Identification and Estimation of Spillover Effects in Randomized Experiments: Supplemental Appendix

Gonzalo Vazquez-Bare*

November 5, 2020

Abstract

This supplemental appendix provides the proofs of the results in the paper and additional discussions and results not included in the paper to conserve space.

Contents

*Department of Economics, University of California, Santa Barbara. gvazquez@econ.ucsb.edu

A1 Endogenous effects and structural models

As described in the literature review, a large literature in economics has studied identification in structural models in which the outcome of a unit depends on the outcomes of neighbors. While Manski (2013a), Manski (2013b) and Angrist (2014) have questioned the relevance of such models from a causal perspective, it is interesting to ask what type of structural model can justify the response functions that I study in this paper. To simplify the discussion, consider a setting with groups of size 2, that is, each unit has one neighbor. Suppose the potential outcomes y_i are generated by the following system, where the arguments are suppressed to simplify notation:

$$\begin{aligned} y_1 &= f(d_1, d_2, y_2, \varepsilon_1) \\ y_2 &= f(d_2, d_1, y_1, \varepsilon_2) \end{aligned}$$

This implies that

$$\begin{aligned} y_1 &= f(d_1, d_2, f(d_2, d_1, y_1, \varepsilon_2), \varepsilon_1) \\ y_2 &= f(d_2, d_1, f(d_1, d_2, y_2, \varepsilon_1), \varepsilon_2) \end{aligned}$$

Depending on the form of the $f(\cdot,\cdot,\cdot)$, the above system may have one, zero or multiple equilibria. Suppose that $f(\cdot,\cdot,\cdot)$ is such that the system has a unique equilibrium. Then, the reduced form is given by:

$$y_i = \varphi(d_i, d_j, \varepsilon_i, \varepsilon_j)$$

Now, define $\phi(d_i, d_j, \varepsilon_i) = \mathbb{E}[\varphi(d_i, d_j, \varepsilon_i, \varepsilon_j)|\varepsilon_i]$, which integrates over ε_j.Then,

$$y_i = \varphi(d_i, d_j, \varepsilon_i) + u_{ij}$$

where $u_{ij} = \varphi(d_i, d_j, \varepsilon_i, \varepsilon_j) - \varphi(d_i, d_j, \varepsilon_i)$. Now, because d_i and d_j are binary, we have that:

$$\begin{aligned} \varphi(d_i, d_j, \varepsilon_i) &= \varphi_i^{00}(1-d_i)(1-d_j) + \varphi_i^{10} d_i(1-d_j) \\ &\quad + \varphi_i^{01}(1-d_i)d_j + \varphi_i^{11} d_i d_j \\ &:= \varphi_i^{00} + (\varphi_i^{10} - \varphi_i^{00})d_i + (\varphi_i^{01} - \varphi_i^{00})d_j \\ &\quad + (\varphi_i^{11} - \varphi_i^{01} - \varphi_i^{10} + \varphi_i^{00})d_i d_j \end{aligned}$$

and therefore

$$y_i = \varphi_i^{00} + (\varphi_i^{10} - \varphi_i^{00})d_i + (\varphi_i^{01} - \varphi_i^{00})d_j + (\varphi_i^{11} - \varphi_i^{01} - \varphi_i^{10} + \varphi_i^{00})d_i d_j + u_{ij}$$

The above expression can be relabeled to match the notation in the paper, ignoring the group subscript:

$$Y_i(d, d^1) = Y_{ij}(0) + \tau_i d + \gamma_i(0)d^1 + (\gamma_i(1) - \gamma_i(0))d \cdot d^1$$

where $Y_{ij,g}(0) = \phi_{ig}^{00} + u_{ij,g}$.

An important difference when allowing for endogenous effects is the presence of an additional term, u_{ij}, which depends on the heterogeneity of all units in the group. The presence of this term will generally introduce correlation between units in the same group. This feature does not affect identification, but has to be taken into account when performing inference.

Importantly, since the treatment indicators are binary, the reduced form can always be written in a fully saturated form, which does not require any assumptions on the structural equations, besides the restrictions that guarantee a unique equilibrium.

As an illustration, consider the following structural function:

$$y_i = f(d_i, y_j, \varepsilon_i) = \alpha_i + \beta_i d_i + \theta_i y_j + \delta_i d_i y_j$$

where $\alpha_i = \alpha(\varepsilon_i)$ and similarly for β_i, θ_i and δ_i. Then,

$$\begin{cases} y_i = \alpha_i + \theta_i(\alpha_j + \theta_j y_i) & \text{if } d_i = 0,\ d_j = 0 \\ y_i = \alpha_i + \beta_i + (\theta_i + \delta_i)(\alpha_j + \theta_j y_i) & \text{if } d_i = 1,\ d_j = 0 \\ y_i = \alpha_i + \theta_i(\alpha_j + \beta_j + (\theta_j + \delta_j)y_i) & \text{if } d_i = 0,\ d_j = 1 \\ y_i = \alpha_i + \beta_i + (\theta_i + \delta_i)(\alpha_j + \beta_j + (\theta_j + \delta_j)y_i) & \text{if } d_i = 1,\ d_j = 1 \end{cases}$$

