Online Supplement For

"*Count Data Models with Social Interactions under Rational Expectations*"

Elysée Aristide Houndetoungan*

S.1 Proof of the convergence of the infinite summations

Many infinite summations appear in the paper (e.g., the expected choice, the infinite summations in Proposition 2.2, Assumption 2.4, and several others used throughout the proofs). In this section, I state and prove a general lemma on the convergence of these infinite sums.

Lemma S.1. *Let h be a continuous function on $\mathbb{R}$ and f_γ be a function defined for any $u \in \mathbb{R}$ as $f_\gamma(u) = \sum_{r=0}^{+\infty} r^\gamma h(u - b_r)$, where $\gamma \geqslant 0$ and $(b_k)_{k\in\mathbb{N}}$ is an increasing positive sequence, such that $\lim\limits_{r\to\infty} r^{-\rho}(b_{r+1} - b_r) > 0$, where $\rho \geqslant 0$. The following statements hold.*

(i) For any $u \in \mathbb{R}$, if $h(x) = o(|x|^{-\kappa})$ at $-\infty$, where $(1+\rho)\kappa > 1 + \gamma$, then $f_\gamma(u) < \infty$.

(ii) If $h(x) = o(|x|^{-\kappa})$ at both $-\infty$ and $+\infty$, where $(1+\rho)\kappa > 1$, then f_0 is bounded on $\mathbb{R}$.

Statement (ii) and Assumption 2.3 ensure that B_c defined in Assumption 2.4 is finite. Statement (i) and Assumption 2.3 also imply that the other infinite summations in the paper are finite.

Proof of Lemma S.1

The proof is done in several steps.

Step 0: I show that if $h(x) = o\left(|x|^{-\kappa}\right)$ at both $-\infty$ and $+\infty$, then $\exists\ M \geqslant 1$, such that $|h(u - b_r)| \leqslant M\left(|u - b_r| + 1\right)^{-\kappa}$. Moreover, this is also true for large r even if $h(x) = o\left(|x|^{-\kappa}\right)$ only at $-\infty$.

The condition $h(x) = o\left(|x|^{-\kappa}\right)$ at both $-\infty$ and $+\infty$ is also equivalent to $|h(x)| = o\left((|x|+1)^{-\kappa}\right)$. Thus, $\exists\ x_0 \in \mathbb{R}_+ \ /\ \forall\ x < -x_0$ or $x > x_0$, $|h(x)| < (|x|+1)^{-\kappa}$. As h is continuous, this implies that there exists $M \geqslant 1$, such that $\forall\ x \in \mathbb{R}$, $|h(x)| \leqslant M\left(|x|+1\right)^{-\kappa}$. As a result, $|h(u - b_r)| \leqslant M\left(|u - b_r| + 1\right)^{-\kappa}$.

Step 1: I prove Statement (i).

Let f^* be the real-valued function defined as $f^*(u) = \sum_{r=0}^{\infty}\left(|u - b_r| + 1\right)^{-\kappa}$, $\forall\ u \in \mathbb{R}$.
The condition $\lim\limits_{r\to\infty} r^{-\rho}(b_{r+1} - b_r) > 0$ implies that there exists $k_0 \in \mathbb{N}$ and $b > 0$, such that $\forall\ r \geqslant k_0$, $r^{-\rho}(b_{r+1} - b_r) \geqslant b$, i.e., $b_{r+1} \geqslant b\sum_{s=k_0}^{r} s^\rho + b_{k_0}$. As $\lim\limits_{r\to\infty} b_r = \infty$, $\forall\ u \in \mathbb{R}$, it is possible to choose k_0 sufficiently large, such that $b_{k_0} > u$. It follows that $\forall\ r > k_0$, $|u - b_r| = b_r - u \geqslant b\sum_{s=k_0}^{r-1} s^\rho + b_{k_0} - u \geqslant 0$, which implies $\left(|u - b_r| + 1\right)^{-\kappa} \leqslant \left(b\sum_{s=k_0}^{r-1} s^\rho + b_{k_0} - u\right)^{-\kappa}$, and thus $\left(|u - b_r| + 1\right)^{-\kappa} \leqslant O\left(r^{-(1+\rho)\kappa}\right)$

*Thema, Cy Cergy Paris Université; Email: aristide.houndetoungan@cyu.fr.

since $\sum_{s=k_0}^{r-1} s^\rho = O(r^{1+\rho})$. Therefore, $f^*(u) < \infty$, $\forall\ u \in \mathbb{R}$. Using the result of the step 0, it follows that $\forall\ u \in \mathbb{R},\ \gamma \geqslant 0,\ r^\gamma h(u - b_r) = O\left(r^{-(1+\rho)\kappa+\gamma)}\right)$. Hence, $f_\gamma(u) < \infty$ if $(1+\rho)\kappa > 1+\gamma$.

Step 2: I prove Statement (ii).

As $|h(u-b_r)| \leqslant M\left(|u-b_r|+1\right)^{-\kappa}$, it is sufficient to prove that f^* is bounded. Moreover, since f^* is a continuous function, this also amounts to proving $\lim\limits_{u\to-\infty} f^*(u)$ and $\lim\limits_{u\to+\infty} f^*(u)$ are finite.
For any $u \leqslant 0$, I have $\left(|u-b_r|+1\right)^{-\kappa} = \left(b_r - u + 1\right)^{-\kappa} \leqslant \left(b_r+1\right)^{-\kappa}$. Thus, $f^*(u) \leqslant f^*(0)$.
Since f^* is a positive function, this implies that $\lim\limits_{u\to-\infty} f^*(u)$ is finite.

Let $k_0 \in \mathbb{N}^*$, such that $\forall\ r,\ r' \geqslant k_0$ with $r > r'$, $b_r - b_{r'} \geqslant b\sum_{s=r'}^{r-1} s^\rho$, for some $b > 0$.
For u positive and sufficiently large, $\exists\ k^* \in \mathbb{N}$ (with k^* depending on u), where $k^* > k_0$ and $\forall\ r < k^*$, $u > b_r$, and $\forall\ r \geqslant k^*$, $u \leqslant b_r$. Thus, $f^*(u)$ can be decomposed as

$$
\begin{aligned}
f^*(u) &= \textstyle\sum_{r=0}^{k_0-1}\left(|u-b_r|+1\right)^{-\kappa} + \sum_{r=k_0}^{k^*-1}\left(|u-b_r|+1\right)^{-\kappa} + \sum_{r=k^*}^{\infty}\left(|u-b_r|+1\right)^{-\kappa},\\
f^*(u) &\leqslant k_0 + \textstyle\sum_{r=k_0}^{k^*-1}\left(u-b_r+1\right)^{-\kappa} + \sum_{r=k^*}^{\infty}\left(b_r-u+1\right)^{-\kappa},\\
f^*(u) &\leqslant k_0 + \textstyle\sum_{r=k_0}^{k^*-1}\left(b_{k^*-1}-b_r+1\right)^{-\kappa} + \sum_{r=k^*}^{\infty}\left(b_r-b_{k^*}+1\right)^{-\kappa},\\
f^*(u) &\leqslant 2 + k_0 + \textstyle\sum_{r=k_0}^{k^*-2}\left(b_{k^*-1}-b_r\right)^{-\kappa} + \sum_{r=k^*+1}^{\infty}\left(b_r-b_{k^*}\right)^{-\kappa}.
\end{aligned}
$$

