

Online Appendix for "Identifying Peer Effects in Networks with Unobserved Effort and Isolated Students"

Aristide Houndetoungan^a, Cristelle Kouame^b, and Michael Vlassopoulos^c

^a*CY Cergy Paris University and Thema*

^b*World Bank Group*

^c*University of Southampton, and IZA*

May 10, 2024

C Online Appendix

C.1 Some Basic Properties

In this section, we state and prove some basic properties used throughout the paper.

P.1 Let $[\mathbf{F}_s, \ell_s^I/\sqrt{n_s^I}, \ell_s^{NI}/\sqrt{n_s^{NI}}]$ be the orthonormal matrix of \mathbf{J}_s , where the columns in \mathbf{F}_s are eigenvectors of \mathbf{J}_s corresponding to the eigenvalue one. $\|\mathbf{F}_s\|_2 = 1$, where $\|\cdot\|_2$ is the operator norm induced by the ℓ^2 -norm.

Proof. $\|\mathbf{F}_s\|_2 = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \sqrt{(\mathbf{F}_s \mathbf{u}_s)'(\mathbf{F}_s \mathbf{u}_s)} = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \sqrt{\mathbf{u}_s' \mathbf{u}_s}$ because $\mathbf{F}_s' \mathbf{F}_s = \mathbf{I}_{n_s-2}$, the identity matrix of dimension $n_s - 2$. Thus, $\|\mathbf{F}_s\|_2 = 1$. \square

P.2 For any $n_s \times n_s$ matrix, $\mathbf{B}_s = [b_{s,ij}]$, $|b_{s,ii}| \leq \|\mathbf{B}_s\|_2$.

Proof. Let \mathbf{u}_s be the n_s -vector of zeros except for the i -th element, which is one. Note that $\|\mathbf{u}_s\|_2 = 1$. The i -th entry of $\mathbf{B}_s \mathbf{u}_s$ is $b_{s,ii}$. As a result, $|b_{s,ii}| \leq \sqrt{\sum_{j=1}^{n_s} b_{s,ji}^2} = \sqrt{(\mathbf{B}_s \mathbf{u}_s)'(\mathbf{B}_s \mathbf{u}_s)} \leq \|\mathbf{B}_s\|_2$. \square

P.3 If \mathbf{B}_s is a symmetric matrix of dimension $n_s \times n_s$, then $\|\mathbf{B}_s\|_2 = \pi_{\max}(\mathbf{B}_s)$, where $\pi_{\max}(\cdot)$ is the largest eigenvalue.

Proof. $\|\mathbf{B}_s\|_2 = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \sqrt{(\mathbf{B}_s \mathbf{u}_s)'(\mathbf{B}_s \mathbf{u}_s)} = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \sqrt{\mathbf{u}_s' \mathbf{B}_s^2 \mathbf{u}_s} = \sqrt{\pi_{\max}(\mathbf{B}_s^2)} = \pi_{\max}(\mathbf{B}_s)$. \square

P.4 If \mathbf{B}_s is a symmetric matrix of dimension $n_s \times n_s$, then $\pi_{\max}(\mathbf{F}_s' \mathbf{B}_s \mathbf{F}_s) \leq \pi_{\max}(\mathbf{B}_s)$.

Proof. $\pi_{\max}(\mathbf{F}_s' \mathbf{B}_s \mathbf{F}_s) = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \mathbf{u}_s' \mathbf{F}_s' \mathbf{B}_s \mathbf{F}_s \mathbf{u}_s = \max_{\mathbf{u}_s' \mathbf{u}_s = 1} (\mathbf{F}_s \mathbf{u}_s)' \mathbf{B}_s (\mathbf{F}_s \mathbf{u}_s)$. As $(\mathbf{F}_s \mathbf{u}_s)'(\mathbf{F}_s \mathbf{u}_s) = 1$, then $\max_{\mathbf{u}_s' \mathbf{u}_s = 1} (\mathbf{F}_s \mathbf{u}_s)' \mathbf{B}_s (\mathbf{F}_s \mathbf{u}_s) \leq \max_{\mathbf{u}_s' \mathbf{u}_s = 1} \mathbf{u}_s' \mathbf{B}_s \mathbf{u}_s = \pi_{\max}(\mathbf{B}_s)$. \square

P.5 Let $\mathbf{B}_{s,1}$ and $\mathbf{B}_{s,2}$ be $n_s \times n_s$ matrices. If $\mathbf{B}_{s,1}$ and $\mathbf{B}_{s,2}$ are absolutely bounded in row and column sums, then $\mathbf{B}_{s,1}\mathbf{B}_{s,2}$ is absolutely bounded in row and column sums.

