Supplement to "Estimating local interactions among many agents who observe their neighbors"

(*Quantitative Economics*, Vol. 11, No. 3, July 2020, 917–956)

Nathan Canen
Department of Economics, University of Houston

Jacob Schwartz
Department of Economics, University of Haifa

Kyungchul Song
Vancouver School of Economics, University of British Columbia

This note consists of four appendices. In Appendix A, we give the mathematical proofs of Theorems 2.1 and 2.2. We also provide details on the best response of a Bayesian decision maker and the result of coincidence of equilibrium strategies and behavioral strategies when the payoff graph has multiple disjoint complete subgraphs. In Appendix B, we explain how estimation of the asymptotic covariance matrix is motivated. Appendix C gives the proof of local identification (Theorem 3.1) and the asymptotic results (Theorem 3.2). Appendix D gives the proofs of the results on the convergence of behavioral strategies to equilibrium strategies from game Γ_∞ in Section 2.4.1. The replication files to this paper (Canen, Schwartz, and Song (2020)) contain additional results: information sharing over time, inference for the model with first-order sophisticated agents, model selection between games Γ_0 and Γ_1, testing for information sharing on unobservables, and empirical results based on the first-order sophisticated agents.

Appendix A: Best responses

A.1 *Proofs of Theorems 2.1–2.2*

Lemma A.1. *Suppose that $\beta_0 \in (-1, 1)$. Then, for any $i \in N$ such that $n_P(i) \geq 1$,*

$$n_P(i) - \beta_0^2 \overline{\lambda}_i \geq \frac{n_P(i)\big(n_P(i) + |\beta_0|\big)\big(1 - |\beta_0|\big)}{n_P(i)\big(1 - |\beta_0|\big) + |\beta_0|},$$

and

$$0 \leq w_{ii}^{[0]} \leq 1 + \frac{\beta_0^2}{1 - \beta_0^2}, \quad \textit{and} \quad \big|w_{ij}^{[0]}\big| \leq \frac{|\beta_0|}{n_P(i)\big(1 - |\beta_0|\big)}\bigg(1 + \frac{\beta_0^2}{1 - \beta_0^2}\bigg). \tag{A.1}$$

Nathan Canen: ncanen@uh.edu
Jacob Schwartz: jschwartz@econ.haifa.ac.il
Kyungchul Song: kysong@mail.ubc.ca

Available at http://qeconomics.org. https://doi.org/10.3982/QE923

PROOF. First, note that for any $j \in N_P(i)$,

$$c_{ij} \leq \frac{n_P(i) - 1}{n_P(i)}. \tag{A.2}$$

Hence

$$\lambda_{ij} \leq \frac{n_P(i)}{n_P(i) - |\beta_0|\big(n_P(i) - 1\big)} = \frac{n_P(i)}{n_P(i)\big(1 - |\beta_0|\big) + |\beta_0|}. \tag{A.3}$$

Therefore,

$$\overline{\lambda}_i \leq \frac{n_P(i)}{n_P(i)\big(1 - |\beta_0|\big) + |\beta_0|}. \tag{A.4}$$

Now, we find that

$$\begin{aligned} n_P(i) - \beta_0^2 \overline{\lambda}_i &\geq n_P(i) - \frac{\beta_0^2 n_P(i)}{n_P(i)\big(1 - |\beta_0|\big) + |\beta_0|} \\ &= \frac{n_P(i)\big(n_P(i) + |\beta_0|\big)\big(1 - |\beta_0|\big)}{n_P(i)\big(1 - |\beta_0|\big) + |\beta_0|}. \end{aligned}$$

This gives the desired lower bound for $n_P(i) - \beta_0^2 \overline{\lambda}_i$.

Let us turn to $w_{ii}^{[0]}$. Since $\lambda_{ij} \geq 0$, we have $w_{ii}^{[0]} \geq 0$ by the previous bound. As for its upper bound, from (A.4), we have

$$w_{ii}^{[0]} \leq 1 + \frac{\beta_0^2}{\big(n_P(i) + |\beta_0|\big)\big(1 - |\beta_0|\big)} \leq 1 + \frac{\beta_0^2}{1 - \beta_0^2},$$

because $n_P(i) \geq 1$. Finally, as for $w_{ij}^{[0]}$, it suffices to note that by (A.2),

$$0 \leq \lambda_{ij} \leq \frac{1}{1 - |\beta_0|}. \qquad \square$$

PROOF OF THEOREM 2.1. The case of $n_P(i) = 0$ is trivial. Let us assume that $n_P(i) \geq 1$. From the optimization of agent i,

$$s_i^{[0]}(\mathcal{I}_{i,0}) = \tau_i + \beta_0 \left(\frac{1}{n_P(i)} \sum_{k \in N_P(i)} \sum_{j \in \overline{N}_P(k)} w_{kj}^i \tau_j \right) + \eta_i. \tag{A.5}$$

Reorganizing the terms, we have

$$\begin{aligned} s_i^{[0]}(\mathcal{I}_{i,0}) = &\left(1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ki}^i \right) \tau_i \\ &+ \beta_0 \sum_{j \in N_I(i)} \frac{1}{n_P(i)} \sum_{k \in N_P(k)} w_{kj}^i 1\{j \in \overline{N}_P(k)\} \tau_j + \eta_i. \end{aligned} \tag{A.6}$$

By setting the coefficient of τ_j to be w_{ij}, we obtain

$$w_{ii} = 1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w^i_{ki}, \tag{A.7}$$

and for all $j \in N_I(i)$,

$$\begin{aligned} w_{ij} &= \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w^i_{kj} 1\{j \in \overline{N}_P(k)\} \\ &= \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w^i_{kj} 1\{j \in N_P(k)\} + \beta_0 \frac{w^i_{jj} 1\{j \in N_P(i)\}}{n_P(i)}, \end{aligned} \tag{A.8}$$

where the last term corresponds to the case $j = k \in N_P(i)$. We now apply our behavioral assumptions to (A.7) and (A.8). First, we apply the behavioral assumption to (A.7) to write

$$\begin{aligned} w_{ii} &= 1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w^i_{ki} 1\{i \in \overline{N}_P(k)\} \\ &= 1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ik} 1\{i \in \overline{N}_P(k)\} \\ &= 1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ik}, \end{aligned} \tag{A.9}$$

where the last line follows by the undirectedness of G_P. Turning to (A.8), we have

$$\begin{aligned} w_{ij} &= w_{ij} \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} 1\{j \in N_P(k)\} + \beta_0 \frac{w_{ii} 1\{j \in N_P(i)\}}{n_P(i)} \\ &= w_{ij} \beta_0 c_{ij} + \beta_0 \frac{w_{ii} 1\{j \in N_P(i)\}}{n_P(i)}. \end{aligned} \tag{A.10}$$

Rearranging (A.10) for w_{ij}, we obtain

$$w_{ij} = \beta_0 w_{ii} \frac{\lambda_{ij} 1\{j \in N_P(i)\}}{n_P(i)}. \tag{A.11}$$

We first solve for w_{ii}. By plugging (A.9) into (A.11) and averaging over $N_P(i)$ we obtain

$$\frac{1}{n_P(i)} \sum_{j \in N_p(i)} w_{ij} = \beta_0 \left(1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ik}\right) \frac{\overline{\lambda}_i}{n_P(i)},$$

which, after rearranging terms, becomes

$$\frac{1}{n_P(i)} \sum_{j \in N_p(i)} w_{ij} = \frac{\beta_0 \overline{\lambda}_i}{n_P(i) - \beta_0^2 \overline{\lambda}_i}. \tag{A.12}$$

Hence, by plugging (A.12) into (A.9), we obtain

$$w_{ii} = 1 + \frac{\beta_0^2 \overline{\lambda}_i}{n_P(i) - \beta_0^2 \overline{\lambda}_i}.$$

(Note that $n_P(i) - \beta_0 \overline{\lambda}_i > 0$ by Lemma A.1.) By plugging this back into (A.11), we have

$$w_{ij} = \frac{\beta_0 w_{ii} \lambda_{ij} 1\{j \in \overline{N}_P(i)\}}{n_P(i)}. \tag{A.13}$$

Thus relevant j's that appear in the best responses are only those j's such that $j \in N_P(i)$. Taking $w_{ii}^{[0]} = w_{ii}$ and $w_{ij}^{[0]} = w_{ij}$, we obtain the desired result. □

Proof of Theorem 2.2. We prove the result by induction. We begin by showing the result holds for $m = 1$. Suppose each agent is the first-order sophisticated type ($m = 1$), that is, each $i \in N$ believes that each $k \neq i$ is a simple type ($m = 0$) and chooses strategies according to

$$s_k^i(\mathcal{I}_{i,0}) = \sum_{j \in \overline{N}_P(k)} \tau_j w_{kj}^{[0]} + \eta_k.$$

The best responses of the first-order sophisticated types are linear because the payoff is quadratic in the player's own actions, and they believe simple types play according to linear strategies. Hence the best response of the first-order sophisticated type takes the form

$$\begin{aligned} s_i^{[1]}(\mathcal{I}_{i,1}) =& \tau_i + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} \left(\sum_{j \in \overline{N}_P(k)} \tau_j w_{kj}^{[0]} \right) + \eta_i \\ =& \left(1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ki}^{[0]} \right) \tau_i + \frac{\beta_0}{n_P(i)} \sum_{j \in N_I(i)} \sum_{k \in N_P(i)} \tau_j w_{kj}^{[0]} 1\{j \in \overline{N}_P(k)\} + \eta_i. \end{aligned}$$