If $k_0 \leqslant r \leqslant k^*-1$, then $b_{k^*-1} - b_r \geqslant b\sum_{s=r}^{k^*-2} s^\rho$. Thus, $(b_{k^*-1}-b_r)^{-\kappa} \leqslant \left(b\sum_{s=r}^{k^*-2} s^\rho\right)^{-\kappa}$.
Analogously, if $k^* \leqslant r$, then $b_r - b_{k^*} \geqslant b\sum_{s=k^*}^{r-1} s^\rho$. Thus, $(b_r - b_{k^*})^{-\kappa} \leqslant \left(b\sum_{s=k^*}^{r-1} s^\rho\right)^{-\kappa}$. Therefore,

$$
\begin{aligned}
f^*(u) &\leqslant 2 + k_0 + \textstyle\sum_{r=k_0}^{k^*-2}\left(b\sum_{s=r}^{k^*-2} s^\rho\right)^{-\kappa} + \sum_{r=k^*+1}^{\infty}\left(b\sum_{s=k^*}^{r-1} s^\rho\right)^{-\kappa},\\
f^*(u) &\leqslant 2 + k_0 + \textstyle\sum_{r=k_0}^{k^*-2}\left(b\sum_{s=k_0}^{r} s^\rho\right)^{-\kappa} + \sum_{r=k^*+1}^{\infty}\left(b\sum_{s=k^*}^{r-1} s^\rho\right)^{-\kappa},\\
f^*(u) &\leqslant 2 + k_0 + 2\textstyle\sum_{r=k_0}^{\infty}\left(\sum_{s=k_0}^{r} s^\rho\right)^{-\kappa}.
\end{aligned}
$$

As $2+k_0+2\sum_{r=k_0}^{\infty}\left(\sum_{s=k_0}^{r} s^\rho\right)^{-\kappa}$ does not depend on u and $\left(\sum_{s=k_0}^{r} s^\rho\right)^{-\kappa} = O(r^{-(1+\rho)\kappa})$, $\lim\limits_{u\to+\infty} f^*(u)$ is finite if $(1+\rho)\kappa > 1$. As a result, f_0 is bounded.

S.2 Can the best response function be linear?

The key point of the identification analysis is to establish that $\tilde{\mathbf{Z}}^* = [\bar{\mathbf{y}}^{e*}\ \mathbf{Z}]$ is a full rank matrix, where $\bar{\mathbf{y}}^{e*} = \mathbf{G}\mathbf{y}^{e*}$. Proposition 2.2 states that $\mathbf{y}^{e*}$ verifies a fixed point equation given by $y_i^e = \sum_{r=1}^{\infty} F_{\varepsilon|\mathcal{I}}\left(\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma} - a_r\right)$. Although this equation is nonlinear, Figure S.1 shows that it can be approximated using a linear equation when $\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ is sufficiently large. Indeed, the red line of Figure S.1 represents y_i^e as a function of $\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma}$. The function looks linear when $\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ is positive and large. In this representation, I set $a_1 = 0$, and $a_r = 1$ for all $r \geqslant 2$, which corresponds to the situation where $\rho = 0$ and $\bar{R} = 1$. This configuration suggests that the cost function $c(.)$ in Equation (1) is quadratic and that the model is similar to the Tobit model under rational expectations. The best response function (BRF) of the Tobit model under rational expectations implies that $y_i^e = \mathbb{E}(\max\{0, \lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma} + \varepsilon_i\}|\mathcal{I})$, which is represented by the green line.

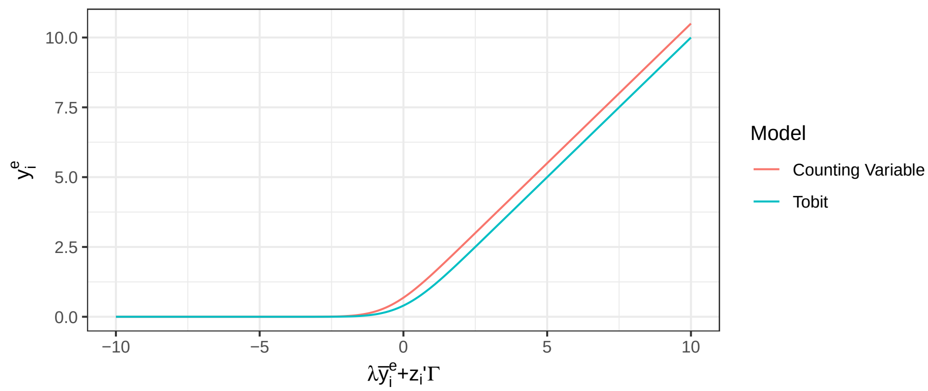

Figure S.1: Expected outcome

The slopes of both BRFs are quite similar. In an old version of the paper, I use the Poisson summation formula to show that $\sum_{r=1}^{\infty} F_{\varepsilon|\mathcal{I}}\left(\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma} - a_r\right)$ can be approximated using a linear function in $\lambda\bar{y}_i^e + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ if $\rho = 0$ and $\bar{R} = 1$

S.3 A More General Identification Analysis

In this section, I relax Assumption 3.2 and conduct a more general identification analysis. I employ Proposition 2 in Manski (1988) on the identification of the parameters in binary response models when the outcome distribution is unknown. For any vector $\mathbf{b}$, b_k denotes its k-th component. I set the following assumption.

Assumption S.1. *(i) For any i, ε_i is independent of $\mathbf{Z}$ and $\mathbf{G}$. (ii) The common density function of ε_i's is positive almost everywhere on $\mathbb{R}$.*

Assumption S.2. *There exists at least one integer $l \in [1, K]$ such that, for almost every value of $\boldsymbol{x}_{-l,i} := (x_{i,1}, \dots, x_{i,l-1}, x_{i,l+1}, \dots, x_{i,K})'$, the conditional distribution of $\beta_l x_{i,l} + \gamma_l \bar{x}_{i,l}$ given $\boldsymbol{x}_{-l,i}$ has everywhere positive density, where $x_{i,l}$ and $\bar{x}_{i,l}$ are respectively the l-th component of $\boldsymbol{x}_i$ and $\bar{\boldsymbol{x}}_i$.*

Condition (i) of Assumption S.1 imposes strict exogeneity for $\mathbf{Z}$ and $\mathbf{G}$. It implies that $F_{\varepsilon|\mathcal{I}} = F_\varepsilon$, where F_ε is the *unconditional* distribution of ε_i. Condition (ii) ensures that there is enough variation in the counting variable y to allow for the identification of the cut points. In particular, the event

$\{y_i > \bar{R}\}$ has a nonzero probability of occurrence. This is important for the identification of $\bar{\delta}$ and ρ. Assumption S.2 originates from Manski (1988) and is adapted to my framework. It imposes unbounded support for $\beta_l x_{i,l} + \gamma_l \bar{x}_{i,l}$ given $\boldsymbol{x}_{-l,i}$, which would help for identifying F_ε.

Proposition S.1. *Under Assumptions 2.1–3.1, 3.3–3.5, S.1, and S.2, λ, $\boldsymbol{\Gamma}$, $\boldsymbol{\delta}$, $\bar{\delta}$, and the distribution function F_ε are identified up to scale, whereas ρ and $\bar{R}$ are point identified.*

The only new free object in the econometric analysis is the distribution function F_ε. Therefore, given Proposition 3.1 it is sufficient to show that F_ε is identified. This result directly comes from Proposition 2 of Manski (1988). If λ, $\boldsymbol{\Gamma} = \left(\alpha, \boldsymbol{\beta}', \boldsymbol{\gamma}'\right)'$, and F_ε are observationally equivalent to $\tilde{\lambda}$, $\tilde{\boldsymbol{\Gamma}} = (\tilde{\alpha}, \tilde{\boldsymbol{\beta}}', \tilde{\boldsymbol{\gamma}}')'$, and $\tilde{F}_\varepsilon$, then $p_{i0} = \tilde{p}_{i0}$, where $p_{i0} = F_\varepsilon(\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma})$, $\tilde{p}_{i0} = \tilde{F}_\varepsilon(\tilde{\lambda} y_i^{e*} + \boldsymbol{z}_i'\tilde{\boldsymbol{\Gamma}})$.[2]