which implies the reduced form:

$$\begin{cases} y_i = \frac{\alpha_i + \theta_i\alpha_j}{1 - \theta_i\theta_j} & \text{if } d_i = 0,\ d_j = 0 \\ y_i = \frac{\alpha_i + \theta_i\alpha_j}{1 - (\theta_i + \delta_i)\theta_j} + \frac{\beta_i}{1 - (\theta_i + \delta_i)\theta_j} & \text{if } d_i = 1,\ d_j = 0 \\ y_i = \frac{\alpha_i + \theta_i\alpha_j}{1 - \theta_i(\theta_j + \delta_j)} + \frac{\beta_j\theta_i}{1 - \theta_i(\theta_j + \delta_j)} & \text{if } d_i = 0,\ d_j = 1 \\ y_i = \frac{\alpha_i + \theta_i\alpha_j}{1 - (\theta_i + \delta_i)(\theta_j + \delta_j)} + \frac{\beta_i}{1 - (\theta_i + \delta_i)(\theta_j + \delta_j)} + \frac{\beta_j\theta_i}{1 - (\theta_i + \delta_i(\theta_j + \delta_j)} & \text{if } d_i = 1,\ d_j = 1 \end{cases}$$

as long as $\theta_i\theta_j$, $(\theta_i + \delta_i)\theta_j$, $\theta_i(\theta_j + \delta_j)$ and $(\theta + \delta_i)(\theta_j + \delta_j)$ are different from 1 almost surely. This expression can be rewritten as before:

$$y_i = \varphi_i^{00} + (\varphi_i^{10} - \varphi_i^{00})d_i + (\varphi_i^{01} - \varphi_i^{00})d_j + (\varphi_i^{11} - \varphi_i^{01} - \varphi_i^{10} + \varphi_i^{00})d_i d_j + u_{ij}$$

where now all the φ_i terms are functions of the structural parameters. Notice that the neighbor's treatment assignment enters the reduced form in levels and with an interaction, even though it does not enter the structural equation. This is so because in the reduced form equation, d_j captures the effect of y_j.

A2 Implications for experimental design

Theorem 4 shows that the accuracy of the standard normal to approximate the distribution of the standardized statistic depends on the treatment assignment mechanism through $\underline{\pi}_n$. The intuition behind this result is that the amount of information to estimate each $\mu(\mathbf{a})$ depends on the number of observations facing assignment $\mathbf{a}$, and this number depends on $\pi(\mathbf{a})$. When the goal is to estimate all the $\mu(\mathbf{a})$ simultaneously, the binding factor will be the number of observations in the smallest cell, controlled by $\underline{\pi}_n$. When an assignment sets a value of $\underline{\pi}_n$ that is very close to zero, the Gaussian distribution may provide a poor approximation to the distribution of the estimators.

When designing an experiment to estimate spillover effects, the researcher can choose distribution of treatment assignments $\pi(\cdot)$. Theorem 4 provides a way to rank different assignment mechanisms based on their rate of the approximation, which gives a principled way to choose between different assignment mechanisms.

To illustrate these issues, consider the case of an exchangeable exposure mapping $\mathcal{A}_n = \{(d, s) : d = 0, 1, s = 0, 1, \ldots, n\}$. The results below compare two treatment assignment mechanisms. The first one, *simple random assignment (SR)*, consists in assigning the treatment independently at the individual level with probability $\mathbb{P}[D_{ig} = 1] = p$. The second mechanism will be two-stage randomization. Although there are several ways to implement a two-stage design, I will focus on the case in which each group is assigned a fixed number of treated units between 0 and $n + 1$ with equal probability. For example, if groups have size 3, then this mechanism assigns each group to receive 0, 1, 2 or 3 treated units with probability 1/4. This mechanism will be referred to as *two-stage randomization with fixed margins (2SR-FM)*.

Corollary A1 (SR) *Under simple random assignment, condition* (4) *holds whenever:*

$$\frac{n+1}{\log G} \to 0. \tag{1}$$

Corollary A2 (2SR-FM) *Under a 2SR-FM mechanism, condition* (4) *holds whenever:*

$$\frac{\log(n+1)}{\log G} \to 0. \tag{2}$$

In qualitative terms, both results imply that estimation and inference for spillover effects require group size to be small relative to the total number of groups. Thus, these results formalize the requirement of "many small groups" that is commonly invoked, for example, when estimating LIM models (see e.g. Davezies, D'Haultfoeuille, and Fougère, 2009; Kline and Tamer, 2019).

Corollary A1 shows that when the treatment is assigned using a simple random assignment, group size has to be small relative to $\log G$. Given the concavity of the log function, this is a strong requirement; for instance, with a sample of $G = 300$ groups, having $n = 5$

neighbors already gives $n+1 > \log G$. Hence, groups have to be very small relative to the sample size for inference to be asymptotically valid. The intuition behind this result is that under a SR, the probability of the tail assignments $(0,0)$ and $(1,n)$ decreases exponentially fast with group size.

