *et al*. 2010a,b).

3. Use methods for (constrained) network simulation (e.g., Sinclair, 1993)

Constructing the Test

Under the null we have, for $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$,

$$\Pr\left(D_{ij} = 1 \middle| \tilde{\mathbf{A}}\right) = \frac{\exp\left(\tilde{A}_i + \tilde{A}_j\right)}{1 + \exp\left(\tilde{A}_i + \tilde{A}_j\right)},$$

which corresponds to the β-model of network formation.

Probability of $\mathbf{D} = \mathbf{d}$ takes the exponential family form

$$P_0\left(\mathbf{d}; \tilde{\mathbf{A}}\right) = c\left(\tilde{\mathbf{A}}\right) \exp\left(\mathbf{d}_+' \tilde{\mathbf{A}}\right)$$

with $\mathbf{d}_+ = (d_{1+}, \ldots, d_{N+})$ equal to the degree sequence of the network.

Constructing the Test (continued)

Let $\mathbb{D}_{N,\mathbf{d}_+}$ denote the set of all undirected $N \times N$ adjacency matrices with degree counts also equal to $\mathbf{d}_+$.

$\left|\mathbb{D}_{N,\mathbf{d}_+}\right|$ denotes the size, or cardinality, of this set.

Constructing the Test (continued)

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

$$P_0\left(\mathbf{d} \middle| \mathbf{D}_+ = \mathbf{d}_+\right) = \frac{1}{\left|\mathbb{D}_{N,\mathbf{d}_+}\right|}.$$

Under the null of no externalities *all networks with identical degree sequences are equally probable*.

This insight will form the basis of our test.

Constructing the Test (continued)

Let $T(\mathbf{d})$ be some statistic of the adjacency matrix $\mathbf{D} = \mathbf{d}$, say its transitivity index.

Test critical function equals

$$\phi(\mathbf{d}) = \begin{cases} 1 & T(\mathbf{d}) > c_\alpha(\mathbf{d}_+) \\ g_\alpha(\mathbf{d}_+) & T(\mathbf{d}) = c_\alpha(\mathbf{d}_+) \\ 0 & T(\mathbf{d}) < c_\alpha(\mathbf{d}_+) \end{cases} .$$

We will reject the null if our statistic exceeds some critical value, $c_\alpha(\mathbf{d}_+)$ and accept it – or fail to reject it – if our statistic falls below this critical value.

Constructing the Test (continued)

The critical value $c_\alpha(\mathbf{d}_+)$ is chosen to set the rejection probability of our test under the null equal to α (i.e., to control size).

In order to find the appropriate value of $c_\alpha(\mathbf{d}_+)$ we need to know the distribution of $T(\mathbf{D})$ under the null.

This distribution is straightforward to characterize if we proceed *conditional* on the degree sequence observed in the network in hand.

Constructing the Test (continued)

Under the null all possible adjacency matrices with degree sequence $\mathbf{d}_+$ are equally probable.

The null distribution of $T(\mathbf{D})$ therefore equals its distribution across all these matrices.

By enumerating all the elements of $\mathbb{D}_{N,\mathbf{d}_+}$ and calculating $T(\mathbf{d})$ for each one, we could directly – and exactly – compute this distribution.

In practice this is not (generally) computationally feasible.

Constructing the Test (continued)

If we could efficiently enumerate the elements of $\mathbb{D}_{N,\mathbf{d}_+}$ we would find $c_\alpha(\mathbf{d}_+)$ by solving

$$1-\alpha = \frac{\sum_{\mathbf{D}\in\mathbb{D}_{\mathbf{N},\mathbf{d}_+}} \mathbf{1}\left(T(\mathbf{D}) \leq c_\alpha(\mathbf{d}_+)\right)}{\left|\mathbb{D}_{N,\mathbf{d}_+}\right|}$$

Alternatively we might instead calculate the p-value:

$$\Pr\left(T(\mathbf{D}) \geq T(\mathbf{d}) \middle| \mathbf{D} \in \mathbb{D}_{\mathbf{N},\mathbf{d}_+}\right) = \frac{\sum_{\mathbf{D}\in\mathbb{D}_{\mathbf{N},\mathbf{d}_+}} \mathbf{1}\left(T(\mathbf{D}) \geq T(\mathbf{d})\right)}{\left|\mathbb{D}_{N,\mathbf{d}_+}\right|}$$

Choosing $T(\mathbf{d})$

Pelican and Graham (2019) show how to choose $T(\mathbf{d})$ to maximize power against local alternatives.