I show in Appendix A.4 that the support of $(y_i^{e*}, \boldsymbol{z}_i)$ cannot be contained in some proper linear subspace of $\mathbb{R}^{2K+2}$. Note that this result does not depend on the identification of F_ε. Even for two F_ε and $\tilde{F}_\varepsilon$, as they yield the same rational expected outcome y_i^{e*}, using the same argument of Appendix A.4, one can show that y_i^{e*} cannot be written as a linear combination of the components $\boldsymbol{z}_i$. Therefore, if $\frac{(\lambda, \boldsymbol{\Gamma}')}{\|(\lambda, \boldsymbol{\Gamma}')\|_2} \neq \frac{(\tilde{\lambda}, \tilde{\boldsymbol{\Gamma}}')}{\|(\tilde{\lambda},, \tilde{\boldsymbol{\Gamma}}')\|_2}$, where $\|.\|_2$ is the euclidean norm, then there exists a strictly positive proportion of agents i, such that

$$\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma} < 0 \leqslant \tilde{\lambda} y_i^{e*} + \boldsymbol{z}_i'\tilde{\boldsymbol{\Gamma}} \quad \text{or} \quad \tilde{\lambda} y_i^{e*} + \boldsymbol{z}_i'\tilde{\boldsymbol{\Gamma}} < 0 \leqslant \lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}. \tag{S.1}$$

This is because $\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ and $\tilde{\lambda} y_i^{e*} + \boldsymbol{z}_i'\tilde{\boldsymbol{\Gamma}}$ cannot be equal for all i and Assumption S.2 implies that $\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ can take all values in $\mathbb{R}$. To be precise, if $\beta_l x_{i,l} + \gamma_l \bar{x}_{i,l}$ increases to ∞, for $\boldsymbol{x}_{-l,i}$ set fixed, then so does y_i^{e*} and $\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}$. Alternatively if $\beta_l x_{i,l} + \gamma_l \bar{x}_{i,l}$ goes to $-\infty$, for $\boldsymbol{x}_{-l,i}$ set fixed, then so does $\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ because y_i^{e*} converges to zero.

I can now use Proposition 2 of Mansky (1988). Under Equation (S.1), the proposition implies that λ, and $\boldsymbol{\Gamma}$ are identified with respect to $\tilde{\lambda}$, and $\tilde{\boldsymbol{\Gamma}}$, i.e., $\frac{(\lambda, \boldsymbol{\Gamma}')}{\|(\lambda, \boldsymbol{\Gamma}')\|_2} = \frac{(\tilde{\lambda}, \tilde{\boldsymbol{\Gamma}}')}{\|(\tilde{\lambda},, \tilde{\boldsymbol{\Gamma}}')\|_2}$. As a result $(\lambda, \boldsymbol{\Gamma}')$ is identified up to scale. Assumption S.1 guarantees the quantile independence required for this proposition. By replacing $(\lambda, \boldsymbol{\Gamma}')$ with $(\tilde{\lambda}, \tilde{\boldsymbol{\Gamma}}')$ in the formula of $\tilde{p}_{i0}$, the condition $p_{i0} = \tilde{p}_{i0}$ would imply that $F_\varepsilon(\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}) = \tilde{F}_\varepsilon(\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma})$. It follows that F_ε is identified up to scale because $\lambda y_i^{e*} + \boldsymbol{z}_i'\boldsymbol{\Gamma}$ can take all values in $\mathbb{R}$. The rest of the proof is similar to Appendix A.4. The parameters λ, $\boldsymbol{\Gamma}$, $\boldsymbol{\delta}$, $\bar{\delta}$ are identified up to scale, whereas ρ and $\bar{R}$ are point identified. Indeed, ρ and $\bar{R}$ are point identified because the scale of F_ε has no impact on them.

[2] Again, the expected outcome y_i^{e*} is necessarily the same for both sets of parameters because $p_{ir} = \tilde{p}_{ir}$.

S.4 Variance of the NPL estimator

I assume that $\hat{\bar{R}} \geqslant \bar{R}^0$. By Proposition 3.2, this implies that $\hat{\boldsymbol{\theta}}(\hat{\bar{R}}) = \boldsymbol{\theta}^0$ and $\hat{\mathbf{y}}^e_n(\hat{\bar{R}}) = \mathbf{y}^{e*}$. I recall that $\boldsymbol{\theta} = \left(\log(\lambda),\ \boldsymbol{\Gamma}', \log(\tilde{\boldsymbol{\delta}}'), \log(\bar{\delta}), \log(\rho)\right)'$. Let $\phi_{i,r} = \phi\left(\lambda\mathbf{g}_i\mathbf{y}^e + \boldsymbol{z}_i'\boldsymbol{\Gamma} - a_r\right)$, $\Phi_{i,r} = \Phi\left(\lambda\mathbf{g}_i\mathbf{y}^e + \boldsymbol{z}_i'\boldsymbol{\Gamma} - a_r\right)$, $\Delta\phi_{i,r} = \phi_{i,r} - \phi_{i,r-1}$, and $\Delta\Phi_{i,r} = \Phi_{i,r} - \Phi_{i,r-1}$ for any $r \geqslant 1$, where $\bar{R} = \bar{R}^0$, $\boldsymbol{\theta} = \boldsymbol{\theta}^0$, and $\mathbf{y}^e = \mathbf{y}^{e*}$. I have

$$\nabla_{\log(\lambda)}\mathcal{L}_n(\boldsymbol{\theta}^0, \mathbf{y}^{e*}) = \lambda\sum_{i=1}^n\sum_{r=0}^\infty d_{ir}\frac{\Delta(\phi_{i,r+1}\check{z}_{i,r})}{\Delta\Phi_{i,r+1}}, \text{ where } \check{z}_{i,r} = \mathbf{g}_i\mathbf{y}^e - r,$$

$$\nabla_{\boldsymbol{\Gamma}}\mathcal{L}_n(\boldsymbol{\theta}^0, \mathbf{y}^{e*}) = \sum_{i=1}^n\sum_{r=0}^\infty d_{ir}\frac{\Delta\phi_{i,r+1}}{\Delta\Phi_{i,r+1}}\boldsymbol{z}_i,$$

$$\nabla_{\log(\tilde{\delta}_k)}\mathcal{L}_{n,i}(\boldsymbol{\theta}^0, \mathbf{y}^{e*}) = -\tilde{\delta}_k\sum_{i=1}^n\sum_{r=k-1}^\infty d_{ir}\frac{\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \tilde{\delta}_k\sum_{i=1}^n\sum_{r=k}^\infty d_{ir}\frac{\phi_{i,r}}{\Delta\Phi_{i,r+1}} \text{ for } 2 \leqslant k \leqslant \bar{R},$$

$$\nabla_{\log(\bar{\delta})}\mathcal{L}_{n,i}(\boldsymbol{\theta}^0, \mathbf{y}^{e*}) = -\bar{\delta}\sum_{i=1}^n\sum_{r=\bar{R}}^\infty d_{ir}\frac{\dot{a}_{\delta,r+1}\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \bar{\delta}\sum_{i=1}^n\sum_{r=\bar{R}+1}^\infty d_{ir}\frac{\dot{a}_{\delta,r}\phi_{i,r}}{\Delta\Phi_{i,r+1}},$$

$$\nabla_{\log(\rho)}\mathcal{L}_n(\boldsymbol{\theta}^0, \mathbf{y}^{e*}) = -\rho\sum_{i=1}^n\sum_{r=\bar{R}}^\infty d_{ir}\frac{\dot{a}_{\rho,r+1}\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \rho\sum_{i=1}^n\sum_{r=\bar{R}+1}^\infty d_{ir}\frac{\dot{a}_{\rho,r}\phi_{i,r}}{\Delta\Phi_{i,r+1}},$$

where $\dot{a}_{\delta,r} = \sum_{k=\bar{R}+1}^r(k-1)^\rho$ and $\dot{a}_{\rho,r} = \bar{\delta}\sum_{k=\bar{R}+1}^r(k-1)^\rho\log(k-1)$ for $r \geqslant \bar{R}+1$.