On the other hand, Corollary A2 shows that a 2SR-FM mechanism reduces the requirement to $\log(n+1)/\log G \approx 0$, so now the log of group size has to be small compared to the log of the number of groups. This condition is much more easily satisfied, which in practical terms implies that a 2SR-FM mechanism can handle larger groups compared to SR. The intuition behind this result is that, by fixing the number of treated units in each group, a 2SR-FM design has better control on how small the probabilities of each assignment can be, hence facilitating the estimation of the tail assignments.

A3 Unequally-sized groups

To explicitly account for different group sizes, let n (the total number of peers in each group) take values in $\mathcal{N} = \{n_1, n_2, \ldots, n_K\}$ where $n_k \geq 1$ for all k and $n_1 < n_2 < \ldots < n_K$. Let the potential outcome be $Y_{ig}(n, d, s(n))$ where $n \in \mathcal{N}$ and $s(n) \in \{0, 1, 2, \ldots, n\}$. Let N_g be the observed value of n_g, $S_{ig}(n) = \sum_{j \neq i}^{n} D_{jg}$ and $S_{ig} = \sum_{k=1}^{K} S_{ig}(n_k)\mathbb{1}(N_g = n_k)$. The independence assumption can be modified to hold conditional on group size:

$$\{Y_{ig}(n, d, s(n)) : d = 0, 1, s(n) = 0, 1, \ldots, n\}_{i=1}^{n} \perp\!\!\!\perp \mathbf{D}_g(n))|N_g = n$$

where $\mathbf{D}_g(n)$ is the vector of all treatment assignments when the group size is $n+1$.

Under this assumption, we have that for $n \in \mathcal{N}$ and $s \leq n$,

$$\mathbb{E}[Y_{ig}|D_{ig} = d, S_{ig} = s, N_g = n] = \mathbb{E}[Y_{ig}(n, d, s)].$$

The average observed outcome conditional on $N_g = n$ can be written as:

$$\begin{aligned}\mathbb{E}[Y_{ig}|D_{ig}, S_{ig}, N_g = n] = \mathbb{E}[Y_{ig}(n, 0, 0)] &+ \tau_0(n) D_{ig} \\ &+ \sum_{s=1}^{n} \theta_0(s, n)\mathbb{1}(S_{ig} = s)(1 - D_{ig}) \\ &+ \sum_{s=1}^{n} \theta_1(s, n)\mathbb{1}(S_{ig} = s) D_{ig}\end{aligned}$$

The easiest approach is to simply run separate analyses for each group size and estimate all the effects separately. In this case, it is possible to test whether spillover effects are different in groups with different sizes. The total number of parameters in this case is given by $\sum_{k=1}^{K}(n_k + 1)$.

In practice, however, there may be cases in which group size has a rich support with only a few groups at each value n_g, so separate analyses may not be feasible. In such a setting, a possible solution is to impose an additivity assumption on group size. According to this assumption, the average direct and spillover effects do not change with group size. For example, the spillover effect of having one treated neighbor is the same in a group with two or three units. Under this assumption,

$$\begin{aligned}\mathbb{E}[Y_{ig}|D_{ig},S_{ig},N_g] = &\sum_{n_g\in\mathcal{N}_g}\alpha(n_g)\mathbb{1}(N_g=n_g)+\tau_0 D_{ig}\\&+\sum_{s=1}^{N_g}\theta_0(s)\mathbb{1}(S_{ig}=s)(1-D_{ig})\\&+\sum_{s=1}^{N_g}\theta_1(s)\mathbb{1}(S_{ig}=s)D_{ig}\end{aligned}$$

where the first sum can be seen in practice as adding group-size fixed effects. Then, the identification results and estimation strategies in the paper are valid after controlling for group-size fixed effects. Note that in this case the total number of parameters to estimate is n_K+K-1 where n_K is the size of the largest group and K is the total number of different group sizes.

Another possibility is to assume that for any constant $c\in\mathbb{N}$, $Y_{ig}(c\cdot n,d,c\cdot s)=Y_{ig}(n,d,s)$. This assumption allows us to rewrite the potential outcomes as a function of the ratio of treated peers, $Y_{ig}(d,s/n)$. Letting $P_{ig}=S_{ig}/N_g$, all the parameters can be estimated by running a regression including D_{ig}, $\mathbb{1}(P_{ig}=p)$ for all possible values of $p>0$ (excluding $p=0$ to avoid perfect collinearity) and interactions. In this case, the total number of parameters can be bounded by $n_1+\sum_{k=2}^{K}(n_k-1)$. Note that assuming that the potential outcomes depend only on the proportion of treated siblings does not justify in any way including the variable P_{ig} linearly, as commonly done in linear-in-means models.

A4 Including covariates

There are several reasons why one may want to include covariates when estimating direct and spillover effects. First, pre-treatment characteristics may help reduce the variability of the estimators and decrease small-sample bias, which is standard practice when analyzing randomly assigned programs. Covariates can also help get valid inference when the assignment mechanisms stratifies on baseline covariates. This can be done by simply augmenting Equation (7) with a vector of covariates $\boldsymbol{\gamma}'\mathbf{x}_{ig}$ which can vary at the unit or at the group level. The covariates can also be interacted with the treatment assignment indicators to explore effect heterogeneity across observable characteristics (for example, by separately estimating effects for males and females.