By setting the coefficient of τ_j to $w_{ij}^{[1]}$, we obtain the weights

$$w_{ii}^{[1]} = 1 + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ki}^{[0]}, \tag{A.14}$$

and for each $j \in N_I(i)$

$$\begin{aligned} w_{ij}^{[1]} &= \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{kj}^{[0]} 1\{j \in \overline{N}_P(k)\} \\ &= \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{kj}^{[0]} 1\{j \in N_P(k)\} + \beta_0 \frac{w_{jj}^{[0]} 1\{j \in N_P(i)\}}{n_P(i)}. \end{aligned} \tag{A.15}$$

Therefore, the weights for $j \in N_I(i)$ involve agents up to $N_{P,2}(i)$. Since

$$s_i^{[1]} = \tau_i w_{ii}^{[1]} + \sum_{j \in N_{P,2}(i)} w_{ij}^{[1]} \tau_j + \eta_i,$$

with the weights for $w_{ii}^{[1]}$ and $w_{ij}^{[1]}$ given in (A.14) and (A.15), we have shown that the result holds for $m=1$. Now suppose that for some m, the best responses are given by

$$s_i^{[m]}(\mathcal{I}_{i,m}) = w_{ii}^{[m]}\tau_i + \sum_{j\in N_{P,m+1}(i)} w_{ij}^{[m]}\tau_j + \eta_i,$$

where $w_{ii}^{[m]}$ and $w_{ij}^{[m]}$ are as defined in the statement of the result. From the optimization problem of $m+1$ types, we can write the best response for these types as

$$\begin{aligned} s_i^{[m+1]}(\mathcal{I}_{i,1}) &= \tau_i + \beta_0\left(\frac{1}{n_P(i)}\sum_{k\in N_P(i)}\sum_{j\in \overline{N}_{P,m+1}(k)} \tau_j w_{kj}^{[m]}\right) + \eta_i \\ &= \left(1+\frac{\beta_0}{n_P(i)}\sum_{k\in N_P(i)} w_{ki}^{[m]}\right)\tau_i \\ &\quad + \frac{\beta_0}{n_P(i)}\sum_{j\in N_I(i)}\sum_{k\in N_P(i)} \tau_j w_{kj}^{[m]} 1\{j\in \overline{N}_{P,m+1}(k)\} + \eta_i. \end{aligned}$$

Note that this expression involves only those $j\in N_I(i)$ up to $N_{P,m+2}(i)$. Hence we conclude that

$$s_i^{[m+1]}(\mathcal{I}_{i,m}) = w_{ii}^{[m+1]}\tau_i + \sum_{j\in N_{P,m+2}(i)} w_{ij}^{[m+1]}\tau_j + \eta_i,$$

with

$$w_{ii}^{[m+1]} = 1 + \frac{\beta_0}{n_P(i)}\sum_{k\in N_P(i)} w_{ki}^{[m]}$$

and

$$w_{ij}^{[m+1]} = \frac{\beta_0}{n_P(i)}\sum_{k\in N_P(i)} w_{kj}^{[m]} 1\{j\in \overline{N}_{P,m+1}(k)\}$$

as required. Therefore, we have shown that the result holds for all $m\geq 1$ by mathematical induction. □

A.2 *The best response of a Bayesian decision maker*

In this section, we elaborate on the points made in Section 2.3.3, and prove the claim that under the prior Q_i satisfying Conditions (a)–(c), the best response of the Bayesian decision maker coincides with quadratic utility with that of a simple type player with belief projection.

To see this, consider the following problem of Bayesian decision maker i:

$$\sup_{y_i\in\mathcal{Y}_i} \mathbf{E}_{Q_i}\left[u_i\left(y_i, s_{-i}(\mathcal{I}_{i,0}; W_{-i}), \tau, \eta_i\right) | W_i = w_i, \mathcal{I}_{i,0}\right].$$

The solution y_i takes the following form:

$$s_i(\mathcal{I}_{i,0}; w_i) = \tau_i + \frac{\beta_0}{n_P(i)} \sum_{k\in N_P(i)} \mathbf{E}_{Q_i}\bigg[\sum_{j\in \overline{N}_I(k)} W'_{kj}\tau_j \bigg| W_i = w_i, \mathcal{I}_{i,0}\bigg]. \tag{A.16}$$

By condition (c), we write

$$\begin{aligned} &\frac{1}{n_P(i)} \sum_{k\in N_P(i)} \mathbf{E}_{Q_i}\bigg[\sum_{j\in \overline{N}_I(k)} W'_{kj}\tau_j \bigg| W_i = w_i, \mathcal{I}_{i,0}\bigg] \\ &= \frac{1}{n_P(i)} \sum_{k\in N_P(i)} \sum_{j\in \overline{N}_P(k)} \mathbf{E}_{Q_i}[W'_{kj}|W_i = w_i, \mathcal{I}_{i,0}]\tau_j. \end{aligned}$$

We plug this back into (A.16). Using restrictions (a) and (b), we can rewrite

$$s_i(\mathcal{I}_{i,0}; w_i) = w_i' h(\tau_i) + \eta_i, \tag{A.17}$$

for some function h of τ_i. We equate this to

$$s_i(\mathcal{I}_{i,0}; w_i) = w_i'\tau_i + \eta_i,$$

as $s_i(\mathcal{I}_{i,0}; w_i)$ is a linear strategy, and find out the weight vector w_i. However, this is precisely how the best response in Theorem 2.1 was derived. Indeed, By restrictions (a) and (b), the right-hand side of (A.6) coincides with that of (A.17).

A.3 *The coincidence of the equilibrium strategies and the behavioral strategies in the case of complete payoff subgraphs*

Here, we show that the equilibrium strategies and the behavioral strategies coincide when the payoff graphs G_P have disjoint multiple subgraphs and each subgraph is a complete graph as in Section 2.3.1. For this, it suffices to derive the Bayesian–Nash Equilibrium (BNE) as in (2.9). Suppose that $\mathbf{y}(i)$ is the $n_P(i)+1$ dimensional column vector of actions that are realized from the BNE of the game. Let $\boldsymbol{\eta}(i)$ and $\boldsymbol{\tau}(i)$ be an $n_P(i)+1$ dimensional column vector whose entries are η_j and τ_j, respectively, for $j \in \overline{N}_P(i)$. Then, from the first-order condition of the expected payoff, we find that

$$\mathbf{y}(i) = \mathbf{y}^*(i) + \boldsymbol{\eta}(i), \tag{A.18}$$

where $\mathbf{y}^*(i)$ is a vector that satisfies

$$\mathbf{y}^*(i) = \beta_0 A(i)\mathbf{y}^*(i) + \boldsymbol{\tau}(i),$$

with

$$A(i) = \frac{1}{n_P(i)}\big(\mathbf{1}(i)\mathbf{1}(i)' - I_{n_P(i)+1}\big).$$

Hence

$$\mathbf{y}^*(i) = \big(I_{n_P(i)+1} - \beta_0 A(i)\big)^{-1}\boldsymbol{\tau}(i), \tag{A.19}$$

where $\mathbf{1}(i)$ is the $n_P(i)+1$ dimensional column vector of ones and $I_{n_P(i)+1}$ is the $n_P(i)+1$ dimensional identity matrix. Using the Woodbury formula (see Hager (1989)), we obtain that

$$\begin{aligned}\big(I_{n_P(i)+1}-\beta_0 A(i)\big)^{-1} = {} & \big(1+\beta_0/n_P(i)\big)^{-1} I_{n_P(i)+1} \\ & + \frac{1}{\big(1+\beta_0/n_P(i)\big)^2-\beta_0\big(1+\beta_0/n_P(i)\big)\big(n_P(i)+1\big)/n_P(i)} \\ & \cdot \frac{\beta_0 \mathbf{1}(i)\mathbf{1}(i)'}{n_P(i)}.\end{aligned}$$

Hence plugging this back to (A.19), we obtain that the entry y_i^* of $\mathbf{y}^*(i)$ that corresponds to agent i takes the following form:

$$\begin{aligned} y_i^* = {} & \left(1+\frac{\beta_0}{n_P(i)}\right)^{-1}\tau_i \\ & + \left(1+\frac{\beta_0}{n_P(i)}\right)^{-1}\frac{\beta_0 n_P(i)}{n_P(i)(1-\beta_0)}\cdot \overline{\tau}(i)\cdot\frac{n_P(i)+1}{n_P(i)},\end{aligned}$$

where

$$\overline{\tau}(i)=\frac{\mathbf{1}(i)'\boldsymbol{\tau}(i)}{n_P(i)+1}.$$

By rearranging the terms, we obtain that

$$\begin{aligned} y_i^* = {} & \left(1+\frac{\beta_0^2}{\big(n_P(i)+\beta_0\big)(1-\beta_0)}\right)\tau_i \\ & + \frac{\beta_0}{\big(n_P(i)+\beta_0\big)(1-\beta_0)}\sum_{j\in N_P(i)}\tau_j.\end{aligned}$$

Thus (A.18) coincides with (2.9).

Appendix B: Estimation of the asymptotic covariance matrix

We now explain our proposal to estimate the asymptotic covariance matrix, given in equation (3.16) for the model with agents of simple type.

