Econometrica Supplementary Material

SUPPLEMENT TO "TESTING FOR DIFFERENCES IN STOCHASTIC NETWORK STRUCTURE"

(*Econometrica*, Vol. 90, No. 3, May 2022, 1205–1223)

ERIC AUERBACH
Department of Economics, Northwestern University

SECTION B contains additional results about the power properties of the tests based on $T_{2\to 2}$ and $S_{\infty\to 1}$. Section C contains results from Monte Carlo experiments. Section D contains details about the applications and extensions described in Sections 2.3 and 6.1 of the main text.

APPENDIX B: ADDITIONAL RESULTS

B.1. *Near-Necessity of Rate Conditions*

The rate conditions $T_{2\to 2}(F_1, F_2)/\tau \to \infty$ and $T_{\infty\to 1}(F_1, F_2)/\sigma \to \infty$ are close to necessary.

THEOREM 3: *For any sequence of positive real numbers $\delta_N \to \infty$*

$$\inf_{\substack{F_1,F_2:\delta_N[T_{2\to 2}(F_1,F_2)/\tau]\to\infty \\ \text{and } \tau/\sqrt{\ln(N)}\to\infty}} P\left((R+1)^{-1}\left(1+\sum_{r\in[R]}\mathbb{1}\left\{T_{2\to 2}\left(D_1^r, D_2^r\right) \geq T_{2\to 2}(D_1, D_2)\right\}\right) \leq \alpha\right)$$
$$\to \alpha$$

and

$$\inf_{\substack{F_1,F_2:\delta_N[T_{\infty\to 1}(F_1,F_2)/\sigma]\to\infty \\ \text{and } \sigma/\sqrt{\ln(N)}\to\infty}} P\left((R+1)^{-1}\left(1+\sum_{r\in[R]}\mathbb{1}\left\{S_{\infty\to 1}\left(D_1^r, D_2^r\right) \geq S_{\infty\to 1}(D_1, D_2)\right\}\right) \leq \alpha\right)$$
$$\to \alpha.$$

The statement of Theorem 3 differs from that of Theorems 1 and 2 in two ways. The first is that the first rate conditions have been changed from $[T_{2\to 2}(F_1, F_2)/\tau] \to \infty$ and $[T_{\infty\to 1}(F_1, F_2)/\sigma] \to \infty$ to $\delta_N[T_{2\to 2}(F_1, F_2)/\tau] \to \infty$ and $\delta_N[T_{\infty\to 1}(F_1, F_2)/\sigma] \to \infty$, respectively. That is, the infima are taken over a (slightly) larger class of sequences in H_1. The second difference is the conclusion that the power of the tests no longer tends to one. In fact, the limiting power of the tests may be no greater than α. To prove the theorem, the main work is in constructing a sequence of alternatives such that $T_{2\to 2}(F_1, F_2)/\tau \to 0$ slower than δ_N^{-1}, and $T_{2\to 2}(D_1, D_2)$ and $T_{2\to 2}(D_1^r, D_2^r)$ converge to the same nondegenerate distribution. Intuitively, Theorem 3 states that the tests proposed in Theorems 1 and 2 cannot detect differences between random graph models that are too similar, in the sense that $T_{2\to 2}(F_1, F_2)$ or $T_{\infty\to 1}(F_1, F_2)$ are too close to 0.

Eric Auerbach: eric.auerbach@northwestern.edu

 https://doi.org/10.3982/ECTA18093

PROOF OF THEOREM 3: I demonstrate the claim only for the test based on the $2 \to 2$ norm since the proof of that based on the $\infty \to 1$ norm is identical. The proof is constructive in that, for any sequence $\delta_N \to \infty$, it specifies a specific sequence of distribution function matrices F_1 and F_2, depending on δ_N, such that $\delta_N T_{2\to2}(F_1, F_2)/\tau \to \infty$ or

$$\delta_N \frac{\max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in[N]} \left(\sum_{j\in[N]} (F_{ij,1}(s) - F_{ij,2}(s)) \varphi_j \right)^2}}{\max_{r\in[R]} \max_{j\in[N]} \sqrt{\sum_{i\in[N]} [F_{ij,1}(s) + F_{ij,2}(s) - 2F_{ij,1}(s)F_{ij,2}(s)]}} \to \infty$$

and

$$P\left((R+1)^{-1} \left(1 + \sum_{r\in[R]} \mathbb{1}\left\{ T_{2\to2}(D_1^r, D_2^r) \geq T_{2\to2}(D_1, D_2) \right\} \right) \leq \alpha \right) \to \alpha.$$

The proof has three steps. The first step is to specify F_1 and F_2. For an arbitrary $\varepsilon > 0$, define $A_{1-\varepsilon} = [\lceil (1-\varepsilon)N \rceil]$ and $A_\varepsilon = [N] \setminus A_{1-\varepsilon}$. That is, let $A_{1-\varepsilon}$ index the first $\lceil (1-\varepsilon)N \rceil$ agents in the sample and A_ε the last $\lfloor \varepsilon N \rfloor$. Suppose $F_{ij,1} = F_{ij,2}$ for $i, j \in A_{1-\varepsilon}$ with $F_{ij,1}$ and $F_{ij,2}$ uniformly bounded away from 0 and 1, $F_{ij,1} = F_{ij,2} = 0$ for $i \in A_\varepsilon$ and $j \in A_{1-\varepsilon}$ (or $i \in A_{1-\varepsilon}$ and $j \in A_\varepsilon$), and $F_{ij,1} = 1 + F_{ij,2}$ for $i, j \in A_\varepsilon$.

The second step is to fix $\varepsilon = (\delta_N N)^{-1/2}$. Since $T_{2\to2}$ is $O(N\varepsilon)$ and τ is $O(\sqrt{N})$ by construction from the first step, it follows that $T_{2\to2}(F_1, F_2)/\tau \to 0$, but $\delta_N T_{2\to2}(F_1, F_2)/\tau \to \infty$.

The third step is to apply the triangle inequality twice. The first application gives

$$\begin{aligned} &\max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in[N]} \left(\sum_{j\in[N]} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s}) \varphi_j \right)^2} \\ &\quad \geq \max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in A_{1-\varepsilon}} \left(\sum_{j\in A_{1-\varepsilon}} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s}) \varphi_j \right)^2} \\ &\quad\quad - \max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in A_\varepsilon} \left(\sum_{j\in A_\varepsilon} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s}) \varphi_j \right)^2}. \end{aligned}$$

The second application gives

$$\begin{aligned} &\max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in[N]} \left(\sum_{j\in[N]} (\mathbb{1}_{D^r_{ij,1}\leq s} - \mathbb{1}_{D^r_{ij,2}\leq s}) \varphi_j \right)^2} \\ &\quad \leq \max_{s\in\mathbb{R}} \max_{\varphi\in\mathcal{S}^N} \sqrt{\sum_{i\in A_{1-\varepsilon}} \left(\sum_{j\in A_{1-\varepsilon}} (\mathbb{1}_{D^r_{ij,1}\leq s} - \mathbb{1}_{D^r_{ij,2}\leq s}) \varphi_j \right)^2} \end{aligned}$$

$$+\max_{s\in\mathbb{R}}\max_{\varphi\in\mathcal{S}^N}\sqrt{\sum_{i\in A_\varepsilon}\Bigl(\sum_{j\in A_\varepsilon}(\mathbb{1}_{D^r_{ij,1}\le s}-\mathbb{1}_{D^r_{ij,2}\le s})\varphi_j\Bigr)^2}.$$

for any $r\in[R]$. Both

$$\max_{s\in\mathbb{R}}\max_{\varphi\in\mathcal{S}^N}\sqrt{\sum_{i\in A_\varepsilon}\Bigl(\sum_{j\in A_\varepsilon}(\mathbb{1}_{D_{ij,1}\le s}-\mathbb{1}_{D_{ij,2}\le s})\varphi_j\Bigr)^2}\quad\text{and}$$

$$\max_{r\in[R]}\max_{s\in\mathbb{R}}\max_{\varphi\in\mathcal{S}^N}\sqrt{\sum_{i\in A_\varepsilon}\Bigl(\sum_{j\in A_\varepsilon}(\mathbb{1}_{D^r_{ij,1}\le s}-\mathbb{1}_{D^r_{ij,2}\le s})\varphi_j\Bigr)^2}$$

are bounded by $N\varepsilon$ by construction, and thus are $o(\sqrt{N})$ by the second step. However,

$$\max_{s\in\mathbb{R}}\max_{\varphi\in\mathcal{S}^N}\sqrt{\sum_{i\in A_{1-\varepsilon}}\Bigl(\sum_{j\in A_{1-\varepsilon}}(\mathbb{1}_{D_{ij,1}\le s}-\mathbb{1}_{D_{ij,2}\le s})\varphi_j\Bigr)^2}/\sqrt{N}\quad\text{and}$$

$$\max_{s\in\mathbb{R}}\max_{\varphi\in\mathcal{S}^N}\sqrt{\sum_{i\in A_{1-\varepsilon}}\Bigl(\sum_{j\in A_{1-\varepsilon}}(\mathbb{1}_{D^r_{ij,1}\le s}-\mathbb{1}_{D^r_{ij,2}\le s})\varphi_j\Bigr)^2}/\sqrt{N}$$

are bounded away from 0 by the lower bound in Lemma 1. Since they are identically distributed and nondegenerate by construction, the result follows. *Q.E.D.*

B.2. *Pointwise Consistency for Regular Alternatives*

Under certain conditions, the power of the tests from Theorems 1 and 2 tend to one whenever the difference between F_1 and F_2 is regular in the sense that there exist two nontrivially sized subcommunities $I_N, J_N \subseteq [N]$ such that the probabilities that any agent in I_N links to any agent in J_N all either increase or decrease with t. The hypothesis of this theorem does not specify rate conditions that depend on operator norms, and so may be easier to interpret and apply in practice. I demonstrate its use with the concrete example of Section 2.