Proof. It is sufficient to show that the entries of $\mathbf{B}_{s,1}\mathbf{B}_{s,2}\mathbf{u}_s$ and $\mathbf{u}'_s\mathbf{B}_{s,1}\mathbf{B}_{s,2}$ are absolutely bounded for all n_s -vector \mathbf{u}_s whose entries take -1 or 1 . Assume that $\mathbf{B}_{s,1}$ is absolutely bounded in row sum by $C_{b,1}$ and absolutely bounded in the row sum by $R_{b,1}$. Assume also that $\mathbf{B}_{s,2}$ is absolutely bounded in the row sum by $C_{b,2}$ and absolutely bounded in row sum by $R_{b,2}$. We have $\mathbf{B}_{s,2}\mathbf{u}_s \leq R_{b,2}\mathbf{1}_{n_s}$ and $\mathbf{B}_{s,1}\mathbf{1}_{n_s} \leq R_{b,1}\mathbf{1}_{n_s}$, where \leq is the pointwise inequality \leq and $\mathbf{1}_{n_s}$ is an n_s -vector of ones. Thus, $\mathbf{B}_{s,1}\mathbf{B}_{s,2}\mathbf{u}_s \leq R_{b,2}\mathbf{B}_{s,1}\mathbf{1}_{n_s} \leq R_{b,1}R_{b,2}\mathbf{1}_{n_s}$. Hence, $\mathbf{B}_{s,1}\mathbf{B}_{s,2}$ is bounded in row sum. Analogously, we have $\mathbf{u}'_s\mathbf{B}_{s,1} \leq C_{b,1}\mathbf{1}'_{n_s}$ and $\mathbf{1}'_{n_s}\mathbf{B}_{s,2} \leq C_{b,2}\mathbf{1}'_{n_s}$. Thus, $\mathbf{u}'_s\mathbf{B}_{s,1}\mathbf{B}_{s,2} \leq C_{b,1}\mathbf{1}'_{n_s}\mathbf{B}_{s,2} \leq C_{b,1}C_{b,2}\mathbf{1}'_{n_s}$. Hence, $\mathbf{B}_{s,1}\mathbf{B}_{s,2}$ is bounded in column sum. \square

P.6 If an $n_s \times n_s$ matrix \mathbf{B}_s is absolutely bounded in both row and column sums, then $|\pi_{\max}(\mathbf{B}_s)| < \infty$ and $\|\mathbf{B}_s\|_2 < \infty$.

Proof. $|\pi_{\max}(\mathbf{B}_s)| < \infty$ is a direct implication of the Gershgorin circle theorem.¹

Besides, $\|\mathbf{B}_s\|_2 = \sqrt{\pi_{\max}(\mathbf{B}'_s\mathbf{B}_s)} < \infty$ because $\mathbf{B}'_s\mathbf{B}_s$ is absolutely bounded in row and column sums by P.5. \square

P.7 Let $\mathbf{B}_s = [b_{ij}]$, $\dot{\mathbf{B}}_s = [\dot{b}_{ij}]$ be $n_s \times n_s$ matrices. Let $\mathbf{G} = \text{diag}(\mathbf{G}_1, \dots, \mathbf{G}_S)$, where diag is the block diagonal operator. Assume that $[\boldsymbol{\eta}_s, \boldsymbol{\varepsilon}_s]$ are independent of \mathbf{G}_s and \mathbf{X}_s . Let $\mu_{2\eta} = \mathbb{E}(\eta_{s,i}^2)$, $\mu_{2\epsilon} = \mathbb{E}(\varepsilon_{s,i}^2)$, $\mu_{4\eta} = \mathbb{E}(\eta_{s,i}^4)$, $\mu_{4\epsilon} = \mathbb{E}(\varepsilon_{s,i}^4)$, $\mu_{22} = \mathbb{E}(\eta_{s,i}^2\varepsilon_{s,i}^2)$, $\mu_{31} = \mathbb{E}(\eta_{s,i}^3\varepsilon_{s,i})$, and $\mu_{13} = \mathbb{E}(\eta_{s,i}\varepsilon_{s,i}^3)$.