We first explain our proposal to estimate Λ consistently for the case of $\beta_0 \neq 0$. Then we later show how the estimator works even for the case of $\beta_0 = 0$. We first write

$$v_i = R_i(\varepsilon)+\eta_i, \tag{B.1}$$

where

$$R_i(\varepsilon)=w_{ii}^{[0]}\varepsilon_i+\frac{\beta_0 w_{ii}^{[0]}}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}\varepsilon_j.$$

Define for $i, j \in N$,

$$e_{ij} = \mathbf{E}\big[R_i(\varepsilon)R_j(\varepsilon)|\mathcal{F}\big]/\sigma_\varepsilon^2, \tag{B.2}$$

where $\sigma_\varepsilon^2 = \mathrm{Var}(\varepsilon_i^2|\mathcal{F})$ denotes the variance of ε_i. It is not hard to see that for all $i \in N$,

$$e_{ii} = \big(w_{ii}^{[0]}\big)^2 + \frac{\beta_0^2\big(w_{ii}^{[0]}\big)^2}{n_P^2(i)} \sum_{j \in N_P(i)} \lambda_{ij}^2, \tag{B.3}$$

and for $i \neq j$ such that $N_P(i) \cap N_P(j) \neq \emptyset$, $e_{ij} = \beta_0 q_{\varepsilon,ij}$, where

$$q_{\varepsilon,ij} = w_{ii}^{[0]} w_{jj}^{[0]} \left(\frac{\lambda_{ji} 1\{i \in N_P(j)\}}{n_P(j)} + \frac{\lambda_{ij} 1\{j \in N_P(i)\}}{n_P(i)} + \frac{\beta_0}{n_P(i) n_P(j)} \sum_{k \in N_P(i) \cap N_P(j)} \lambda_{ik}\lambda_{jk} \right).$$

Thus, we write

$$\begin{gathered} \frac{1}{n^*} \sum_{i \in N^*} \mathbf{E}\big[v_i^2|\mathcal{F}\big] = a_\varepsilon \sigma_\varepsilon^2 + \sigma_\eta^2, \quad \text{and} \\ \frac{1}{n^*} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} \mathbf{E}[v_i v_j|\mathcal{F}] = \beta_0 b_\varepsilon \sigma_\varepsilon^2, \end{gathered} \tag{B.4}$$

where σ_η^2 denotes the variance of η_i,

$$a_\varepsilon = \frac{1}{n^*} \sum_{i \in N^*} e_{ii}, \quad \text{and} \quad b_\varepsilon = \frac{1}{n^*} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} q_{\varepsilon,ij}.$$

(Note that since not all agents in $N_P(i)$ are in N^* for all $i \in N^*$, the set $N_P(i) \cap N^*$ does not necessarily coincide with $N_P(i)$.) When $\beta_0 \neq 0$, the solution takes the following form:

$$\begin{gathered} \sigma_\varepsilon^2 = \frac{1}{n^* \beta_0 b_\varepsilon} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} \mathbf{E}[v_i v_j|\mathcal{F}] \quad \text{and} \\ \sigma_\eta^2 = \frac{1}{n^*} \sum_{i \in N^*} \mathbf{E}\big[v_i^2|\mathcal{F}\big] - \frac{a_\varepsilon}{n^* \beta_0 b_\varepsilon} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} \mathbf{E}[v_i v_j|\mathcal{F}]. \end{gathered} \tag{B.5}$$

In other words, when $\beta_0 \neq 0$, that is, when there is strategic interaction among the players, we can "identify" σ_ε^2 and σ_η^2 by using the variances and covariances of residuals v_i's. The intuition is as follows. Since the source of cross-sectional dependence of v_i's is due to the presence of ε_i's, we can identify first σ_ε^2 using covariance between v_i and v_j for linked pairs i, j, and then identify σ_η^2 by subtracting from the variance of v_i the contribution from ε_i.

In order to obtain a consistent estimator of Λ which does not require that $\beta_0 \neq 0$, we derive its alternative expression. Let us first write

$$\Lambda = \Lambda_1 + \Lambda_2, \tag{B.6}$$

where

$$\Lambda_1 = \frac{1}{n^*} \sum_{i \in N^*} \mathbf{E}\big[v_i^2 | \mathcal{F}\big] \tilde{\varphi}_i \tilde{\varphi}_i', \quad \text{and}$$

$$\Lambda_2 = \frac{1}{n^*} \sum_{i \in N^*} \sum_{j \in N^*_{-i}} \mathbf{E}[v_i v_j | \mathcal{F}] \tilde{\varphi}_i \tilde{\varphi}_j',$$

where $N^*_{-i} = N^* \backslash \{i\}$. Using (B.1) and (B.5), we can rewrite

$$\begin{aligned}
\Lambda_2 &= \frac{1}{n^*} \sum_{i \in N^*} \sum_{j \in N^*_{-i} : N_P(i) \cap N_P(j) \neq \emptyset} e_{ij} \sigma_\varepsilon^2 \tilde{\varphi}_i \tilde{\varphi}_j' \\
&= \frac{\beta_0}{n^*} \sum_{i \in N^*} \sum_{j \in N^*_{-i} : N_P(i) \cap N_P(j) \neq \emptyset} q_{\varepsilon, ij} \sigma_\varepsilon^2 \tilde{\varphi}_i \tilde{\varphi}_j' \\
&= \frac{s_\varepsilon}{n^*} \sum_{i \in N^*} \sum_{j \in N^*_{-i} : N_P(i) \cap N_P(j) \neq \emptyset} q_{\varepsilon, ij} \tilde{\varphi}_i \tilde{\varphi}_j',
\end{aligned}$$

where

$$s_\varepsilon = \frac{\displaystyle\sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} \mathbf{E}[v_i v_j | \mathcal{F}]}{\displaystyle\sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} q_{\varepsilon, ij}}.$$

Now, it is clear that with this expression for Λ_2, the definition of Λ is well-defined regardless of whether $\beta_0 = 0$ or $\beta_0 \neq 0$. We can then find the estimator of Λ, $\hat{\Lambda}$, by using the empirical analogues to the above, as shown in the main text.

Appendix C: Local identification and estimation

C.1 *Preliminary results for local identification*

In this section, we give computations of derivatives that are used for the proof of local identification. Define

$$v_i(\beta, \rho) = Y_i - q_i(\beta) \bigg(X_i + \frac{\beta}{n_P(i)} \sum_{j \in N_P(i)} \lambda_{ij}(\beta) X_j \bigg)' \rho, \tag{C.1}$$

where

$$q_i(\beta) = \frac{n_P(i)}{n_P(i) - \beta^2 \overline{\lambda}_i}, \qquad \lambda_{ij}(\beta) = \frac{1}{1 - \beta c_{ij}}, \quad \text{and} \quad \overline{\lambda}_i(\beta) = \frac{1}{n_P(i)} \sum_{j \in N_P(i)} \lambda_{ij}(\beta).$$

The relevant derivatives of (C.1) are given by

$$\frac{\partial v_i(\beta,\rho)}{\partial \beta} = -q_i(\beta)\frac{1}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^2(\beta)X_j'\rho - q_i'(\beta)\left(X_i' + \frac{\beta}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}(\beta)X_j'\right)\rho. \tag{C.2}$$

$$\frac{\partial^2 v_i(\beta,\rho)}{\partial \beta^2} = -\frac{2q_i(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^3(\beta)c_{ij}X_j'\rho - \frac{2q_i'(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^2(\beta)X_j'\rho - q_i''(\beta)\left(X_i' + \frac{\beta}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}(\beta)X_j'\right)\rho, \tag{C.3}$$

$$\frac{\partial^3 v_i(\beta,\rho)}{\partial \beta^3} = -\frac{6q_i(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^4(\beta)c_{ij}^2X_j'\rho - \frac{6q_i'(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^3(\beta)c_{ij}X_j'\rho - \frac{3q_i''(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^2(\beta)X_j'\rho - q_i'''(\beta)\left(X_i' + \frac{\beta}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}(\beta)X_j'\right)\rho, \tag{C.4}$$

and

$$\frac{\partial^3 v_i(\beta,\rho)}{\partial \beta^2\,\partial\rho} = \left(\frac{\partial^3 v_i(\beta,\rho)}{\partial \beta^2\,\partial\rho_1}, \dots, \frac{\partial^3 v_i(\beta,\rho)}{\partial \beta^2\,\partial\rho_d}\right)', \tag{C.5}$$

where for each $s = 1, \dots, d$, we have

$$\frac{\partial^3 v_i(\beta,\rho)}{\partial \beta^2\,\partial\rho_s} = -\frac{2q_i(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^3(\beta)c_{ij}X_{j,s} - \frac{2q_i'(\beta)}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}^2(\beta)X_{j,s} - q_i''(\beta)\left(X_{i,s} + \frac{\beta}{n_P(i)}\sum_{j\in N_P(i)} \lambda_{ij}(\beta)X_{j,s}\right).$$

We will focus on showing (C.2), (C.3), and (C.4), since (C.5) follows immediately from (C.3). We provide expressions for $q_i'(\beta)$, $q_i''(\beta)$, and $q_i'''(\beta)$ in Section C.1.4 later.

For the derivations, we will repeatedly make use of the following result.

LEMMA C.1. *For any integer $m \geq 1$,*

$$\frac{\partial \lambda_{ij}^m}{\partial \beta} = m\lambda_{ij}^{m+1}c_{ij}. \tag{C.6}$$

PROOF. First, observe that

$$\begin{aligned}\frac{\partial\lambda_{ij}}{\partial\beta}=\frac{\partial}{\partial\beta}(1-\beta c_{ij})^{-1}&=(-1)(1-\beta c_{ij})^{-2}(-c_{ij})\\&=\frac{c_{ij}}{(1-\beta c_{ij})^2}=\lambda_{ij}^2c_{ij},\end{aligned}$$

Then we note that equation (C.6) with m implies that equation (C.6) also holds for $m+1$:

$$\begin{aligned}\frac{\partial\lambda_{ij}^{m+1}}{\partial\beta}&=\frac{\partial\lambda_{ij}\lambda_{ij}^m}{\partial\beta}\\&=\lambda_{ij}\frac{\partial\lambda_{ij}^m}{\partial\beta}+\lambda_{ij}^m\frac{\partial\lambda_{ij}}{\partial\beta}=\lambda_{ij}\big(m\lambda_{ij}^{m+1}c_{ij}\big)+\lambda_{ij}^m\lambda_{ij}^2c_{ij}=(m+1)\lambda_{ij}^{m+2}c_{ij}.\end{aligned}$$

It follows from the claim above that for each integer $m\geq 1$:

$$\frac{\partial\bar{\lambda}_{ij}^m}{\partial\beta}=\frac{1}{n_P(i)}\sum_{j\in N_P(i)}m\lambda_{ij}^{m+1}c_{ij}. \tag{C.7}$$

□

Below we give derivatives of various functions. The derivations are found in the replication file (Canen, Schwartz, and Song (2020)).