I define the following notations: $\mathbf{A}_i^{\lambda\lambda} = \lambda^2\sum_{r=0}^\infty\dfrac{\phi_{i,r+1}^2\check{z}_{i,r}^2 - 2\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r}\check{z}_{i,r-1} + \phi_{i,r}^2\check{z}_{i,r-1}^2}{\Delta\Phi_{i,r+1}}$,

$$\mathbf{A}_i^{\Gamma\Gamma} = \sum_{r=0}^\infty\frac{\phi_{i,r+1}^2 - 2\phi_{i,r}\phi_{i,r+1} + \phi_{i,r}^2}{\Delta\Phi_{i,r+1}},$$

$$\mathbf{A}_i^{\delta_k\delta_k} = \tilde{\delta}_k^2\Big(\sum_{r=k-1}^\infty\frac{\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} - 2\sum_{r=k}^\infty\frac{\phi_{i,r}\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \sum_{r=k}^\infty\frac{\phi_{i,r}^2}{\Delta\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\bar{\delta}\bar{\delta}} = \bar{\delta}^2\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\delta,r+1}^2\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} - 2\sum_{r=\bar{R}+1}^\infty\frac{a_{\delta,r}a_{\delta,r+1}\phi_{i,r}\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{a_{\delta,r}^2\phi_{i,r}^2}{\Delta\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\rho\rho} = \rho^2\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\rho,r+1}^2\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} - 2\sum_{r=\bar{R}+1}^\infty\frac{a_{\rho,r}a_{\rho,r+1}\phi_{i,r}\phi_{i,r+1}}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{a_{\rho,r}^2\phi_{i,r}^2}{\Delta\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\lambda\Gamma} = \lambda\sum_{r=0}^\infty\frac{\phi_{i,r+1}^2\check{z}_{i,r} - \phi_{i,r+1}\phi_{i,r}\check{z}_{i,r-1} - \phi_{i,r}\phi_{i,r+1}\check{z}_{i,r} + \phi_{i,r}^2\check{z}_{i,r-1}}{\Delta\Phi_{i,r+1}},$$

$$\mathbf{A}_i^{\lambda\delta_k} = \lambda\tilde{\delta}_k\Big(\sum_{r=k-1}^\infty\frac{\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r-1} - \phi_{i,r+1}^2\check{z}_{i,r}}{\Delta\Phi_{i,r+1}} + \sum_{r=k}^\infty\frac{\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r} - \phi_{i,r}^2\check{z}_{i,r-1}}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\lambda\bar{\delta}} = \lambda\bar{\delta}\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\delta,r+1}\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r-1} - \dot{a}_{\delta,r+1}\phi_{i,r+1}^2\check{z}_{i,r}}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{\dot{a}_{\delta,r}\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r} - \dot{a}_{\delta,r}\phi_{i,r}^2\check{z}_{i,r-1}}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\lambda\rho} = \lambda\rho\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\rho,r+1}\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r-1} - \dot{a}_{\rho,r+1}\phi_{i,r+1}^2\check{z}_{i,r}}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{\dot{a}_{\rho,r}\phi_{i,r}\phi_{i,r+1}\check{z}_{i,r} - \dot{a}_{\rho,r}\phi_{i,r}^2\check{z}_{i,r-1}}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\Gamma\delta_k} = \tilde{\delta}_k\Big(\sum_{r=k-1}^\infty\frac{\phi_{i,r}\phi_{i,r+1} - \phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} + \sum_{r=k}^\infty\frac{\phi_{i,r}\phi_{i,r+1} - \phi_{i,r}^2}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\Gamma\bar{\delta}} = \bar{\delta}\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\delta,r+1}\phi_{i,r}\phi_{i,r+1} - \dot{a}_{\delta,r+1}\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{\dot{a}_{\delta,r}\phi_{i,r}\phi_{i,r+1} - \dot{a}_{\delta,r}\phi_{i,r}^2}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\Gamma\rho} = \rho\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\rho,r+1}\phi_{i,r}\phi_{i,r+1} - \dot{a}_{\rho,r+1}\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} + \sum_{r=\bar{R}+1}^\infty\frac{\dot{a}_{\rho,r}\phi_{i,r}\phi_{i,r+1} - \dot{a}_{\rho,r}\phi_{i,r}^2}{\Phi_{i,r+1}}\Big),$$

$$\mathbf{A}_i^{\delta_k\delta_{k'}} = -\tilde{\delta}_k\mathbf{A}_i^{\Gamma\delta_k'} \text{ for } 2 \leqslant k < k' \leqslant \bar{R}, \quad \mathbf{A}_i^{\delta_k\bar{\delta}} = -\tilde{\delta}_k\mathbf{A}_i^{\Gamma\bar{\delta}}, \quad \mathbf{A}_i^{\delta_k\rho} = -\tilde{\delta}_k\mathbf{A}_i^{\Gamma\rho},$$

$$\mathbf{A}_i^{\bar{\delta}\rho} = \bar{\delta}\rho\Big(\sum_{r=\bar{R}}^\infty\frac{\dot{a}_{\delta,r+1}\dot{a}_{\rho,r+1}\phi_{i,r+1}^2}{\Delta\Phi_{i,r+1}} - \sum_{r=\bar{R}+1}^\infty\frac{\dot{a}_{\delta,r}\dot{a}_{\rho,r+1}\phi_{i,r}\phi_{i,r+1} + \dot{a}_{\delta,r+1}\dot{a}_{\rho,r}\phi_{i,r}\phi_{i,r+1} - \dot{a}_{\delta,r}\dot{a}_{\rho,r}\phi_{i,r}^2}{\Delta\Phi_{i,r+1}}\Big).$$

Let $\mathbf{\Sigma}_{n,i} := \mathbb{V}(\nabla_{\boldsymbol{\theta}}\mathcal{L}_{n,i}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})|\boldsymbol{\chi}_n)$. It follows that

$$\mathbf{\Sigma}_{n,i} = -\begin{pmatrix} \mathbf{A}_i^{\lambda\lambda} & \mathbf{A}_i^{\lambda\Gamma}\boldsymbol{z}_i' & \mathbf{A}_i^{\lambda\delta_2} & \dots & \mathbf{A}_i^{\lambda\rho} \\ \mathbf{A}_i^{\lambda\Gamma}\boldsymbol{z}_i & \mathbf{A}_i^{\Gamma\Gamma}\boldsymbol{z}_i\boldsymbol{z}_i' & \mathbf{A}_i^{\Gamma\delta_2}\boldsymbol{z}_i & \dots & \mathbf{A}_i^{\Gamma\rho}\boldsymbol{z}_i \\ \mathbf{A}_i^{\lambda\delta_2} & \mathbf{A}_i^{\Gamma\delta_2}\boldsymbol{z}_i' & \mathbf{A}_i^{\delta_2\delta_2} & \dots & \mathbf{A}_i^{\delta_2\rho} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_i^{\lambda\rho} & \mathbf{A}_i^{\Gamma\rho}\boldsymbol{z}_i' & \mathbf{A}_i^{\delta_2\rho} & \dots & \mathbf{A}_i^{\rho\rho}. \end{pmatrix}$$

By the law of large numbers (LLN), $\mathbf{\Sigma}_0$ is the limit of $(1/n)\sum_{i=1}^n \mathbf{\Sigma}_{n,i}$ as n grows to infinity.