THEOREM 4: *Suppose there exists $I_N, J_N \subseteq [N]$ with $\liminf_{N\to\infty}\frac{|I_N|\wedge|J_N|}{N}>0$, $s\in\mathbb{R}$, and $\rho_N>0$ such that either for all $i\in I_N$ and $j\in J_N$, $(F_{ij,1}(s)-F_{ij,2}(s))>\rho_N$ or for all $i\in I_N$ and $j\in J_N$, $(F_{ij,1}(s)-F_{ij,2}(s))<-\rho_N$. Then the power of the test from Theorem 1 converges to one if $\rho_N N/\ln(N)\to\infty$. The power of the test from Theorem 2 converges to one if $\rho_N N\to\infty$.*

The benefit of Theorem 4 relative to Theorems 1 and 2 is that its hypothesis does not use rate conditions that depend on operator norms. To illustrate its use, I sketch two simple testing problems with models based on the concrete example from Section 2. Recall for this example that $F_{ij,t}(s)=1-G_{ij,t}(f_t(\alpha_{i,t},\alpha_{j,t},w_{ij,t}))$. Suppose, for example, that the idiosyncratic errors are identically distributed for all agent pairs and networks, the agent-pair attributes are the same for the two networks, and the community link function has the form $f(\alpha_{i,t},\alpha_{j,t},w_{ij})=\Lambda(\alpha_{i,t}+\alpha_{j,t}+w_{ij}\beta)$ for some unknown vector β and strictly monotonic function Λ. Then the hypothesis of Theorem 4 is satisfied if there exists an I_N

with $\liminf_{N\to\infty} I_N/N > 0$ such that $|\alpha_{i,1} - \alpha_{i,2}| > \rho_N$ for all $i \in I_N$. That is, under these conditions, the tests of Theorems 1 and 2 eventually (correctly) reject the null hypothesis that the two networks have the same collection of agent-specific effects when it is false.

Alternatively, suppose that the idiosyncratic errors are identically distributed, the agent-specific effects and agent-pair attributes are the same for the two networks, and the community link function has the form $f_t(\alpha_i, \alpha_j, w_{ij}) = \Lambda_t(\alpha_i, \alpha_j) + w_{ij}\beta$ for some functions $\{\Lambda_t\}_{t\in[2]}$ and vector β. Then the hypothesis of Theorem 4 is satisfied if $\Lambda_1(\alpha_i, \alpha_j)$ and $\Lambda_2(\alpha_i, \alpha_j)$ disagree on $I_N \times J_N$ with $\liminf_{N\to\infty}(I_N \wedge J_N)/N > 0$. That is, under these conditions, the tests from Theorems 1 and 2 eventually (correctly) reject the null hypothesis that the two networks have the same community link function when it is false.

PROOF OF THEOREM 4: The claim is proven by checking the hypotheses of Theorems 1 and 2. This is done in three steps. The first step is to demonstrate that the assumption that there exists $I_N, J_N \subseteq [N]$ with $\liminf_{N\to\infty} \frac{|I_N|\wedge|J_N|}{N} > 0$ and $\rho_N > 0$ such that for all $i \in I_N, j \in J_N$ there exists s such that $(F_{ij,1}(s) - F_{ij,2}(s)) > \rho_N$ or for all $i \in I_N, j \in J_N$ there exists s such that $(F_{ij,1}(s) - F_{ij,2}(s)) < -\rho_N$ implies that $T_{2\to2}(F_1, F_2) \geq \rho_N N$ and $T_{\infty\to1}(F_1, F_2) \geq \rho_N N^2$ eventually ($N \to \infty$). Write $\delta = \liminf_{N\to\infty} \frac{|I_N|\wedge|J_N|}{N} > 0$. Then

$$
\begin{aligned}
T_{2\to2}(F_1, F_2) &= \max_{s\in\mathbb{R}} \max_{\varphi:\|\varphi\|_2=1} \left(\sum_{i\in[N]}\left(\sum_{j\in[N]}\big(F_{ij,1}(s) - F_{ij,2}(s)\big)\varphi_j\right)^2\right)^{1/2} \\
&\geq \left(\sum_{i\in I_N}\left(\sum_{j\in J_N}\frac{\big(F_{ij,1}(s) - F_{ij,2}(s)\big)}{\sqrt{\sum_{j\in J_N} 1}}\right)^2\right)^{1/2} \geq \rho_N N\delta^2
\end{aligned}
$$

eventually and

$$
\begin{aligned}
T_{\infty\to1}(\rho_N F_1, \rho_N F_2) &= \max_{s\in\mathbb{R}} \max_{\varphi:\|\varphi\|_\infty=1} \sum_{i\in[N]}\left|\sum_{j\in[N]}(F_{ij,1}(s) - F_{ij,2}(s)\varphi_j\right| \\
&= \max_{s\in\mathbb{R}} \max_{\varphi:\|\varphi\|_\infty=1} \max_{\psi:\|\psi\|_\infty=1} \sum_{i\in[N]}\sum_{j\in[N]}\big(F_{ij,1}(s) - F_{ij,2}(s)\big)\varphi_j\psi_i \\
&\geq \max_{s\in\mathbb{R}} \max_{\varphi\in\{0,1\}^N} \max_{\psi\in\{0,1\}^N} \left|\sum_{i\in[N]}\sum_{j\in[N]}\big(F_{ij,1}(s) - F_{ij,2}(s)\big)\varphi_j\psi_i\right| \\
&\geq \max_{s\in\mathbb{R}}\left|\sum_{i\in I_N}\sum_{j\in J_N}\big(F_{ij,1}(s) - F_{ij,2}(s)\big)\right| \geq \rho_N N^2\delta^2
\end{aligned}
$$

eventually where the first inequality follows from the fact that for any $N \times N$ dimensional matrix X_s,

$$
\begin{aligned}
&\max_{\varphi\in\{0,1\}^N} \max_{\psi\in\{0,1\}^N}\left|\sum_{i\in[N]}\sum_{j\in[N]} X_{ij,s}\varphi_j\psi_i\right| \\
&= \max_{\varphi\in\{-1,1\}^N} \max_{\psi\in\{-1,1\}^N}\left|\sum_{i\in[N]}\sum_{j\in[N]} X_{ij,s}\left(\frac{\varphi_j + 1}{2}\right)\left(\frac{\psi_i + 1}{2}\right)\right|
\end{aligned}
$$

$$
\begin{aligned}
&= \max_{\varphi\in\{-1,1\}^N}\max_{\psi\in\{-1,1\}^N}\left|\sum_{i\in[N]}\sum_{j\in[N]}[X_{ij,s}\varphi_j\psi_i + X_{ij,s} + X_{ij,s}\varphi_j + X_{ij,s}\psi_i]/4\right| \\
&\le \max_{\varphi\in\{-1,1\}^N}\max_{\psi\in\{-1,1\}^N}\sum_{i\in[N]}\sum_{j\in[N]} X_{ij,s}\varphi_j\psi_i = \max_{\varphi:\|\varphi\|_\infty=1}\max_{\psi:\|\psi\|_\infty=1}\sum_{i\in[N]}\sum_{j\in[N]} X_{ij,s}\varphi_j\psi_i.
\end{aligned}
$$