Second, exogenous covariates can be used to relax the mean-independence assumption in observational studies. More precisely, if $\mathbf{X}_g$ is a matrix of covariates, a conditional mean-independence assumption would be

$$\mathbb{E}[Y_{ig}(d,\mathbf{d}_g)|\mathbf{X}_g,\mathbf{D}_g]=\mathbb{E}[Y_{ig}(d,\mathbf{d}_g)|\mathbf{X}_g]$$

which is a version of the standard unconfoundeness condition. The vector of covariates can include both individual-level and group-level characteristics.

Third, covariates can be included to make an exposure mapping more likely to be correctly specified. For instance, the exchangeability assumption can be relaxed by assuming it holds after conditioning on covariates, so that for any pair of treatment assignments $\mathbf{d}_g$ and $\tilde{\mathbf{d}}_g$ with the same number of ones,

$$\mathbb{E}[Y_{ig}(d,\mathbf{d}_g)|\mathbf{X}_g]=\mathbb{E}[Y_{ig}(d,\tilde{\mathbf{d}}_g)|\mathbf{X}_g]$$

As an example, exchangeability can be assumed to hold for all siblings with the same age, gender or going to the same school.

All the identification results in the paper can be adapted to hold after conditioning on covariates. In terms of implementation, when the covariates are discrete the parameters of interest can be estimated at each possible value of the matrix $\mathbf{X}_g$, although this strategy can worsen the dimensionality problem. Alternatively, covariates can be included in a regression framework after imposing parametric assumptions, for example, assuming the covariates enter linearly.

A5 Additional theoretical results

Lemma A1 (Assignment probabilities) *Let $\hat{\pi}(\mathbf{a}):=\sum_g\sum_i \mathbb{1}_{ig}(\mathbf{a})/G(n+1)$. Then under the assumptions of Lemma 2,*

$$\max_{\mathbf{a}\in\mathcal{A}_n}\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})}-1\right|\to_{\mathbb{P}}0.$$

A6 Proofs of additional results

Proof of Lemma A1 Take $\varepsilon>0$, then

$$\begin{aligned}\mathbb{P}\left[\max_{\mathbf{a}\in\mathcal{A}_n}\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})}-1\right|>\varepsilon\right]&\leq\sum_{\mathbf{a}\in\mathcal{A}_n}\mathbb{P}\left[\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})}-1\right|>\varepsilon\right]\leq|\mathcal{A}_n|\max_{\mathbf{a}\in\mathcal{A}_n}\mathbb{P}\left[|\hat{\pi}(\mathbf{a})-\pi(\mathbf{a})|>\varepsilon\pi(\mathbf{a})\right]\\&=|\mathcal{A}_n|\max_{\mathbf{a}\in\mathcal{A}_n}\mathbb{P}\left[|N(\mathbf{a})-\mathbb{E}[N(\mathbf{a})]|>\varepsilon\mathbb{E}[N(\mathbf{a})]\right]\end{aligned}$$

Now,

$$N(\mathbf{a})-\mathbb{E}[N(\mathbf{a})]=\sum_g\sum_i \mathbb{1}_{ig}(\mathbf{a})-G(n+1)\pi(\mathbf{a})=\sum_g W_g$$

where $W_g = \sum_i \mathbb{1}_{ig}(\mathbf{a}) - (n+1)\pi(\mathbf{a}) = N_g(\mathbf{a}) - \mathbb{E}[N_g(\mathbf{a})]$. Note that the W_g are independent and:

$$\begin{aligned}
\mathbb{E}[W_g] &= 0 \\
|W_g| &\leq (n+1)\max\{\pi(\mathbf{a}), 1-\pi(\mathbf{a})\} \\
\mathbb{V}[W_g] &= \mathbb{V}\left[\sum_i \mathbb{1}_{ig}(\mathbf{a})\right] = \sum_i \mathbb{V}[\mathbb{1}_{ig}(\mathbf{a})] + 2\sum_i \sum_{j>i} \mathbb{C}\mathrm{ov}(\mathbb{1}_{ig}(\mathbf{a}), \mathbb{1}_{jg}(\mathbf{a})) \\
&= (n+1)\pi(\mathbf{a})(1-\pi(\mathbf{a})) + (n+1)(n+2)\{\mathbb{E}[\mathbb{1}_{ig}(\mathbf{a})\mathbb{1}_{jg(\mathbf{a})}] - \pi(\mathbf{a})^2\} \\
&\leq (n+1)\pi(\mathbf{a})(1-\pi(\mathbf{a})) + (n+1)(n+2)\pi(\mathbf{a})(1-\pi(\mathbf{a})) \\
&= (n+1)(n+3)\pi(\mathbf{a})(1-\pi(\mathbf{a}))
\end{aligned}$$