This is hard because one must work with the likelihood of the network under the alternative (which is incomplete).

In practice – as is common with randomization tests – can also pick a test statistic intuitively.

For example $T(\mathbf{d})$ might be the transitivity index.

Testing: Intuition

If the probability that measured transitivity, in a network randomly drawn from the null distribution, lies above observed transitivity is very low...

...we take that as evidence against the β-model and “reject”.

Testing

- This approach to testing is
  - very precise about its description of the null hypothesis;
  - exact.

- We have motivated this test via a particular alternative (and can optimize power vis-a-vis it), but rejection may occur for many reasons.

- ...at minimum the choice of statistic should be guided by researcher intuitions about what departures from the null model are of particular concern.

Sampling from $\mathbb{D}_{N,\mathbf{d}_+}$

- Direct enumeration of all the elements of $\mathbb{D}_{N,\mathbf{d}_+}$ is generally not feasible.
- Need a method of sampling from $\mathbb{D}_{N,\mathbf{d}_+}$ *uniformly* and also estimating its size.
- We will implement an approximation of the ideal test.

Sampling from $\mathbb{D}_{N,\mathbf{d}_+}$ (continued)

- Blitzstein and Diaconis (2011) develop a sequential importance sampling algorithm for (effectively) uniformly sampling from $\mathbb{D}_{N,\mathbf{d}_+}$

- Two challenges:
  - how to generate a random draw from $\mathbb{D}_{N,\mathbf{d}_+}$;
  - how to do so uniformly (importance weights).

Graphical Integer Sequences

- To construct $\mathbf{D}$ we begin with a matrix of zeros and sequentially add links to it until its rows and columns sum to the target degree sequence.

- Problem is that unless links are added carefully it is easy to get "stuck" (cf., Snijders, 1991).

- The key is to check whether residual degree sequences are graphical as you add links (avoid dead ends).

- $\mathbf{D}_{+} = (2, 2, 1)$ is not graphic

Graphical Integer Sequences (continued)

- Erdos and Gallai (1961) showed $\mathbf{D}_+$ is graphical if and only if $\sum_{i=1}^{N} D_{i+}$ is even and

$$\sum_{i=1}^{k} D_{i+} \leq k\,(k-1) + \sum_{i=k+1}^{N} \min\left(k, D_{i+}\right) \ \text{for each}\ k \in \{1, \ldots, N\}\,.$$

Graphical Integer Sequences (continued)

Necessity:

- even: if i is linked to j, then the link is counted in both D_{i+} and D_{j+}.

- For any set S of k agents, there can be at most $\binom{k}{2} = \frac{1}{2}k(k-1)$ links between them (first term).

- For the $N - k$ agents $i \notin S$, there can be at most $\min(k, D_{i+})$ links from i to agents in S.

Graphical Integer Sequences (continued)

Sufficiency of the condition is (evidently) much harder to show.

Erdos and Gallai Theorem provides a simple test for graphicality of a degree sequence.

The next theorem, due to Havel (1955) and Hakimi (1962), shows that this test may be applied recursively.

A Recursive Test

Theorem: *(Havel-Hakimi) Let $D_{i+} > 0$, if $\mathbf{D}_+$ does not have at least D_{i+} positive entries other than $\boldsymbol{i}$ it is not graphical. Assume this condition holds. Let $\tilde{\mathbf{D}}_+$ be a degree sequence of length $N-1$ obtained by*

[i] deleting the i^{th} entry of $\mathbf{D}_+$ and

[ii] subtracting 1 from each of the D_{i+} highest elements in $\mathbf{D}_+$ (aside from the i^{th} one).

$\mathbf{D}_+$ is graphical if and only if $\tilde{\mathbf{D}}_+$ is graphical. If $\mathbf{D}_+$ is graphical, then it has a realization where agent $\boldsymbol{i}$ is connected to any of the D_{i+} highest degree agents (other than $\boldsymbol{i}$).