C.1.1 *First derivative, $\partial v_i(\beta,\rho)/\partial\beta$*

$$\frac{\partial v_i(\beta,\rho)}{\partial\beta}=-q_i(\beta)\frac{1}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^2(\beta)X_j'\rho-q_i'(\beta)\bigg(X_i'+\frac{\beta}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}(\beta)X_j'\bigg)\rho.$$

C.1.2 *Second derivative, $\partial^2 v_i(\beta,\rho)/\partial\beta^2$*

$$\begin{aligned}\frac{\partial^2 v_i(\beta,\rho)}{\partial\beta^2}=&-q_i(\beta)\frac{2}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^3(\beta)c_{ij}X_j'\rho-q_i'(\beta)\frac{2}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^2(\beta)X_j'\rho\\&-q_i''(\beta)\bigg(X_i'+\frac{\beta}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}(\beta)X_j'\bigg)\rho.\end{aligned}$$

C.1.3 *Third derivative, $\partial^3 v_i(\beta,\rho)/\partial\beta^3$*

$$\begin{aligned}\frac{\partial^3 v_i(\beta,\rho)}{\partial\beta^3}=&-q_i(\beta)\frac{6}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^4(\beta)c_{ij}^2X_j'\rho-q_i'(\beta)\frac{6}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^3(\beta)c_{ij}X_j'\rho\\&-q_i''(\beta)\frac{3}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^2(\beta)X_j'\rho-q_i'''(\beta)\bigg(X_i'+\frac{\beta}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}(\beta)X_j'\bigg)\rho.\end{aligned}$$

C.1.4 *Expressions for derivatives of $q_i(\beta)$*

$$q_i'(\beta)=\frac{\beta}{n_P(i)}q_i^2(\beta)\big(2\overline{\lambda}_i+\beta\overline{\lambda}_i'\big), \tag{C.8}$$

$$q_i''(\beta) = \frac{\beta}{n_P(i)} q_i(\beta)^2 \big(3\overline{\lambda}_i' + \beta\overline{\lambda}_i''\big) + \frac{q_i(\beta)}{n_P(i)}\big(2\overline{\lambda}_i + \beta\overline{\lambda}_i'\big)\big(q_i(\beta) + 2\beta q_i'(\beta)\big), \quad \text{and} \tag{C.9}$$

$$\begin{aligned} q_i'''(\beta) = {} & \frac{\beta}{n_P(i)} q_i(\beta)^2 \big(4\overline{\lambda}_i'' + \beta\overline{\lambda}_i'''\big) + \frac{q_i(\beta)}{n_P(i)}\big(6\overline{\lambda}_i + 2\beta\overline{\lambda}_i''\big)\big(q_i(\beta) + 2\beta q_i'(\beta)\big) \\ & + \frac{2\big(2\overline{\lambda}_i + \beta\overline{\lambda}_i'\big)}{n_P(i)}\big(q_i(\beta)q_i'(\beta) + \beta q_i(\beta)q_i''(\beta) + q_i'(\beta)\big(q_i(\beta) + \beta q_i'(\beta)\big)\big), \end{aligned} \tag{C.10}$$

where we have from (C.7) that:

$$\overline{\lambda}_i'(\beta) = \frac{1}{n_P(i)} \sum_{j\in N_P(i)} c_{ij}\lambda_{ij}^2 \tag{C.11}$$

$$\overline{\lambda}_i''(\beta) = \frac{2}{n_P(i)} \sum_{j\in N_P(i)} c_{ij}^2\lambda_{ij}^3, \quad \text{and} \quad \overline{\lambda}_i'''(\beta) = \frac{6}{n_P(i)} \sum_{j\in N_P(i)} c_{ij}^3\lambda_{ij}^4. \tag{C.12}$$

C.2 *Proof of Theorem 3.1*

Let us recall our notation first. Let $\theta = [\beta, \rho']'$ and $\theta_0 = [\beta_0, \rho_0']'$. As in Theorem 3.1, we assume that φ_i does not depend on θ. Let $\lambda_{ij}(\beta) = 1/(1 - \beta c_{ij})$, and

$$\overline{\lambda}_i(\beta) = \frac{1}{n_P(i)} \sum_{j\in N_P(i)} \lambda_{ij}(\beta).$$

Let

$$Z_i(\beta) = \left(1 + \frac{\beta^2\overline{\lambda}_i(\beta)}{n_P(i) - \beta^2\overline{\lambda}_i(\beta)}\right)\left(X_i + \frac{\beta}{n_P(i)} \sum_{j\in N_P(i)} \lambda_{ij}(\beta)X_j\right).$$

Then we define

$$v_i(\theta) = Y_i - Z_i(\beta)'\rho.$$

We let

$$G_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^n \mathbf{E}\big[v_i(\theta)\varphi_i | G_P\big] \quad \text{and}$$

$$\hat{G}_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^n v_i(\theta)\varphi_i.$$

We also let for $m = 1, \dots, M$

$$G_{n,m}(\theta) \equiv \frac{1}{n}\sum_{i=1}^n \mathbf{E}\big[v_i(\theta)\varphi_{i,m} | G_P\big],$$

where we recall that $\varphi_{i,m}$ is the mth entry of φ_i.

Lemma C.2. *Suppose that Assumption* 3.2(ii) *holds and that there exists* $\varepsilon > 0$ *such that* $\Theta = \overline{B}(\theta_0; \varepsilon)$. *Then there exists a constant* $C > 0$ *such that for each* $m = 1, \dots, M$ *and for all* $n \geq 2$,

$$\frac{1}{n}\sum_{i \in N} \mathbf{E}\left[\sup_{\theta \in \Theta}\left\|\frac{\partial v_i(\theta)}{\partial \theta}\right\|^2 \varphi_{i,m}^2 \middle| G_P\right] \leq C,$$

and for all $k_1, k_2, k_3 = 1, \dots, d$,

$$\frac{1}{n}\sum_{i \in N} \mathbf{E}\left[\sup_{\theta \in \Theta}\left|\frac{\partial^3 v_i(\theta)}{\partial \theta_{k_1} \partial \theta_{k_2} \partial \theta_{k_2}}\right| |\varphi_{i,m}| \middle| G_P\right] \leq C.$$

Proof. By the assumption that $\beta \in [-1+\nu, 1-\nu]$ for some $\nu > 0$, whenever β is such that $(\beta, \rho) \in \Theta$, and that Θ is compact, and $0 \leq c_{ij} \leq 1$, we have

$$\frac{1}{2-\nu} \leq \lambda_{ij}(\beta) \leq \frac{1}{\nu}.$$

for all $\beta \in [-1+\nu, 1-\nu]$ such that $(\beta, \rho) \in \Theta$. Hence the results immediately follow from Assumption 3.2(ii) and the derivatives that we computed previously. □

Proof of Theorem 3.1. Fix a small $\varepsilon > 0$ and take $\Theta = \overline{B}(\theta_0, \varepsilon)$. Note that $\|\hat{G}_n(\theta)\|$ is continuous in $\theta \in \Theta$ for every realization of the payoff graph G_P and every realization of $(Y_i, X_{i,1}, X_{i,2})_{i \in N}$. Since Θ is compact, the minimizer of $\|\hat{G}_n(\theta)\|$ over Θ exists in Θ. Let us take

$$\hat{\theta} \in \operatorname*{argmin}_{\theta \in \Theta}\|\hat{G}_n(\theta)\|.$$

It suffices to show that $\hat{\theta}$ is consistent for θ_0. For this, we prove the following two claims:

Claim 1. *There exists* $\bar{\delta} > 0$ *such that for any* $\delta \in (0, \bar{\delta}]$, *there exists* $\varepsilon_\delta > 0$ *such that for all* $n \geq 1$,

$$\inf_{\theta \in \Theta: \|\theta - \theta_0\| > \delta}\|G_n(\theta)\| > \varepsilon_\delta.$$

Claim 2. $\sup_{\theta \in \Theta}\|\hat{G}_n(\theta) - G_n(\theta)\| = o_P(1)$.

Then the consistency of $\hat{\theta}$ over $\Theta = \overline{B}(\theta_0, \varepsilon)$ follows as in the proof of Corollary 3.2 of Pakes and Pollard (1989).