The second step is to observe that the assumption that there exists $I_N, J_N \subseteq [N]$ with $\liminf_{N\to\infty}\frac{|I_N|\wedge|J_N|}{N} > 0$ and $\rho_N > 0$ such that for all $i \in I_N$, $j \in J_N$ there exists s such that $(F_{ij,1}(s) - F_{ij,2}(s)) > \rho_N$ or for all $i \in I_N$, $j \in J_N$ there exists s such that $(F_{ij,1}(s) - F_{ij,2}(s)) < -\rho_N$ implies that

$$
\delta\sqrt{\rho_N N} \le \tau \le \sqrt{2}\sqrt{\rho_N N} \quad \text{and} \quad \delta\sqrt{\rho_N N^3} \le \sigma \le \sqrt{2}\sqrt{\rho_N N^3}
$$

eventually. Without loss of generality, suppose that $(F_{ij,1}(s) - F_{ij,2}(s)) > \rho_N$. The two upper bounds then follow from the fact that $F_{ij,t}(s)$ is bounded in $[0, 1]$. The first lower bound follows from

$$
\begin{aligned}
\tau &= \max_{s\in\mathbb{R}}\max_{i\in[N]}\sqrt{\sum_{j\in[N]}\bigl(F_{ij,1}(s) + F_{ij,2}(s) - 2F_{ij,1}(s)F_{ij,2}(s)\bigr)} \\
&\ge \max_{s\in\mathbb{R}}\max_{i\in I_N}\sqrt{\sum_{j\in J_N}\bigl(\bigl[F_{ij,1}(s) - F_{ij,2}(s)\bigr] + 2F_{ij,2}(s)\bigl(1 - F_{ij,1}(s)\bigr)\bigr)}
\end{aligned}
$$

and the fact that $2F_{ij,2}(s)(1 - F_{ij,1}(s))$ is not negative. Similarly,

$$
\begin{aligned}
\sigma &= \max_{s\in\mathbb{R}}\sum_{i\in[N]}\sqrt{\sum_{j\in[N]}\bigl(F_{ij,1}(s) + F_{ij,2}(s) - 2F_{ij,1}(s)F_{ij,2}(s)\bigr)} \\
&\ge \max_{s\in\mathbb{R}}\sum_{i\in I_N}\sqrt{\sum_{j\in J_N}\bigl(\bigl[F_{ij,1}(s) - F_{ij,2}(s)\bigr] + 2F_{ij,2}(s)\bigl(1 - F_{ij,1}(s)\bigr)\bigr)}
\end{aligned}
$$

implies the second lower bound.

The third step is to observe that steps 1 and 2 imply that

$$
T_{2\to 2}(F_1, F_2)/\tau \ge \sqrt{\rho_N N}\delta^2/\sqrt{2} \quad \text{and} \quad T_{\infty\to 1}(F_1, F_2)/\sigma \ge \sqrt{\rho_N N}\delta^2/\sqrt{2}
$$

eventually so that $T_{2\to 2}(F_1, F_2)/\tau \to \infty$ and $T_{\infty\to 1}(F_1, F_2)/\sigma \to \infty$ so long as $\rho_N N \to \infty$. Since

$$
\sigma/\sqrt{\ln(N)} \ge \delta\sqrt{\rho_N N^3/\ln(N)}
$$

eventually, $\rho_N N \to \infty$ also implies that $\sigma/\sqrt{\ln(N)} \to \infty$ and so the hypothesis of Theorem 2 is satisfied. Since

$$
\tau/\sqrt{\ln(N)} \ge \delta\sqrt{\rho_N N/\ln(N)},
$$

eventually, strengthening the rate condition to $\rho_N N/\ln(N) \to \infty$ implies that $\tau/\sqrt{\ln(N)} \to \infty$ so that the hypothesis of Theorem 1 is also satisfied. This demonstrates the claim. *Q.E.D.*

B.3. *Bounds on Power*

The arguments underlying the proofs of Theorems 1–2 can be modified to provide bounds on the power of the two tests. The upper bounds are relevant to the case in which the effect size is small relative to the reference distribution and the lower bounds are relevant to the case in which the effect size is large relative to the reference distribution. The upper bounds do not depend on α because the (lower) bounds on the reference distribution given by Lemmas 1 and 2 used to construct them hold exactly.

The bounds depend on the parameters $E[T_{2\to 2}(D_1, D_2)]$, $E[T_{\infty\to 1}(D_1, D_2)]$, τ, and σ as defined in Section 4.2.1 of the main text. The properties of these parameters depend on F_1 and F_2. I give the bounds associated with two relatively simple examples in Sections B.3.1 and B.3.2 below.

THEOREM 5: *Let* $(x)_+^2 = x^2 \mathbb{1}_{x>0}$, $\gamma \in [0, 1/2]$ *be arbitrary, and* $K = 1.783$. *The power of the* α*-sized test based on* $T_{2\to 2}$ *(given in Theorem* 1*) is bounded from above by*

$$\exp\left(-\left(\tau - \sqrt[4]{-\frac{N}{2}\ln\left(\frac{\gamma}{N^3}\right)} - E\left[T_{2\to 2}(D_1, D_2)\right]\right)_+^2 /2\right) + \gamma$$

and bounded from below by

$$1 - \exp\left(-\left(E\left[T_{2\to 2}(D_1, D_2)\right] - \sqrt{-2\ln\left(\frac{\alpha}{N}\right)} - (1+\gamma)2\tau - \frac{6(1+\gamma)}{\sqrt{\ln(1+\gamma)}}\sqrt{\ln(N)}\right)_+^2 /2\right).$$

The power of the α*-sized test based on* $S_{\infty\to 1}$ *(given in Theorem* 2*) is bounded from above by*

$$\exp\left(-\left(\sigma - \sqrt[4]{-\frac{N^5}{2}\ln\left(\frac{\gamma}{N^3}\right)} - KE\left[T_{\infty\to 1}(D_1, D_2)\right]\right)_+^2 /2\right) + \gamma$$

and bounded from below by

$$1 - \exp\left(-\left(E\left[T_{\infty\to 1}(D_1, D_2)\right] - \left(\sqrt{-2\ln\left(\frac{\alpha}{N}\right)} + 4\sigma\right)/K\right)_+^2 /2\right).$$

PROOF OF THEOREM 5: The claim follows the logic of Theorems 1 and 2, and so only a sketch is provided here. The probability that either test statistic exceeds the $1-\alpha$ quantile of its reference distribution is bounded from above following

$$\begin{aligned} P\left(T(D_1, D_2) \geq Q_{1-\alpha}\left(D_1^r, D_2^r\right)\right) &= P\left(U(D_1, D_2) \geq \left(Q_{1-\alpha}\left(D_1^r, D_2^r\right) - E\left[T(D_1, D_2)\right]\right)\right) \\ &\leq \exp\left(-\left(Q_{1-\alpha}\left(D_1^r, D_2^r\right) - E\left[T(D_1, D_2)\right]\right)_+^2/2\right) \end{aligned}$$

and bounded from below following

$$\begin{aligned} &P\left(T(D_1, D_2) \geq Q_{1-\alpha}\left(D_1^r, D_2^r\right)\right) \\ &\quad = 1 - P\left(-U(D_1, D_2) \geq \left(E\left[T(D_1, D_2)\right] - Q_{1-\alpha}\left(D_1^r, D_2^r\right)\right)\right) \\ &\quad \geq 1 - \exp\left(-\left(E\left[T(D_1, D_2)\right] - Q_{1-\alpha}\left(D_1^r, D_2^r\right)\right)_+^2/2\right), \end{aligned}$$

where $Q_{1-\alpha}(D_1^r, D_2^r)$ refers to the $(1-\alpha)$-quantile of $\{T(D_1^r, D_2^r)\}_{r\in[R]}$, $U(D_1, D_2) = (T(D_1, D_2) - E[T(D_1, D_2)])$, T may refer to one of $T_{2\to 2}$ or $T_{\infty\to 1}$, $(x)_+^2 = x^2\mathbb{1}_{x>0}$, and the inequality is due to Talagrand (see Boucheron, Lugosi, and Massart (2013), Theorem 6.10) since $T_{2\to 2}$ and $T_{\infty\to 1}$ are both convex Lipschitz functions.

Applying the bounds from Lemma 1 to the test based on $T_{2\to 2}$ gives

$$\begin{aligned} &P\big(T_{2\to 2}(D_1, D_2) \ge Q_{2\to 2,1-\alpha}\big(D_1^r, D_2^r\big)\big) \\ &\quad\le \exp\big(-\big(Q_{2\to 2,1-\alpha}\big(D_1^r, D_2^r\big) - E\big[T_{2\to 2}(D_1, D_2)\big]\big)_+^2/2\big) \\ &\quad\le \exp\big(-\big(\hat{\tau} - E\big[T_{2\to 2}(D_1, D_2)\big]\big)_+^2/2\big) \end{aligned}$$

and

$$\begin{aligned} &P\big(T_{2\to 2}(D_1, D_2) \ge Q_{2\to 2,1-\alpha}\big(D_1^r, D_2^r\big)\big) \\ &\quad\ge 1 - \exp\big(-\big(E\big[T_{2\to 2}(D_1, D_2)\big] - Q_{2\to 2,1-\alpha}\big(D_1^r, D_2^r\big)\big)_+^2/2\big) \\ &\quad\ge 1 - \exp\left(-\left(E\big[T_{2\to 2}(D_1, D_2)\big] - \sqrt{-2\ln\left(\frac{\alpha}{N}\right)}\right.\right. \\ &\qquad\left.\left. - (1+\gamma)2\tau - \frac{6(1+\gamma)}{\sqrt{\ln(1+\gamma)}}\sqrt{\ln(N)}\right)_+^2/2\right) \end{aligned}$$

for any $\gamma \in [0, 1/2]$, where $Q_{2\to 2,1-\alpha}(D_1^r, D_2^r)$ refers to the $(1-\alpha)$-quantile of $\{T_{2\to 2}(D_1^r, D_2^r)\}_{r\in[R]}$,

$$\begin{aligned} \nu_{ij}(s) &= F_{ij,1}(s) + F_{ij,2}(s) - 2F_{ij,1}(s)F_{ij,2}(s), \\ \hat{\tau} &= \max_{s\in\mathbb{R}}\max_{i\in[N]}\sqrt{\sum_{j\in[N]}(\mathbb{1}_{D_{ij,1}\le s} - \mathbb{1}_{D_{ij,2}\le s})^2}, \quad \text{and} \\ \tau &= \max_{s\in\mathbb{R}}\max_{i\in[N]}\sqrt{\sum_{j\in[N]}\nu_{ij}(s)}. \end{aligned}$$