Blitzstein and Diaconis Procedure

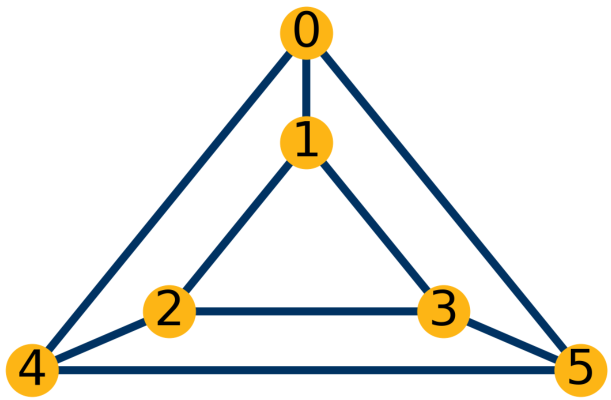
- Start with lowest degree agent (with at least one link).

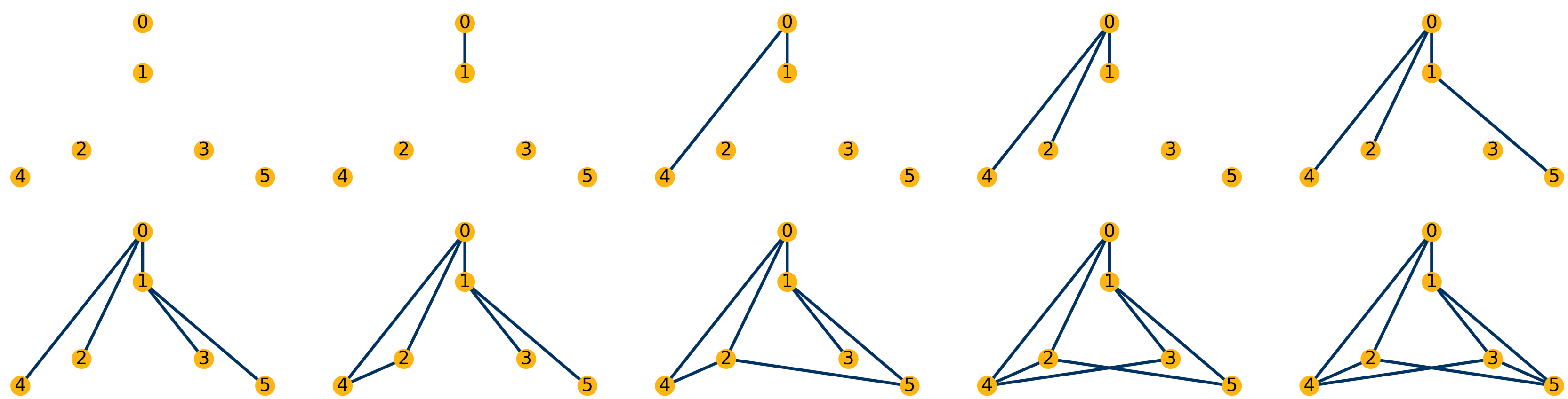
- (Randomly) Link this agent with high degree agents.

- A one is subtracted from the lowest degree agent's, as well as her chosen partners', degree counts.

- Continue until the **residual degree sequence** is zero.

3-regular (i.e., cubic graph)




0
1
2
3
4
5

Blitzstein and Diaconis Procedure (continued)

- Consider the example

$$
\begin{aligned}
(3,3,3,3,3,3) &\to (2,2,3,3,3,3) \to (1,2,3,3,2,3) \to (0,2,2,3,2,3) \\
&\to (0,1,2,3,2,2) \to (0,0,2,2,2,2) \to (0,0,1,2,1,2) \\
&\to (0,0,0,2,1,1) \to (0,0,0,1,0,1) \to (0,0,0,0,0,0)\,.
\end{aligned}
$$

- Now imagine that in the 8th step instead of linking agent 3 with agent 4, agents 4 and 5 were linked.

Blitzstein and Diaconis Procedure (continued)

- This would have resulted in a residual degree sequence of $(0, 0, 0, 2, 0, 0)$, which is *not* graphic.

- Algorithm doesn't allow this to occur by checking for whether the residual degree sequence associated with a candidate link is graphical.