On the other hand, I have

$\mathbf{H}_{1,n} := \nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}\mathcal{L}_n(\dot{\boldsymbol{\theta}}_n(\hat{\bar{R}}), \dot{\mathbf{y}}_n^e(\hat{\bar{R}}))$ and $\mathbf{H}_{2,n} := \nabla_{\boldsymbol{\theta}\mathbf{y}^{e\prime}}\mathcal{L}_n(\dot{\boldsymbol{\theta}}_n(\hat{\bar{R}}), \dot{\mathbf{y}}_n^e(\hat{\bar{R}}))\nabla_{\boldsymbol{\theta}'}\dot{\mathbf{y}}^{e\prime}(\hat{\bar{R}})$ for some point $\dot{\boldsymbol{\theta}}_n(\hat{\bar{R}})$ between $\hat{\boldsymbol{\theta}}_n(\hat{\bar{R}})$ and $\boldsymbol{\theta}^0$, such that $\dot{\mathbf{y}}_n^e(\hat{\bar{R}}) = \mathbf{L}(\dot{\boldsymbol{\theta}}_n(\hat{\bar{R}}), \dot{\mathbf{y}}_n^e(\hat{\bar{R}}))$. As $\hat{\boldsymbol{\theta}}_n(\hat{\bar{R}})$ converges to $\boldsymbol{\theta}^0$, by the LLN, $\mathbf{H}_{1,n} \to \mathbf{H}_{1,0} := \lim_{n\to\infty}(1/n)\sum_{i=1}^n \mathbb{E}(\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}\mathcal{L}_{i,n}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})|\boldsymbol{\chi}_n)$ and

$\mathbf{H}_{2,n} \to \mathbf{H}_{2,0} := \lim_{n\to\infty}(1/n)\sum_{i=1}^n \mathbb{E}(\nabla_{\boldsymbol{\theta}\mathbf{y}^{e\prime}}\mathcal{L}_{i,n}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})|\boldsymbol{\chi}_n)\nabla_{\boldsymbol{\theta}'}\mathbf{y}^{e*}$.

On can verify that $\mathbb{E}(\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}'}\mathcal{L}_{i,n}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})|\boldsymbol{\chi}_n) = -\mathbf{\Sigma}_{n,i}$. Thus, $\mathbf{H}_{1,0} = -\mathbf{\Sigma}_0$.

Besides, $\mathbb{E}(\nabla_{\boldsymbol{\theta}\mathbf{y}^{e\prime}}\mathcal{L}_{i,n}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})|\boldsymbol{\chi}_n) = \lambda\big(\mathbf{A}_i^{\lambda\Gamma},\ \mathbf{A}_i^{\Gamma\Gamma}\boldsymbol{z}_i',\ \mathbf{A}_i^{\Gamma\delta_2},\ \dots,\ \mathbf{A}_i^{\Gamma\rho}\big)'\mathbf{g}_i$.

$\nabla_{\boldsymbol{\theta}}\mathbf{y}^{e*}$ can be computed using the implicit definition of $\mathbf{y}^{e*}$; that is $\mathbf{y}^{e*} = \mathbf{L}(\boldsymbol{\theta}^0, \mathbf{y}^{e*})$. This implies that $\nabla_{\boldsymbol{\theta}}\mathbf{y}^{e*} = \mathbf{S}^{-1}\mathbf{B}$, where $\mathbf{S} = \mathbf{I}_n - \lambda\mathbf{D}\mathbf{G}$, $\mathbf{I}_n$ is the identity matrix of dimension n, $\mathbf{D} = \operatorname{diag}\big(\sum_{r=1}^{\infty}\phi_{1,r},\ \dots,\ \sum_{r=1}^{\infty}\phi_{n,r}\big)$, and $\mathbf{B} = \big(\mathbf{B}^1,\ \mathbf{D}\mathbf{Z},\ \mathbf{B}^2\big)$. The component $\mathbf{B}^1$ is an n-vector whose i-th element is $\lambda\sum_{r=1}^{\infty}\phi_{i,r}\check{z}_{i,r-1}$. The component $\mathbf{B}^2$ is a matrix of n rows, where the i-th row is $\mathbf{B}_i^2 = \Big(-\tilde{\delta}_2\sum_{r=2}^{\infty}\phi_{i,r},\ \dots, -\tilde{\delta}_{\bar{R}}\sum_{r=\bar{R}}^{\infty}\phi_{i,r},\ -\bar{\delta}\sum_{r=\bar{R}+1}^{\infty}\phi_{i,r}\dot{a}_{\delta,r},\ -\rho\sum_{r=\bar{R}+1}^{\infty}\phi_{i,r}\dot{a}_{\rho,r}\Big)$.

I assume that $\lim_{n\to\infty}\dfrac{\lambda}{n}\sum_{i=1}^n\big(\mathbf{A}_i^{\lambda\Gamma},\ \mathbf{A}_i^{\Gamma\Gamma}\boldsymbol{z}_i',\ \mathbf{A}_i^{\Gamma\delta_2},\ \dots,\ \mathbf{A}_i^{\Gamma\rho}\big)'\mathbf{g}_i\mathbf{S}^{-1}\mathbf{B}$ exists and is equal to $\mathbf{\Omega}_0$. As a result, $\mathbf{H}_{1,0} = -\mathbf{\Sigma}_0$, $\mathbf{H}_{2,0} = \mathbf{\Omega}_0$, where $\mathbf{\Sigma}_0$ is the limit of $(1/n)\sum_i^n \mathbf{\Sigma}_{n,i}$ as n grows to infinity.

S.5 Asymptotics in the case of endogenous networks

Let $\hat{\boldsymbol{\mu}}_n = (\hat{\mu}_1, \dots, \hat{\mu}_n)'$, $\hat{\boldsymbol{\nu}}_n = (\hat{\nu}_1, \dots, \hat{\nu}_n)'$, $\bar{\hat{\mu}}_i = \sum_{j=1}^n \mathbf{g}_{ij}\hat{\mu}_i$, $\bar{\hat{\nu}}_i = \sum_{j=1}^n \mathbf{g}_{ij}\hat{\nu}_i$, $\hat{\boldsymbol{\chi}}_n = (\hat{\boldsymbol{\mu}}_n',\ \hat{\boldsymbol{\nu}}_n')'$, and $\boldsymbol{\chi}_n^0$ be the true value of $((\mu_1, \dots, \mu_n)',\ (\nu_1, \dots, \nu_n)')'$. As I have new regressors that are estimated, I define the following notations. For any $\bar{R}$, $\boldsymbol{\theta}^*(\bar{R})$ is the vector of new parameters to be estimated. $\mathbf{\Theta}^*(\bar{R})$ is the space of $\boldsymbol{\theta}^*(\bar{R})$. The mapping $\mathbf{L}$ is redefined as $\mathbf{L}^*(\boldsymbol{\theta}^*, \mathbf{y}^e) = \sum_{r=0}^{\infty}\Phi(\lambda\mathbf{g}_i\mathbf{y}^e + \boldsymbol{z}_i'\mathbf{\Gamma} + \hat{h}_{\psi,i} - a_r)$, where $\hat{h}_{\psi,i} = \sum_k^T(\theta_{1,k}\hat{\mu}_i^k + \theta_{2,k}\hat{\nu}_i^k + \theta_{3,k}\bar{\hat{\mu}}_i^k + \theta_{4,k}\bar{\hat{\nu}}_i^k)$ is assumed to be a consistent approximation of $h_{\psi}(\mu_i, \nu_i, \bar{\mu}_i, \bar{\nu}_i)$. $\mathcal{L}_{n,i}^*(\boldsymbol{\theta}, \mathbf{y}^e) = \sum_{r=0}^{\infty} d_{ir}\log(\Phi(\lambda\mathbf{g}_i\mathbf{y}^e + \boldsymbol{z}_i'\mathbf{\Gamma} + \hat{h}_{\psi,i} - a_r) - \Phi(\lambda\mathbf{g}_i\mathbf{y}^e + \boldsymbol{z}_i'\mathbf{\Gamma} + \hat{h}_{\psi,i} - a_{r+1}))$. $\mathcal{L}_0^*(\boldsymbol{\theta}, \mathbf{y}^e) = \mathbb{E}(\mathcal{L}_n^*(\boldsymbol{\theta}, \mathbf{y}^e)|\boldsymbol{\chi}_n, \hat{\boldsymbol{\chi}}_n)$, $\tilde{\boldsymbol{\theta}}^*(\mathbf{y}^e, \bar{R}) = \arg\max_{\boldsymbol{\theta}\in\mathbf{\Theta}(\bar{R})}\mathcal{L}_0^*(\boldsymbol{\theta}, \mathbf{y}^e)$, $\boldsymbol{\phi}_0^*(\mathbf{y}^e, \bar{R}) = \mathbf{L}^*(\tilde{\boldsymbol{\theta}}^*(\mathbf{y}^e, \bar{R}), \mathbf{y}^e)$, and $\mathcal{A}_0^*(\bar{R}) = \{(\boldsymbol{\theta}^*, \mathbf{y}^e) \in \mathbf{\Theta}^*(\bar{R}) \times [0, \overset{o}{y}]^n$, such that $\boldsymbol{\theta}^* = \tilde{\boldsymbol{\theta}}_0^*(\mathbf{y}^e, \bar{R})$ and $\mathbf{y}^e = \boldsymbol{\phi}_0^*(\mathbf{y}^e, \bar{R})\}$. Let also $\boldsymbol{\theta}^{*0}$ be the true value of $\boldsymbol{\theta}^*$ and $\mathbf{y}_{\boldsymbol{\chi}}^{e*0} \in \mathbb{R}^n$, such that $\mathbf{y}_{\boldsymbol{\chi}}^{e*0} = \mathbf{L}^*(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$.

