Online Appendix: Identifying Socially Disruptive Policies

Eric Auerbach* Yong Cai †

July 27, 2025

B Auxiliary lemmas

Lemma B1 (Lusin): For any measurable $f : [0,1]^2 \to \mathbb{R}$ and $\epsilon > 0$ there exists a compact $E_\epsilon \subseteq [0,1]^2$ with Lebesgue measure at least $1-\epsilon$ such that f is continuous when restricted to E_ϵ. See Dudley (2002) Theorem 7.5.2.

Lemma B2 (Spectral): Let $f : [0,1]^2 \to \mathbb{R}$ be a bounded symmetric measurable function and $T_f : L_2([0,1]) \to L_2([0,1])$ the associated integral operator $(T_f g)(u) = \int f(u,\tau)g(\tau)d\tau$. T_f admits the spectral decomposition $f(u,v) = \sum_{r=1}^{\infty} \lambda_r \phi_r(u), \phi_r(v)$ in the sense that $(T_f g)(u) = \int f(u,\tau)g(\tau)d\tau = \sum_{r=1}^{\infty} \lambda_r \phi_r(u) \int \phi_r(\tau)g(\tau)d\tau$ for any $g \in L_2([0,1])$. Each (λ_r, ϕ_r) pair satisfies $\int f(u,\tau)\phi_r(\tau)d\tau = \lambda_r \phi_r(u)$ where $\{\lambda_r\}_{r=1}^{\infty}$ is a multiset of bounded real numbers with 0 as its only limit point and $\{\phi_r\}_{r=1}^{\infty}$ is an orthogonal basis of $L_2([0,1])$. See Birman and Solomjak (2012) equation (5) preceding Theorem 4 in Chapter 9.2.

The spectral decomposition in Lemma B2 is related to the finite-dimensional version that is commonly used for matrices which is if Y is an $N \times N$ dimensional symmetric real-valued matrix then $Y_{ij} = \sum_{r=1}^{N} \lambda_r \phi_{ir}\phi_{jr}$. Each (λ_r, ϕ_r) pair satisfies $\sum_{j=1}^{N} Y_{ij}\phi_{jr} = \lambda_r \phi_{ir}$ where $\{\lambda_r\}_{r=1}^{N}$ is a multiset of real numbers and $\{\phi_{ir}\}_{i,r=1}^{N}$ is an $N \times N$ orthogonal matrix with rth column denoted by ϕ_r.

*Department of Economics, Northwestern University. E-mail: eric.auerbach@northwestern.edu.

†Department of Economics, Northwestern University. E-mail: yongcai2023@u.northwestern.edu.

Lemma B3 (Continuity): Let $f, g : [0,1]^2 \to \mathbb{R}$ be bounded symmetric measurable functions with positive eigenvalues $\{\lambda_r^+(f), \lambda_r^+(g)\}_{r\in\mathbb{N}}$ and negative eigenvalues $\{\lambda_r^-(f), \lambda_r^-(g)\}_{r\in\mathbb{N}}$ both ordered to be decreasing in absolute value. Suppose $\left(\int\int (f(u,v) - g(u,v))^2\, dudv\right)^{1/2} \leq \epsilon$. Then $|\lambda_r^+(f) - \lambda_r^+(g)| \leq \epsilon$ and $|\lambda_r^-(f) - \lambda_r^-(g)| \leq \epsilon$ for every $r \in \mathbb{N}$.
See Birman and Solomjak (2012), equation (19) following Theorem 8 in Chapter 9.2.

In Lemma 3 of Appendix Section A.3 in the main text, we use an implication of Lemma B3 and Theorem 368 of Hardy et al. (1952) (Lemma B5 below) that $\left(\sum_{r\in[R]} (\lambda_r(f) - \lambda_r(g))^2\right)^{1/2} \leq \sqrt{R}\left(\int\int (f(u,v) - g(u,v))^2\, dudv\right)^{1/2}$ where $\{\lambda_r(f), \lambda_r(g)\}_{r\in[R]}$ are the R largest in absolute value eigenvalues of f and g ordered to be decreasing. This result is an analog of the Hoffman-Wielandt inequality for matrices (Lemma B6 below) which, in a previous version of our paper, we refined in Proposition 4 of the main text.

Lemma B4 (Birkhoff): For every $M \in \mathcal{D}_n^+$ there exists an $m \in \mathbb{N}$, $\alpha_1, ..., \alpha_m > 0$, and $P_1, ..., P_m \in \mathcal{P}_n$ such that $\sum_{t=1}^m \alpha_t = 1$ and $M_{ij} = \sum_{t=1}^m \alpha_t P_{ij,t}$. See Birkhoff (1946).

Lemma B5 (Hardy-Littlewood-Polya Theorem 368): For any $m \in \mathbb{N}$ and $g, h \in \mathbb{R}^m$ we have $\sum_{r=1}^m g_{(r)} h_{(m-r+1)} \leq \sum_{r=1}^m g_r h_r \leq \sum_{r=1}^m g_{(r)} h_{(r)}$ where $g_{(r)}$ is the rth order statistic of g. See Hardy et al. (1952), Section 10.2, Theorem 368.

A version of Lemma B5 also holds for elements of $L^2([0,1])$, see Hardy et al. (1952), Section 10.13, Theorem 378. Specifically, for any $g, h \in L^2([0,1])$ we have $\int g^+(u)h^+(1-u)du \leq \int g(u)h(u)du \leq \int g^+(u)h^+(u)du$ where g^+ is the quantile function of g. This result is also used in the second proof of Theorems 2.1 and 2.5 in Whitt (1976).

Lemma B6 (Hoffman-Wielandt): Let $\{\lambda_r(F)\}_{r\in[n]}$ and $\{\lambda_r(G)\}_{r\in[n]}$ be the eigenvalues of two $n \times n$ real symmetric matrices F and G, ordered to be decreasing. Then $\sum_{r=1}^n (\lambda_r(F) - \lambda_r(G))^2 \leq \sum_{i=1}^n \sum_{j=1}^n (F_{ij} - G_{ij})^2$. See Hoffman and Wielandt (1953).

Lemma B7 (Lovász Corollary 10.35(a)): Let $f, g : [0,1]^2 \to \mathbb{R}$ be bounded symmetric measurable functions and $G = (V, E)$ be an arbitrary graph where V is a finite set of vertices and $E \subseteq V \times V$ be an arbitrary subset of vertex-pairs. Define the graph homomorphism $t(G, f) = \int_{u_1,...,u_{|V|}} \prod_{ij\in E} f(u_i, u_j) du_1, \ldots, du_{|V|}$. Then $t(G, f) = t(G, g)$ for every graph G if and only if f and g are equivalent up to a measure preserving transformation (see Definition 2 in Section 4.1). See Lovász (2012), Section 10.7, Corollary 10.35(a).[1]

Lemma B8 (Whitt Lemma 2.7): For any cdf H on $\mathbb{R}^n$ and uniform random variable U, there exits a measurable $x : [0,1] \to \mathbb{R}^n$ such that $x(U)$ has cdf H. See Whitt (1976), Lemma 2.7.

C Quadratic assignment problem

The quadratic assignment problem described in Section 5.1 of the main text is neither analytically solvable nor directly computable in general. However, there are special cases for which there is a simple analytical solution. In this section we give two examples. In both cases, the outer bounds we propose in Proposition 2 of Section 5.2 agree with the analytical solutions, which means that, in these examples, our outer bounds are sharp.

C.1 Diagonal graph functions

In this example, the graph functions are diagonal. That is, $h_s(u,v) = \alpha_s(u)\delta_u(v)$ where δ refers to the Dirac delta function. Following Proposition 1, the lower bound on the identified set for the DPO is

$$\min_{\psi_0,\psi_1 \in \mathcal{M}} \int\int \prod_{s\in\{0,1\}} h_s(\psi_s(u),\psi_s(v))dudv = \min_{\varphi_0,\varphi_1 \in \mathcal{M}} \int \prod_{s\in\{0,1\}} \alpha_s(\psi_s(u))du = \int \alpha_0^+(u)\alpha_1^+(1-u)du$$

where the first equality is due to the definition of the Dirac delta function and the second equality is due to the functional version of the Hardy-Littlewood-Polya Theorem 368 (Lemma B5 in Online Appendix Section B). By the same arguments, the upper bound is

$$\max_{\psi_0,\psi_1 \in \mathcal{M}} \int\int \prod_{t\in\{0,1\}} h_s(\psi_s(u),\psi_s(v))dudv = \int \alpha_0^+(u)\alpha_1^+(u)du.$$

[1]When Lovász (2012) states this corollary, he does not explicitly state the condition that $t(G,f) = t(G,g)$ for every graph G. He instead only writes that f and g are weakly isomorphic. The former condition is the definition of a weak isomorphism, which he gives after Proposition 7.10 in Section 7.3.

These bounds are equivalent to the bounds we propose in Proposition 2. This is because, in this example, $h_s(u, v)$ is diagonal, and so the lower bound can be represented

$$\int \alpha_0^+(u)\alpha_1^+(1-u)du = \sum_r \lambda_{r0}\lambda_{s(r)1}$$

and the upper bound can be represented

$$\int \alpha_0^+(u)\alpha_1^+(u)du = \sum_r \lambda_{r0}\lambda_{r1}$$

which are weakly outside the Proposition 2 bounds. But since these bounds are sharp, it follows that the Proposition 2 bounds must be sharp as well.