Let us first prove Claim 1. By $\beta \in [-1+\nu, 1-\nu]$ with $(\beta, \rho) \in \Theta$ and $0 \leq c_{ij} \leq 1$, $G_n(\theta)$ is infinite times differentiable over $\theta \in \Theta$. For each $m = 1, \dots, M$ and $\theta, \theta_0 \in \Theta$, there exists a point θ_m^* on the line segment between θ and θ_0 such that

$$\begin{aligned} G_{n,m}^2(\theta) &= \frac{\partial G_{n,m}^2(\theta_0)}{\partial \theta'}(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)'\frac{\partial^2 G_{n,m}^2(\theta_m^*)}{\partial \theta \partial \theta'}(\theta - \theta_0) \\ &= \frac{1}{2}(\theta - \theta_0)'\frac{\partial^2 G_{n,m}^2(\theta_m^*)}{\partial \theta \partial \theta'}(\theta - \theta_0), \end{aligned}$$

because $G_{n,m}(\theta_0) = 0$. As for the last term, by Lemma C.2, there exists a constant $C_1 > 0$ such that for all $n \geq 2$,

$$\frac{1}{2}(\theta - \theta_0)' \frac{\partial^2 G^2_{n,m}(\theta^*)}{\partial\theta\,\partial\theta'}(\theta - \theta_0) \geq \frac{1}{2}(\theta - \theta_0)' \frac{\partial^2 G^2_{n,m}(\theta_0)}{\partial\theta\,\partial\theta'}(\theta - \theta_0) - C_1\|\theta - \theta_0\|^3.$$

Note that (again, from $G_{n,m}(\theta_0) = 0$)

$$\frac{\partial^2 G^2_{n,m}(\theta_0)}{\partial\theta\,\partial\theta'} = 2\left(\frac{1}{n}\sum_{i=1}^n H_{i,m}(\theta_0)\right)\left(\frac{1}{n}\sum_{i=1}^n H_{i,m}(\theta_0)\right)'.$$

Hence

$$\begin{aligned}\|G_n(\theta)\|^2 &= G_n(\theta)'G_n(\theta) = \sum_{m=1}^M G^2_{n,m}(\theta) \\ &\geq \sum_{m=1}^M (\theta - \theta_0)'\left(\frac{1}{n}\sum_{i=1}^n H_{i,m}(\theta_0)\right)\left(\frac{1}{n}\sum_{i=1}^n H_{i,m}(\theta_0)\right)'(\theta - \theta_0) - C_1 M\|\theta - \theta_0\|^3.\end{aligned}$$

By Assumption 3.2(iii), with the constant $c > 0$ there, we obtain that for all $n \geq 1$,

$$\|G_n(\theta)\| \geq c\|\theta - \theta_0\|^2 - C_1\|\theta - \theta_0\|^3.$$

We can find $C_2 > 0$ and $\bar{\delta} > 0$ such that for all $0 < \delta \leq \bar{\delta}$, $c\delta^2 - C_1\delta^3 > C_2\delta^2$. Thus, we obtain Claim 1.

Let us turn to the proof of Claim 2. Let $G^*_P = (N, E^*_P)$ be a graph on N such that $ij \in E^*_P$ if and only if $N_P(i) \cap N_P(j) \neq \emptyset$. Then the maximum degree of G^*_P is bounded by $\max_{i\in N} n^2_P(i)$. Note that $v_i(\theta)\varphi_i$ is a function of Y_i and X_j's, and that Y_i's have G^*_P as a conditional dependency graph given $\mathcal{F}$.[1] Since $\mathcal{F}$ involves the σ-field of $X = (X_i)_{i\in N}$, $v_i(\theta)\varphi_i$ has G^*_P as a conditional dependency graph given $\mathcal{F}$. For the proof, we use Lemma 3.4 of Lee and Song (2019). We first show that there exists $C > 0$ such that for all $n \geq 1$ and all $\theta, \tilde{\theta} \in \Theta$,

$$\sqrt{\frac{1}{n}\sum_{i\in N}\mathbf{E}\left[\left(v_i(\theta) - v_i(\tilde{\theta})\right)^2\varphi^2_{i,m}|G_P\right]} \leq C\|\theta - \tilde{\theta}\|. \tag{C.13}$$

[1] Random variables $(W_i)_{i\in N}$ having G^*_P as a conditional dependency graph given $\mathcal{F}$ means that for any set $A \subset N$, $(W_i)_{i\in A}$ and $(W_i)_{i\in N\setminus\overline{N}_P(A)}$ are conditionally independent given $\mathcal{F}$, where $\overline{N}_P(A)$ is the union of $\overline{N}_P(j)$ over $j \in A$.

By Assumption 3.2(ii) and Lemma C.2, there exist constants $C_1, C_2 > 0$ such that for all $m = 1, \dots, M$, and for all $n \geq 1$,

$$\frac{1}{n}\sum_{i\in N}\mathbf{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\partial v_i(\theta)}{\partial\theta}\right\|^2 \varphi_{i,m}^2 \middle| G_P\right] \leq C_1, \quad \text{and}$$
$$\frac{1}{n}\sum_{i\in N}\mathbf{E}\left[\sup_{\theta\in\Theta} v_i^2(\theta)\varphi_{i,m}^2 \middle| G_P\right] \leq C_2. \tag{C.14}$$

The first statement of (C.14) immediately yields (C.13) by the first-order Taylor expansion. Combining this with the second statement of (C.14), and noting that Θ is compact in a finite dimensional Euclidean space, we find from Lemma 3.4 of Lee and Song (2019) that there exists $C > 0$ such that for all $n \geq 1$,

$$\mathbf{E}\left[\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{i=1}^{n}\left(v_i(\theta)\varphi_{i,m} - \mathbf{E}\left[v_i(\theta)\varphi_{i,m}|G_P\right]\right)\right| \middle| G_P\right] \leq C\left(1 + \max_{i\in N} n_P^2(i)\right)/\sqrt{n}.$$

By Assumption 3.2(iv), we obtain Claim 2, which completes the proof. □

C.3 *Proof of Theorem 3.2*

Throughout the proofs, we use the notation C_1 and C_2 to represent a constant which does not depend on n or n^*. Without loss of generality, we also assume that N^* is $\mathcal{F}$-measurable. This loses no generality because due to Condition A of the sampling process in the paper, the same proof goes through if we redefine $\mathcal{F}$ to be the σ-field generated by both $\mathcal{F}$ and N^*.

We introduce auxiliary lemmas which are used for proving Theorem 3.2.

Lemma C.3. *For any array of numbers $\{a_{ij}\}_{i,j\in N}$ and a sequence $\{b_i\}_{i\in N}$ of numbers, we have for any subsets $A, B \subset N$ and for any undirected graph $G = (N, E)$,*

$$\sum_{i\in B}\sum_{j\in N(i)\cap A} a_{ij}b_j = \sum_{i\in A}\left(\sum_{j\in N(i)\cap B} a_{ji}\right)b_i,$$

where $N(i) = \{i \in N : ij \in E\}$.

Proof. Since the graph G is undirected, that is, $1\{j \in N(i)\} = 1\{i \in N(j)\}$, we write the left-hand side sum as

$$\sum_{i\in B}\sum_{j\in A} 1\{j \in N(i)\}a_{ij}b_j = \sum_{j\in A}\sum_{i\in B} 1\{i \in N(j)\}a_{ij}b_j.$$

Interchanging the index notation i and j gives the desired result. □

Lemma C.4. *Suppose that the conditions of Theorem 3.2 hold. Then*

$$\Lambda^{-1/2}\frac{1}{\sqrt{n^*}}\sum_{i\in N^*}\tilde{\varphi}_i v_i \to_d N(0, I_M).$$

Proof. Choose any vector $b \in \mathbf{R}^M$ such that $\|b\| = 1$ and let $\tilde{\varphi}_{i,b} = b'\tilde{\varphi}_i$. Recall that

$$v_i = w_{ii}^{[0]}\varepsilon_i + \sum_{j \in N_P(i)} w_{ij}^{[0]}\varepsilon_j + \eta_i.$$

Define

$$a_i = w_{ii}^{[0]}\tilde{\varphi}_{i,b}1\{i \in N^*\} + \sum_{j \in N_P(i) \cap N^*} \tilde{\varphi}_{j,b} w_{ji}^{[0]}.$$

Then, from (A.1),

$$|a_i| \le \left(1 + \frac{\beta_0^2}{1-\beta_0^2}\right)\left(|\tilde{\varphi}_{i,b}|1\{i \in N^*\} + |\beta_0| \sum_{j \in N_P(i) \cap N^*} \frac{|\tilde{\varphi}_{j,b}|}{n_P(j)(1-|\beta_0|)}\right).$$

Using Lemma C.3, we can write

$$\frac{1}{\sqrt{n^*}} \sum_{i \in N^*} \tilde{\varphi}_{i,b} v_i = \sum_{i \in N^\circ} \xi_i, \tag{C.15}$$

where we recall $N^\circ = \bigcup_{i \in N^*} \overline{N}_P(i)$, and

$$\xi_i = \left(a_i\varepsilon_i + \tilde{\varphi}_{i,b}\eta_i 1\{i \in N^*\}\right)/\sqrt{n^*}.$$