S.5.1 Consistency

Under Assumptions A.1–A.2 adapted to the new regressors and $\mathbf{\Theta}^*(\bar{R})$, Results A.1–A.2 can be extended to the new pseudo-likelihood. Thus, $\mathcal{L}_n^*$ uniformly converges to $\mathcal{L}_0^*$. Moreover, $\mathcal{L}_0^*$ has a unique maximizer $(\check{\boldsymbol{\theta}}_n^*(\bar{R}), \check{\mathbf{y}}_n^{e*}(\bar{R}))$, such that $\check{\mathbf{y}}_n^{e*}(\bar{R})) = \mathbf{L}(\check{\boldsymbol{\theta}}_n^*(\bar{R}), \check{\mathbf{y}}_n^{e*}(\bar{R})))$. As for the case of exogenous network, under Assumptions A.1–A.2 and Assumptions A.3–A.5 adapted to the new maximizer $(\check{\boldsymbol{\theta}}_n^*(\bar{R}), \check{\mathbf{y}}_n^{e*}(\bar{R}))$, the new NPL estimator $\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$ converges in probability to $\check{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$. However, Gibbs' inequality cannot be applied because $h_\psi(\mu_i, \nu_i, \bar{\mu}_i, \bar{\nu}_i)$ is replaced by its estimator. However, as the estimator is assumed to be consistent, Gibbs' inequality can be applied as n grows to infinity. It then follows that $\lim\limits_{n\to\infty} \check{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}) = \boldsymbol{\theta}^{*0}$ if $\hat{\bar{R}} \geqslant \bar{R}^0$. As a result, the NPL estimator converges to $\boldsymbol{\theta}^{*0}$ if $\bar{R} \geqslant \bar{R}^0$.

S.5.2 Asymptotic normality

I assume for simplicity that $\hat{\bar{R}} \geqslant \bar{R}^0$.

By applying the mean value theorem (MVT) to $\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^*(\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}), \hat{\mathbf{y}}_n^{e*}(\hat{\bar{R}}))$ between $\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$ and $\boldsymbol{\theta}^{*0}$, I get

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}) - \boldsymbol{\theta}^{*0}) = -\left(\mathbf{H}_{1,n}^* + \mathbf{H}_{2,n}^*\right)^{-1} \sqrt{n}\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^*(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0}) \tag{S.2}$$

for some point $\dot{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$ between $\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$ and $\check{\boldsymbol{\theta}}_n^*(\hat{\bar{R}})$, where $\mathbf{H}_{1,n}^* := \nabla_{\boldsymbol{\theta}^*\boldsymbol{\theta}^{*\prime}}\mathcal{L}_n(\dot{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}), \dot{\mathbf{y}}_n^{e*}(\hat{\bar{R}}))$, $\mathbf{H}_{2,n}^* := \nabla_{\boldsymbol{\theta}^*\mathbf{y}^{e\prime}}\mathcal{L}_n(\dot{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}), \dot{\mathbf{y}}_n^{*e}(\hat{\bar{R}}))\nabla_{\boldsymbol{\theta}^{*\prime}}\dot{\mathbf{y}}^{*e\prime}(\hat{\bar{R}})$, and $\dot{\mathbf{y}}_n^{*e}(\hat{\bar{R}}) = \mathbf{L}^*(\dot{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}), \dot{\mathbf{y}}_n^{*e}(\hat{\bar{R}}))$.

I apply the MVT a second time to $\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_{n,i}^*(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ between $\hat{\boldsymbol{\chi}}_n$ and $\boldsymbol{\chi}_n^0$. I have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}) - \boldsymbol{\theta}^{*0}) = -\left(\mathbf{H}_{1,n}^* + \mathbf{H}_{2,n}^*\right)^{-1}\left(\sqrt{n}\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0}) + n^{-1/2}\textstyle\sum_{i=1}^n \mathcal{E}_{i,n}\Delta\hat{\boldsymbol{\chi}}_n\right), \tag{S.3}$$

where $\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ is the value of $\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^*(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ when $\hat{\boldsymbol{\mu}}_n$ and $\hat{\boldsymbol{\nu}}_n$ are equal to their true values, $\Delta\hat{\boldsymbol{\chi}}_n = \hat{\boldsymbol{\chi}}_n - \boldsymbol{\chi}_n^0$, and $\mathcal{E}_{i,n}$ is the derivative of $\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_{n,i}^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ with respect to $\hat{\boldsymbol{\chi}}_n'$, applied to some point between $\hat{\boldsymbol{\chi}}_n$ and $\boldsymbol{\chi}_n^0$. I set the following regulatory condition.

Assumption S.1. *$n^{-1/2}\sum_{i=1}^n \mathcal{E}_{i,n}\Delta\hat{\boldsymbol{\chi}}_n$ is $o_p(1)$.*

A similar assumption is also set in the case of the control function approach.[3] Assumption S.1 requires $\Delta\hat{\boldsymbol{\chi}}_n$ to converge to zero at some rate. Yan et al. (2019) show that $\|\Delta\hat{\boldsymbol{\chi}}_n\|_\infty = O_p\left((\log(n)/n)^{1/2}\right)$. Thus, a sufficient condition for this assumption to hold is that $(1/n)\|\sum_{i=1}^n \mathcal{E}_{i,n}\|_\infty = o_p(\zeta_n^{1/2})$, where $\zeta_n\log(n)$ converges to zero as n grows to infinity. This condition is realistic because, if $\mathcal{E}_{i,n}$'s were independent across i, then $(1/n)\|\sum_{i=1}^n \mathcal{E}_{i,n}\|_\infty = o_p(\zeta_n^{1/2})$ would be true for any $\zeta_n = n^{-\zeta}$ where $\zeta \in (0,\ 1)$. Therefore, if the dependence is not too strong, Assumption S.1 will be verified.