C.2 Block graph functions

In this example, the graph functions have a block structure. That is, $h_s(u, v) = f_s(u)f_s(v)$ where $f_s(u) = \mathbb{1}\{u \in A_s\}$ for some Lebesgue measurable sets $A_s \subseteq [0, 1]$ and $s \in \{0, 1\}$. Following Proposition 1, the lower bond on the identified set for the DPO is, for $y_s \in [0, 1)$

$$\begin{aligned}\min_{\psi_0,\psi_1\in\mathcal{M}} \int\int \prod_{s\in\{0,1\}} h_s(\psi_s(u), \psi_s(v))dudv &= \min_{\psi_0,\psi_1\in\mathcal{M}} \int\int \prod_{s\in\{0,1\}} f_s(\psi_t(u))f_s(\psi_s(v))dudv \\ &= \int f_0^+(u)f_1^-(u)du = \max(|A_0| + |A_1| - 1, 0)\end{aligned}$$

where $|A_s| = \int f_s(\tau)d\tau$ refers to the measure of A_s, the second equality is due to the functional version of the Hardy-Littlewood-Polya Theorem 368 (Lemma B5 in Online Appendix Section B). By the same arguments, the upper bound is

$$\max_{\psi_0,\psi_1\in\mathcal{M}} \int\int \prod_{t\in\{0,1\}} h_s(\psi_s(u), \psi_s(v))dudv = \int f_0^+(u)f_1^+(u)du = \min(|A_0|, |A_1|).$$

These bounds are equivalent to the bounds we propose in Proposition 2. This is because h_s has one non-zero eigenvalue that is equal to $\sqrt{|A_s|}$. As a result, the above lower bound can be rewritten

$$\max(|A_0| + |A_1| - 1, 0) = \max\left(\sum_r (\lambda_{r0}^2 + \lambda_{r1}^2) - 1, 0\right)$$

and the above upper bound can be rewritten

$$\min(|A_0|, |A_1|) = \min\left(\sum_r \lambda_{r0}^2, \sum_r \lambda_{r1}^2\right)$$

which are weakly outside the Proposition 2 bounds. But since these bounds are sharp, it follows that the Proposition 2 bounds must be sharp as well.

D Additional results and details

D.1 Asymmetric outcome matrices

Our identification arguments in Section 4 and bounds in Section 5 of the main text apply to undirected unipartite networks. These networks are represented by symmetric adjacency matrices where the rows and columns of the matrix are indexed by the same community of agents. $Y_{ij}(s)$ and $Y_{ji}(s)$ describe the magnitude of a connection between agents i and j under policy s. We require symmetry here because, in our proof of Proposition 2, when we take a spectral decomposition of the histogram approximation to the graph function, we assume that the right and left eigenvectors are the same. This is only the case when the graph function is symmetric.[2]

Directed networks connecting two or more types of agents (sometimes called bipartite or multipartite networks in the literature) can be incorporated through symmetrization, which is commonly used in the literature on U-statistics, going back to at least Hoeffding (1948) (see his equation 3.3), A textbook reference is Section 5.1.1 of Serfling (2009). Formally, we consider a community of N agents with K types where T_i takes value $k \in 1, \ldots, K$ if agent i is of type k. The types are assumed to be mutually exclusive and

[2]While other spectral decompositions may be applied when the graph function is asymmetric, these decompositions must, by definition, have different right and left eigenvectors when the graph function is asymmetric. Our current proof strategy does not directly apply to this decomposition.

collectively exhaustive. The researcher conducts a completely randomized balanced experiment with $2N_k$ agents of type k where $\sum_{k=1}^{K} N_k = N$. The agents are randomly assigned to two communities as given by the vector D drawn uniformly at random from the set $\{d \in \{0,1\}^N : \sum_{i=1}^{2N} \mathbb{1}\{T_i = 1\}d_i = N_1, \ldots, \sum_{i=1}^{N} \mathbb{1}\{T_i = K\}d_i = N_K\}$. We model the potential connection from agent i of type $T_i = k$ to agent j of type $T_j = l$ with the model $Y_{ij}(s) = g_{s,kl}(\varphi_s(w_i), \varphi_s(w_j), \eta_{ij})$, where w_i and η_{ij} are defined as in Section 3.2. We define the graph function as in Section 4.1: $h_{s,kl}(a,b) = \int \mathbb{1}\{g_{s,kl}(\varphi_s(a), \varphi_s(b), w) \leq y_s\}dw$. Relative to the link function specified in Section 3.2 and the graph function specified in Section 4.1, the link and graph functions here are allowed to vary with the agent types k and l and may be asymmetric. We focus on the DPO for the connections from agents of type k to agents of type l. This is the parameter $\mathbb{P}\left(Y_{ij}(1) \leq y_1, Y_{ij}(0) \leq y_0 | T_i = k, T_j = l\right)$ which, following the logic of Section 4.2, can be rewritten as $\int\int h_{1,kl}(u,v)h_{0,kl}(u,v)dudv$.

To apply the results of Section 4 and 5, the symmetrization strategy works by replacing the graph functions $h_{0,kl}$ and $h_{1,kl}$ with the functions $h^{\dagger}_{0,kl}$ and $h^{\dagger}_{1,kl}$ that are symmetric. Specifically, for $a, b \in [0,1]$ and $s \in \{0,1\}$, we define

$$h^{\dagger}_{s,kl}(a,b) = \left[h_{s,kl}(2a, 2b-1)\mathbb{1}\{a \leq 1/2, b > 1/2\} + h_{s,kl}(2b, 2a-1)\mathbb{1}\{a > 1/2, b \leq 1/2\}\right]/2.$$

The function $h^{\dagger}_{s,kl}$ is symmetric and satisfies $\mathbb{P}\left(Y_{ij}(1) \leq y_1, Y_{ij}(0) \leq y_0 | T_i = k, T_j = l\right) = \int\int h_{1,kl}(u,v)h_{0,kl}(u,v)dudv = \int\int h^{\dagger}_{1,kl}(u,v)h^{\dagger}_{0,kl}(u,v)dudv$. Since $h^{\dagger}_{s,kl}$ is symmetric, Proposition 1 in Section 4.2 and Propositions 2 and 3 of Section 5.2 can be applied. Furthermore, since $h^{\dagger}_{s,kl}$ is a simple function of $h_{s,kl}$ the researcher can use any estimate of the latter to construct an estimate of the former.

D.2 Row and column heterogeneity

In practice, the bounds in Propositions 2 and 3 may be wide when there is nontrivial heterogeneity in the row and column means of the graph functions. In such cases, we propose an adjustment building on Section 5 of Finke et al. (1987) that can, in some cases, lead to tighter bounds. Specifically, we write $h_s(u,v) = \alpha_s(u) + \alpha_s(v) + \epsilon_s(u,v)$ where $\alpha_s(u) = \int h_s(u,v)dv - \frac{1}{2}\iint h_s(u,v)dudv$ and $\epsilon_s(u,v) = h_s(u,v) - \alpha_s(u) - \alpha_s(v)$. Since h_s is identified up to a measure preserving transformation by Proposition 1 in the main text, $\alpha_s(u)$, the projection of h_s onto its first argument, is identified up to a measure preserving transformation. It follows that ϵ_s, a linear function of h_s and α_s, is also identified up to a measure preserving transformation.

The DPO is then

$$F(y_1, y_0) = \int\int \prod_{s\in\{0,1\}} (\alpha_s(u) + \alpha_s(v) + \epsilon_s(u,v))\, dudv$$
$$= \int\int \prod_{s\in\{0,1\}} (\alpha_s(u) + \alpha_s(v))\, dudv + \int\int \prod_{s\in\{0,1\}} \epsilon_s(u,v) dudv.$$

We bound the two summands separately. Specifically, the upper bound is

$$F(y_1, y_0) \leq \max_{\psi_1,\psi_0\in\mathcal{M}} \left[\int\int \prod_{s\in\{0,1\}} (\alpha_s(\psi_s(u)) + \alpha_s(\psi_s(v)))\, dudv + \int\int \prod_{s\in\{0,1\}} \epsilon_s(\psi_s(u), \psi_s(v)) dudv\right]$$
$$\leq \max_{\psi_1,\psi_0\in\mathcal{M}} \left[\int\int \prod_{s\in\{0,1\}} (\alpha_s(\psi_s(u)) + \alpha_s(\psi_s(v)))\, dudv\right] + \max_{\psi_1,\psi_0\in\mathcal{M}} \left[\int\int \prod_{s\in\{0,1\}} \epsilon_s(\psi_s(u), \psi_s(v)) dudv\right].$$