$$\mathbb{V}(\boldsymbol{\eta}'_s\mathbf{B}_s\boldsymbol{\eta}_s) = (\mu_{4\eta} - 3\mu_{2\eta}^2) \sum_{i=1}^{n_s} b_{ii}^2 + \mu_{2\eta}^2 (\text{Tr}(\mathbf{B}_s\mathbf{B}'_s) + \text{Tr}(\mathbf{B}_s^2)),$$

$$\mathbb{V}(\boldsymbol{\varepsilon}'_s\mathbf{B}_s\boldsymbol{\varepsilon}_s) = (\mu_{4\epsilon} - 3\mu_{2\epsilon}^2) \sum_{i=1}^{n_s} b_{ii}^2 + \mu_{2\epsilon}^2 (\text{Tr}(\mathbf{B}_s\mathbf{B}'_s) + \text{Tr}(\mathbf{B}_s^2)),$$

$$\mathbb{V}(\boldsymbol{\varepsilon}'_s\mathbf{B}_s\boldsymbol{\eta}_s) = (\mu_{22} - 3\mu_{2\eta}\mu_{2\epsilon}) \sum_{i=1}^{n_s} b_{ii}^2 + \mu_{2\eta}\mu_{2\epsilon} ((1 - \rho^2)(\text{Tr}(\mathbf{B}_s))^2 + \text{Tr}(\mathbf{B}_s\mathbf{B}'_s) + \rho^2 \text{Tr}(\mathbf{B}_s^2)),$$

$$\text{Cov}(\boldsymbol{\eta}'_s\mathbf{B}_s\boldsymbol{\eta}_s, \boldsymbol{\varepsilon}'_s\dot{\mathbf{B}}_s\boldsymbol{\eta}_s) = (\mu_{31} - 3\rho\sigma_\eta^3\sigma_\epsilon) \sum_{i=1}^{n_s} b_{ii}\dot{b}_{ii} + \rho\sigma_\eta^3\sigma_\epsilon (\text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}'_s) + \text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}_s)),$$

$$\text{Cov}(\boldsymbol{\varepsilon}'_s\mathbf{B}_s\boldsymbol{\varepsilon}_s, \boldsymbol{\eta}'_s\dot{\mathbf{B}}_s\boldsymbol{\varepsilon}_s) = (\mu_{13} - 3\rho\sigma_\eta\sigma_\epsilon^3) \sum_{i=1}^{n_s} b_{ii}\dot{b}_{ii} + \rho\sigma_\eta\sigma_\epsilon^3 (\text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}'_s) + \text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}_s)),$$

$$\text{Cov}(\boldsymbol{\eta}'_s\mathbf{B}_s\boldsymbol{\eta}_s, \boldsymbol{\varepsilon}'_s\dot{\mathbf{B}}_s\boldsymbol{\varepsilon}_s) = (\mu_{22} - 2\rho^2\mu_{2\eta}\mu_{2\epsilon} - \mu_{2\eta}\mu_{2\epsilon}) \sum_{i=1}^{n_s} b_{ii}\dot{b}_{ii} + \rho^2\mu_{2\eta}\mu_{2\epsilon} (\text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}'_s) + \text{Tr}(\mathbf{B}_s\dot{\mathbf{B}}_s)).$$

The proof of the lemma is straightforward using the expression of variance and covariance.

C.2 Supplementary Results on the Identification of the Variance Parameters

In this section, we use different notations for the parameters and their true values; that is their values in the data-generating process. We denote by ψ_0 , $\sigma_{0\eta}$, $\sigma_{0\epsilon}$, and ρ_0 the true values of ψ , σ_η , σ_ϵ , and ρ , respectively. We must show that $\mathbb{V}(\hat{\sigma}_\epsilon(\tau, \rho)|\mathbf{G}) = o_p(1)$.