Similarly, applying Lemma 2 and the inequalities $T_{\infty\to 1} \le S_{\infty\to 1} \le KT_{\infty\to 1}$ to the test based on $S_{\infty\to 1}$ gives

$$\begin{aligned} &P\big(S_{\infty\to 1}(D_1, D_2) \ge Q_{\infty\to 1,1-\alpha}\big(D_1^r, D_2^r\big)\big) \\ &\quad\le \exp\big(-\big(Q_{\infty\to 1,1-\alpha}\big(D_1^r, D_2^r\big) - E\big[S_{\infty\to 1}(D_1, D_2)\big]\big)_+^2/2\big) \\ &\quad\le \exp\big(-\big(\hat{\sigma} - KE\big[T_{\infty\to 1}(D_1, D_2)\big]\big)_+^2/2\big) \end{aligned}$$

and

$$\begin{aligned} &P\big(S_{\infty\to 1}(D_1, D_2) \ge Q_{\infty\to 1,1-\alpha}\big(D_1^r, D_2^r\big)\big) \\ &\quad\ge 1 - \exp\big(-\big(E\big[S_{\infty\to 1}(D_1, D_2)\big] - Q_{\infty\to 1,1-\alpha}\big(D_1^r, D_2^r\big)\big)_+^2/2\big) \end{aligned}$$

$$\geq 1 - \exp\left(-\left(E\left[S_{\infty\to 1}(D_1, D_2)\right] - \sqrt{-2\ln\left(\frac{\alpha}{N}\right)} - 4\sigma\right)_+^2 /2\right),$$

where $Q_{\infty\to 1,1-\alpha}(D_1^r, D_2^r)$ refers to the $(1-\alpha)$-quantile of $\{S_{\infty\to 1}(D_1^r, D_2^r)\}_{r\in[R]}$, $\nu_{ij}(s) = F_{ij,1}(s) + F_{ij,2}(s) - 2F_{ij,1}(s)F_{ij,2}(s)$, $\hat{\sigma} = \max_{s\in\mathbb{R}} \sum_{i\in[N]} \sqrt{\sum_{j\in[N]} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s})^2}$, and $\sigma = \max_{s\in\mathbb{R}} \sum_{i\in[N]} \sqrt{\sum_{j\in[N]} \nu_{ij}(s)}$.

The last step of the proof is to replace $\hat{\tau}$ with τ and $\hat{\sigma}$ with σ in the upper bounds. To do this, write for any $i \in [N]$, $s \in \mathbb{R}$, and $t > 0$,

$$\begin{aligned} &P\left(\sqrt{\sum_{j\in[N]} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s})^2} \leq \sqrt{\sum_{j\in[N]} \nu_{ij}(s)} - \sqrt{t}\right) \\ &\quad \leq P\left(\sum_{j\in[N]} (\mathbb{1}_{D_{ij,1}\leq s} - \mathbb{1}_{D_{ij,2}\leq s})^2 \leq \sum_{j\in[N]} \nu_{ij}(s) - t\right) \leq \exp\left(\frac{-2t^2}{N}\right), \end{aligned}$$

where the second inequality is due to Hoeffding (see Boucheron, Lugosi, and Massart (2013), Theorem 2.8). The union bound implies

$$P(\hat{\tau} \leq \tau - \sqrt{t}) \leq N^3 \exp\left(\frac{-2t^2}{N}\right)$$

and

$$P(\hat{\sigma} \leq \sigma - N\sqrt{t}) \leq N^3 \exp\left(\frac{-2t^2}{N}\right)$$

which when combined with the upper bounds from before give

$$\begin{aligned} &P\left(T_{2\to 2}(D_1, D_2) \geq Q_{2\to 2,1-\alpha}\left(D_1^r, D_2^r\right)\right) \\ &\quad \leq \exp\left(-\left(\hat{\tau} - E\left[T_{2\to 2}(D_1, D_2)\right]\right)_+^2/2\right) \\ &\quad \leq \exp\left(-\left(\tau - \sqrt{t} - E\left[T_{2\to 2}(D_1, D_2)\right]\right)_+^2/2\right) + N^3 \exp\left(\frac{-2t^2}{N}\right) \end{aligned}$$

and

$$\begin{aligned} &P\left(S_{\infty\to 1}(D_1, D_2) \geq Q_{\infty\to 1,1-\alpha}\left(D_1^r, D_2^r\right)\right) \\ &\quad \leq \exp\left(-\left(\hat{\sigma} - KE\left[T_{\infty\to 1}(D_1, D_2)\right]\right)_+^2/2\right) \\ &\quad \leq \exp\left(-\left(\sigma - N\sqrt{t} - KE\left[T_{\infty\to 1}(D_1, D_2)\right]\right)_+^2/2\right) + N^3 \exp\left(\frac{-2t^2}{N}\right) \end{aligned}$$

by the law of total probability. The claim follows by choosing t so that $\gamma = N^3 \exp(-2t^2/N)$. *Q.E.D.*

What remains is to characterize the parameters $E[T_{2\to 2}(D_1, D_2)]$, $E[T_{\infty\to 1}(D_1, D_2)]$, τ, and σ. Since these are context specific, I consider two examples in the subsections below.

The first example is a degree experiment where the treatment increases or decreases the probability that every pair of agents forms a link by a constant amount. The second example is a cluster experiment where the treatment alters agent assignment to cliques or clusters, but does not alter link probabilities directly.

B.3.1. *Example 1: Degree Experiment*

The first example is a degree experiment. The treatment increases or decreases the probability that every pair of agents forms a link by some fixed amount. In this example, D_1 and D_2 are the adjacency matrices corresponding to two unweighted and undirected networks with no loops. That is, D_1 and D_2 are symmetric, binary, and hollow matrices. The entries within a network are also identically distributed: $P(D_{ij,t} = 1) = p_t \in (0, 1)$. Setting $F_t(s) = p_t$ in the above definitions for this example gives that $\tau = \sqrt{(p_1 + p_2 - 2p_1p_2)N}$ and $\sigma = \sqrt{(p_1 + p_2 - 2p_1p_2)N^3}$. Since the matrix norms $T_{2\to2}$ and $T_{\infty\to1}$ are convex, we also have $E[T_{2\to2}(D_1, D_2)] \geq T_{2\to2}(E[D_1], E[D_2]) = N|p_1 - p_2|$ and $E[S_{\infty\to1}(D_1, D_2)] \geq E[T_{\infty\to1}(D_1, D_2)] \geq T_{\infty\to1}(E[D_1], E[D_2]) = N^2|p_1 - p_2|$.

It follows from Theorem 5 that the power of the test based on $T_{2\to2}$ is bounded from above by

$$\exp\big(-\big(\sqrt{(p_1 + p_2 - 2p_1p_2)N} - N|p_1 - p_2| + o(\sqrt{N})\big)_+^2/2\big) + o(1)$$

and bounded from below by

$$1 - \exp\big(-\big(N|p_1 - p_2| - 2\sqrt{(p_1 + p_2 - 2p_1p_2)N} + o(\sqrt{N})\big)_+^2/2\big).$$

The power of the test based on $S_{\infty\to1}$ is bounded from above by

$$\exp\big(-\big(\sqrt{(p_1 + p_2 - 2p_1p_2)N^3} - KN^2|p_1 - p_2| + o(\sqrt{N})\big)_+^2/2\big) + o(1)$$

and bounded from below by

$$1 - \exp\big(-\big(N^2|p_1 - p_2| - 4\sqrt{(p_1 + p_2 - 2p_1p_2)N^3}/K + o(\sqrt{N})\big)_+^2/2\big).$$

Intuitively, the upper (lower) bounds decrease (increase) with the expected entrywise difference between D_1 and D_2 (given by $|p_1 - p_2|$) and increase (decrease) with the variance of the entrywise difference between D_1 and D_2 (given by $(p_1 + p_2 - 2p_1p_2)$). The remaining terms are asymptotically negligible.

B.3.2. *Example 2: Cluster Experiment*

The second example is a cluster experiment where agents belong to one of two groups. Agents within a group are more (or less) likely to form a link than agents across groups. In this experiment, the treatment only changes the agent's group assignments. Unlike the first example, the expected change in the network has mean zero, and so the power of the test comes from the operator norms "picking out" the subset of agents that switch from both being in the same group to being in different groups or switch from being in different groups to being in the same group.