The first summand is bounded from above by

$$2 \max_{\psi_1,\psi_0\in\mathcal{M}} \left[\int \alpha_1(\psi_1(u))\alpha_0(\psi_0(u))du\right] + 2\bar{\alpha}_1\bar{\alpha}_0 \leq 2\int \alpha_1^+(u)\alpha_0^+(u)du + 2\bar{\alpha}_1\bar{\alpha}_0$$

where $\bar{\alpha}_s = \int \alpha_s(u)du$ and α_s^+ is the quantile function of α_s. See the functional version of the Hardy-Littlewood-Polya Theorem 368 (Lemma B5 in Online Appendix Section B). Following Proposition 2 of the main text, the second summand is bounded from above by $\min\left(\sum_r \lambda_{r1}^2, \sum_r \lambda_{r0}^2, \sum_r \lambda_{r1}\lambda_{r0}\right)$ where λ_{rs} refers to the rth eigenvalue of ϵ_s and the sums are as defined as in Section 4.1.2 of the main text. Together, the bounds imply that

$$F(y_1, y_0) \leq 2\int \alpha_1^+(u)\alpha_0^+(u)du + 2\bar{\alpha}_1\bar{\alpha}_0 + \min\left(\sum_r \lambda_{r1}^2, \sum_r \lambda_{r0}^2, \sum_r \lambda_{r1}\lambda_{r0}\right).$$

By the same logic, the lower bound on the DPO is

$$F(y_1, y_0) \geq 2\int \alpha_1^+(u)\alpha_0^+(1-u)du + 2\bar{\alpha}_1\bar{\alpha}_0 + \max\left(\sum_r \left(\lambda_{r1}^2 + \lambda_{r0}^2\right) - 1, \sum_r \lambda_{r1}\lambda_{s(r)0}, 0\right).$$

We emphasize that α_s^+, $\bar{\alpha}_s$, and λ_{rs} are all invariant to measure preserving transformations of α_s and ϵ_s. Since these functions are identified up to a measure preserving transformation, is follows that α_s^+, $\bar{\alpha}_s$, and λ_{rs} are all point identified, and so the upper and lower bounds provided above are point identified. Bounds on the DTE can be constructed from those on the DPO following the logic of Proposition 3 of the main text.

D.3 Estimation and inference

In this section, we provide details about how we estimated the bounds and constructed confidence intervals in Table 2 of Section 6 in the main text. As described in Section 5.3, we first estimate the graph functions h_1 and h_0, compute the eigenvalues of the estimated functions, and then plug the eigenvalues into the relevant bounds. A large econometrics and statistics literature considers the problem of estimating graph functions for dyadic data, under a wide variety of conditions, see broadly, the handbook chapter by Graham (2020). It is impossible for us to cover this entire literature here so instead we focus on the class of dyadic regression models described by Graham (2020) in his Section 4. We focus on this class of models for two reasons. First, dyadic regression models are popular in the network economics literature, being used in nearly all of the motivating examples we referenced in the introduction of the main text. Second, we can build on the estimation and inference results provided by Graham (2020) in that section.

D.3.1 Dyadic regression model

We start by summarizing the class of dyadic regression models in Section 4 of the handbook chapter by Graham (2020). In his setting, the researcher observes network connections between N agents. For each agent $i = 1, \ldots, N$, they observe a vector of covariates $X_i \in \mathbb{X}$ where $\mathbb{X}$ is a compact subset of $\mathbb{R}^L$. For each of the $N(N-1)$ pairs of agents $i, j = 1, \ldots, \binom{N}{2}$ with $i \neq j$ they observe the network connection Y_{ij} determined by the model

$$Y_{ij} = g\left(X_i, X_j, \varepsilon_{ij}\right)$$

where ε_{ij} is an unobserved error. The error ε_{ij} is decomposed into $\varepsilon_{ij} = \{U_i, U_j, V_{ij}\}$ where $\{X_i, U_i\}_{i=1}^N$ and $\{V_{ij}, V_{ji}\}_{i,j=1}^N$ are iid, have mutually independent entries, and the marginal distributions of U_i and V_{ij} are normalized to be standard uniform. Graham (2020) further normalizes the errors $\{U_i\}_{i=1}^N$ so that the entries of $\{X_i\}_{i=1}^N$ and $\{U_i\}_{i=1}^N$ are mutually independent. These normalizations are made at the bottom of his page 141 in Section 4.1.

Under this model, the outcomes $\{Y_{ij}\}_{i,j=1}^N$ are jointly exchangeable (i.e. their distribution is invariant to relabelings of the agent indices) and the outcomes Y_{ij} and $Y_{i'j'}$ are independent if $\{i, j\}$ and $\{i', j'\}$ do not share an index. In contrast to our general model given in Section 3.2.2, the link function g is not indexed by N in this setting, so the network is either dense or empty in the limit. Graham (2020) assumes that the marginal density of Y_{ij} conditional on X_i and X_j belongs to a parametric family $\mathcal{F} = \{f_{Y_{12}|X_1,X_2}(Y_{12}|X_1, X_2; \theta) : \theta \in \Theta \subseteq \mathbb{R}^K)\}$ in his Section 4.2, using θ_0 to denote the parameter that describes the data. We will use f_0 for the corresponding element of $\mathcal{F}$.

A concrete example of a dyadic regression model is the one Fafchamps and Gubert (2007) specify for informal risk sharing that we described in Example 1 of Section 3.2.2 in the main text. Additional examples applying the model to international trade, supply chain linkages across firms, R&D collaborations and more can be found in Section 4 of Graham (2020).

D.3.2 Estimation of θ_0

We now review the estimation strategy proposed by Graham (2020). Setting $l_{ij}(\theta) = \ln f_{Y_{12}|X_1,X_2}(y_{ij}|X_i, X_j; \theta)$, Graham (2020) considers the estimator that maximizes the composite likelihood

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} \frac{1}{N(N-1)} \sum_{i \neq j} l_{ij}(\theta).$$

The objective function on the right-hand side of this problem is called the composite likelihood, and not the log-likelihood function, for the connections given the covariates because it ignores the fact that the connections are dependent across pairs of agents that share an index. Still, Graham (2020) writes that, under certain regularity conditions, the resulting estimator

$\hat{\theta}$ is consistent and asymptotically normal in his Section 4.3. Specifically, he finds that

$$\sqrt{N}(\hat{\theta} - \theta_0) \to_d \mathcal{N}(0, V).$$

(his equation 33 in Section 4.3) where $V = 4\left(\Gamma_0 \Sigma_1^{-1} \Gamma_0\right)^{-1}$, Γ_0 is the probability limit of the Hessian evaluated at θ, $H_N(\theta) = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{\partial^2 l_{ij}(\theta)}{\partial \theta \partial \theta'}$, and Σ_1 is the variance of the projections of the symmetrized score $\frac{s_{ij}+s_{ji}}{2}$ where $s_{ij} = \frac{\partial l_{ij}(\theta)}{\partial \theta}$. Graham (2020) defines these parameters after his equation 28 in Section 4.3.

To estimate the variance V, Graham (2020) proposes $\hat{V} = \left(\hat{\Gamma}\hat{\Omega}\hat{\Gamma}\right)^{-1}$ in equation 43 of his Section 4.4 where $\hat{\Gamma} = H_N(\hat{\theta})$, the Hessian evaluated at $\hat{\theta}$ (equation 44), and

$$\hat{\Omega} = 4\hat{\Sigma}_1 + \frac{2}{N-1}\left(\widehat{\Sigma_2 + \Sigma_3} - 2\hat{\Sigma}_1\right)$$

where

$$\hat{\Sigma}_1 = \binom{N}{3}^{-1} \sum_{i<j<k} \frac{1}{3}\left\{\left(\frac{\hat{s}_{ij}+\hat{s}_{ji}}{2}\right)\left(\frac{\hat{s}_{ik}+\hat{s}_{ki}}{2}\right)' + \left(\frac{\hat{s}_{ij}+\hat{s}_{ji}}{2}\right)\left(\frac{\hat{s}_{jk}+\hat{s}_{kj}}{2}\right)' \right.$$
$$\left. + \left(\frac{\hat{s}_{ik}+\hat{s}_{ki}}{2}\right)\left(\frac{\hat{s}_{jk}+\hat{s}_{kj}}{2}\right)'\right\},$$

$$\widehat{\Sigma_2 + \Sigma_3} = \binom{N}{2}^{-1} \sum_{i<j} \left(\frac{\hat{s}_{ij}+\hat{s}_{ji}}{2}\right)\left(\frac{\hat{s}_{ij}+\hat{s}_{ji}}{2}\right)',$$

and $\hat{s}_{ij}$ is the score s_{ij} evaluated at $\hat{\theta}$ (equation 45). Graham (2020) also proposes a jackknife estimator for the variance and considers bootstrap inference in his Section 4.5 following Menzel (2017).

D.3.3 Potential outcomes and the parameter of interest

In this subsection, we incorporate the dyadic regression model into our framework, following the logic of our Section 3.2.3. Specifically, we suppose that the potential connections under each policy $s \in \{0,1\}$ are described by a dyadic regression model

$$Y_{ij}(s) = g_s(X_{i,s}, X_{j,s}, \varepsilon_{ij,s}).$$

where $X_{i,s} \in \mathbb{X}$ is the vector of covariates observed under policy $s \in \{0,1\}$. Importantly, we do not assume that $X_{i,1}$ and $X_{i,0}$ are equal, so that the covariate value for agent i depends on the policy that was implemented. Following Section 3.2.3, we allow g_1 and g_0 to be any function (so long as the conditional density of $Y_{ij}(s)$ given $X_{i,s}$ and $X_{j,s}$ belongs to an element of $\mathcal{F}$ defined in Section D.3.1) and we model the dependence between $X_{i,1}$ and $X_{i,0}$ as $X_{i,s} = \varphi_s(w_i)$ where $\{w_i\}_{i=1}^N$ is an unobserved iid random variable with standard uniform marginals and the function φ_s is an unknown measure preserving transformation, except φ_s now takes values in $\mathbb{X}$ instead of $[0,1]$. We also assume that $\{\varepsilon_{ij,1}\}_{i,j=1}^N$ and $\{\varepsilon_{ij,0}\}_{i,j=1}^N$ have mutually independent entries. For $s \in \{0,1\}$ use $f_s \in \mathcal{F}$ to describe the data under policy s, determined by $\theta_s \in \Theta$.