We have $\hat{\sigma}_\epsilon^2(\tau, \rho) = \sum_{s=1}^S \frac{((\mathbf{I}_{n_s} - \lambda_0\mathbf{G}_s)\boldsymbol{\eta}_s + \delta^2\boldsymbol{\varepsilon}_s)' \mathbf{F}_s \boldsymbol{\Omega}_s^{-1}(\lambda_0, \tau, \rho) \mathbf{F}'_s ((\mathbf{I}_{n_s} - \lambda_0\mathbf{G}_s)\boldsymbol{\eta}_s + \delta^2\boldsymbol{\varepsilon}_s)}{n - 2S}$. Thus,

¹See Horn, R. A. and C. R. Johnson (2012): *Matrix analysis*, Cambridge university press.

$$\begin{aligned}
\mathbb{V}(\hat{\sigma}_\epsilon^2(\tau, \rho) | \mathbf{G}) &= \frac{1}{(n-2S)^2} \sum_{s=1}^S (\mathbb{V}(\boldsymbol{\eta}'_s \ddot{\mathbf{M}}_s \boldsymbol{\eta}_s | \mathbf{G}) + 4\delta^4 \mathbb{V}(\boldsymbol{\eta}'_s \dot{\mathbf{M}}_s \boldsymbol{\epsilon}_s | \mathbf{G}) + \delta^8 \mathbb{V}(\boldsymbol{\epsilon}'_s \mathbf{M}_s \boldsymbol{\epsilon}_s | \mathbf{G}) + \\
&\quad 4\delta^2 \mathbb{Cov}(\boldsymbol{\eta}'_s \ddot{\mathbf{M}}_s \boldsymbol{\eta}_s, \boldsymbol{\eta}'_s \dot{\mathbf{M}}_s \boldsymbol{\epsilon}_s | \mathbf{G}) + 2\delta^4 \mathbb{Cov}(\boldsymbol{\eta}'_s \ddot{\mathbf{M}}_s \boldsymbol{\eta}_s, \boldsymbol{\epsilon}'_s \mathbf{M}_s \boldsymbol{\epsilon}_s | \mathbf{G}) + \\
&\quad 4\delta^6 \mathbb{Cov}(\boldsymbol{\epsilon}'_s \mathbf{M}_s \boldsymbol{\epsilon}_s, \boldsymbol{\eta}'_s \dot{\mathbf{M}}_s \boldsymbol{\epsilon}_s | \mathbf{G})),
\end{aligned} \tag{C.12}$$

where $\mathbf{M}_s = \mathbf{F}_s \boldsymbol{\Omega}_s^{-1}(\lambda_0, \tau, \rho) \mathbf{F}'_s$, $\dot{\mathbf{M}}_s = (\mathbf{I}_{n_s} - \lambda_0 \mathbf{G}_s)' \mathbf{M}_s$, and $\ddot{\mathbf{M}}_s = \dot{\mathbf{M}}_s (\mathbf{I}_{n_s} - \lambda_0 \mathbf{G}_s)$.

As $\pi_{\min}(\boldsymbol{\Omega}_s(\lambda_0, \tau, \rho))$ is bounded away from zero (Assumption A.2), we have $|\pi_{\max}(\boldsymbol{\Omega}_s^{-1}(\lambda_0, \tau, \rho))| = O_p(1)$. Thus, $\max_s \|\boldsymbol{\Omega}_s^{-1}(\lambda_0, \tau, \rho)\|_2 = O_p(1)$ by P.3. This implies that $\max_s \|\mathbf{M}_s\|_2 = O_p(1)$, $\max_s \|\dot{\mathbf{M}}_s\|_2 = O_p(1)$, and $\max_s \|\ddot{\mathbf{M}}_s\|_2 = O_p(1)$ because $\|\mathbf{F}_s\|_2 = 1$ and $\|\mathbf{I}_{n_s} - \lambda_0 \mathbf{G}_s\|_2 = O_p(1)$ by P.6.

We now need to show that the sum over s of each term of the variance (C.12) is $o_p((n-2S)^2)$. By P.2, the trace of any product of matrices chosen among \mathbf{M}_s , $\dot{\mathbf{M}}_s$, and $\ddot{\mathbf{M}}_s$ is $O_p(n_s)$ and thus, $o_p((n-2S)^2)$. For example, $|\text{Tr}(\mathbf{M}_s \dot{\mathbf{M}}_s)| \leq n_s \|\mathbf{M}_s \dot{\mathbf{M}}_s\|_2 \leq n_s \|\mathbf{M}_s\|_2 \|\dot{\mathbf{M}}_s\|_2 = O_p(n_s) = o_p((n-2S)^2)$. On the other hand, $\sum_{s=1}^S (\text{Tr}(\mathbf{M}_s))^2 = O_p(\sum_{s=1}^S n_s^2) = o_p((n-2S)^2)$. Moreover, $\sum_{i=1}^{n_s} m_{ii}^2 \leq n_s \|\mathbf{M}_s\|_2^2 = O_p(n_s) = o_p((n-2S)^2)$ by P.2. Analogously, $\sum_{i=1}^{n_s} m_{ii} \dot{m}_{ii} = o_p((n-2S)^2)$. As a result, $\mathbb{V}(\hat{\sigma}_\epsilon^2(\tau, \rho) | \mathbf{G}) = o_p(1)$.