Then, by Bernstein's inequality,

$$\begin{aligned}
\mathbb{P}\left[|W_g| > \varepsilon \mathbb{E}[N(\mathbf{a})]\right] &\leq 2\exp\left\{-\frac{\mathbb{E}[N(\mathbf{a})]^2\varepsilon^2}{\sum_g \mathbb{V}[W_g] + \frac{1}{3}(n+1)\max\{\pi(\mathbf{a}), 1-\pi(\mathbf{a})\}\mathbb{E}[N\mathbf{a})]\varepsilon}\right\} \\
&= 2\exp\left\{-\frac{\frac{1}{2}G^2(n+1)^2\pi(\mathbf{a})^2\varepsilon^2}{G(n+1)(n+3)\pi(\mathbf{a})(1-\pi(\mathbf{a})) + \frac{1}{3}G(n+1)^2\pi(\mathbf{a})\max\{\pi(\mathbf{a}), 1-\pi(\mathbf{a})\}\varepsilon}\right\} \\
&= 2\exp\left\{-\frac{\frac{1}{2}G\pi(\mathbf{a})\varepsilon^2}{\frac{n+3}{n+1}(1-\pi(\mathbf{a})) + \frac{1}{3}\max\{\pi(\mathbf{a}), 1-\pi(\mathbf{a})\varepsilon\}}\right\} \\
&\leq 2\exp\left\{-\frac{\frac{1}{2}G\pi(\mathbf{a})\varepsilon^2}{\frac{n+3}{n+1} + \frac{\varepsilon}{3}}\right\}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{P}\left[\max_{\mathbf{a}\in\mathcal{A}_n}\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \varepsilon\right] &\leq |\mathcal{A}_n| \max_{\mathbf{a}\in\mathcal{A}_n} \mathbb{P}\left[|N(\mathbf{a}) - \mathbb{E}[N(\mathbf{a})]| > \varepsilon\mathbb{E}[N(\mathbf{a})]\right] \\
&\leq 2\exp\left\{-G\underline{\pi}_n\left(\frac{\frac{1}{2}\varepsilon^2}{\frac{n+3}{n+1} + \frac{\varepsilon}{3}} - \frac{\log|\mathcal{A}_n|}{G\underline{\pi}_n}\right)\right\} \to 0.
\end{aligned}$$

as required. □

Proof of Corollary A1 Under exchangeability $\pi(\mathbf{a}) = \pi(d,s) = p^d(1-p)^{1-d}\binom{n}{s}p^s(1-p)^{n-s} = \binom{n}{s}p^{s+d}(1-p)^{n+1-s-d}$. This function is minimized at $\underline{\pi}_n = \underline{p}^{n+1} \propto \underline{p}^n$ where $\underline{p} = \min\{p, 1-p\}$. Thus,

$$\frac{\log|\mathcal{A}_n|}{G\underline{p}^n} = \exp\left\{-\log G\left(1 - \frac{n+1}{\log G}2\log\underline{p} - \frac{\log\log|\mathcal{A}_n|}{\log G}\right)\right\}$$

and since $|\mathcal{A}_n| = 2(n+1)$, this term converges to zero when $(n+1)/\log G \to 0$. □

Proof of Corollary A2 Under exchangeability, $\pi(\mathbf{a}) = \pi(d,s) = q_{d+s}\left(\frac{s+1}{n+1}\right)^d \times \left(1-\frac{s}{n+1}\right)^{1-d}$ which in this case equals $\frac{1}{n+1}\left(\frac{s+1}{n+1}\right)^d \left(1-\frac{s}{n+1}\right)^{1-d}$. This function has two minima, one at $(d,s) = (0,n)$ and one at $(d,s) = (1,0)$, giving the same minimized value of $\underline{\pi}_n = (n+1)^{-2}$. Then,

$$\frac{\log|\mathcal{A}_n|}{G\underline{\pi}_n} = \exp\left\{-\log G\left(1 - \frac{\log(n+1)}{\log G}4 - \frac{\log\log 2(n+1)}{\log G} + o(1)\right)\right\} \to 0$$

if $\log(n+1)/\log G \to 0$. $\square$

A7 Proofs of main results

Proof of Lemma 1 If $\mathbb{P}[D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}] > 0$,

$$\begin{aligned}
\mathbb{E}[Y_{ig}|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}] &= \sum_{\mathbf{h}_0} \mathbb{E}[Y_{ig}|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}, \mathbf{H}^0_{ig} = \mathbf{h}_0] \\
&\quad \times \mathbb{P}[\mathbf{H}^0_{ig} = \mathbf{h}_0|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}] \\
&= \sum_{\mathbf{h}_0} \mathbb{E}[Y_{ig}(d,\mathbf{h}_0)|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}, \mathbf{H}^0_{ig} = \mathbf{h}_0] \\
&\quad \times \mathbb{P}[\mathbf{H}^0_{ig} = \mathbf{h}^0|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}] \\
&= \sum_{\mathbf{h}_0} \mathbb{E}[Y_{ig}(d,\mathbf{h}_0)]\mathbb{P}[\mathbf{H}^0_{ig} = \mathbf{h}_0|D_{ig} = d, \mathbf{H}_{ig} = \mathbf{h}]
\end{aligned}$$

where the first equality follows from the law if iterated expectations, the second equality follows by definition of the observed outcomes and the third equality follows from random assignment of the treatment vector given that both $\mathbf{H}^0_{ig}$ and $\mathbf{H}_{ig}$ are deterministic functions of $\mathbf{D}_g$. Finally, if $h_0(\cdot)$ is coarser than $h(\cdot)$, then $\mathbf{H}_{ig} = \mathbf{h}$ uniquely determines the value of $\mathbf{H}^0_{ig}$ and the result follows. $\square$