[3] See the Lipschitz condition set in Assumption 8 of Johnsson, I. and H. R. Moon (2021): "Estimation of peer effects in endogenous social networks: control function approach," *Review of Economics and Statistics*, 103, 328–345.

Under Assumption S.1, Equation (S.3) implies that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}) - \boldsymbol{\theta}^{*0}) = -\left(\mathbf{H}_{1,n}^* + \mathbf{H}_{2,n}^*\right)^{-1}\sqrt{n}\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0}) + o_p(1).$$

The CLT can be applied to $\sqrt{n}\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ since $\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$ is a sum of n independent bounded variables conditionally on $\boldsymbol{\chi}_n$. It then follows that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n^*(\hat{\bar{R}}) - \boldsymbol{\theta}^{*0}) \xrightarrow{d} \mathcal{N}(0, \left(\mathbf{H}_{1,0}^* + \mathbf{H}_{2,0}^*\right)^{-1}\boldsymbol{\Sigma}_0^*\left(\mathbf{H}_{1,0}^{*\prime} + \mathbf{H}_{2,0}^{*\prime}\right)^{-1}), \tag{S.4}$$

where $\mathbf{H}_{1,0}^*$ and $\mathbf{H}_{2,0}^*$ are the limits of $\mathbf{H}_{1,n}^*$ and $\mathbf{H}_{2,n}^*$ as n grows to infinity, and $\boldsymbol{\Sigma}_0^*$ is a consistent estimator of the variance of $\sqrt{n}\nabla_{\boldsymbol{\theta}^*}\mathcal{L}_n^{*0}(\boldsymbol{\theta}^{*0}, \mathbf{y}_{\boldsymbol{\chi}}^{e*0})$.

S.6 Marginal effects

The parameters of the counting variable model cannot be interpreted directly. Policymakers are interested in the marginal effect of the explanatory variables on the expected outcome.
Let us recall that $\boldsymbol{\theta} = \left(\log(\lambda),\ \boldsymbol{\Gamma}', \log(\tilde{\boldsymbol{\delta}}'), \log(\bar{\delta}), \log(\rho)\right)'$, where $\tilde{\boldsymbol{\delta}} = (\tilde{\delta}_2, \dots, \tilde{\delta}_{\bar{R}})$, and $\tilde{\delta}_r = \delta_r - \lambda$. Let $\tilde{\boldsymbol{z}}_i' = (\mathbf{g}_i\mathbf{y}^e,\ \boldsymbol{z}_i')$ and $\boldsymbol{\Lambda} = (\lambda,\ \boldsymbol{\Gamma}')'$. For any $k = 1, \dots, \dim(\boldsymbol{\Lambda})$, let λ_k and $\tilde{z}_{ik}$ be the k-th component in $\boldsymbol{\Lambda}$ and $\tilde{\boldsymbol{z}}_i$, respectively. The marginal effect of the explanatory variable $\tilde{z}_{ik}$ on y_i^e is given by

$$\delta_{ik}(\boldsymbol{\theta}) = \frac{\partial y_i^e}{\partial \tilde{z}_{ik}} = \lambda_k \textstyle\sum_{r=1}^{\infty} \phi\left(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r\right). \tag{S.5}$$

Using the Delta method, I show that

$$\frac{1}{n}\textstyle\sum_{i=1}^n \delta_{ik}(\hat{\boldsymbol{\theta}}_n(\hat{\bar{R}})) \overset{a}{\sim} \mathcal{N}\left(\delta_{ik}(\boldsymbol{\theta}_0), \mathbf{Q}_0^*\,\mathbb{V}(\hat{\boldsymbol{\theta}}_n(\hat{\bar{R}})|\boldsymbol{\chi}_n)\mathbf{Q}_0^{*\prime}\right),$$

where $\mathbf{Q}_0^* = (1/n)\sum_{i=1}^n \nabla_{\boldsymbol{\theta}'}\delta_{ik}(\boldsymbol{\theta}_0)$,
$\nabla_{\log(\lambda)}\delta_{ik}(\boldsymbol{\theta}) = \mathbf{1}(k=1)\lambda\sum_{r=1}^{\infty}\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r) - \lambda\lambda_k\sum_{r=1}^{\infty}(\mathbf{g}_i\mathbf{y}^e - r)(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)$,
$\nabla_{\boldsymbol{\Gamma}'}\delta_{ik}(\boldsymbol{\theta}) = \mathbf{e}_k\sum_{r=1}^{\infty}\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r) - \lambda_k\boldsymbol{z}_i'\sum_{r=1}^{\infty}(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)$, where $\mathbf{e}_k$ is a $\dim(\boldsymbol{\Gamma})$-dimensional row vector with zero everywhere except the $(k-1)$-th term which equals one if $k \geqslant 2$,
$\nabla_{\log(\delta_l)}\delta_{ik}(\boldsymbol{\theta}) = \tilde{\delta}_l\lambda_k\sum_{r=l}^{\infty}(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r) \quad$ for $2 \leqslant l < \bar{R}$,
$\nabla_{\log(\bar{\delta})}\delta_{ik}(\boldsymbol{\theta}) = \bar{\delta}\lambda_k\sum_{r=\bar{R}+1}^{\infty}\dot{a}_{\delta,r}(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)$,
$\nabla_{\log(\rho)}\delta_{ik}(\boldsymbol{\theta}) = \rho\lambda_k\sum_{r=l}^{\infty}\dot{a}_{\rho,r}(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)\phi(\tilde{\boldsymbol{z}}_i'\boldsymbol{\Lambda} - a_r)$.

S.7 Empirical results controlling for network endogeneity

Table S.1: Empirical results controlling for network endogeneity (NE). Unobserved attributes are treated as random effects and estimated using a Bayesian approach