By the Berry–Esseen lemma for independent random variables (see, e.g., Shorack (2000, p. 259)),

$$\sup_{t \in \mathbf{R}} \left| P\left\{ \sum_{i \in N^\circ} \frac{\xi_i}{\sigma_{\xi,i}} \le t \,\middle|\, \mathcal{F} \right\} - \Phi(t) \right| \le \frac{9\mathbf{E}\left[\sum_{i \in N^\circ} |\xi_i|^3 \,\middle|\, \mathcal{F}\right]}{\left(\sum_{i \in N^\circ} \sigma_{\xi,i}^2\right)^{3/2}}, \tag{C.16}$$

where $\sigma_{\xi,i}^2 = \mathrm{Var}(\xi_i|\mathcal{F})$. It suffices to show that the last bound vanishes in probability as $n^* \to \infty$. First, observe that

$$\sum_{i \in N^\circ} \sigma_{\xi,i}^2 = \frac{1}{n^*} \sum_{i \in N^\circ} \left(a_i^2\sigma_\varepsilon^2 + \tilde{\varphi}_{i,b}^2\sigma_\eta^2 1\{i \in N^*\}\right) \ge \frac{\sigma_\eta^2}{n^*} \sum_{i \in N^*} \tilde{\varphi}_{i,b}^2 = \sigma_\eta^2 > 0,$$

because $\frac{1}{n^*}\sum_{i \in N^*} \tilde{\varphi}_{i,b}^2 = 1$. Observe that

$$\begin{aligned}
\mathbf{E}\left[\sum_{i \in N^\circ} |\xi_i|^3 \,\middle|\, \mathcal{F}\right] &\le \frac{4 \max_{i \in N} \mathbf{E}\left[|\varepsilon_i|^3|\mathcal{F}\right]}{(n^*)^{3/2}} \sum_{i \in N^\circ} |\tilde{\varphi}_{i,b}|^3 |a_i|^3 \\
&\quad + \frac{4 \max_{i \in N} \mathbf{E}\left[|\eta_i|^3|\mathcal{F}\right]}{(n^*)^{3/2}} \sum_{i \in N^\circ} |\tilde{\varphi}_{i,b}|^3 \\
&\le \frac{C_1 \max_{i \in N} \mathbf{E}\left[|\varepsilon_i|^3|\mathcal{F}\right]}{(n^*)^{3/2}} \sum_{i \in N^\circ} |a_i|^3 + \frac{C_1 n^\circ \max_{i \in N} \mathbf{E}\left[|\eta_i|^3|\mathcal{F}\right]}{(n^*)^{3/2}},
\end{aligned} \tag{C.17}$$

for some constant $C_1 > 0$, by Assumption 3.4. Now, using the fact that $|\tilde{\varphi}_{i,b}| \le C$, for some constant $C > 0$, we bound the leading term as (for some constants $C_2, C_3 > 0$)

$$\frac{C_2}{n^*} \sum_{i \in N^\circ} |a_i|^3 \le C\left(1 + \frac{\beta_0^2}{1-\beta_0^2}\right)^3 \frac{1}{n^*} \sum_{i \in N^\circ} \left(|\tilde{\varphi}_{i,b}| 1\{i \in N^*\} + |\beta_0| \sum_{j \in N_P(i) \cap N^*} \frac{|\tilde{\varphi}_{j,b}|}{n_P(j)(1-|\beta_0|)} \right)^3$$
$$\le \frac{C_3}{(1-|\beta_0|)^3}\left(1 + \frac{\beta_0^2}{1-\beta_0^2}\right)^3,$$

by Assumption 3.5. Therefore, for some constant $C_2 > 0$,

$$\mathbf{E}\left[\sum_{i \in N^\circ} |\xi_i|^3 \Big| \mathcal{F}\right] \le \frac{C_2}{\sqrt{n^*}(1-|\beta_0|)^6} \max_{i \in N^*} \mathbf{E}\left[|\varepsilon_i|^3 | \mathcal{F}\right] + \frac{C_2 n^\circ}{(n^*)^{3/2}} \max_{i \in N^*} \mathbf{E}\left[|\eta_i|^3 | \mathcal{F}\right].$$

Thus we conclude that the bound in (C.16) is $O_P((n^*)^{-1/2} + n^\circ (n^*)^{-3/2})$, where $n^\circ = |N^\circ|$. However, for some constant $C > 0$,

$$n^\circ \le \sum_{i \in N^*} \left|\overline{N}_P(i)\right| \le C n^*,$$

by Assumption 3.5. Hence we obtain the desired result. □

LEMMA C.5. *Suppose that the conditions of Theorem* 3.2 *hold. Then*

$$\mathbf{E}\left[\|S_{\tilde{\varphi}v}\|^2 | \mathcal{F}\right] = O\left((n^*)^{-1}\right), \quad \textit{and} \quad \mathbf{E}\left[\|S_{Z^* v}\|^2 | \mathcal{F}\right] = O\left((n^*)^{-1}\right),$$

where

$$Z_i^* = \sum_{j \in N_P(i) \cap N^*} Z_j, \; S_{\tilde{\varphi}v} = \frac{1}{n^*} \sum_{i \in N^*} \tilde{\varphi}_i v_i, \quad \textit{and} \quad S_{Z^* v} = \frac{1}{n^*} \sum_{i \in N^*} Z_i^* v_i.$$

PROOF. Recall the definitions of e_{ij} and e_{ii} in (B.2) and (B.3). First, observe that

$$e_{ii} \le \left(1 + \frac{\beta_0^2}{1-\beta_0^2}\right)^2 \left(1 + \frac{\beta_0^2}{(1-|\beta_0|)^2}\right), \quad \text{and}$$
$$|e_{ij}| \le \frac{2+|\beta_0|}{(1-|\beta_0|)^2}\left(1 + \frac{\beta_0^2}{1-\beta_0^2}\right)^2. \tag{C.18}$$

Note that

$$\mathbf{E}\left[\|S_{\tilde{\varphi}v}\|^2 | \mathcal{F}\right] \le \frac{\sigma_\varepsilon^2}{(n^*)^2} \sum_{i \in N^*} \sum_{j \in N_{-i}^* : N_P(i) \cap N_P(j) \ne \emptyset} |e_{ij}| \|\tilde{\varphi}_i\| \|\tilde{\varphi}_j\|$$
$$+ \frac{1}{(n^*)^2} \sum_{i \in N^*} \left(|e_{ii}| \sigma_\varepsilon^2 + \sigma_\eta^2\right) \|\tilde{\varphi}_i\|^2.$$

However, since $\|\tilde{\varphi}_i\| \le C$ by Assumption 3.4, we use (C.18) to obtain that $\mathbf{E}[\|S_{\tilde{\varphi}v}\|^2 | \mathcal{F}] = O((n^*)^{-1})$.

Let us turn to the second bound. Observe that by Assumption 3.4, we have some $C > 0$ such that for all $i \in N^*$, $\|Z_i^*\| \le C$. Following the same proof as before, we obtain the desired result for $\mathbf{E}[\|S_{Z^*v}\|^2|\mathcal{F}]$ as well. □

LEMMA C.6. *Suppose that the conditions of Theorem* 3.2 *hold. Then the following holds*:

(i) $\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i^2 - v_i^2)\tilde{\varphi}_i\tilde{\varphi}_i' = O_P(1/\sqrt{n^*})$.

(ii) $\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_i\tilde{v}_j - v_iv_j)\tilde{\varphi}_i\tilde{\varphi}_j' = O_P(1/n^*)$.

(iii) $\frac{1}{n^*}\sum_{i\in N^*}(v_i^2 - \mathbf{E}[v_i^2|\mathcal{F}])\tilde{\varphi}_i\tilde{\varphi}_i' = O_P(1/\sqrt{n^*})$.

(iv) $\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}(v_iv_j - \mathbf{E}[v_iv_j|\mathcal{F}])\tilde{\varphi}_i\tilde{\varphi}_j' = O_P(1/\sqrt{n^*})$.

PROOF. (i) First, write $\tilde{v} - v = -Z(\tilde{\rho} - \rho_0)$, where $\tilde{\rho} - \rho_0 = [S_{Z\tilde{\varphi}}S_{Z\tilde{\varphi}}']^{-1}S_{Z\tilde{\varphi}}S_{\tilde{\varphi}v}$. Hence

$$\left\|\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i - v_i)^2\tilde{\varphi}_i\tilde{\varphi}_i'\right\| \le \frac{C_1}{n^*}\sum_{i\in N^*}(\tilde{v}_i - v_i)^2,$$

for some constant $C_1 > 0$. As for the last term, note that

$$\begin{aligned}&\frac{1}{n^*}\sum_{i\in N^*}\mathbf{E}\big[(\tilde{v}_i - v_i)^2|\mathcal{F}\big]\\ &= \frac{1}{n^*}\operatorname{tr}\big(S_{Z\tilde{\varphi}}'\big[S_{Z\tilde{\varphi}}S_{Z\tilde{\varphi}}'\big]^{-1}S_{ZZ}\big[S_{Z\tilde{\varphi}}S_{Z\tilde{\varphi}}'\big]^{-1}S_{Z\tilde{\varphi}}\Lambda\big) = O_P\left(\frac{1}{n^*}\right),\end{aligned} \tag{C.19}$$

by the definition of Λ in (3.5) and by Lemma C.5. However, we need to deal with

$$\left|\frac{1}{n^*}\sum_{i\in N^*}\big(\tilde{v}_i^2 - v_i^2\big)\right| \le \sqrt{\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i - v_i)^2}\sqrt{\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i + v_i)^2}. \tag{C.20}$$

Note that

$$\begin{aligned}\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i + v_i)^2 &\le \frac{2}{n^*}\sum_{i\in N^*}(\tilde{v}_i - v_i)^2 + \frac{8}{n^*}\sum_{i\in N^*}v_i^2\\ &= O_P\left(\frac{1}{n^*}\right) + \frac{8}{n^*}\sum_{i\in N^*}v_i^2,\end{aligned}$$

by (C.19). As for the last term,

$$\frac{1}{n^*}\sum_{i\in N^*}\mathbf{E}\big[v_i^2|\mathcal{F}\big] \le \frac{2}{n^*}\sum_{i\in N^*}\mathbf{E}\big[R_i^2(\varepsilon)|\mathcal{F}\big] + \frac{2}{n^*}\sum_{i\in N^*}\mathbf{E}\big[\eta_i^2|\mathcal{F}\big].$$

The last term is bounded by $2\sigma_\eta^2$, and the first term on the right-hand side is bounded by

$$\frac{2\sigma_\varepsilon^2}{n^*}\sum_{i\in N^*}e_{ii} \le C,$$

by (C.18). Combining this with (C.19) and (C.20), we obtain the desired result.