As in Section 4.1, we define the graph function to be

$$h_s(a, b; \theta_s) = \mathbb{P}\left(Y_{ij}(s) \le y_s | w_i = a, w_j = b; \theta_s\right)$$

where we continue to suppress the dependence of the parameter on the value of y_s, but have now added the parameter θ_s that determines f_s. In this dyadic regression setting, we note that the graph function can be written as a functional of f_s, i.e.

$$h_s(a, b; \theta_s) = \int f_{Y_{12}|X_1,X_2}(\tau|\varphi_s(a), \varphi_s(b); \theta_s) \mathbb{1}\{\tau \le y_s\} d\tau.$$

We assume that $\mathcal{F}$ is defined so that h_s is twice continuously differentiable in θ_s with a uniformly bounded derivative.

Following the logic of Section 4.2, the DPO is

$$F(y_1, y_0) = \mathbb{P}\left(Y_{ij}(1) \le y_1, Y_{ij}(0) \le y_0\right) = \iint h_1(u, v; \theta_1) h_0(u, v; \theta_0) du dv.$$

By Lemma 1 in the main text, the functions h_1 and h_0 are identified up to a measure-preserving transformation. The identified set for the DPO is given by Proposition 1.

Remark 1. We emphasize that even though in the dyadic regression setting the researcher observes covariates and assumes that the marginal distribution of connections conditional on the covariates belongs to a parametric family, this additional information has no effect on the identification analysis of Section 4. This is because, in that analysis, the graph functions were identified up to a measure preserving transformation. While the covariates and parametric assumptions further restrict the shape of the graph functions, they cannot distinguish between graph functions that are equivalent up to a measure preserving transformation because there is no restriction on how the covariates across the two policies are related. And so the identified set is unchanged.

Intuitively, this means that, in terms of identification, the covariates and parametric restriction are completely unnecessary because they only reveal information about the DPO that, according to our Proposition 1 in the main text, is already known from the distribution of the data. The additional information may be relevant for estimation and inference, however, which is why we make use of it in this section.

D.3.4 Outer bounds on the identified set

We specify outer bounds on the identified set adjusting for row and column heterogeneity as described in Section D.2. Specifically, we decompose $h_s(u,v) = \alpha_s(u) + \alpha_s(v) + \epsilon_s(u,v)$ where $\alpha_s(u) = \int h_s(u,v)dv - \frac{1}{2}\iint h_s(u,v)dudv$ and $\epsilon_s(u,v) = h_s(u,v) - \int h_s(u,v)du - \int h_s(u,v)dv + \iint h_s(u,v)dudv$. Following the logic of Online Appendix Section D.2, the upper bound on the identified set for the DPO is

$$U = 2\int \alpha_1^+(u)\alpha_0^+(u)du + 2\int \alpha_1^+(u)du \int \alpha_0^+(u)du + \min\left(\sum_r \lambda_{r0}^2, \sum_r \lambda_{r1}^2, \sum_r \lambda_{r0}\lambda_{r1}\right)$$

and the lower bound on the identified set is

$$L = 2\int \alpha_1^+(u)\alpha_0^+(1-u)du + 2\int \alpha_1^+(u)du \int \alpha_0^+(u)du + \max\left(\sum_r \left(\lambda_{r0}^2 + \lambda_{r1}^2\right) - 1, \sum_r \lambda_{r0}\lambda_{\rho(r)1}, 0\right)$$

where α_s^+ is the quantile function associated with α_s, λ_{rs} is the rth eigenvalue (ordered to be decreasing) of ϵ_s, and the infinite sums are as defined in Section 4.1.2 of the main text.

As pointed out in that section, the DPO, $F(y_1, y_0) \in [L, U]$. In addition, α_s^+ and λ_{rs} are invariant to measure preserving transformations of h_s and so are point identified because h_s is identified up to a measure preserving transformation by Lemma 1 of the main text.

D.3.5 Estimation of the outer bounds

For each $s \in \{0, 1\}$, we separately estimate θ_s by maximizing the composite likelihood of Graham (2020) using the data on the network connections and covariates for the agents assigned to community s, and call this estimator $\hat{\theta}_s$. We then use $\hat{\theta}_s$ to construct a plug-in estimator for the graph function h_s, $\tilde{h}_s(a, b) = \int f_{Y_{12}|X_1,X_2}(\tau|a, b; \hat{\theta}_s)\mathbb{1}\{\tau \leq y_s\}d\tau$.

The estimator $\tilde{h}_s$ differs from its estimand h_s in two ways. First, it is defined using $\hat{\theta}_s$ instead of θ_s. Second, it omits the measure preserving transformations φ_s. The first difference results in an estimation error which we account for in our confidence interval below. The second difference does not result in an estimation error. This is because our outer bounds proposed in Section D.3.2 only depend on the quantiles of α_s and the eigenvalues of ϵ_s, both of which are invariant to measure preserving transformations of the graph function. It follows that using the estimated graph function $\tilde{h}_s$ leads to bounds that are equivalent to the infeasible estimate $\hat{h}_s(a, b) = \int f_s(\tau|\varphi_s(a), \varphi_s(b); \hat{\theta}_s)\mathbb{1}\{\tau \leq y_s\}d\tau$.

We use $\hat{\alpha}_s(u) = \int \tilde{h}_s(u, v)dv - \frac{1}{2}\int\int \tilde{h}_s(u, v)dudv$ as our estimator for $\alpha_s(u)$ and $\hat{\epsilon}_s(u, v) = \tilde{h}_s(u, v) - \int \tilde{h}_s(u, v)du - \int \tilde{h}_s(u, v)dv + \int\int \tilde{h}_s(u, v)dudv$ as our estimator for $\epsilon_s(u, v)$. Our estimators of the bounds described in Online Appendix Section D.3.2 above are then

$$\hat{U} = 2\int \hat{\alpha}_1^+(u)\hat{\alpha}_0^+(u)du + 2\int \hat{\alpha}_1^+(u)du \int \hat{\alpha}_0^+(u)du + \min\left(\sum_r \hat{\lambda}_{r0}^2, \sum_r \hat{\lambda}_{r1}^2, \sum_r \hat{\lambda}_{r0}\hat{\lambda}_{r1}\right), \text{ and}$$

$$\hat{L} = 2\int \hat{\alpha}_1^+(u)\hat{\alpha}_0^+(1-u)du + 2\int \hat{\alpha}_1^+(u)du \int \hat{\alpha}_0^+(u)du + \max\left(\sum_r \left(\hat{\lambda}_{r0}^2 + \hat{\lambda}_{r1}^2\right) - 1, \sum_r \hat{\lambda}_{r0}\hat{\lambda}_{\rho(r)1}, 0\right)$$

where $\hat{\alpha}_s^+$ is the quantile function of $\hat{\alpha}_s$ and $\hat{\lambda}_{rs}$ is the rth eigenvalue (ordered to be decreasing) of $\hat{\epsilon}_s$.

Since these bounds are equivalent to those bounds based on the infeasible estimator $\hat{h}_s$, it is without loss to conduct our statistical analysis below as though we had used $\hat{h}_s$ instead of $\tilde{h}_s$ in their construction. See Section D.4 below for a discussion.

D.3.6 Confidence interval

For a fixed $\alpha > 0$, we propose the confidence interval

$$S_\alpha = \left[\hat{L} - N^{-1/2} C_\alpha^L, \hat{U} + N^{-1/2} C_\alpha^U\right]$$

where C_α^L and C_α^U are chosen such that

$$\mathbb{P}\left(\sum_{j=1}^{6}\left(\xi' \hat{\Omega}_j^L \xi\right)^{1/2} \le C_\alpha^L, \sum_{j=1}^{4}\left(\xi' \hat{\Omega}_j^U \xi\right)^{1/2} \le C_\alpha^U\right) \ge 1-\alpha, \tag{1}$$

$\xi \sim \mathcal{N}(0, I_K)$, $\hat{\Omega}_j^L = \hat{V}^{1/2} W_j^L \hat{V}^{1/2}$, $\hat{\Omega}_j^U = \hat{V}^{1/2} W_j^U \hat{V}^{1/2}$, and W_j^L and W_j^U are defined in Online Appendix Section D.3.5 below. It follows from Proposition D3 below that if C_α^L and C_α^U satisfy condition (1) then $\liminf_{N\to\infty} \mathbb{P}([L, U] \subseteq S_\alpha) \ge 1 - \alpha$. Since our confidence interval is valid for the identified set, it is potentially conservative for the DPO, see broadly Section 4.3.1 of Molinari (2020). The arguments of Imbens and Manski (2004); Stoye (2009) could be applied here to, in some cases, produce a shorter interval. However, we do not formally describe such a refinement here.[3]