In this example, D_1 and D_2 are the adjacency matrices corresponding to two unweighted and undirected networks with no loops. That is, D_1 and D_2 are symmetric, binary, and hollow matrices. Agents are assigned to one of two groups in each period as denoted by the variable $Z_{i,t} \in \{1, 2\}$. For any time period, all agents in the same group are linked the probability p. Agents in different time periods are linked with probability q. That is, $P(D_{ij,t} = 1 | Z_{i,t}, Z_{j,t}) = p\mathbb{1}_{Z_{i,t}=Z_{j,t}} + q\mathbb{1}_{Z_{i,t}\neq Z_{j,t}}$. I assume that $p \neq q$.

When $t = 1$, half of the agents are assigned to group one and the other half are assigned to group two. That is without loss, $Z_{i,1} = 1 + \mathbb{1}_{i>N/2}$. When $t = 2$, agents switch groups independently and randomly with probability $\pi \in (0, 1)$. Let $\rho = 2\pi(1 - \pi)$ denote the probability for any agent-pair ij, either i switches groups in time period 2 or j switches groups, but not both. Then direct calculation of the variances for each of the four possible group combinations yields $\tau = \sqrt{((p(1-p) + q(1-q))(1-\rho) + 2(p+q-2pq)\rho)N}$ and $\sigma = \sqrt{((p(1-p) + q(1-q))(1-\rho) + 2(p+q-2pq)\rho)N^3}$. Intuitively, $(p(1-p) + q(1-q))(1-\rho)$ gives the variance of $(D_{ij,1} - D_{ij,2})$ when $|Z_{i,1} - Z_{j,1}| = |Z_{i,2} - Z_{j,2}|$ (neither agents switch or both agents switch) and $2(p+q-2pq)\rho$ gives the variance of $(D_{ij,1} - D_{ij,2})$ when $|Z_{i,1} - Z_{j,1}| \neq |Z_{i,2} - Z_{j,2}|$ (either one of the agents switch but not both). As in the first example, convexity of the matrix norms gives $E[T_{2\to 2}(D_1, D_2)] \geq T_{2\to 2}(E[D_1], E[D_2]) = N|p-q|\rho$ and $E[T_{\infty\to 1}(D_1, D_2)] \geq T_{\infty\to 1}(E[D_1], E[D_2]) = N^2|p-q|\rho$.

It follows from Theorem 5 that the power of the test based on $T_{2\to 2}$ is bounded from above by

$$\exp\Big(-\Big(2\sqrt{\big((p(1-p) + q(1-q)\big)(1-\rho) + 2(p+q-2pq)\rho\big)N} - N|p-q|\rho + o(\sqrt{N})\Big)_+^2/2\Big) + o(1)$$

and bounded from below by

$$\geq 1 - \exp\Big(-\Big(N|p-q|\rho - 2\sqrt{\big((p(1-p) + q(1-q)\big)(1-\rho) + 2(p+q-2pq)\rho\big)N} + o(\sqrt{N})\Big)_+^2/2\Big).$$

The power of the test based on $S_{\infty\to 1}$ is bounded from above by

$$\exp\Big(-\Big(2\sqrt{\big((p(1-p) + q(1-q)\big)(1-\rho) + 2(p+q-2pq)\rho\big)N^3} - KN^2|p-q|\rho + o(\sqrt{N})\Big)_+^2/2\Big) + o(1)$$

and bounded from below by

$$1 - \exp\Big(-\Big(N^2|p-q|\rho - 4\sqrt{\big((p(1-p) + q(1-q)\big)(1-\rho) + 2(p+q-2pq)\rho\big)N^3}/K + o(\sqrt{N})\Big)_+^2/2\Big).$$

Intuitively, the bounds increase with the mean of the entrywise difference between the within and across group linking probabilities (given by $|p-q|$) and the probability that one but not both agents switch groups (given by $\rho = 2\pi(1-\pi)$). It decreases with the variances of the within and across group links (given by $p(1-p)$ and $q(1-q)$), and the variance of their difference (given by $p+q-2pq$). The remaining terms are asymptotically negligible.

APPENDIX C: SIMULATION EVIDENCE

Section 4.2 in the main text predicts that the test based on $S_{\infty\to 1}$ is potentially more powerful than that based on $T_{2\to 2}$ for sparse and degree-heterogeneous alternatives. This subsection provides supporting evidence from two Monte Carlo experiments. It considers the case of unweighted unipartite networks with no loops (symmetric, binary, and hollow adjacency matrices) for simplicity. The purpose of this section is not to simulate data that mimics real-world networks (see instead Section 5 of the main text), but rather to assess the predictions in a controlled environment.

C.1. *The Sparse Experiment*

Sparsity is a common feature of social and economic networks. For example, in many social surveys it is common for agents to report only a handful of connections. To examine the impact of network sparsity on the power of the two tests, I consider two Erdős–Renyi graph models. In these models, the adjacency matrices are $\{0, 1\}$-valued with $P(D_{ij,t} = 1) = 1 - F_{ij,1}(0) = \frac{8}{N}$ and $1 - F_{ij,2}(0) = \frac{5}{N}$ for every $i, j \in [N]$. Agents in the first network have approximately 60% more links than agents in the second network, violating H_0. Applying the two tests to data simulated from the models with $N = 50/100$ and $R = 10{,}000$ yields an average p-value for the test based on the $2 \to 2$ norm of approximately 0.070/0.020 and an average p-value for the test based on the $\infty \to 1$ norm of approximately 0.049/0.013. The test based on the $\infty \to 1$ norm is more powerful, but not dramatically so.

C.2. *The Degree Heterogeneous Experiment*

Degree heterogeneity is another common feature of social and economic networks. For example, in many production and collaboration networks it is common for a small number of agents to have an order of magnitude more links than the median agent. To examine the impact of degree heterogeneity on the power of the two tests, I consider two second-order stochastic block models. In these models, $P(D_{ij,t} = 1) = 1 - F_{ij,t}(0)$, with $F_{1j,1}(0) = F_{1j,2}(0) = 0.5$ for all $j \in [N]$ and $1 - F_{ij,1}(0) = 0.02$ and $1 - F_{ij,2}(0) = 0.08$ for any $i, j \in [N] \setminus [1]$. Agents in the first network have approximately 400% percent more links than in the second network, violating H_0. However, the high degree agent, agent 1, has approximately the same number of links. Applying the two tests to data simulated from the models with $N = 50/100$ and $R = 10{,}000$ yields an average p-value for the test based on the $2 \to 2$ norm of approximately 0.521/0.204 and an average p-value for the test based on the $\infty \to 1$ norm of approximately 0.001/0.000. The test based on the $\infty \to 1$ norm is much more powerful.

APPENDIX D: DETAILS ABOUT THE APPLICATIONS AND EXTENSIONS

D.1. *Application 1: A Test of Link Stationarity*

Goyal, Van Der Leij, and Moraga-González (2006) observe coauthorships between economists over time and argue that the profession has become more interconnected in response to new research technologies such as the internet. The framework of Section 2 can be used to evaluate whether the changes in network structure are statistically significant. Let $D_{ij,t}$ describe the existence of a coauthorship between economists i and j in time period t. Suppose that the researcher observes the coauthorship data for M time periods. Then $H_0: F_1 = F_2 = \cdots = F_M$ is the hypothesis of link stationarity that the differences between coauthorships over time can be explained by M draws from the same link formation model.

One can extend the randomization test of Section 3 to this testing problem by independently permuting all of the M links associated with each agent pair. For the choice of test statistic, I recommend the maximum or average difference over all $\binom{M}{2}$ pairs of networks using the semidefinite approximation to the $\infty \to 1$-norm: $\max_{t,t' \in [M]} S_{\infty\to 1}(D_t, D_{t'})$ or $\sum_{t,t' \in [M]} S_{\infty\to 1}(D_t, D_{t'})$. The test proposed in the hypothesis of Theorem 2 corresponds to the case of $M = 2$. Theorem 2 applies to the case of $M > 2$ mutatis mutandis.

Failure to reject H_0 using the network data $D_1, D_2, \ldots, D_M$ suggests that the observed changes in network interconnectedness are not statistically significant. The first example in Section 5 demonstrates this application to testing link stationarity.

D.2. *Application 2: A Test for Link Heterogeneity*

Banerjee, Chandrasekhar, Duflo, and Jackson (2013) collect data on a dozen social and economic ties between villagers in Karnataka, India. Jackson, Rogers, and Zenou (2017) suggest that this data on multiple types of connections between villagers "encode richer information than simply identifying whether two people are close or not." The framework of Section 2 can be used to evaluate this hypothesis. Let $D_{ij,1}$ denote whether agents i and j have one type of connection (they report being friends) and $D_{ij,2}$ denote whether agents i and j have another type of connection (they report having borrowed or lent money to one another). Then $H_0: F_1 = F_2$ is the hypothesis of link homogeneity that the differences between the networks can be explained by draws from the same link formation model. Failure to reject H_0 using the network data D_1 and D_2 suggests that the observed differences between the friendship and lending networks are not statistically significant. The second example in Section 5 demonstrates this application to testing link homogeneity.