The proof implies, by Chebyshev inequality, that $\hat{\sigma}_\epsilon^2(\tau, \rho) - \mathbb{E}(\hat{\sigma}_\epsilon^2(\tau, \rho) | \mathbf{G}_1, \dots, \mathbf{G}_S)$ converges in probability to zero. The convergence is uniform in the space of (τ, ρ) because $\mathbb{E}(\hat{\sigma}_\epsilon^2(\tau, \rho) | \mathbf{G}_1, \dots, \mathbf{G}_S)$ and $\hat{\sigma}_\epsilon^2(\tau, \rho)$ can be expressed as a polynomial function in (τ, ρ) . Thus, $\frac{1}{n}(L_c(\tau, \rho) - L_c^*(\tau, \rho))$ converges uniformly to zero. This proof also implies that $\text{plim } \hat{\sigma}_\epsilon^2(\tau_0, \rho_0) = \sigma_{0\epsilon}^2$, where $\tau_0 = \sigma_{0\eta}/\sigma_{0\epsilon}$.

C.3 Necessary Conditions for the Identification of $(\sigma_\epsilon^2, \tau, \rho)$

As $\lambda \neq 0$ (Assumption 3.2) and is identified, $\mathbb{E}(\mathbf{v}_s \mathbf{v}'_s | \mathbf{G}_s)$ implies a unique $(\sigma_\eta, \sigma_\epsilon, \rho)$ if $\mathbf{J}_s, \mathbf{J}_s(\mathbf{G}_s + \mathbf{G}'_s)\mathbf{J}_s$ and $\mathbf{J}_s \mathbf{G}_s \mathbf{G}'_s \mathbf{J}_s$ are linearly independent. We present a simple subnetwork structure that verifies this condition.

Let \mathbf{C}_s be an arbitrary $n_s \times n_s$ matrix. Unless otherwise stated, we use $\mathbf{C}_{s,ij}$ to denote the (i, j) -th entry of \mathbf{C}_s . Assume that i and j are from the subset of students who have friends in the school s . The (i, j) -th entry of $\mathbf{J}_s \mathbf{C}_s \mathbf{J}_s$ is $\mathbf{C}_{s,ij} - \hat{\mathbf{C}}_{s,\bullet j} - \hat{\mathbf{C}}_{s,i\bullet} + \hat{\mathbf{C}}_{s,\bullet\bullet}$, where $\hat{\mathbf{C}}_{s,\bullet j} = (1/n_s^{NI}) \sum_{k \in \mathcal{V}_s^{NI}} \mathbf{C}_{s,kj}$, $\hat{\mathbf{C}}_{s,i\bullet} = (1/n_s^{NI}) \sum_{l \in \mathcal{V}_s^{NI}} \mathbf{C}_{s,il}$, and $\hat{\mathbf{C}}_{s,\bullet\bullet} = (1/(n_s^{NI})^2) \sum_{k,l \in \mathcal{V}_s^{NI}} \mathbf{C}_{s,kl}$.

Let $\tilde{\mathbf{G}}_s = \mathbf{G}_s \mathbf{G}'_s$ and i_1, \dots, i_4 be four students from \mathcal{V}_s^{NI} who are not directly linked and where only two of them have common friends. Without loss of generality, assume that i_1 and i_3 have common friends. For any $i \in \{i_1, i_2\}$ and $j \in \{i_3, i_4\}$, $\mathbf{J}_{s,ij} = -1/n_s^{NI}$, $\mathbf{G}_{s,ij} = 0$, and $\mathbf{G}'_{s,ij} = 0$. Moreover, $\tilde{\mathbf{G}}_{s,ij} = 0$ except for the pair (i_i, i_3) , who have common friends. Let $\mathbf{L}_s = b_1 \mathbf{J}_s + b_2 \mathbf{J}_s(\mathbf{G}_s + \mathbf{G}'_s)\mathbf{J}_s + b_3 \mathbf{J}_s \mathbf{G}_s \mathbf{G}'_s \mathbf{J}_s = 0$ for some $b_1, b_2, b_3 \in \mathbb{R}$. We have $\mathbf{L}_{s,ij} = -b_1/n_s^{NI} - b_2(\mathbf{G}_{s,ij} - \mathbf{G}_{s,\bullet j} - \mathbf{G}_{s,i\bullet} + \mathbf{G}_{s,\bullet\bullet} + \mathbf{G}'_{s,ij} - \mathbf{G}'_{s,\bullet j} - \mathbf{G}'_{s,i\bullet} + \mathbf{G}'_{s,\bullet\bullet}) + b_3(\tilde{\mathbf{G}}_{s,ij} - \tilde{\mathbf{G}}_{s,\bullet j} - \tilde{\mathbf{G}}_{s,i\bullet} + \tilde{\mathbf{G}}_{s,\bullet\bullet})$. This implies that $\mathbf{L}_{s,i_1 i_3} + \mathbf{L}_{s,i_2 i_4} -$