Blitzstein and Diaconis Procedure (continued)

- Let $\left(\oplus_{i_1,\ldots,i_k}\mathbf{D}_+\right)$ be the vector obtained by adding a one to the $i_1, \ldots, i_k$ elements of $\mathbf{D}_+$:

$$\left(\oplus_{i_1,\ldots,i_k}\mathbf{D}_+\right)_j = \begin{cases} D_{j+} + 1 & \text{for } j \in \{i_1, \ldots, i_k\} \\ D_{j+} & \text{otherwise} \end{cases}$$

- Let $\left(\ominus_{i_1,\ldots,i_k}\mathbf{D}_+\right)$ be the vector obtained by subtracting one from the $i_1, \ldots, i_k$ elements of $\mathbf{D}_+$:

$$\left(\ominus_{i_1,\ldots,i_k}\mathbf{D}_+\right)_j = \begin{cases} D_{j+} - 1 & \text{for } j \in \{i_1, \ldots, i_k\} \\ D_{j+} & \text{otherwise} \end{cases}$$

Blitzstein and Diaconis Procedure (continued)

Algorithm: A sequential algorithm for constructing a random graph with degree sequence $\mathbf{d}_+ = (d_{1+}, \ldots, d_{N+})'$ is

1. Let $\mathbf{D}$ be an empty adjacency matrix.

2. If $\mathbf{D}_+ = \mathbf{0}$ terminate with output $\mathbf{D}$

3. Choose the agent i with minimal positive degree d_{i+}.

4. Construct a list of candidate partners

$$J = \left\{ j \neq i \,:\, \mathbf{D}_{ij} = \mathbf{D}_{ji} = 0 \text{ and } \ominus_{i,j} \mathbf{d}_+ \text{ graphical} \right\}.$$

5. Pick a partner $j \in J$ with probability proportional to its degree in $\mathbf{d}_+$.

6. Set $\mathbf{D}_{ij} = \mathbf{D}_{ji} = 1$ and update $\mathbf{d}_+$ to $\ominus_{i,j}\mathbf{d}_+$.

7. Repeat steps 4 to 6 until the degree of agent i is zero.

8. Return to step 2.

The input for the algorithm is the target degree sequence $\mathbf{d}_+$ and the output is an undirected adjacency matrix $\mathbf{D}$ with $\mathbf{D}'\iota = \mathbf{d}_+$.

Importance Weights

- The Blitzstein and Diaconis (2010) procedure delivers a random draw from $\mathbb{D}_{N,\mathbf{d}_+}$, but not a *uniform* random draw.

- Construct importance weights in order to compute expectations using the correct reference distribution.

- Let $\mathbb{Y}_{N,\mathbf{d}_+}$ denote the set of all possible sequences of links generated by the algorithm given input $\mathbf{d}_+$.

Importance Weights (continued)

- Let $\mathcal{D}(Y)$ be the adjacency matrix induced by link sequence Y.
  - Let Y and Y' are equivalent if $\mathcal{D}(Y) = \mathcal{D}(Y')$.

- We can partition $\mathbb{Y}_{N,\mathbf{d}_+}$ into a set of equivalence classes whose number coincides with the cardinality of $\mathbb{D}_{N,\mathbf{d}_+}$.

Importance Weights (continued)

- Let $c(Y)$ denote the number of possible link sequences produced by the algorithm that produce Y's end point adjacency matrix.

- Let $i_1, i_2, \ldots, i_M$ be the sequence of agents chosen in step 3 of the algorithm in which Y is the output.

Importance Weights (continued)

- Let $a_1, \ldots, a_m$ be the degrees of $i_1, \ldots, i_M$ at the time when each agent was *first* selected in step 3.

- Blitzstein and Diaconis show that:

$$c\left(Y\right) = \prod_{k=1}^{M} a_k!$$

Importance Weights (continued)

Consider two equivalent link sequences Y and Y'.

Because links are added to vertices by minimal degree (see Step 3), the sequences $i_1, i_2, \ldots, i_M$ coincide for Y and Y'.

This means that *the exact same links*, albeit perhaps in a different order, are added at each "stage" of the algorithm (i.e., when the algorithm iterates through steps 4 to 7 repeatedly for a given agent).

The number of different ways to add agent i_k's links during such a "stage" is simply $a_k!$ and hence $c(Y) = \prod_{k=1}^{M} a_k!$

Importance Weights (continued)

- Let $\sigma(Y)$ be the probability that the algorithm produces link sequence Y.
- $\sigma(Y)$ is easy to compute:
  - each time a link in step 5 is chosen we record the probability with which it was chosen.
  - this equals the residual degree of the chosen agent divided by the sum of the residual degrees of all agents in the choice set.
  - the product of all these probabilities equals $\sigma(Y)$.