D.3. *Application 3: A Test of No Treatment Effects*

Rose (2004) analyzes yearly aggregate international trade data and argues that participation in trade agreements such as the World Trade Organization (WTO) does not significantly alter the level of trade between countries. The framework of Section 2 can be used to evaluate this hypothesis. Let $D_{ij,t}$ describe the logarithm of the total value of trade between countries i and j in year t, and $X_{ij,t}$ be an indicator for whether country i or country j are members of the WTO in year t. Let N denote the number of countries and M denote the number of time periods. A nonparametric version of Rose (2004)'s gravity model of trade is

$$D_{ij,t} = \alpha_{i,t} + \alpha_{j,t} + \beta_{ij} + \gamma_{ij,t} X_{ij,t} + \varepsilon_{ij,t},$$

where $\alpha_{i,t}$ and $\alpha_{j,t}$ are country-specific determinants of trade that may vary over time such as GDP or population, β_{ij} are country-pair-specific determinants of trade that do not vary over time such as physical distance, and $\varepsilon_{ij,t}$ is an independent and identically distributed idiosyncratic error. The null hypothesis of no treatment effects is $H_0: \gamma_{ij,t} = 0$ for all $i, j \in [N]$ and $t \in [M]$. Rose (2004) also allows for two observed country-pair determinants of trade that vary over time: indicators for whether the two countries share the same currency or one is a colony of the other. One can restrict the randomization to be conditional on the value of these binary variables.

The hypothesis H_0 can be tested using the framework of Section 2 by taking triple differences so that

$$\begin{aligned}
&\big(\big[(D_{ij,t} - D_{i1,t}) - (D_{2j,t} - D_{21,t})\big] - \big[(D_{ij,t'} - D_{i1,t'}) - (D_{2j,t'} - D_{21,t'})\big]\big) \\
&\quad = \big(\big[(\gamma_{ij,t}X_{ij,t} - \gamma_{i1,t}X_{i1,t}) - (\gamma_{2j,t}X_{2j,t} - \gamma_{21,t}X_{21,t})\big] \\
&\qquad - \big[(\gamma_{ij,t'}X_{ij,t'} - \gamma_{i1,t'}X_{i1,t'}) - (\gamma_{2j,t'}X_{2j,t'} - \gamma_{21,t'}X_{21,t'})\big]\big) \\
&\qquad + \big(\big[(\varepsilon_{ij,t} - \varepsilon_{i1,t}) - (\varepsilon_{2j,t} - \varepsilon_{21,t})\big] - \big[(\varepsilon_{ij,t'} - \varepsilon_{i1,t'}) - (\varepsilon_{2j,t'} - \varepsilon_{21,t'})\big]\big)
\end{aligned}$$

for every $t, t' \in [M]$ and $i, j, \in [N]$. That is, H_0 implies that $G_{ij,t} = G_{ij,t'}$ for every $t, t' \in [M]$ and $i, j, \in [N]$ where $G_{ij,t}$ refers to the marginal distribution of $[(D_{ij,t} - D_{i1,t}) - (D_{2j,t} - D_{21,t})]$. This implication can be tested using the framework of Sections 3 and 4, applied to the differenced data $[(D_{ij,t} - D_{i1,t}) - (D_{2j,t} - D_{21,t})]$ instead of $D_{ij,t}$. Failure to reject H_0 using this data suggests that the observed differences in trade over time can be explained by the gravity model of trade with no treatment effects in that any differences in trade for country pairs across years with and without participation in the WTO are not statistically significant.

To be sure, the Rose (2004) model of trade does not treat participation in the WTO as endogenous or allow agent-specific parameters to be multilateral terms coming from general equilibrium. Such complications may require alternative testing procedures.

D.4. *Application 4: A Test for Endogenous Link Formation*

Goldsmith-Pinkham and Imbens (2013) consider a joint model of student GPA and link formation in a high-school social network.[1] They hypothesize that a determinant of GPA also drives variation in network links, and propose a one-sample parametric test for such endogenous link formation. The framework of Section 2 can be used to specify a nonparametric two-sample test.

Let $D_{ij,t}$ denote whether students i and j report a friendship in school year t and $\eta_{i,t}$ describe the social characteristic of agent i in time period t thought to drive link formation. For example, $\eta_{i,t}$ might be a measure of agent i's participation in an extracurricular activity or membership in a social clique. In the setting of Goldsmith-Pinkham and Imbens (2013), $\eta_{i,t}$ is the residual from a linear-in-means model of network peer effects. Network endogeneity may then refer to the idea that the distribution of $D_{ij,t}$ varies with the social proximity of the agents in the social characteristics space as measured by $|\eta_{i,t} - \eta_{j,t}|$. Define

$$\begin{aligned}
D^{\dagger}_{ij,1} &= D_{ij,1}\mathbb{1}_{|\eta_{i,1}-\eta_{j,1}|>|\eta_{i,2}-\eta_{j,2}|} + D_{ij,2}\mathbb{1}_{|\eta_{i,1}-\eta_{j,1}|<|\eta_{i,2}-\eta_{j,2}|} \quad \text{and} \\
D^{\dagger}_{ij,2} &= D_{ij,2}\mathbb{1}_{|\eta_{i,1}-\eta_{j,1}|>|\eta_{i,2}-\eta_{j,2}|} + D_{ij,1}\mathbb{1}_{|\eta_{i,1}-\eta_{j,1}|<|\eta_{i,2}-\eta_{j,2}|}.
\end{aligned}$$

In words, $D^{\dagger}_{ij,1}$ is an indicator for whether agents i and j are linked when they are (relatively) farther apart in the social characteristics space and $D^{\dagger}_{ij,2}$ is an indicator for whether agents i and j are linked when they are (relatively) closer in the social characteristics space. Let $F_{ij,t}$ refer to the marginal distribution of $D^{\dagger}_{ij,t}$, conditional on the collection of social characteristics η_1 and η_2.

[1]See also Hsieh and Lee (2014), Johnsson and Moon (2021), Arduini, Patacchini, and Rainone (2015), Auerbach (2019).

Then $H_0 : F_1 = F_2$ is the hypothesis of exogenous link formation that the differences in friendship links across school years are unrelated to student proximity in the social characteristic space. Failure to reject H_0 using the network and social characteristic data suggests that any relationship between the students' social proximity and the formation of network links is not statistically significant.

D.5. *Application 5: A Test for Network Externalities*

Pelican and Graham (2020) consider a model of link formation in which the propensity for agents to form a link may depend on the existence of other links in the network. They propose a one-sample parametric test for such network externalities. The framework of Section 2 can be used to specify a nonparametric two-sample test.

I illustrate the application with the following nonparametric version of a model motivated by Bloch and Jackson (2007) (see Graham (2015), Section 2)

$$D_{ij,t} = \mathbb{1}\left\{\alpha_{ij} + \gamma_{ij}\sum_{k=1}^{N} D_{ik,t}D_{jk,t} - \varepsilon_{ij,t} \geq 0\right\},$$

where $\varepsilon_{ij,t}$ is independent, identically distributed, and mean-zero. The parameters α_{ij} and γ_{ij} do not vary with t. In this model, agents with many friends in common are more likely to become friends, and the agents first draw $\{\varepsilon_{ij,t}\}_{i\neq j}$ and then choose links so that the link formation rule is satisfied for every ij-pair. The use of $\sum_{k=1}^{N} D_{ik,t}D_{jk,t}$ on the right-hand side is arbitrary and can be replaced by any network statistic.

The hypothesis of no network externalities corresponds to $H_0 : \gamma_{ij} = 0$ for all $i, j \in [N]$. Under this hypothesis, D_1 and D_2 are drawn from the same random graph model (in the sense of Section 2.1), and so the randomization test proposed in Section 3 controls size in finite samples. Notice that under the null hypothesis of no network externalities, the joint distribution of network links does not have multiple or no equilibria and so this complication that makes estimation of γ_{ij} difficult does not impact the validity of the proposed test procedure.

When α_{ij} and γ_{ij} are thought to also vary across the two networks, the researcher might instead test the more general hypothesis $H_0 : \alpha_{ij,1} = \alpha_{ij,2}$ and $\gamma_{ij,1} = \gamma_{ij,2} = 0$ for every $i, j \in [N]$. Failure to reject H_0 using the network data D_1 and D_2 suggests that the distribution of network links can be explained by draws from a model without network externalities and so any network externalities are not statistically significant.

D.6. *Application 6: A Test of Link Reciprocity*

Calvó-Armengol, Patacchini, and Zenou (2009) specify a model of network peer effects in which any nomination of a friendship from one agent to another indicates a social tie between agents. The framework of Section 2 can be used to detect potential asymmetries in link nominations. Let $D_{ij,1} = D_{ji,1}$ be an indicator for whether i nominates j and $D_{ij,2} = D_{ji,2}$ be an indicator for whether j nominates i when surveyed.[2]

To illustrate the application, suppose that, in contrast to the null hypothesis of nomination symmetry, high out-degree agents (agents who make many nominations) are thought to nominate differently than low out-degree agents (agents who make few nominations).

[2] I thank Vincent Boucher for suggesting the example.