$\mathbf{L}_{s,i_2i_3} - \mathbf{L}_{s,i_1i_4} = b_3 \tilde{\mathbf{G}}_{s,i_1i_3}$. Thus, if the combination \mathbf{L}_s is zero, then $b_3 = 0$.

Let j_1, \dots, j_4 be four students from \mathcal{V}_s^{NI} , where only two of them are directly linked (mutually or not), and the others are not directly linked. Without loss of generality, assume that only j_1 to j_3 are linked, that is, for any $i \in \{j_1, j_2\}$ and $j \in \{j_3, j_4\}$, $\mathbf{G}_{s,ij} = 0$ and $\mathbf{G}'_{s,ij} = 0$ except for the pairs (j_1, j_3) and (j_3, j_1) . As $b_3 = 0$, we have $\mathbf{L}_{s,j_1j_3} + \mathbf{L}_{s,j_2j_4} - \mathbf{L}_{s,j_2j_3} - \mathbf{L}_{s,j_1j_4} = b_2(\mathbf{G}_{s,j_1j_3} + \mathbf{G}'_{s,j_1j_3})$. Thus if \mathbf{L}_s is zero, then $b_2 = 0$, and it follows that $b_1 = 0$.

As a result, $\mathbf{J}_s, \mathbf{J}_s(\mathbf{G}_s + \mathbf{G}'_s)\mathbf{J}_s$, and $\mathbf{J}_s\mathbf{G}_s\mathbf{G}'_s\mathbf{J}_s$ are linearly independent if, in some school s , there are four students from \mathcal{V}_s^{NI} who are not directly linked and only two of them have common friends, and if in some school s , there are four students from \mathcal{V}_s^{NI} , where only two of them are linked.

We present an example of this condition by adding three nodes to Figure 1 with two additional links (see Figure C.1). There are no links within the nodes i_1, i_4, i_5 , and i_6 , and only i_5 and i_6 have common a friends (i_7). Besides, only i_5 and i_7 are linked within the nodes i_1, i_2, i_5 , and i_7 .

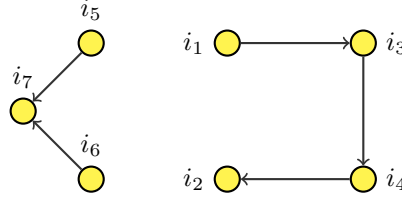


Figure C.1: Illustration of the identification

Note: \rightarrow means that the node on the right side is a friend of the node on the left side.

Many other situations lead to $b_1 = b_2 = b_3 = 0$. In practice, one can easily verify if $\mathbf{J}_s, \mathbf{J}_s(\mathbf{G}_s + \mathbf{G}'_s)\mathbf{J}_s$ and $\mathbf{J}_s\mathbf{G}_s\mathbf{G}'_s\mathbf{J}_s$ are linearly independent.