There are typically multiple choices of C_α^L and C_α^U that satisfy condition (1). One way to choose these parameters is to first take R draws $\{\xi_r\}_{r=1}^R$ from $\mathcal{N}(0, I_K)$, where R is a large positive integer like 10000, compute the $1-\alpha$ isoquant of the empirical joint distribution of $\left\{\left(\sum_{j=1}^{6}\left(\xi_r' \hat{\Omega}_j^L \xi_r\right)^{1/2}, \sum_{j=1}^{4}\left(\xi_r' \hat{\Omega}_j^U \xi_r\right)^{1/2}\right)\right\}_{r=1}^R$, and find the point on the isoquant such that the sum of the inputs are minimized. This strategy is designed to approximate an interval with the smallest possible length. For our empirical results in Section 6 of the main text we use a conservative choice of C_α^L and C_α^U that avoids computing the isoquant of a joint distribution function. That is, we choose C_α^L to be the $1-\alpha/2$ quantile of $\left\{\sum_{j=1}^{6}\left(\xi_r' \hat{\Omega}_j^L \xi_r\right)^{1/2}\right\}_{r=1}^R$ and C_α^U to be the $1-\alpha/2$ quantile of $\left\{\sum_{j=1}^{4}\left(\xi_r' \hat{\Omega}_j^U \xi_r\right)^{1/2}\right\}_{r=1}^R$. This choice is conservative because $\mathbb{P}\left(\sum_{j=1}^{6}\left(\xi' \hat{\Omega}_j^L \xi\right)^{1/2} \le C_\alpha^L, \sum_{j=1}^{4}\left(\xi' \hat{\Omega}_j^U \xi\right)^{1/2} \le C_\alpha^U\right) \ge$

[3] This literature typically focuses on uniform validity results, which we do not provide here. We can show that our intervals are uniformly valid if the limiting results provided by Graham (2020) in Section D.3.2 hold uniformly. However, since Graham (2020) only states pointwise results in his handbook chapter, we do the same here. We conjecture that it is straightforward to, under certainty additional conditions, extend his results to be uniform, but we leave the details to future work.

$\mathbb{P}\left(\sum_{j=1}^{6}\left(\xi'\hat{\Omega}_j^L\xi\right)^{1/2} \leq C_\alpha^L\right)+\mathbb{P}\left(\sum_{j=1}^{4}\left(\xi'\hat{\Omega}_j^U\xi\right)^{1/2} \leq C_\alpha^U\right) - 1 = 1-\alpha.$

D.3.7 List of weight matrices

Our confidence interval described above depends on several weight matrices which we describe here. We use the following definitions: $\hat{h}'_{sk}(u,v) = \partial h_s(u,v;\theta)/\partial\theta_{sk}$ where θ_{sk} is the kth entry of θ_s, $\bar{\hat{\alpha}}_s = \int \hat{\alpha}_s^+(u)du$, and $\bar{\alpha}_s = \int \alpha_s^+(u)du$. For the upper bound weights, $(\tau,\tau') \in \operatorname{argmin}_{s,s'\in\{0,1\}} \sum_r \lambda_{rs}\lambda_{rs'}$. The upper bound weights are

$$
\begin{aligned}
W_{kl,1}^U &= 4\int\left(\int \hat{h}'_{1k}(u,v)dv + \iint \hat{h}'_{1k}(u,v)dudv\right) \\
&\qquad \left(\int \hat{h}'_{1l}(u,v)dv + \iint \hat{h}'_{1l}(u,v)dudv\right) du \left(\int \hat{\alpha}_0(u)^2 du\right) \\
W_{kl,2}^U &= 4\left(\int \left(\hat{\alpha}_1(u)+\bar{\hat{\alpha}}_1\right)^2 du\right)\int\left(\int \hat{h}'_{0k}(u,v)dv - \frac{1}{2}\int\int \hat{h}'_{0k}(u,v)dudv\right) \\
&\qquad \left(\int \hat{h}'_{0l}(u,v)dv - \frac{1}{2}\int\int \hat{h}'_{0l}(u,v)dudv\right) du \\
W_{kl,3}^U &= \int\int\left(\hat{h}'_{\tau k}(u,v) - \int \hat{h}'_{\tau k}(u,v)du - \int \hat{h}'_{\tau k}(u,v)dv\right. \\
&\quad \left. + \int\int \hat{h}'_{\tau k}(u,v)dudv\right)\left(\hat{h}'_{\tau l}(u,v) - \int \hat{h}'_{\tau l}(u,v)dv\right. \\
&\quad \left. - \int \hat{h}'_{\tau l}(u,v)du + \int\int \hat{h}'_{\tau l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}_{r\tau'}^2\right) \\
W_{kl,4}^U &= \int\int\left(\hat{h}'_{\tau' k}(u,v) - \int \hat{h}'_{\tau' k}(u,v)du - \int \hat{h}'_{\tau' k}(u,v)dv\right. \\
&\quad \left. + \int\int \hat{h}'_{\tau' k}(u,v)dudv\right)\left(\hat{h}'_{\tau' l}(u,v) - \int \hat{h}'_{\tau' l}(u,v)dv\right. \\
&\quad \left. - \int \hat{h}'_{\tau' l}(u,v)du + \int\int \hat{h}'_{\tau' l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}_{r\tau}^2\right).
\end{aligned}
$$

For the lower bound weights,

$$
(\tau,\tau') \in \operatorname*{argmax}_{t,t'\in\{0,1\}}\left(\left(\sum_r \lambda_{rt}\lambda_{s(r)t'}\right)\mathbb{1}\{t\neq t'\} + \left(\sum_r \lambda_{rt}\lambda_{rt'} + \sum_r \lambda_{r(1-t)}\lambda_{r(1-t')} - 1\right)\mathbb{1}\{t=t'\}\right)_+
$$

and $\sigma \in \{0,1\}$. The lower bound weights are

$$
\begin{aligned}
W^L_{kl,1} = {} & 4\int \left(\int \hat{h}'_{1k}(u,v)dv + \iint \hat{h}'_{1k}(u,v)dudv\right) \\
& \left(\int \hat{h}'_{1l}(u,v)dv + \iint \hat{h}'_{1l}(u,v)dudv\right) du \left(\int \hat{\alpha}_0(u)^2 du\right) \\
W^L_{kl,2} = {} & 4\left(\int \left(\hat{\alpha}_1(u) + \bar{\hat{\alpha}}_1\right)^2 du\right) \int \left(\int \hat{h}'_{0k}(u,v)dv - \frac{1}{2}\int\int \hat{h}'_{0k}(u,v)dudv\right) \\
& \left(\int \hat{h}'_{0l}(u,v)dv - \frac{1}{2}\int\int \hat{h}'_{0l}(u,v)dudv\right) du \\
W^L_{kl,3} = {} & \left(\int\int \left(\hat{h}'_{\tau'k}(u,v) - \int \hat{h}'_{\tau'k}(u,v)du - \int \hat{h}'_{\tau'k}(u,v)dv + \int\int \hat{h}'_{\tau'k}(u,v)dudv\right)\right. \\
& \left(\hat{h}'_{\tau'l}(u,v) - \int \hat{h}'_{\tau'l}(u,v)dv - \int \hat{h}'_{\tau'l}(u,v)du\right. \\
& \left.\left.+ \int\int \hat{h}'_{\tau'l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}^2_{r\tau}\right)\right) \mathbb{1}\{\tau \neq \tau'\} \\
W^L_{kl,4} = {} & \left(\int\int \left(\hat{h}'_{\tau'k}(u,v) - \int \hat{h}'_{\tau'k}(u,v)du - \int \hat{h}'_{\tau'k}(u,v)dv + \int\int \hat{h}'_{\tau'k}(u,v)dudv\right)\right. \\
& \left(\hat{h}'_{\tau l}(u,v) - \int \hat{h}'_{\tau l}(u,v)dv - \int \hat{h}'_{\tau l}(u,v)du\right. \\
& \left.\left.+ \int\int \hat{h}'_{\tau l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}^2_{r\tau'}\right)\right) \mathbb{1}\{\tau \neq \tau'\} \\
W^L_{kl,5+\sigma} = {} & 4\left(\int\int \left(\hat{h}'_{\sigma k}(u,v) - \int \hat{h}'_{\sigma k}(u,v)du - \int \hat{h}'_{\sigma k}(u,v)dv + \int\int \hat{h}'_{\sigma k}(u,v)dudv\right)\right. \\
& \left(\hat{h}'_{\sigma l}(u,v) - \int \hat{h}'_{\sigma l}(u,v)dv - \int \hat{h}'_{\sigma l}(u,v)du\right. \\
& \left.\left.+ \int\int \hat{h}'_{\sigma l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}^2_{r\sigma}\right)\right) \mathbb{1}\{\tau = \tau'\}.
\end{aligned}
$$

D.4 Proof of the consistency claim

Our main justification for the confidence interval S_α is Proposition D3 in Online Appendix Section D.4.2 below. Its proof relies on Lemma D2 which we state and demonstrate first.