Proof of Theorem 1 Follows from Lemma 1 letting $h(\cdot)$ be a constant function, using the fact that by construction $\beta_{\mathsf{D}} = \mathbb{E}[Y_{ig}|D_{ig} = 1] - \mathbb{E}[Y_{ig}|D_{ig} = 0]$. $\square$

Proof of Theorem 2 The coefficients from Equation (2) are characterized by the minimization problem:

$$\min_{(\alpha_\ell,\beta_\ell,\gamma_\ell)} \mathbb{E}\left[\left(Y_{ig} - \alpha_\ell - \beta_\ell D_{ig} - \gamma_\ell \bar{D}^{(i)}_g\right)^2\right].$$

The objective function can be rewritten as:

$$\mathbb{E}\left[\left(Y_{ig} - \alpha_\ell - \gamma_\ell \bar{D}^{(i)}_g\right)^2 (1 - D_{ig})\right] + \mathbb{E}\left[\left(Y_{ig} - \alpha_\ell - \beta_\ell - \gamma_\ell \bar{D}^{(i)}_g\right)^2 D_{ig}\right]$$

which can be reparameterized as

$$\mathbb{E}\left[\left(Y_{ig}-\alpha_0-\gamma_\ell \bar{D}_g^{(i)}\right)^2(1-D_{ig})\right]+\mathbb{E}\left[\left(Y_{ig}-\alpha_1-\gamma_\ell \bar{D}_g^{(i)}\right)^2 D_{ig}\right]$$

where $\alpha_0=\alpha_\ell$ and $\alpha_1=\alpha_\ell+\beta_\ell$. The first-order condition for α_0 and α_1 are:

$$\begin{aligned}
0&=\mathbb{E}\left[\left(Y_{ig}-\alpha_0-\gamma_\ell \bar{D}_g^{(i)}\right)(1-D_{ig})\right]\\
0&=\mathbb{E}\left[\left(Y_{ig}-\alpha_1-\gamma_\ell \bar{D}_g^{(i)}\right)D_{ig}\right]
\end{aligned}$$

from which

$$\begin{aligned}
\alpha_0&=\mathbb{E}[Y_{ig}|D_{ig}=0]-\gamma_\ell\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=0]\\
\alpha_1&=\mathbb{E}[Y_{ig}|D_{ig}=1]-\gamma_\ell\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=1].
\end{aligned}$$

But $\beta_\ell=\alpha_1-\alpha_0$ and thus:

$$\begin{aligned}
\beta_\ell&=\mathbb{E}[Y_{ig}|D_{ig}=1]-\mathbb{E}[Y_{ig}|D_{ig}=0]-\gamma_\ell(\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=1]-\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=0])\\
&=\mathbb{E}[Y_{ig}|D_{ig}=1]-\mathbb{E}[Y_{ig}|D_{ig}=0]-\frac{\gamma_\ell}{n_g}(\mathbb{E}[S_{ig}|D_{ig}=1]-\mathbb{E}[S_{ig}|D_{ig}=0]).
\end{aligned}$$

The first-order condition for γ_ℓ is:

$$\begin{aligned}
0&=\mathbb{E}\left[\left(Y_{ig}-\alpha_0-\gamma_\ell \bar{D}_g^{(i)}\right)\bar{D}_g^{(i)}(1-D_{ig})\right]+\mathbb{E}\left[\left(Y_{ig}-\alpha_1-\gamma_\ell \bar{D}_g^{(i)}\right)\bar{D}_g^{(i)}D_{ig}\right]\\
&=\mathbb{E}\left[\left(Y_{ig}-\mathbb{E}[Y_{ig}|D_{ig}=0]-\gamma_\ell(\bar{D}_g^{(i)}-\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=0])\right)\bar{D}_g^{(i)}(1-D_{ig})\right]\\
&+\mathbb{E}\left[\left(Y_{ig}-\mathbb{E}[Y_{ig}|D_{ig}=1]-\gamma_\ell(\bar{D}_g^{(i)}-\mathbb{E}[\bar{D}_g^{(i)}|D_{ig}=1])\right)\bar{D}_g^{(i)}D_{ig}\right]\\
&=\mathbb{C}\mathrm{ov}(Y_{ig},\bar{D}_g^{(i)}|D_{ig}=0)\mathbb{P}[D_{ig}=0]+\mathbb{C}\mathrm{ov}(Y_{ig},\bar{D}_g^{(i)}|D_{ig}=1)\mathbb{P}[D_{ig}=1]\\
&-\gamma_\ell\left(\mathbb{V}[\bar{D}_g^{(i)}|D_{ig}=0]\mathbb{P}[D_{ig}=0]+\mathbb{V}[\bar{D}_g^{(i)}|D_{ig}=1]\mathbb{P}[D_{ig}=1]\right)
\end{aligned}$$

from which:

$$\gamma_\ell=\frac{\mathbb{C}\mathrm{ov}(Y_{ig},\bar{D}_g^{(i)}|D_{ig}=0)\mathbb{P}[D_{ig}=0]+\mathbb{C}\mathrm{ov}(Y_{ig},\bar{D}_g^{(i)}|D_{ig}=1)\mathbb{P}[D_{ig}=1]}{\mathbb{V}[\bar{D}_g^{(i)}|D_{ig}=0]\mathbb{P}[D_{ig}=0]+\mathbb{V}[\bar{D}_g^{(i)}|D_{ig}=1]\mathbb{P}[D_{ig}=1]}.$$