This choice of network statistic is arbitrary: any other network statistic can be used as a substitute for out-degrees to construct the test. Let $N_i = \sum_j D_{ij,1}$ describe the out-degree of agent i. Define

$$D^{\dagger}_{ij,1} = D_{ij,1}\mathbb{1}_{N_i>N_j} + D_{ij,2}\mathbb{1}_{N_i<N_j} \quad \text{and}$$
$$D^{\dagger}_{ij,2} = D_{ij,1}\mathbb{1}_{N_i<N_j} + D_{ij,2}\mathbb{1}_{N_i>N_j}.$$

In words, $D^{\dagger}_{ij,1}$ is an indicator for whether i nominates j when i makes more nominations than j or j nominates i when j makes more nominations than i. $D^{\dagger}_{ij,2}$ is an indicator for whether i nominates j when i makes less nominations than j or j nominates i when j makes less nominations than i. Let $F_{ij,t}$ refer to the marginal distribution of $D^{\dagger}_{ij,t}$. In contrast to the testing problem in Application 4, the distribution of $F_{ij,t}$ is not conditional on N_i and N_j. That is, when constructing the randomization test for this application, the number of nominations are to be recomputed with each simulation.

Then $H_0 : F_1 = F_2$ is the hypothesis of link reciprocity that the differences in nominations between pairs of agents are explained by draws from the same link formation model. Failure to reject H_0 using the data D_1 and D_2 suggests that any asymmetry in linking behavior (e.g., associated with the agent out-degrees) is not statistically significant.

D.7. *Extension 1: A Completely Randomized Experiment*

Banerjee, Chandrasekhar, Duflo, and Jackson (2018) collect data on social connections between villagers in 75 villages before and after a microfinance agency offers loans to villagers in 43 of the villages. They find that villages in which the microfinance agency entered were associated with relatively lower densities and argue that access to microfinance disincentivized the formation of some types of connections in the network. The framework of Section 2 can be extended to test the hypothesis that the changes in the network structure for the treatment villages are statistically significant.

Let $D_{ij,t,v}$ describe whether villagers i and j in village v report a social connection in time period t, X_v be a binary indicator for whether village v was assigned to the treatment group, N_v represent the number of agents in village v, V_1 be the number of treatment villages, V_0 be the number of control villages, and $V = V_1 + V_0$ be the total number of villages. Following Banerjee et al. (2018), each villager is assigned to exactly one village, each village is assigned to one of two treatment statuses, villagers do not form social connections across villages, and there are two time periods. Time period $t = 1$ is the before period in which no treatment has been assigned to either the treatment or control villages. Time period $t = 2$ is the after period in which treatment has been assigned to the treatment ($X_v = 1$) but not the control ($X_v = 0$) villages.

To illustrate the extension, I suppose that the microfinance agency selected V_1 villages uniformly at random from the collection of V villages for treatment. This treatment assignment mechanism corresponds to a "completely randomized experiment" in the terminology of Imbens and Rubin (2015). Different treatment assignment mechanisms require different inference strategies. The hypothesis to be tested is the null of $H_0 : D_{ij,t,v}(0) = D_{ij,t,v}(1)$ for every $i, j \in [N_v]$, $v \in [V]$, and $t \in [2]$ (see Imbens and Rubin (2015), Chapter 5), where $D_{ij,t,v}(\tau)$ is the potential outcome (network) associated with treatment ($\tau = 1$) or no treatment ($\tau = 0$).

For this problem, a test statistic is any real-valued function of the collection village adjacency matrices

$$T\big(\{D_{t,v}\}_{t\in[T],v\in[V]:X_v=1}, \{D_{t,v}\}_{t\in[T],v\in[V]:X_v=0}\big).$$

Since the villages may all be defined on communities of different sizes, it is assumed that the test statistic is well-defined on matrices of arbitrary dimension. Test statistics based on the usual network statistics such as density, clustering, eigenvector centrality, etc. satisfy this property. For the reasons outlined in the main text of the paper, I recommend choosing T to be the average squared difference between the entrywise differences of the two adjacency matrices as measured by the semidefinite approximation to the $\infty \to 1$ norm between the treatment and control groups. That is,

$$T\big(\{D_{t,v}\}_{t\in[T],v\in\{[V]:X_v=1\}}, \{D_{t,v}\}_{t\in[T],v\in\{[V]:X_v=0\}}\big)$$
$$= \sum_{v\in\{[V]:X_v=1\},v'\in\{[V]:X_{v'}=0\}} \big(S_{\infty\to 1}(D_{1,v}, D_{2,v}) - S_{\infty\to 1}(D_{1,v'}, D_{2,v'})\big)^2.$$

An appealing feature of this test statistic is that it has a difference-in-differences-like structure in that it measures the difference in the change in the network for the treated villages relative to the change in the control villages. The use of average squared loss here to compare treatment and control villages is arbitrary. I propose the use of the $S_{\infty\to 1}$ statistic to measure changes within a village over time.

One way to construct a critical value for this test is to generate a reference distribution by re-randomizing the treatment assignment X_v (see Imbens and Rubin (2015), Chapter 5). That is, let $\mathcal{X} := \{x \in \{0, 1\}^V : \|x\|_1 = V_1\}$ describe the set of possible counterfactual treatment assignments (the set of all subsets of $[V]$ of size V_1) and let $\{X^r\}_{r\in[R]}$ be R independent and uniformly distributed draws from $\mathcal{X}$. Then the idea is to use

$$\big\{T\big(\{D_{t,v}\}_{t\in[T],v\in[V]:X^r_v=1}, \{D_{t,v}\}_{t\in[T],v\in[V]:X^r_v=0}\big)\big\}_{r\in[R]}$$

as a reference distribution for testing H_0. The construction of the test based on this reference distribution follows exactly Section 3 in the main text. The test controls size by construction. I suspect that it is straightforward to derive power properties of the test based on the $\infty \to 1$ norm suggested above using the arguments of Theorems 2, 3, and 5, but leave this to future work.

D.8. *Extension 2: A One-Sample Test of Independence*

Fafchamps and Gubert (2007) study link formation in a risk-sharing network and argue that the surveyed network connections are unrelated to the respondents' occupations. The framework of Section 2 can be extended to test that the network links and the agent occupations are unrelated in the following sense. Let D_{ij} describe whether agents i and j report a network connection, X_i describe the occupation of agent i, and N be the number of agents in the community. Then the hypothesis to be tested is that $\{D_{ij}\}_{i,j\in[N]}$ and $\{(X_i, X_j)\}_{i,j\in[N]}$ have mutually independent entries.

Let $T(D, X)$ be a real-valued test statistic defined on the matrix of network connections and vector of occupation assignments. For this testing problem, I suggest a randomization test based on rerandomizing the occupation assignments. Let Π be the set of permutations on $[N]$ and $\{\pi^r\}_{r\in[R]}$ be a collection of independent and uniformly distributed draws from Π. In words, π^r_i refers to another agent in the community that is randomly assigned to agent i and $X^r_i := X_{\pi^r_i}$ refers to the occupation associated with agent π^r_i. Then under the null hypothesis, $T(D, X)$ and $T(D, X^r)$ have the same distribution and so $\{T(D, X^r)\}_{r\in[R]}$ can be used as a reference distribution to test the null hypothesis, exactly as in Section 3 of the main text.

For the reasons outlined in the main text of the paper, I recommend choosing T to be the semidefinite approximation to the $\infty \to 1$ norm of the Hadamard (entrywise) product of the matrices D and $W := \{\mathbb{1}_{X_i = X_j}\}_{i,j\in[N]}$. That is,

$$S_\infty(D \cdot W, 0) = \frac{1}{2} \max_{s\in\mathbb{R}} \max_{X\in\mathcal{X}_{2N}} \left\langle \begin{bmatrix} 0_{N\times N} & D \cdot W \\ D \cdot W & 0_{N\times N} \end{bmatrix}, X \right\rangle,$$

where $\cdot$ refers to the Hadamard product and $0_{N\times N}$ is an $N \times N$ matrix of 0s. The intuition behind this test statistic is that if agents with similar occupations are more likely to form a link, then one would expect $D \cdot W$ to be larger (in the sense that its matrix norm is bigger) than $D \cdot W^r$, where $W^r := \{\mathbb{1}_{X_i^r = X_j^r}\}_{i,j\in[N]}$ is the matrix of occupations assigned at random. The test controls size by construction. A study of its power properties is left to future work.

D.9. *Extension 3: A One-Sample Specification Test*

Jackson and Rogers (2007) argue that real-world social networks are connected in ways that are poorly approximated by, for example, an Erdős–Renyi model of link formation. The framework of Section 2 can be extended to test whether some observed network data can be explained by a particular parametric model of link formation. Let D_{ij} be the observed networks links described by the (true) unknown model $F_{ij,1}$. Let $F_{ij,2}$ be the distribution of links associated with a model chosen by the researcher. Then $H_0 : F_{ij,1} = F_{ij,2}$ is the hypothesis that the distribution of links given by the model that generated the data and those given by the model chosen by the researcher are the same. The network formation model F_2 does not need to have a closed-form representation (it only needs to be simulatable).