D.4.1 Consistency lemma

Lemma D2: (i) $\sup_{u\in[0,1]} |\hat{\alpha}^+_s(u) - \alpha^+_s(u)| = O_p(N^{-1/2})$,
(ii) $\sup_{i\in\mathbb{N}} \left|\hat{\lambda}^+_{is} - \lambda^+_{is}\right| = O_p(N^{-1/2})$, and (iii) $\sup_{i\in\mathbb{N}} \left|\hat{\lambda}^-_{is} - \lambda^-_{is}\right| = O_p(N^{-1/2})$ where

$\hat{\lambda}_{is}^{+}$ and λ_{is}^{+} are the ith positive eigenvalue of $\hat{\epsilon}_s$ and ϵ_s, and $\hat{\lambda}_{is}^{-}$ and λ_{is}^{-} are the ith negative eigenvalue of $\hat{\epsilon}_s$ and ϵ_s. The eigenvalues are all ordered to be decreasing in magnitude.

Proof of Lemma D2: Since h_s is continuously differentiable with a uniformly bounded derivative and X_s is bounded, it follows from Assumption D1 and the mean value theorem that $\sup_{u,v}\left|\left(\hat{h}_s(u,v) - h_s(u,v)\right)\right| = O_p\left(|\hat{\theta}-\theta|\right) = O_p(N^{-1/2})$. The claim then follows from the fact that $\hat{\alpha}_s^{+}(u)$, $\hat{\lambda}_{is}^{+}$, and $\hat{\lambda}_{is}^{-}$ are Lipschitz continuous functions of $\hat{h}_s$, and $\alpha_s^{+}(u)$, λ_{is}^{+}, and λ_{is}^{-} are Lipschitz continuous functions of h_s (see Lemma B3 in Section B above). $\square$

D.4.2 Result

Proposition D3: $\liminf_{N\to\infty}\inf_{\theta\in[L,U]}\mathbb{P}\left([L,U]\subseteq S_\alpha\right)\geq 1-\alpha$

Proof of Proposition D3: We break up the proof of Proposition D3 into three parts. The first two parts derive bounds on the estimation error of $\hat{U}$ and $\hat{L}$. The third part combines the bounds from the first two parts to demonstrate the claim.

D.4.3 Step 1: Estimation error for the upper bound

We first bound the estimation error of $\hat{U}$.

$$\begin{aligned}
\hat{U}-U &= 2\int\left(\hat{\alpha}_1^{+}(u)+\bar{\hat{\alpha}}_1\right)\hat{\alpha}_0^{+}(u)du - 2\int\left(\alpha_1^{+}(u)+\bar{\alpha}_1\right)\alpha_0^{+}(u)du \\
&\qquad + \min_{t,t'\in\{0,1\}}\sum_r \hat{\lambda}_{rt}\hat{\lambda}_{rt'} - \min_{t,t'\in\{0,1\}}\sum_r \lambda_{rt}\lambda_{rt'} \\
&= 2\int\left(\hat{\alpha}_1^{+}(u)+\bar{\hat{\alpha}}_1-\alpha_1^{+}(u)-\bar{\alpha}_1\right)\hat{\alpha}_0^{+}(u)du + 2\int\left(\alpha_1^{+}(u)+\bar{\alpha}_1\right)\left(\hat{\alpha}_0^{+}(u)-\alpha_0^{+}(u)\right)du \\
&\qquad + \sum_r\left(\hat{\lambda}_{r\tau}-\lambda_{r\tau}\right)\hat{\lambda}_{r\tau'} + \sum_r \lambda_{r\tau}\left(\hat{\lambda}_{r\tau'}-\lambda_{r\tau'}\right)
\end{aligned}$$

where $\bar{\hat{\alpha}}_t = \int \hat{\alpha}_t^{+}(u)du$, $\bar{\alpha}_t = \int \alpha_t^{+}(u)du$, $(\tau,\tau')\in\operatorname{argmin}_{t,t'\in\{0,1\}}\sum_r \lambda_{rt}\lambda_{rt'}$, and the second equality holds eventually because $\sum_r \hat{\lambda}_{r\tau}\left(\hat{\lambda}_{r\tau'}-\lambda_{r\tau'}\right) = \min_{t,t'\in\{0,1\}}\sum_r \lambda_{rt}\lambda_{rt'}$ by definition of (τ,τ') and $\sum_r\left(\hat{\lambda}_{r\tau}-\lambda_{r\tau}\right)\hat{\lambda}_{r\tau'} = \min_{t,t'\in\{0,1\}}\sum_r \hat{\lambda}_{rt}\hat{\lambda}_{rt'}$ eventually by Lemma D2 and the continuous mapping theorem. We analyze the four summands in the second line of the displayed equation separately. The first summand can be bounded

$$
\begin{aligned}
&\left|2\int\left(\hat{\alpha}_1^+(u)+\bar{\hat{\alpha}}_1-\alpha_1^+(u)-\bar{\alpha}_1\right)\hat{\alpha}_0^+(u)du\right| \\
&\quad\leq 2\left(\int\left(\hat{\alpha}_1^+(u)-\alpha_1^+(u)+\bar{\hat{\alpha}}_1-\bar{\alpha}_1\right)^2\right)^{1/2}\left(\int\hat{\alpha}_0(u)^2du\right)^{1/2} \\
&\quad\leq 2\left(\int\left(\hat{\alpha}_1(u)-\alpha_1(u)+\bar{\hat{\alpha}}_1-\bar{\alpha}_1\right)^2\right)^{1/2}\left(\int\hat{\alpha}_0(u)^2du\right)^{1/2} \\
&\quad= 2\left(\int\left(\left(\int\hat{h}_1(u,v)-h_1(u,v)dv\right)+\left(\iint\hat{h}_1(u,v)-h_1(u,v)dudv\right)\right)^2du\right)^{1/2}\left(\int\hat{\alpha}_0(u)^2du\right)^{1/2} \\
&\quad= 2\left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)\int\left(\left(\int\hat{h}'_{1k}(u,v)dv+\iint\hat{h}'_{1k}(u,v)dudv\right)\right.\right. \\
&\qquad\left.\left.\left(\int\hat{h}'_{1l}(u,v)dv+\iint\hat{h}'_{1l}(u,v)dudv\right)\right)^2du\right)^{1/2}\left(\int\hat{\alpha}_0(u)^2du\right)^{1/2}+o_p\left(N^{-1/2}\right) \\
&\quad= \left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)W^U_{kl,1}\right)^{1/2}+o_p\left(N^{-1/2}\right)
\end{aligned}
$$

where $\hat{h}'_{sk}(u,v) = h'_{sk}(\hat{\theta})$, $h'_{sk}(\theta) = \partial h_s(u,v;\theta)/\partial\theta_{sk}$ refers to the partial derivative of h_s with respect to the kth element of θ_s at the point (u,v) and θ, the first inequality is due to Cauchy-Schwarz, the second inequality is due to the functional version of Hardy-Littlewood-Polya Theorem 368 (see Lemma B5 in Online Appendix Section B.1), the first equality is due to plugging in the definition of $\hat{\alpha}_1$, α_1, $\bar{\hat{\alpha}}_1$, and $\bar{\alpha}_1$, and algebra, and the second equality is due to the definition of $\hat{h}_1$ and h_1, a first-order Taylor expansion, Proposition D3, and more algebra. The same exact arguments can be applied to the second summand

$$
\begin{aligned}
&\left|2\int\left(\alpha_1^+(u)+\bar{\alpha}_1\right)\left(\hat{\alpha}_0^+(u)-\alpha_0^+(u)\right)du\right| \\
&\leq 2\left(\int\left(\alpha_1(u)+\bar{\alpha}_1\right)^2du\right)^{1/2}\left(\int\left(\hat{\alpha}_0(u)-\alpha_0(u)\right)^2du\right)^{1/2} \\
&= 2\left(\int\left(\hat{\alpha}_1(u)+\bar{\hat{\alpha}}_1\right)^2du\right)^{1/2}\left(\int\left(\hat{\alpha}_0(u)-\alpha_0(u)\right)^2du\right)^{1/2}+o_p\left(N^{-\gamma}\right) \\
&= 2\left(\sum_{k,l=1}^{K}(\hat{\theta}_k-\theta_k)(\hat{\theta}_l-\theta_l)\left(\int\left(\hat{\alpha}_1(u)+\bar{\hat{\alpha}}_1\right)^2du\right)\int\left(\int\hat{h}'_{0k}(u,v)dv-\frac{1}{2}\int\int\hat{h}'_{0k}(u,v)dudv\right)\right. \\
&\quad\left.\left(\int\hat{h}'_{0l}(u,v)dv-\frac{1}{2}\int\int\hat{h}'_{0l}(u,v)dudv\right)du\right)^{1/2}+o_p\left(N^{-1/2}\right) \\
&= \left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)W^U_{kl,2}\right)^{1/2}+o_p\left(N^{-1/2}\right).
\end{aligned}
$$