C.4 Asymptotic Normality in the Case of Endogenous Networks

The specification controlling for network endogeneity is:

$$y_{s,i} = \kappa_s^{NI} \ell_{s,i}^{NI} + \kappa_s^I (1 - \ell_{s,i}^{NI}) + \mathbf{g}_{s,i} \mathbf{y}_s + \mathbf{x}'_{s,i} \tilde{\boldsymbol{\beta}} + \mathbf{g}_{s,i} \mathbf{X}_s \tilde{\boldsymbol{\gamma}} + h_{s,i} + \tilde{v}_{s,i}, \quad (\text{C.13})$$

where $h_{s,i} = h^{out}(\mu_{s,i}^{out}) + h^{in}(\mu_{s,i}^{in})$. We replace $\mu_{s,i}^{out}$ and $\mu_{s,i}^{in}$ with their estimator and approximate the functions h^{out} and h^{in} with cubic B-spline approximations. Specifically, we approximate $h^{out}(\mu_{s,i}^{out})$ by cubic polynomials on ten different intervals covering the range of $\mu_{s,i}^{out}$. The intervals are defined so that each comprises approximately the same share of observations. We also apply this approach to $h^{in}(\mu_{s,i}^{in})$. Given the number of intervals and the degree of the polynomials, this approach results in approximating $h_{s,i}$ by a combination of 26 variables, called bases, that are computed from the estimates of $\mu_{s,i}^{out}$ and $\mu_{s,i}^{in}$.

Let $\dot{\mathbf{X}}_s$ be the matrix of the new 26 bases. The approximation of $h_{s,i}$ is $\dot{\mathbf{x}}'_{s,i}\beta_h$, where $\dot{\mathbf{x}}_{s,i}$ is the i -th row of $\dot{\mathbf{X}}_s$ and β_h is a parameter to be estimate. Let $\hat{\mathbf{R}}_s = [\mathbf{R}_s, \mathbf{J}_s \dot{\mathbf{X}}_s]$ be the new design matrix. We keep the same instrument matrix $\mathbf{J}_s \mathbf{G}_s^2 \mathbf{X}_s$ for $\mathbf{J}_s \mathbf{G}_s \mathbf{y}_s$. We define $\hat{\mathbf{Z}}_s = [\mathbf{J}_s \mathbf{G}_s^2 \mathbf{X}_s, \tilde{\mathbf{X}}_s, \mathbf{J}_s \dot{\mathbf{X}}_s]$, $\hat{\mathbf{R}}' \hat{\mathbf{Z}} = \sum_{s=1}^S \hat{\mathbf{R}}'_s \hat{\mathbf{Z}}_s$, $\hat{\mathbf{Z}}' \hat{\mathbf{Z}} = \sum_{s=1}^S \hat{\mathbf{Z}}'_s \hat{\mathbf{Z}}_s$, and $\hat{\mathbf{Z}}' \mathbf{y} = \sum_{s=1}^S \hat{\mathbf{Z}}'_s \mathbf{J}_s \mathbf{y}_s$. Let $\hat{\mathbf{\Gamma}}$ be the estimator of the coefficients associated with $\hat{\mathbf{R}}_s$; i.e., $\hat{\mathbf{\Gamma}} = ((\hat{\mathbf{R}}' \hat{\mathbf{Z}})(\hat{\mathbf{Z}}' \hat{\mathbf{Z}})^{-1}(\hat{\mathbf{R}}' \hat{\mathbf{Z}})')^{-1}(\hat{\mathbf{R}}' \hat{\mathbf{Z}})(\hat{\mathbf{Z}}' \hat{\mathbf{Z}})^{-1}(\hat{\mathbf{Z}}' \mathbf{y})$.

The regularity assumption we need for the asymptotic normality is $\sum_{s=1}^S \hat{\mathbf{Z}}'_s (\mathbf{h}_s - \dot{\mathbf{X}}_s \hat{\beta}_h) / \sqrt{n} = o_p(1)$, where $\mathbf{h}_s = (h_{s,1}, \dots, h_{s,n_s})'$ and $\hat{\beta}_h$ is the estimator of the coefficients associated with $\dot{\mathbf{X}}_s$. A similar condition is also imposed by Johnsson and Moon (2021) (see Lipschitz condition in their Assumption 8). It holds if the approximation error of $h_{s,i}$ by $\dot{\mathbf{x}}'_s \hat{\beta}_h$ converges at some rate to zero. Under this condition $\hat{\mathbf{\Gamma}}$ is normally distributed with the asymptotic distribution $\frac{\tilde{\mathbf{B}}^{-1} \tilde{\mathbf{D}} \tilde{\mathbf{B}}^{-1}}{n}$. The matrices $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{D}}$ are defined as the original \mathbf{B} and \mathbf{D} , where \mathbf{R}_s and \mathbf{Z}_s are replaced by $\hat{\mathbf{R}}_s$ and $\hat{\mathbf{Z}}_s$.