(ii) Let us first write

$$\begin{aligned}
&\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_i\tilde{v}_j - v_iv_j)\\
&\quad=\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_i - v_i)(\tilde{v}_j - v_j)\\
&\qquad+\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_i - v_i)v_j\\
&\qquad+\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*}v_i(\tilde{v}_j - v_j) = A_{n,1}+A_{n,2}+A_{n,3}, \quad \text{say.}
\end{aligned}$$

As for the leading term, by Cauchy–Schwarz inequality,

$$|A_{n,1}| = \sqrt{\frac{1}{n^*}\sum_{i\in N^*}(\tilde{v}_i - v_i)^2}\sqrt{\frac{1}{n^*}\sum_{i\in N^*}\left(\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_j - v_j)\right)^2}.$$

Note that

$$\begin{aligned}
&\frac{1}{n^*}\sum_{i\in N^*}\mathbf{E}\left[\left(\sum_{j\in N_P(i)\cap N^*}(\tilde{v}_j - v_j)\right)^2\Big|\mathcal{F}\right]\\
&\quad\le\frac{1}{n^*}\sum_{i\in N^*}\left|N_P(i)\cap N^*\right|\sum_{j\in N_P(i)\cap N^*}\mathbf{E}\left[(\tilde{v}_j - v_j)^2|\mathcal{F}\right]\\
&\quad=\frac{1}{n^*}\sum_{i\in N^*}\left(\sum_{j\in N_P(i)\cap N^*}\left|N_P(j)\cap N^*\right|\right)\mathbf{E}\left[(\tilde{v}_i - v_i)^2|\mathcal{F}\right],
\end{aligned}$$

where the inequality above uses Jensen's inequality and the equality above uses Lemma C.3. Hence the last term is bounded by

$$\frac{\max_{i\in N^*}\left|N_P(i)\cap N^*\right|^2}{n^*}\sum_{i\in N^*}\mathbf{E}\left[(\tilde{v}_i - v_i)^2|\mathcal{F}\right]\le O_P\left(\frac{1}{n^*}\right),$$

by (C.19). Thus we conclude that

$$|A_{n,1}| = O_P\left(\frac{1}{n^*}\right).$$

Now, let us turn to $A_{n,2}$. Observe that

$$\begin{aligned}
A_{n,2} &= -\frac{1}{n^*}\sum_{i\in N^*}Z_i'\sum_{j\in N_P(i)\cap N^*}v_j(\tilde{\rho}-\rho_0)\\
&= -\left(\frac{1}{n^*}\sum_{i\in N^*}Z_i^{*'}v_i\right)(\tilde{\rho}-\rho_0) = -S_{Z^*v}(\tilde{\rho}-\rho_0)
\end{aligned}$$

using Lemma C.3. From the proof of (i), we obtain that

$$\tilde{\rho} - \rho_0 = O_P\left(\frac{1}{\sqrt{n^*}}\right).$$

Hence combined with Lemma C.5, we have

$$|A_{n,2}| = O_P\left(\frac{1}{n^*}\right).$$

Since by Lemma C.3, $A_{n,2} = A_{n,3}$, the proof of (ii) is complete.

(iii) Note that

$$\mathrm{Var}\left(\frac{1}{n^*}\sum_{i\in N^*} R_i^2(\varepsilon)\Big|\mathcal{F}\right) \leq \frac{2}{(n^*)^2}\sum_{i\in N^*}\mathrm{Var}\left(\left(w_{ii}^{[0]}\right)^2\varepsilon_i^2|\mathcal{F}\right) + \frac{2}{(n^*)^2}\sum_{i\in N^*}\mathrm{Var}\left(\left(\frac{\beta_0 w_{ii}^{[0]}}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}\varepsilon_j\right)^2\Big|\mathcal{F}\right).$$

The leading term is $O_P((n^*)^{-1})$. The last term is bounded by

$$\frac{2}{(n^*)^2}\sum_{i\in N^*}\frac{\beta_0^4\left(w_{ii}^{[0]}\right)^4}{n_P(i)}\sum_{j\in N_P(i)}\lambda_{ij}^4\mathbf{E}\left[\varepsilon_j^4|\mathcal{F}\right] = O_P\left((n^*)^{-1}\right).$$

Since $v_i = R_i(\varepsilon) + \eta_i$ and ε_i's and η_i's are independent, we obtain the desired rate.

(iv) For simplicity of notation, define

$$V_{ij} = \left(v_i v_j - \mathbf{E}[v_i v_j|\mathcal{F}]\right)\tilde{\varphi}_i\tilde{\varphi}_j'.$$

Then we write

$$\mathbf{E}\left[\left(\frac{1}{n^*}\sum_{i\in N^*}\sum_{j\in N_P(i)\cap N^*} V_{ij}\right)^2\Big|\mathcal{F}\right] = \frac{1}{(n^*)^2}\sum_{i_1\in N^*}\sum_{j_1\in N_P(i)\cap N^*}\sum_{i_2\in N^*}\sum_{j_2\in N_P(i)\cap N^*}\mathbf{E}[V_{i_1j_1}V_{i_2j_2}|\mathcal{F}].$$

The last expection is zero, whenever (i_2, j_2) is away from (i_1, j_1) by more than two edges. Hence we can bound the last term by (using Assumption 3.5))

$$\frac{C_1}{n^*}\max_{i\in N}\mathbf{E}\left[v_i^2|\mathcal{F}\right] \leq \frac{C_2}{n^*}$$

for some constants C_1, C_2 which do not depend on n. □

LEMMA C.7. *Suppose that the conditions of Theorem* 3.2 *hold. Then*

$$\hat{\Lambda} - \Lambda = O_P\left(\frac{1}{\sqrt{n^*}}\right).$$

PROOF. We write

$$\hat{\Lambda}_1 - \Lambda_1 = \frac{1}{n^*} \sum_{i \in N^*} \left(\tilde{v}_i^2 - \mathbf{E}\left[v_i^2 | \mathcal{F}\right]\right) \tilde{\varphi}_i \tilde{\varphi}_i' \quad \text{and}$$

$$\hat{\Lambda}_2 - \Lambda_2 = \frac{\hat{s}_\varepsilon - s_\varepsilon}{n^*} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} q_{\varepsilon,ij} \tilde{\varphi}_i \tilde{\varphi}_j'.$$

By Assumption 3.3 and Lemma C.6(ii)(iv), we have

$$\hat{s}_\varepsilon - s_\varepsilon = O_P\left(1/\sqrt{n^*}\right).$$

The desired result follows from this and applying Lemma C.6(i) and (iii) to $\hat{\Lambda}_1 - \Lambda_1$. □

LEMMA C.8. *Suppose that the conditions of Theorem* 3.2 *hold. Then the following holds*:

(i) $\frac{1}{n^*} \sum_{i \in N^*} (\hat{v}_i^2 - v_i^2) \tilde{\varphi}_i \tilde{\varphi}_i' = O_P(1/\sqrt{n^*})$.

(ii) $\frac{1}{n^*} \sum_{i \in N^*} \sum_{j \in N_P(i) \cap N^*} (\hat{v}_i \hat{v}_j - v_i v_j) \tilde{\varphi}_i \tilde{\varphi}_j' = O_P(1/n^*)$.

PROOF. First, write $\hat{v} - v = -Z(\hat{\rho} - \rho_0)$, where

$$\hat{\rho} - \rho_0 = \left[S_{Z\tilde{\varphi}} \hat{\Lambda}^{-1} S_{Z\tilde{\varphi}}'\right]^{-1} S_{Z\tilde{\varphi}} \hat{\Lambda}^{-1} S_{\tilde{\varphi} v}. \tag{C.21}$$

Following the same arguments as in the proof of Lemma C.6(i) and (ii) and Lemma C.7, we obtain the desired result. □

PROOF OF THEOREM 3.2. Let us consider the first statement. We write

$$\begin{aligned} \frac{1}{\sqrt{n^*}} \hat{\Lambda}^{-1/2} \tilde{\varphi}' \hat{v} &= \frac{1}{\sqrt{n^*}} \hat{\Lambda}^{-1/2} \tilde{\varphi}' (\hat{v} - v) + \frac{1}{\sqrt{n^*}} \hat{\Lambda}^{-1/2} \tilde{\varphi}' v \\ &= -\frac{1}{\sqrt{n^*}} \hat{\Lambda}^{-1/2} \tilde{\varphi}' Z(\hat{\rho} - \rho_0) + \frac{1}{\sqrt{n^*}} \hat{\Lambda}^{-1/2} \tilde{\varphi}' v = \sqrt{n^*}(I - P) \hat{\Lambda}^{-1/2} S_{\tilde{\varphi} v}, \end{aligned}$$

using (C.21), where

$$P = \hat{\Lambda}^{-1/2} S_{Z\tilde{\varphi}}' \left[S_{Z\tilde{\varphi}} \hat{\Lambda}^{-1} S_{Z\tilde{\varphi}}'\right]^{-1} S_{Z\tilde{\varphi}} \hat{\Lambda}^{-1/2}.$$

Note that P is a projection matrix from $\mathbf{R}^M$ onto the range space of $\hat{\Lambda}^{-1/2} S_{Z\tilde{\varphi}}'$. Hence combining Lemmas C.4 and C.7. We obtain the desired result. The second result follows from Lemma C.4 and equation (C.21). □

APPENDIX D: CONVERGENCE OF BEHAVIORAL STRATEGIES TO EQUILIBRIUM STRATEGIES

PROOF OF THEOREM 2.3. Our proof is in two steps. First, we show the convergence of the behavioral strategies to the equilibrium strategies in a game without private information (without η_i). In the second step, we use this first result and show it also holds when we extend the game to allow for η_i.