The third summand can be bounded

$$
\begin{aligned}
&\left|\sum_r\left(\hat{\lambda}_{r\tau}-\lambda_{r\tau}\right)\hat{\lambda}_{r\tau'}\right| \leq \left(\sum_r\left(\hat{\lambda}_{r\tau}-\lambda_{r\tau}\right)^2\right)^{1/2}\left(\sum_r\hat{\lambda}^2_{r\tau'}\right)^{1/2} \\
&\leq \left(\int\int\left(\hat{\epsilon}_\tau(u,v)-\epsilon_\tau(u,v)\right)^2dudv\right)^{1/2}\left(\sum_r\hat{\lambda}^2_{r\tau'}\right)^{1/2} \\
&= \left(\int\int\left(\left(\hat{h}_\tau(u,v)-h_\tau(u,v)\right)-\int\left(\hat{h}_\tau(u,v)-h_\tau(u,v)\right)dv-\int\left(\hat{h}_\tau(u,v)-h_\tau(u,v)\right)du\right.\right. \\
&\quad\left.\left.+\iint\left(\hat{h}_\tau(u,v)-h_\tau(u,v)\right)dudv\right)^2dudv\right)^{1/2}\left(\sum_r\hat{\lambda}^2_{r\tau'}\right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\sum_{k,l=1}^{K}(\hat{\theta}_k-\theta_k)(\hat{\theta}_l-\theta_l)\int\int\left(\hat{h}'_{\tau k}(u,v)-\int\hat{h}'_{\tau k}(u,v)du-\int\hat{h}'_{\tau k}(u,v)dv+\int\int\hat{h}'_{\tau k}(u,v)dudv\right)\right. \\
&\quad\left.\left(\hat{h}'_{\tau l}(u,v)-\int\hat{h}'_{\tau l}(u,v)dv-\int\hat{h}'_{\tau l}(u,v)du+\int\int\hat{h}'_{\tau l}(u,v)dudv\right)dudv\left(\sum_r\hat{\lambda}^2_{r\tau'}\right)\right)^{1/2}+o_p\left(N^{-1/2}\right) \\
&= \left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)W^U_{kl,3}\right)^{1/2}+o_p\left(N^{-1/2}\right)
\end{aligned}
$$

where the first inequality is due to Cauchy-Schwarz, the second inequality is due to our Proposition 4, the first equality is due to plugging in the definition of $\hat{\epsilon}_\tau$ and ϵ_τ, and the second equality is due to the definition of $\hat{\mu}_\tau$ and μ_τ, a first-order Taylor expansion, Proposition D2, and some algebra. The same exact arguments can be applied to bound the fourth summand

$$
\begin{aligned}
&\left|\sum_r \lambda_{r\tau}\left(\hat{\lambda}_{r\tau'} - \lambda_{r\tau'}\right)\right| \\
&\leq \Bigg(\sum_{k,l=1}^{K} (\hat{\theta}_k - \theta_k)(\hat{\theta}_l - \theta_l) \int\int \left(\hat{h}'_{\tau'k}(u,v) - \int \hat{h}'_{\tau'k}(u,v)du - \int \hat{h}'_{\tau'k}(u,v)dv + \int\int \hat{h}'_{\tau'k}(u,v)dudv\right) \\
&\quad \left(\hat{h}'_{\tau'l}(u,v) - \int \hat{h}'_{\tau'l}(u,v)dv - \int \hat{h}'_{\tau'l}(u,v)du + \int\int \hat{h}'_{\tau'l}(u,v)dudv\right) dudv \left(\sum_r \hat{\lambda}^2_{r\tau}\right)\Bigg)^{1/2} + o_p\left(N^{-1/2}\right) \\
&= \left(\sum_{k,l=1}^{K} \left(\hat{\theta}_k - \theta_k\right)\left(\hat{\theta}_l - \theta_l\right) W^U_{kl,4}\right)^{1/2} + o_p\left(N^{-1/2}\right).
\end{aligned}
$$

Combining the bounds for all four results, it follows that

$$
N^{1/2}|\hat{U} - U| \leq \sum_{j=1}^{4} \left(\sum_{k,l=1}^{K} N^{1/2}\left(\hat{\theta}_k - \theta_k\right) N^{1/2}\left(\hat{\theta}_l - \theta_l\right) W^U_{kl,j}\right)^{1/2} + o_p\left(1\right).
$$

D.4.4 Step 2: Estimation error for the lower bound

We now bound the estimation error of $\hat{L}$.

$$\begin{aligned}
\hat{L} - L = {} & 2\int \left(\hat{\alpha}_1^+(u) + \bar{\hat{\alpha}}_1\right)\hat{\alpha}_0^+(1-u)du - 2\int \left(\alpha_1^+(u) + \bar{\alpha}_1\right)\alpha_0^+(1-u)du \\
& + \max_{t,t' \in \{0,1\}} \left(\left(\sum_r \hat{\lambda}_{rt}\hat{\lambda}_{s(r)t'} \right) \mathbb{1}\{t \neq t'\} + \left(\sum_r \hat{\lambda}_{rt}\hat{\lambda}_{rt'} + \sum_r \hat{\lambda}_{r(1-t)}\hat{\lambda}_{r(1-t')} - 1 \right) \mathbb{1}\{t = t'\} \right)_+ \\
& - \max_{t,t' \in \{0,1\}} \left(\left(\sum_r \lambda_{rt}\lambda_{s(r)t'} \right) \mathbb{1}\{t \neq t'\} + \left(\sum_r \lambda_{rt}\lambda_{rt'} + \sum_r \lambda_{r(1-t)}\lambda_{r(1-t')} - 1 \right) \mathbb{1}\{t = t'\} \right)_+ \\
= {} & 2\int \left(\hat{\alpha}_1^+(u) + \bar{\hat{\alpha}}_1 - \alpha_1^+(u) - \bar{\alpha}_1\right)\hat{\alpha}_0^+(1-u)du \\
& + 2\int \left(\alpha_1^+(u) + \bar{\alpha}_1\right)\left(\hat{\alpha}_0^+(1-u) - \alpha_0^+(1-u)\right)du \\
& + \left(\left(\sum_r \hat{\lambda}_{r\tau}\hat{\lambda}_{s(r)\tau'} \right)_+ - \left(\sum_r \lambda_{rt}\lambda_{s(r)t'} \right)_+ \right) \mathbb{1}\{\tau \neq \tau'\} \\
& + \left(\left(\sum_r \hat{\lambda}_{r\tau}\hat{\lambda}_{r\tau'} + \sum_r \hat{\lambda}_{r(1-\tau)}\hat{\lambda}_{r(1-\tau')} - 1 \right)_+ \right. \\
& \qquad \left. - \left(\sum_r \lambda_{r\tau}\lambda_{r\tau'} + \sum_r \lambda_{r(1-\tau)}\lambda_{r(1-\tau')} - 1 \right)_+ \right) \mathbb{1}\{\tau = \tau'\}
\end{aligned}$$

where $\bar{\hat{\alpha}}_t = \int \hat{\alpha}_t^+(u)du$, $\bar{\alpha}_t = \int \alpha_t^+(u)du$,

$$(\tau, \tau') \in \operatorname*{argmax}_{t,t' \in \{0,1\}} \left(\left(\sum_r \lambda_{rt}\lambda_{s(r)t'} \right) \mathbb{1}\{t \neq t'\} + \left(\sum_r \lambda_{rt}\lambda_{rt'} + \sum_r \lambda_{r(1-t)}\lambda_{r(1-t')} - 1 \right) \mathbb{1}\{t = t'\} \right)_+$$

and the second equality holds eventually by Lemma D2 and the continuous mapping theorem. We analyze the four summands in the second line of the displayed equation separately.

The first two summands can be bounded uniformly over $\mathcal{P}$

$$
\begin{aligned}
&\left|2\int\left(\hat{\alpha}_1^+(u)+\bar{\hat{\alpha}}_1-\alpha_1^+(u)-\bar{\alpha}_1\right)\hat{\alpha}_0^+(1-u)du\right| \\
&\leq 2\left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)\int\left(\left(\int\hat{h}'_{1k}(u,v)dv+\iint\hat{h}'_{1k}(u,v)dudv\right)\right.\right. \\
&\quad\left.\left.\left(\int\hat{h}'_{1l}(u,v)dv+\iint\hat{h}'_{1l}(u,v)dudv\right)\right)^2du\right)^{1/2}\left(\int\hat{\alpha}_0(u)^2du\right)^{1/2}+o_p\left(N^{-1/2}\right) \\
&=\left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)W^L_{kl,1}\right)^{1/2}+o_p\left(N^{-1/2}\right)\text{ and}
\end{aligned}
$$

$$
\begin{aligned}
&\left|2\int\left(\alpha_1^+(u)+\bar{\alpha}_1\right)\left(\hat{\alpha}_0^+(1-u)-\alpha_0^+(1-u)\right)du\right| \\
&\leq 2\left(\sum_{k,l=1}^{K}(\hat{\theta}_k-\theta_k)(\hat{\theta}_l-\theta_l)\left(\int\left(\hat{\alpha}_1(u)+\bar{\hat{\alpha}}_1\right)^2du\right)\right. \\
&\quad\left.\int\left(\int\hat{h}'_{0k}(u,v)-\frac{1}{2}\int\int\hat{h}'_{0k}(u,v)dudv\right)\left(\int\hat{h}'_{0l}(u,v)dv-\frac{1}{2}\int\int\hat{h}'_{0l}(u,v)dudv\right)du\right)^{1/2}+o_p\left(N^{-1/2}\right) \\
&=\left(\sum_{k,l=1}^{K}\left(\hat{\theta}_k-\theta_k\right)\left(\hat{\theta}_l-\theta_l\right)W^L_{kl,2}\right)^{1/2}+o_p\left(N^{-1/2}\right)
\end{aligned}
$$

where h' is the first derivative of h, following the arguments of Step 1 from Section D.4.3. The second two summands can also be bounded