Next,

$$\mathbb{V}[\bar{D}_g^{(i)}|D_{ig}=d]=\frac{1}{n_g^2}\mathbb{V}[S_{ig}|D_{ig}=d]$$

and

$$
\begin{aligned}
\mathbb{C}\text{ov}(Y_{ig}, \bar{D}_g^{(i)}|D_{ig} = d) &= \frac{1}{n_g}\mathbb{C}\text{ov}(Y_{ig}, S_{ig}|D_{ig} = d) \\
&= \frac{1}{n_g}\mathbb{C}\text{ov}(\mathbb{E}[Y_{ig}|D_{ig} = d, S_{ig}], S_{ig}|D_{ig} = d) \\
&= \frac{1}{n_g}\sum_{s=0}^{n_g}\mathbb{E}[Y_{ig}|D_{ig} = d, S_{ig} = s]\mathbb{C}\text{ov}(\mathbb{1}(S_{ig} = s), S_{ig}|D_{ig} = d).
\end{aligned}
$$

But

$$
\begin{aligned}
\mathbb{C}\text{ov}(\mathbb{1}(S_{ig} = s), S_{ig}|D_{ig} = d) &= \mathbb{E}[\mathbb{1}(S_{ig} = s)S_{ig}|D_{ig} = d] - \mathbb{E}[\mathbb{1}(S_{ig} = s)|D_{ig} = d]\mathbb{E}[S_{ig}|D_{ig} = d] \\
&= (s - \mathbb{E}[S_{ig}|D_{ig} = d])\mathbb{P}[S_{ig} = s|D_{ig} = d].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\gamma_\ell &= \frac{\frac{1}{n_g}\sum_{s=0}^{n_g}\mathbb{E}[Y_{ig}|D_{ig} = 0, S_{ig} = s](s - \mathbb{E}[S_{ig}|D_{ig} = 0])\mathbb{P}[S_{ig} = s|D_{ig} = 0]\mathbb{P}[D_{ig} = 0]}{\frac{1}{n_g^2}\mathbb{V}[S_{ig}|D_{ig} = 0]\mathbb{P}[D_{ig} = 0] + \frac{1}{n_g^2}\mathbb{V}[S_{ig}|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]} \\
&\quad + \frac{\frac{1}{n_g}\sum_{s=0}^{n_g}\mathbb{E}[Y_{ig}|D_{ig} = 1, S_{ig} = s](s - \mathbb{E}[S_{ig}|D_{ig} = 1])\mathbb{P}[S_{ig} = s|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]}{\frac{1}{n_g^2}\mathbb{V}[S_{ig}|D_{ig} = 0]\mathbb{P}[D_{ig} = 0] + \frac{1}{n_g^2}\mathbb{V}[S_{ig}|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]} \\
&= \frac{\sum_{s=0}^{n_g}\mathbb{E}[Y_{ig}|D_{ig} = 0, S_{ig} = s]n_g\mathbb{P}[D_{ig} = 0](s - \mathbb{E}[S_{ig}|D_{ig} = 0])\mathbb{P}[S_{ig} = s|D_{ig} = 0]}{\mathbb{V}[S_{ig}|D_{ig} = 0]\mathbb{P}[D_{ig} = 0] + \mathbb{V}[S_{ig}|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]} \\
&\quad + \frac{\sum_{s=0}^{n_g}\mathbb{E}[Y_{ig}|D_{ig} = 1, S_{ig} = s]n_g\mathbb{P}[D_{ig} = 1](s - \mathbb{E}[S_{ig}|D_{ig} = 1])\mathbb{P}[S_{ig} = s|D_{ig} = 1]}{\mathbb{V}[S_{ig}|D_{ig} = 0]\mathbb{P}[D_{ig} = 0] + \mathbb{V}[S_{ig}|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]} \\
&= \sum_{s=0}^{n_g}\phi_0(s)\mathbb{E}[Y_{ig}|D_{ig} = 0, S_{ig} = s] + \sum_{s=0}^{n_g}\phi_1(s)\mathbb{E}[Y_{ig}|D_{ig} = 1, S_{ig} = s]
\end{aligned}
$$

where

$$
\phi_d(s) = \frac{n_g\mathbb{P}[D_{ig} = d]\mathbb{P}[S_{ig} = s|D_{ig} = d]}{\mathbb{V}[S_{ig}|D_{ig} = 0]\mathbb{P}[D_{ig} = 0] + \mathbb{V}[S_{ig}|D_{ig} = 1]\mathbb{P}[D_{ig} = 1]} \cdot (s - \mathbb{E}[S_{ig}|D_{ig} = d]).
$$