	Count data model			Tobit model		
Parameters	**Coef.**	**Marginal effects**		**Coef.**	**Marginal effects**	
λ	0.046	0.084	(0.020)	0.307	0.249	(0.019)
Own effects						
Age	−0.044	−0.080	(0.008)	−0.045	−0.037	(0.005)
Male	−0.163	−0.298	(0.016)	−0.265	−0.215	(0.011)
Hispanic	−0.006	−0.012	(0.025)	0.116	0.094	(0.017)
Race						
Black	0.208	0.381	(0.031)	0.444	0.360	(0.02)
Asian	0.209	0.383	(0.034)	0.660	0.535	(0.023)
Other	0.027	0.050	(0.028)	0.137	0.111	(0.018)
Years at school	0.030	0.054	(0.007)	0.089	0.072	(0.005)
With both par.	0.073	0.134	(0.019)	0.146	0.118	(0.012)
Mother educ.						
<High	−0.052	−0.094	(0.023)	−0.033	−0.027	(0.015)
>High	0.202	0.369	(0.019)	0.382	0.310	(0.013)
Missing	0.028	0.052	(0.031)	0.211	0.172	(0.021)
Mother job						
Professional	0.134	0.245	(0.024)	0.235	0.191	(0.016)
Other	0.035	0.064	(0.020)	0.053	0.043	(0.013)
Missing	−0.041	−0.074	(0.028)	−0.069	−0.056	(0.019)
μ^1	0.480	0.879	(0.110)	0.917	0.744	(0.073)
μ^2	−1.065	−1.949	(0.428)	−2.122	−1.721	(0.286)
μ^3	−0.705	−1.290	(1.608)	1.050	0.851	(1.080)
μ^4	1.121	2.052	(2.088)	0.615	0.499	(1.404)
ν^1	0.286	0.524	(0.083)	0.338	0.274	(0.056)
ν^2	0.668	1.224	(0.348)	1.810	1.468	(0.232)
ν^3	−0.551	−1.009	(1.504)	0.606	0.492	(1.009)
ν^4	0.575	1.053	(1.736)	0.684	0.555	(1.170)
Contextual effects						
Age	−0.021	−0.039	(0.004)	−0.075	−0.061	(0.003)
Male	−0.050	−0.092	(0.030)	−0.044	−0.035	(0.019)
Hispanic	−0.067	−0.123	(0.042)	−0.077	−0.062	(0.027)
Race						
Black	0.059	0.109	(0.039)	0.016	0.013	(0.026)
Asian	−0.011	−0.020	(0.052)	−0.238	−0.193	(0.035)
Other	−0.101	−0.185	(0.052)	−0.254	−0.206	(0.035)
Years at school	0.016	0.030	(0.011)	−0.007	−0.006	(0.007)
With both par.	0.134	0.246	(0.036)	0.164	0.133	(0.024)
Mother educ.						
<High	−0.116	−0.213	(0.043)	−0.160	−0.130	(0.028)
>High	0.168	0.307	(0.038)	0.197	0.160	(0.025)
Missing	−0.056	−0.102	(0.060)	−0.132	−0.107	(0.040)
Mother job						
Professional	0.165	0.302	(0.048)	0.221	0.180	(0.031)
Other	0.029	0.053	(0.038)	0.019	0.016	(0.025)
Missing	−0.014	−0.026	(0.054)	0.025	0.020	(0.036)
μ^1	0.022	0.041	(0.203)	−0.200	−0.162	(0.135)
μ^2	−0.638	−1.169	(0.894)	−0.889	−0.721	(0.597)
μ^3	−1.894	−3.468	(4.116)	−2.241	−1.818	(2.752)
μ^4	0.967	1.771	(5.839)	0.871	0.706	(3.906)
ν^1	0.597	1.092	(0.171)	0.906	0.735	(0.114)
ν^2	0.705	1.290	(1.099)	0.881	0.715	(0.733)
ν^3	−0.739	−1.354	(4.220)	−1.183	−0.960	(2.821)
ν^4	−0.270	−0.494	(4.465)	−1.191	−0.966	(2.986)
σ				2.420		

For the count data model, $\hat{\bar{R}} = 12$. The estimates of $\delta_2, \ldots, \delta_{\hat{\bar{R}}}$ are 1.555, 0.523, 0.452, 0.385, 0.320, 0.264, 0.218, 0.174, 0.130, 0.100, 0.086. The estimate of $\bar{\delta}$ is $1.2e^{-5}$.

Table S.2: Empirical results controlling for NE. Unobserved attributes are treated as fixed effects

	Count data model			**Tobit model**		
Parameters	**Coef.**	**Marginal effects**		**Coef.**	**Marginal effects**	
λ	0.046	0.084	(0.024)	0.304	0.246	(0.025)
Own effects						
Age	−0.044	−0.081	(0.008)	−0.050	−0.041	(0.005)
Male	−0.160	−0.293	(0.017)	−0.259	−0.210	(0.011)
Hispanic	−0.013	−0.023	(0.025)	0.108	0.088	(0.017)
Race						
Black	0.224	0.411	(0.032)	0.524	0.425	(0.021)
Asian	0.206	0.378	(0.034)	0.641	0.520	(0.023)
Other	0.027	0.050	(0.028)	0.127	0.103	(0.018)
Years at school	0.034	0.062	(0.007)	0.101	0.082	(0.005)
With both par.	0.074	0.137	(0.019)	0.150	0.122	(0.012)
Mother educ.						
<High	−0.055	−0.102	(0.023)	−0.040	−0.033	(0.015)
>High	0.206	0.378	(0.020)	0.390	0.317	(0.013)
Missing						
Mother job	0.027	0.049	(0.031)	0.205	0.166	(0.021)
Professional	0.135	0.249	(0.024)	0.242	0.196	(0.016)
Other	0.037	0.069	(0.020)	0.060	0.049	(0.013)
Missing	−0.039	−0.072	(0.028)	−0.066	−0.053	(0.019)
μ^1	0.126	0.232	(0.137)	0.300	0.244	(0.154)
μ^2	−0.082	−0.151	(0.751)	−0.339	−0.275	(1.440)
μ^3	−0.589	−1.081	(1.851)	−2.132	−1.729	(7.031)
μ^4	−0.452	−0.829	(2.074)	−1.424	−1.156	(18.735)
μ^5	−0.006	−0.010	(0.857)	1.021	0.828	(27.358)
μ^6				0.418	0.339	(20.495)
μ^7				−0.449	−0.364	(6.145)
ν^1	0.093	0.170	(0.014)	0.167	0.135	(0.010)
ν^2	0.061	0.113	(0.055)	0.233	0.189	(0.058)
ν^3	0.048	0.088	(0.074)	−0.356	−0.289	(0.119)
ν^4	−0.109	−0.200	(0.188)	−0.063	−0.051	(0.377)
ν^5	−0.144	−0.265	(0.163)	1.914	1.553	(0.312)
ν^6				0.070	0.056	(0.782)
ν^7				−1.499	−1.216	(0.535)
Contextual effects						
Age	−0.011	−0.021	(0.010)	−0.061	−0.049	(0.007)
Male	−0.044	−0.081	(0.030)	−0.030	−0.024	(0.020)
Hispanic	−0.048	−0.087	(0.041)	−0.059	−0.048	(0.027)
Race						
Black	0.123	0.226	(0.042)	0.074	0.060	(0.029)
Asian	−0.017	−0.032	(0.053)	−0.243	−0.197	(0.036)
Other	−0.105	−0.193	(0.052)	−0.259	−0.210	(0.035)
Years at school	0.003	0.005	(0.012)	−0.021	−0.017	(0.008)
With both par.	0.134	0.246	(0.037)	0.173	0.141	(0.024)
Mother educ.						
<High	−0.126	−0.231	(0.043)	−0.179	−0.145	(0.029)
>High	0.181	0.333	(0.040)	0.216	0.175	(0.026)
Missing	−0.065	−0.119	(0.060)	−0.143	−0.116	(0.041)
Mother job						
Professional	0.176	0.322	(0.048)	0.241	0.195	(0.032)
Other	0.036	0.066	(0.038)	0.034	0.028	(0.025)
Missing	−0.011	−0.019	(0.055)	0.036	0.029	(0.036)
μ^1	0.164	0.301	(0.241)	0.633	0.514	(0.213)
μ^2	0.138	0.253	(2.465)	2.213	1.795	(3.868)
μ^3	−0.019	−0.035	(11.064)	2.720	2.207	(36.684)
μ^4	−0.187	−0.344	(22.434)	−0.265	−0.215	(184.271)
μ^5	0.017	0.031	(16.807)	−0.539	−0.438	(498.076)
μ^6				0.787	0.639	(684.420)
μ^7				−0.370	−0.301	(374.754)
ν^1	0.032	0.060	(0.018)	0.113	0.092	(0.0160)
ν^2	0.021	0.038	(0.077)	−0.025	−0.021	(0.088)
ν^3	−0.034	−0.063	(0.109)	−0.651	−0.528	(0.168)
ν^4	0.075	0.138	(0.301)	0.609	0.494	(0.635)
ν^5	0.110	0.202	(0.280)	1.872	1.518	(0.476)
ν^6				−0.555	−0.450	(1.439)
ν^7				−1.341	−1.088	(0.962)
σ				2.424		

For the count data model, $\hat{\bar{R}} = 12$. The estimates of $\delta_2, \ldots, \delta_{\hat{\bar{R}}}$ are 1.551, 0.521, 0.450, 0.384, 0.319, 0.263, 0.218, 0.174, 0.130, 0.100, 0.086. The estimate of $\bar{\delta}$ is $1.2e^{-5}$.