$$
\begin{aligned}
&\left|\left(\left(\sum_r\hat{\lambda}_{r\tau}\hat{\lambda}_{s(r)\tau'}\right)_+-\left(\sum_r\lambda_{rt}\lambda_{s(r)t'}\right)_+\right)\mathbb{1}\{\tau\neq\tau'\}\right| \\
&\leq\left|\left(\sum_r\left(\hat{\lambda}_{r\tau}-\lambda_{r\tau}\right)\hat{\lambda}_{s(r)\tau'}+\sum_r\lambda_{r\tau}\left(\hat{\lambda}_{s(r)\tau'}-\lambda_{s(r)\tau}\right)\right)\mathbb{1}\{\tau\neq\tau'\}\right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \Bigg(\Bigg(\sum_{k,l=1}^{K} (\hat{\theta}_k - \theta_k)(\hat{\theta}_l - \theta_l) \int\int \Bigg(\hat{h}'_{\tau k}(u,v) - \int \hat{h}'_{\tau k}(u,v) du \\
&\quad - \int \hat{h}'_{\tau k}(u,v) dv + \int\int \hat{h}'_{\tau k}(u,v) du dv \Bigg) \\
&\quad \Bigg(\hat{h}'_{\tau l}(u,v) - \int \hat{h}'_{\tau l}(u,v) dv - \int \hat{h}'_{\tau l}(u,v) du + \int\int \hat{h}'_{\tau l}(u,v) du dv \Bigg) du dv \left(\sum_r \hat{\lambda}^2_{r\tau'} \right) \Bigg)^{1/2} \\
&+ \Bigg(\sum_{k,l=1}^{K} (\hat{\theta}_k - \theta_k)(\hat{\theta}_l - \theta_l) \int\int \Bigg(\hat{h}'_{\tau' k}(u,v) - \int \hat{h}'_{\tau' k} du - \int \hat{h}'_{\tau' k}(u,v) dv + \int\int \hat{h}'_{\tau' k}(u,v) du dv \Bigg) \\
&\quad \Bigg(\hat{h}'_{\tau' l}(u,v) - \int \hat{h}'_{\tau' l}(u,v) dv - \int \hat{h}'_{\tau' l}(u,v) du \\
&\quad + \int\int \hat{h}'_{\tau' l}(u,v) du dv \Bigg) du dv \left(\sum_r \hat{\lambda}^2_{r\tau} \right) \Bigg)^{1/2} + o_p\left(N^{-\gamma}\right) \Bigg) \mathbb{1}\{\tau \neq \tau'\} \\
&= \left(\sum_{k,l=1}^{K} \left(\hat{\theta}_k - \theta_k\right) \left(\hat{\theta}_l - \theta_l\right) W^L_{kl,3} \right)^{1/2} + \left(\sum_{k,l=1}^{K} \left(\hat{\theta}_k - \theta_k\right) \left(\hat{\theta}_l - \theta_l\right) W^L_{kl,4} \right)^{1/2} + o_p\left(N^{-1/2}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
&\left| \left(\left(\sum_r \hat{\lambda}_{r\tau} \hat{\lambda}_{r\tau'} + \sum_r \hat{\lambda}_{r(1-\tau)} \hat{\lambda}_{r(1-\tau')} - 1 \right)_+ - \left(\sum_r \lambda_{r\tau} \lambda_{r\tau'} + \sum_r \lambda_{r(1-\tau)} \lambda_{r(1-\tau')} - 1 \right)_+ \right) \mathbb{1}\{\tau = \tau'\} \right| \\
&\leq \left| \sum_{\sigma \in \{0,1\}} \left(\sum_r \left(\hat{\lambda}_{r\sigma} - \lambda_{r\sigma} \right) \hat{\lambda}_{r\sigma} + \sum_r \lambda_{r\sigma} \left(\hat{\lambda}_{r\sigma} - \lambda_{r\sigma} \right) \right) \mathbb{1}\{\tau = \tau'\} \right| \\
&\leq \Bigg(\sum_{\sigma \in \{0,1\}} 2 \Bigg(\sum_{k,l=1}^{K} (\hat{\theta}_k - \theta_k)(\hat{\theta}_l - \theta_l) \int\int \Bigg(\hat{h}'_{\sigma k}(u,v) - \int \hat{h}'_{\sigma k}(u,v) du \\
&\quad - \int \hat{h}'_{\sigma k}(u,v) dv + \int\int \hat{h}'_{\sigma k}(u,v) du dv \Bigg) \Bigg(\hat{h}'_{\sigma l}(u,v) - \int \hat{h}'_{\sigma l}(u,v) dv - \int \hat{h}'_{\sigma l}(u,v) du \\
&\quad + \int\int \hat{h}'_{\sigma l}(u,v) du dv \Bigg) du dv \left(\sum_r \hat{\lambda}^2_{r\sigma} \right) \Bigg)^{1/2} + o_p\left(N^{-1/2}\right) \Bigg) \mathbb{1}\{\tau = \tau'\} \\
&= \left(\sum_{k,l=1}^{K} \left(\hat{\theta}_k - \theta_k\right) \left(\hat{\theta}_l - \theta_l\right) W^L_{kl,5} \right)^{1/2} + \left(\sum_{k,l=1}^{K} \left(\hat{\theta}_k - \theta_k\right) \left(\hat{\theta}_l - \theta_l\right) W^L_{kl,6} \right)^{1/2} + o_p\left(N^{-1/2}\right)
\end{aligned}
$$

following exactly the arguments of Step 1 from Section D.4.3. It follows that

$$N^{1/2}|\hat{L}-L| \leq \sum_{j=1}^{6}\left(\sum_{k,l=1}^{K} N^{1/2}\left(\hat{\theta}_k-\theta_k\right) N^{1/2}\left(\hat{\theta}_l-\theta_l\right) W^{U}_{kl,j}\right)^{1/2} + o_p\left(1\right).$$

D.4.5 Step 3: Proof of Proposition D3

From Steps 1 and 2, for any $\epsilon > 0$ there exists an $M \in \mathbb{N}$ such that for all $N > M$,

$$\begin{aligned}
\mathbb{P}\left([L,U] \subseteq S_\alpha\right) &\geq \mathbb{P}\left(L \geq \hat{L} - N^{-1/2}C^{L}_{\alpha}, U \leq \hat{U} + N^{-1/2}C^{U}_{\alpha}\right) \\
&\geq \mathbb{P}\left(\left|\hat{L}-L\right| \leq N^{-1/2}C^{L}_{\alpha}, \left|\hat{U}-U\right| \leq N^{-1/2}C^{U}_{\alpha}\right) \\
&\geq \mathbb{P}\left(\sum_{j=1}^{6}\left(\xi'\hat{\Omega}^{L}_{j}\xi\right)^{1/2} \leq C^{L}_{\alpha}+\epsilon, \sum_{j=1}^{4}\left(\xi'\hat{\Omega}^{U}_{j}\xi\right)^{1/2} \leq C^{U}_{\alpha}+\epsilon\right) - \epsilon \\
&\geq 1-\alpha-\epsilon
\end{aligned}$$

where the third inequality follows from the bounds from Steps 1 and 2, and the asymptotic normality result for $\hat{\theta}$ in Section D.3.2. Since the choice of $\epsilon > 0$ was arbitrary, the claim follows.

References

Birkhoff, Garrett, "Three observations on linear algebra," *Univ. Nac. Tacuman, Rev. Ser. A*, 1946, *5*, 147–151.

Birman, Michael Sh and Michael Z Solomjak, *Spectral theory of self-adjoint operators in Hilbert space*, Vol. 5, Springer Science & Business Media, 2012.

Dudley, Richard M, *Real analysis and probability*, Vol. 74, Cambridge University Press, 2002.

Fafchamps, Marcel and Flore Gubert, "Risk sharing and network formation," *American Economic Review*, 2007, *97* (2), 75–79.

Finke, Gerd, Rainer E Burkard, and Franz Rendl, "Quadratic assignment problems," in "North-Holland Mathematics Studies," Vol. 132, Elsevier, 1987, pp. 61–82.

Graham, Bryan S, "Network data," in "Handbook of Econometrics," Vol. 7, Elsevier, 2020, pp. 111–218.

Hardy, Godfrey, John Littlewood, and George Pólya, *Inequalities*, Cambridge university press, 1952.

Hoeffding, Wassily, "A Class of Statistics with Asymptotically Normal Distribution," *The Annals of Mathematical Statistics*, 1948, pp. 293–325.

Hoffman, AJ and HW Wielandt, "The variation of the spectrum of a normal matrix," *Duke Math J*, 1953, *20*, 37–39.

Imbens, Guido W and Charles F Manski, "Confidence intervals for partially identified parameters," *Econometrica*, 2004, *72* (6), 1845–1857.

Lovász, László, *Large networks and graph limits*, Vol. 60, American Mathematical Soc., 2012.

Menzel, Konrad, "Bootstrap with Clustering in Two or More Dimensions," *arXiv preprint arXiv:1703.03043*, 2017.

Molinari, Francesca, "Microeconometrics with partial identification," *Handbook of econometrics*, 2020, *7*, 355–486.

Serfling, Robert J, *Approximation theorems of mathematical statistics*, Vol. 162, John Wiley & Sons, 2009.

Stoye, Jörg, "More on confidence intervals for partially identified parameters," *Econometrica*, 2009, *77* (4), 1299–1315.

Whitt, Ward, "Bivariate distributions with given marginals," *The Annals of statistics*, 1976, *4* (6), 1280–1289.