Importance Weights (continued)

We have that $\mathbb{E}\left[\frac{\pi(\mathcal{D}(Y))}{c(Y)\sigma(Y)}\mathbf{1}\left(T\left(\mathcal{D}\left(Y\right)\right) > T\left(\mathbf{d}\right)\right)\right]$ equals

$$
\begin{aligned}
&= \sum_{y \in \mathbb{Y}_{N,\mathbf{d}}} \frac{\pi\left(\mathcal{D}\left(y\right)\right)}{c\left(y\right)\sigma\left(y\right)}\mathbf{1}\left(T\left(\mathcal{D}\left(Y\right)\right) > T\left(\mathbf{d}\right)\right)\sigma\left(y\right) \\
&= \sum_{y \in \mathbb{Y}_{N,\mathbf{d}}} \frac{\pi\left(\mathcal{D}\left(y\right)\right)}{c\left(y\right)}\mathbf{1}\left(T\left(\mathcal{D}\left(Y\right)\right) > T\left(\mathbf{d}\right)\right) \\
&= \sum_{\mathbf{D} \in \mathbb{D}_{N,\mathbf{d}_+}} \sum_{\{y\,:\,\mathcal{D}(y)=\mathbf{D}\}} \frac{\pi\left(\mathbf{D}\right)}{c\left(y\right)}\mathbf{1}\left(T\left(\mathbf{D}\right) > T\left(\mathbf{d}\right)\right) \\
&= \sum_{\mathbf{D} \in \mathbb{D}_{N,\mathbf{d}_+}} \pi\left(\mathbf{D}\right)\mathbf{1}\left(T\left(\mathbf{D}\right) > T\left(\mathbf{d}\right)\right) \\
&= \mathbb{E}_{\pi}\left[\mathbf{1}\left(T\left(\mathbf{D}\right) > T\left(\mathbf{d}\right)\right)\right].
\end{aligned}
$$

Importance Weights (continued)

Here $\pi(\mathbf{D})$ is the probability attached to the adjacency matrix $\mathbf{D} \in \mathbb{D}_{N,\mathbf{d}_+}$ in the target distribution over $\mathbb{D}_{N,\mathbf{d}_+}$.

The ratio $\pi(\mathcal{D}(Y))/c(Y)\sigma(Y)$ is called the likelihood ratio or the *importance weight*.

We would like $\pi(\mathbf{D}) = 1/\left|\mathbb{D}_{N,\mathbf{d}_+}\right|$ for all $\mathbf{D} \in \mathbb{D}_{N,\mathbf{d}_+}$.

If we set $\pi(\mathbf{D}) = 1$ we see that $\mathbb{E}\left[\frac{1}{c(Y)\sigma(Y)}\right] = \left|\mathbb{D}_{N,\mathbf{d}_+}\right|$. This suggests the analog estimator for $\left|\mathbb{D}_{N,\mathbf{d}_+}\right|$ of

$$\left|\hat{\mathbb{D}_{N,\mathbf{d}_+}}\right| = \left[\frac{1}{B}\sum_{b=1}^{B}\frac{1}{c(Y_b)\,\sigma(Y_b)}\right] \tag{4}$$

Importance Weights (continued)

These results suggest we estimate a p-value for our test by

$$\hat{\rho}_{T(\mathbf{G})} = \left[\frac{1}{B}\sum_{b=1}^{B}\frac{1}{c\left(Y_b\right)\sigma\left(Y_b\right)}\right]^{-1} \times \left[\frac{1}{B}\sum_{b=1}^{B}\frac{1}{c\left(Y_b\right)\sigma\left(Y_b\right)}\mathbf{1}\left(T\left(\mathbf{D}_b\right) > T\left(\mathbf{d}\right)\right)\right]$$

An attractive feature is that the importance weights need only be estimated up to a constant.

This feature is useful when dealing with numerical overflow issues that can arise when $\left|\mathbb{D}_{N,\mathbf{d}_+}\right|$ is too large to estimate.