To extend the framework of Section 2 to this testing problem, I propose converting it to two-sample problem by first drawing network data D' from F_2. Let $T(D, D')$ be an arbitrary real-valued test statistic as described in Section 3 evaluated on D and D'. Then I propose constructing a reference distribution for $T(D, D')$ by drawing additional simulations from F_2. That is, I suggest independently drawing $2R$ collections of networks from F_2, collecting them into two groups $\{D^r\}_{r\in[R]}$ and $\{D'^r\}_{r\in[R]}$, and using $\{T(D^r, D'^r)\}_{r\in[R]}$ as a reference distribution for $T(D, D')$. As motivated in the main text, I propose using $S_{\infty\to 1}(D, D')$ for the choice of test statistic. The test controls size by construction. I leave a study of the power properties to future work.

In many cases, the network formation model F_2 chosen by the researcher may depend on unknown parameters. For example, the researcher may hypothesize that D is drawn from an Erdős–Renyi model with some unknown value of link probability $\theta \in [0, 1]$. One way to test this hypothesis is to test each parameter individually (or each parameter in a representative subcollection) and reject the null hypothesis only if the test rejects at every value. This test controls size by construction, but may have low power. Another way to test this hypothesis is to estimate the parameters of the model under H_0 by some method, and use the estimated value of the parameters to specify F_2 in the above test. I suspect that under certain conditions such a test will control size asymptotically, but leave this to future work.

APPENDIX E: INDEPENDENCE ASSUMPTION

Section 2 assumes that for every ij-pair, the two random variables $D_{ij,1}$ and $D_{ij,2}$ are independently drawn from $F_{ij,1}$ and $F_{ij,2}$, respectively. This assumption can be omitted without loss in two key settings. First, the independence condition is unnecessary if $D_{ij,1}$ and $D_{ij,2}$ are $\{0,1\}$-valued (i.e., D_1 and D_2 represent unweighted networks). Second, the independence condition is unnecessary if the test statistic can be written as a functional of $\Delta(\cdot) = \mathbb{1}\{D_1 \leq \cdot\} - \mathbb{1}\{D_2 \leq \cdot\}$. This second setting includes the class of test statistics based on operator norms proposed in Section 4.1.

To see why, note that the validity of the randomization test outlined in Section 3 follows so long as $T(D_1, D_2)$ and $T(D_1^r, D_2^r)$ have the same distribution under the null hypothesis. Since $\{D_{ij,1}, D_{ij,2}\}_{i>j\in[N]}$ are independent across agent pairs, validity follows if $\{D_{ij,1}, D_{ij,2}\}$ and $\{D^r_{ij,1}, D^r_{ij,2}\}$ are equal in distribution under the null hypothesis. That is, $D_{ij,1}$ and $D_{ij,2}$ are exchangeable: $(D_{ij,1}, D_{ij,2}) =_d (D_{ij,2}, D_{ij,1})$. Exchangeability is generally weaker than the assumption that $D_{ij,1}$ and $D_{ij,2}$ are both independent and identically distributed and stronger than the assumption that $D_{ij,1}$ and $D_{ij,2}$ are just identically distributed. However, in the special case that $D_{ij,1}$ and $D_{ij,2}$ are $\{0,1\}$-valued, the assumption that $D_{ij,1}$ and $D_{ij,2}$ are just identically distributed is equivalent to the assumption $D_{ij,1}$ and $D_{ij,2}$ are exchangeable. And so in the special case that $D_{ij,1}$ and $D_{ij,2}$ are $\{0,1\}$-valued, the assumption that $D_{ij,1}$ and $D_{ij,2}$ are independent can be omitted without loss.

By the same logic, the independence assumption is also unnecessary if the test statistic only depends on D_1 and D_2 through the functional $\Delta(\cdot) = \mathbb{1}\{D_1 \leq \cdot\} - \mathbb{1}\{D_2 \leq \cdot\}$. This is because for any $s \in \mathbb{R}$ the variable $\mathbb{1}\{D_{ij,1} \leq s\}$ is also $\{0,1\}$-valued, and so the null hypothesis that $D_{ij,1}$ and $D_{ij,2}$ are identically distributed implies that $\mathbb{1}\{D_{ij,1} \leq s\}$ and $\mathbb{1}\{D_{ij,2} \leq s\}$ are exchangeable. And so the tests based on the operator norms from Section 4 are valid in finite samples even if $D_{ij,1}$ and $D_{ij,2}$ are real-valued and not independent, because they only depend on $D_{ij,1}$ and $D_{ij,2}$ through the collection of $\{0,1\}$-valued random variables $\mathbb{1}\{D_{ij,2} \leq \cdot\}$ and $\mathbb{1}\{D_{ij,2} \leq \cdot\}$.

In addition, none of the results about the power properties of the two tests based on operator norms (i.e., Theorems 1–5B) require the condition that $D_{ij,1}$ and $D_{ij,2}$ are independent. These results rely on the concentration of $\mathbb{1}\{D_t \leq s\} - F_t$ for $t \in [2]$, which only requires independence of $\{D_{ij,1}, D_{ij,2}\}_{i>j\in[N]}$ across agent pairs. See, in particular, Lemmas 1 and 2 in Appendix A of the main text.

A place where the assumption that $D_{ij,1}$ and $D_{ij,2}$ are independent is used is in the Section 3 claim that a generic test statistic $T(D_1, D_2)$ produces a test that is valid in finite samples. This includes the tests based on the network statistics (e.g., agent degree, eigenvector centrality, clustering, etc.), which are still valid when independence is weakened to exchangeability under the null hypothesis but potentially invalid with just the assumption that $D_{ij,1}$ and $D_{ij,2}$ are identically distributed un the null hypothesis.

REFERENCES

ARDUINI, TIZIANO, ELEONORA PATACCHINI, AND EDOARDO RAINONE (2015): "Parametric and Semiparametric iv Estimation of Network Models With Selectivity," Technical report, Einaudi Institute for Economics and Finance (EIEF). [13]

AUERBACH, ERIC (2019): "Identification and Estimation of a Partially Linear Regression Model Using Network Data," arXiv preprint. arXiv:1903.09679. [13]

BANERJEE, ABHIJIT, ARUN G. CHANDRASEKHAR, ESTHER DUFLO, AND MATTHEW O. JACKSON (2013): "The Diffusion of Microfinance," *Science*, 341 (6144), 1236498. [12]

BANERJEE, ABHIJIT V., ARUN G. CHANDRASEKHAR, ESTHER DUFLO, AND MATTHEW O. JACKSON (2018): "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets," Available at SSRN 3245656. [15]

BLOCH, FRANCIS, AND MATTHEW O. JACKSON (2007): "The Formation of Networks With Transfers Among Players," *Journal of Economic Theory*, 133 (1), 83–110. [14]

BOUCHERON, STÉPHANE, GÁBOR LUGOSI, AND PASCAL MASSART (2013): *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press. [7,8]

CALVÓ-ARMENGOL, ANTONI, ELEONORA PATACCHINI, AND YVES ZENOU (2009): "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 76 (4), 1239–1267. [14]
FAFCHAMPS, MARCEL, AND FLORE GUBERT (2007): "Risk Sharing and Network Formation," *American Economic Review*, 97 (2), 75–79. [16]
GOLDSMITH-PINKHAM, PAUL, AND GUIDO W. IMBENS (2013): "Social Networks and the Identification of Peer Effects," *Journal of Business & Economic Statistics*, 31 (3), 253–264. [13]
GOYAL, SANJEEV, MARCO J. VAN DER LEIJ, AND JOSÉ LUIS MORAGA-GONZÁLEZ (2006): "Economics: An Emerging Small World," *Journal of Political Economy*, 114 (2), 403–412. [11]
GRAHAM, BRYAN S. (2015): "Methods of Identification in Social Networks," *Annual Review of Economics*, 7 (1), 465–485. [14]
HSIEH, CHIH-SHENG, AND LUNG FEI LEE (2014): "A Social Interactions Model With Endogenous Friendship Formation and Selectivity," *Journal of Applied Econometrics*, 31 (2), 301–319. [13]
IMBENS, GUIDO W., AND DONALD B. RUBIN (2015): *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press. [15,16]
JACKSON, MATTHEW O., AND BRIAN W. ROGERS (2007): "Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review*, 97 (3), 890–915. [17]
JACKSON, MATTHEW O., BRIAN W. ROGERS, AND YVES ZENOU (2017): "The Economic Consequences of Social-Network Structure," *Journal of Economic Literature*, 55 (1), 49–95. [12]
JOHNSSON, IDA, AND HYUNGSIK ROGER MOON (2021): "Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach," *Review of Economics and Statistics*, 103 (2), 328–345. [13]
PELICAN, ANDRIN, AND BRYAN S. GRAHAM (2020): "An Optimal Test for Strategic Interaction in Social and Economic Network Formation Between Heterogeneous Agents," Technical report, National Bureau of Economic Research. [14]
ROSE, ANDREW K. (2004): "Do We Really Know That the WTO Increases Trade?" *American Economic Review*, 94 (1), 98–114. [12,13]

Co-editor Guido Imbens handled this manuscript.

Manuscript received 7 February, 2020; final version accepted 6 September, 2021; available online 2 November, 2021.