Let us first consider the game without private information (i.e., $\eta_i = 0$ for all $i \in N$). We denote the behavioral strategies in this complete information game by $\tilde{s}_i^{[m]}$, and the equilibrium strategies as $\tilde{s}_i^{\mathsf{BNE}}$. This notation will allow us to differentiate these strategies from the case with incomplete information. From Theorem 2.2, without η_i's, we have that:

$$\begin{aligned}
\tilde{s}_i^{[m+1]}(\mathcal{I}_{i,m+1}) &= \left(\frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ki}^{[m]} + 1\right)\tau_i \\
&\quad + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} \sum_{j \in N_{P,m+2}(i)} w_{kj}^{[m]} 1\{j \in \overline{N}_{P,m}(k)\} \\
&= \left(\frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} w_{ki}^{[m]} + 1\right)\tau_i \\
&\quad + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} \left(\sum_{j \in N_{P,m+1}(k) \setminus \{i\}} w_{kj}^{[m]} \tau_j + w_{kk}^{[m]} \tau_k \right) \\
&= \tau_i + \frac{\beta_0}{n_P(i)} \sum_{k \in N_P(i)} s_k^{[m]}(\mathcal{I}_{k,m}).
\end{aligned} \tag{D.1}$$

Thus we find that for any $m, m' > 0$,

$$\begin{aligned}
&\left|\tilde{s}_i^{[m+1]}(\mathcal{I}_{i,m+1}) - \tilde{s}_i^{[m'+1]}(\mathcal{I}_{i,m'+1})\right| \\
&\leq |\beta_0| \frac{1}{n_P(i)} \sum_{k \in N_P(i)} \left|\tilde{s}_k^{[m]}(\mathcal{I}_{k,m}) - \tilde{s}_k^{[m']}(\mathcal{I}_{k,m'})\right|.
\end{aligned} \tag{D.2}$$

Let $\mathcal{F}$ be the collection of all the $\mathcal{I}$-measurable $\mathbf{R}^n$-valued maps $f = (f_i)_{i \in N}$ such that $\mathbf{E}[f_i^2] < \infty$ for each $i \in N$. We endow $\mathcal{F}$ with a pseudo metric: for $f = (f_i)_{i \in N}$ and $g = (g_i)_{i \in N}$,

$$\|f - g\|_2 = \max_{1 \leq i \leq n} \sqrt{\mathbf{E}\big[(f_i - g_i)^2\big]}. \tag{D.3}$$

As usual, we view $(\mathcal{F}, \|\cdot\|_2)$ as a collection of equivalence classes on which $d(f, g) \equiv \|f - g\|_2$ is a metric. Since

$$\sqrt{\frac{1}{n} \sum_{i \in N} \mathbf{E}\big[(f_i - g_i)^2\big]} \leq \|f - g\|_2 \leq \sqrt{\sum_{i \in N} \mathbf{E}\big[(f_i - g_i)^2\big]}, \tag{D.4}$$

the metric space $(\mathcal{F}, \|\cdot\|_2)$ is complete, a property inherited from the completeness of an L_2 space.

Each strategy profile $\tilde{s}(\mathcal{I}_m)^{[m]}(\omega)$ from game Γ_m belongs to $(\mathcal{F}, \|\cdot\|_2)$. Consider a sequence of best response strategy profiles $\{\tilde{s}^{[m]}(\mathcal{I}_m)\}_{m=1}^{\infty}$. Certainly by (D.2) and the fact that $|\beta_0| < 1$, the sequence $\{\tilde{s}^{[m]}(\mathcal{I}_m)\}_{m=1}^{\infty}$ is Cauchy in $(\mathcal{F}, \|\cdot\|_2)$, and has a limit, say, $\tilde{s}_{\infty}(\mathcal{I})$ in $\mathcal{F}$ by its completeness. Now, for the first step of the proof, it remains to show that $\tilde{s}_\infty(\mathcal{I})$ is identical to $\tilde{s}^{\mathsf{BNE}}(\mathcal{I})$ almost everywhere, where $\tilde{s}^{\mathsf{BNE}}(\mathcal{I})$ is defined as a fixed point to

$$\tilde{s}_i^{\mathsf{BNE}}(\mathcal{I}) = \tau_i + \beta_0 \frac{1}{n_P(i)} \sum_{k \in N_P(i)} \tilde{s}_k^{\mathsf{BNE}}(\mathcal{I}). \tag{D.5}$$

To see this, let us view $\tilde{s}^{[m+1]}(\mathcal{I}_{m+1})$ as an n-dimensional column vector of $\tilde{s}_i^{[m+1]}(\mathcal{I}_{i,m+1})$, $i \in N$ and A an $n \times n$ matrix whose (i,j)-th entry is given by $1\{j \in N_P(i)\}/n_P(i)$. Then we can rewrite (D.1) as

$$\tilde{s}^{[m]}(\mathcal{I}_m) = \tau + \beta_0 A \tilde{s}^{[m-1]}(\mathcal{I}_{m-1}),$$

where $\tau = (\tau_i)_{i \in N}$. This implies that

$$\tilde{s}_\infty(\mathcal{I}) - \big(\tau + \beta_0 A \tilde{s}_\infty(\mathcal{I})\big) = \tilde{s}_\infty(\mathcal{I}) - \tilde{s}^{[m]}(\mathcal{I}_m) + \beta_0 A\big(\tilde{s}^{[m-1]}(\mathcal{I}_{m-1}) - \tilde{s}_\infty(\mathcal{I})\big).$$

Thus we have

$$\begin{aligned}
&\big\|\tilde{s}_\infty(\mathcal{I}) - \big(\tau + \beta_0 A \tilde{s}_\infty(\mathcal{I})\big)\big\|_2 \\
&\quad \le \big\|\tilde{s}_\infty(\mathcal{I}) - \tilde{s}^{[m]}(\mathcal{I}_m)\big\|_2 + |\beta_0| \|A\| \big\|\tilde{s}^{[m-1]}(\mathcal{I}_{m-1}) - \tilde{s}_\infty(\mathcal{I})\big\|_2,
\end{aligned}$$

where $\|A\| = \sqrt{\mathrm{tr}(A'A)}$. Note that $\|A\| < \infty$ and does not depend on m. Hence by sending $m \to \infty$, we have

$$\big\|\tilde{s}_\infty(\mathcal{I}) - \big(\tau + \beta_0 A \tilde{s}_\infty(\mathcal{I})\big)\big\|_2 = 0. \tag{D.6}$$

Since $|\beta_0| < 1$ and A is row normalized, the matrix $I - \beta_0 A$ is invertible and the row sums of $(I - \beta_0 A)^{-1}$ are uniformly bounded (e.g., see Lee (2002, p. 257)). Therefore, if we define

$$\tilde{s}^*(\mathcal{I}) = (I - \beta_0 A)^{-1}\tau,$$

we have $\|\tilde{s}^*(\mathcal{I})\|_2 < \infty$ by (2.11). On the other hand, it is not hard to see that $\tilde{s}^*(\mathcal{I})$ is almost everywhere identical to the equilibrium strategy profile $\tilde{s}^{\mathsf{BNE}}(\mathcal{I})$. Also, by (D.6), $\tilde{s}^*(\mathcal{I})$ is almost everywhere identical to $\tilde{s}_\infty(\mathcal{I})$. The first part of the proof follows by (D.4) and the fact that

$$\mathbf{E}\Big[\max_{i \in N}(f_i - g_i)^2\Big] \le \sum_{i \in N} \mathbf{E}\big[(f_i - g_i)^2\big].$$

As a result, we have the convergence of behavioral strategies to equilibrium strategies in the complete information analogue to our incomplete information game.

To complete our proof, we note that the actual behavioral strategies with incomplete information (with potentially nonzero η_i's) are given by

$$s_i^{[m]} = \tilde{s}_i^{[m]} + \eta_i. \tag{D.7}$$

This follows immediately from using equation (2.3), Theorem 2.2 and Assumption 3.1. Analogously, the equilibrium strategies from game Γ_∞ are given by

$$s_i^{\mathsf{BNE}} = \tilde{s}_i^{\mathsf{BNE}} + \eta_i. \tag{D.8}$$

As a result, convergence of $\tilde{s}_i^{[m]}$ to $\tilde{s}_i^{\mathsf{BNE}}$ implies convergence of $s_i^{[m]}$ to s_i^{BNE}, which completes the proof. □

References

Canen, N. Schwartz, J., and Song, K. (2020), "Supplement to 'Estimating local interactions among many agents who observe their neighbors'." *Quantitative Economics Supplemental Material*, 11, https://doi.org/10.3982/QE923. [1, 11]

Hager, W. W. (1989), "Updating the inverse of a matrix." *SIAM Review*, 31, 221–239. [7]

Lee, J. H. and K. Song (2019), "Stable limit theorems for empirical processes under conditional neighborhood dependence." *Bernoulli*, 25, 1189–1224. [14, 15]

Lee, L.-F. (2002), "Consistency and efficiency of least squares estimation for mixed regressive. Spatial autoregressive models." *Econometric Theory*, 18, 252–277. [23]

Pakes, A. and D. Pollard (1989), "Simulation and the asymptotics of optimization estimators." *Econometrica*, 57 (5), 1027–1057. [13]

Shorack, G. R. (2000), *Probability for Statistics*. Springer, New York, NY. [16]

Co-editor Andres Santos handled this manuscript.

Manuscript received 11 July, 2017; final version accepted 17 December, 2019; available online 11 February, 2020.