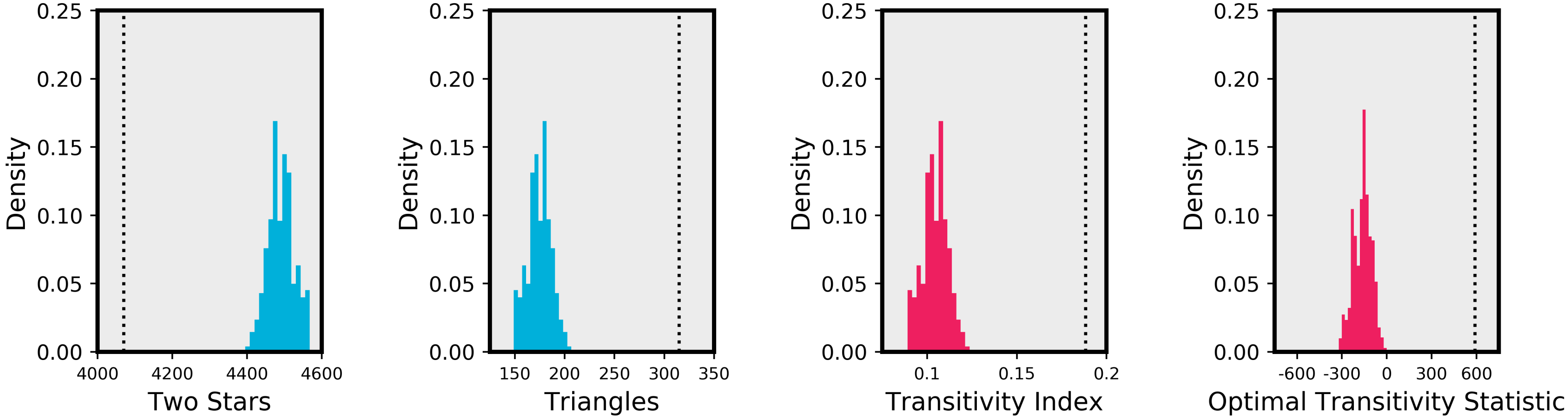
Importance Weights (continued)

- The ratio $\pi\left(\mathbf{D}\left(Y\right)\right)/c\left(Y\right)\sigma\left(Y\right)$ is called the likelihood ratio or the **importance weight.**

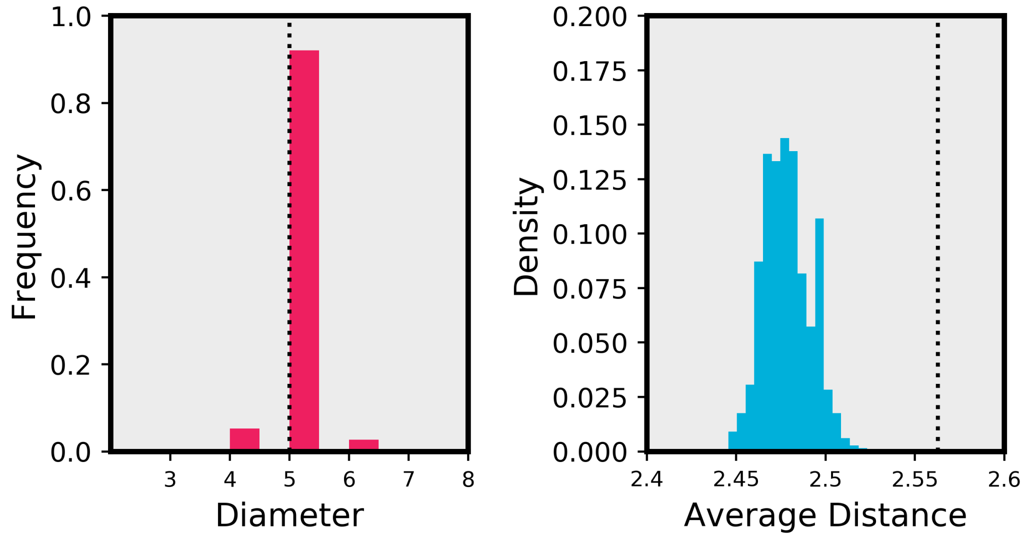
- Our random network draws are not uniform from the set of interest.

- The importance weights correct for the fact that we are sampling from the wrong distribution.

Nyakatoke Example



Nyakatoke Example (continued)



Wrap-Up

- While using the β-model as a reference model is restrictive it
  - is a natural starting point for hypothesis testing;
  - suggests that an investment in computation skills is likely to be valuable to anyone doing empirical work.

- It might be of interest to condition on additional features of the network in hand...

- ...see Pelican and Graham (2019).