Also note that:

$$
\begin{aligned}
\sum_{s=0}^{n_g}\phi_d(s) &= n_g\mathbb{P}[D_{ig} = d]\sum_{s=0}^{n_g}(s - \mathbb{E}[S_{ig}|D_{ig} = d])\mathbb{P}[S_{ig} = s|D_{ig} = d] \\
&= n_g\mathbb{P}[D_{ig} = d]\mathbb{E}\left[S_{ig} - \mathbb{E}[S_{ig}|D_{ig} = d]|D_{ig} = d\right] \\
&= 0.
\end{aligned}
$$

This implies that:

$$
\begin{aligned}
\sum_{s=0}^{n_g} \phi_d(s)\mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=s] &= \sum_{s=0}^{n_g} \phi_d(s)(\mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=s] - \mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=0]) \\
&+ \sum_{s=0}^{n_g} \phi_d(s)\mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=0] \\
&= \sum_{s=0}^{n_g} \phi_d(s)(\mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=s] - \mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=0]) \\
&+ \mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=0] \sum_{s=0}^{n_g} \phi_d(s) \\
&= \sum_{s=1}^{n_g} \phi_d(s)(\mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=s] - \mathbb{E}[Y_{ig}|D_{ig}=d, S_{ig}=0])
\end{aligned}
$$

which gives the required result. □

Proof of Corollary 1 When the true model satisfies exchangeability, potential outcomes have the form $Y_{ig}(d, s)$. From Theorem 2,

$$
\begin{aligned}
\gamma_\ell &= \sum_{s=1}^{n_g} \phi_0(s)(\mathbb{E}[Y_{ig}|D_{ig}=0, S_{ig}=s] - \mathbb{E}[Y_{ig}|D_{ig}=0, S_{ig}=0]) \\
&+ \sum_{s=1}^{n_g} \phi_1(s)(\mathbb{E}[Y_{ig}|D_{ig}=1, S_{ig}=s] - \mathbb{E}[Y_{ig}|D_{ig}=1, S_{ig}=0]) \\
&= \sum_{s=1}^{n_g} \phi_0(s)\mathbb{E}[Y_{ig}(0,s) - Y_{ig}(0,0)] + \sum_{s=1}^{n_g} \phi_1(s)\mathbb{E}[Y_{ig}(1,s) - Y_{ig}(1,0)]
\end{aligned}
$$

By linearity, $\mathbb{E}[Y_{ig}(d,s) - Y_{ig}(d,0)] = s\kappa_d$ and thus:

$$
\begin{aligned}
\gamma_\ell &= \sum_{s=1}^{n_g} \phi_0(s)\mathbb{E}[Y_{ig}(0,s) - Y_{ig}(0,0)] + \sum_{s=1}^{n_g} \phi_1(s)\mathbb{E}[Y_{ig}(1,s) - Y_{ig}(1,0)] \\
&= \kappa_0 \sum_{s=1}^{n_g} s\phi_0(s) + \kappa_1 \sum_{s=1}^{n_g} s\phi_1(s).
\end{aligned}
$$

But

$$
\begin{aligned}
\sum_{s=1}^{n_g} s\phi_d(s) &= \frac{n_g \mathbb{P}[D_{ig}=d] \sum_{s=1}^{n_g} s(s - \mathbb{E}[S_{ig}|D_{ig}=d])\mathbb{P}[S_{ig}=s|D_{ig}=d]}{\mathbb{V}[S_{ig}|D_{ig}=0]\mathbb{P}[D_{ig}=0] + \mathbb{V}[S_{ig}|D_{ig}=1]\mathbb{P}[D_{ig}=1]} \\
&= \frac{n_g \mathbb{P}[D_{ig}=d]\mathbb{V}[S_{ig}|D_{ig}=d]}{\mathbb{V}[S_{ig}|D_{ig}=0]\mathbb{P}[D_{ig}=0] + \mathbb{V}[S_{ig}|D_{ig}=1]\mathbb{P}[D_{ig}=1]}.
\end{aligned}
$$

Hence,

$$\begin{aligned}\gamma_\ell &= \kappa_0 \sum_{s=1}^{n_g} s\phi_0(s) + \kappa_1 \sum_{s=1}^{n_g} s\phi_1(s) \\ &= n_g\kappa_1\lambda + n_g\kappa_0(1-\lambda)\end{aligned}$$

where

$$\lambda = \frac{\mathbb{P}[D_{ig}=d]\mathbb{V}[S_{ig}|D_{ig}=d]}{\mathbb{V}[S_{ig}|D_{ig}=0]\mathbb{P}[D_{ig}=0]+\mathbb{V}[S_{ig}|D_{ig}=1]\mathbb{P}[D_{ig}=1]}.$$

But $n_g\kappa_d = \mathbb{E}[Y_{ig}(d,n_g) - Y_{ig}(d,0)]$ which gives the result for γ_ℓ.

On the other hand, from Theorem 1, the difference in means is:

$$\begin{aligned}\beta_{\mathsf{D}} &= \mathbb{E}[Y_{ig}|D_{ig}=1] - \mathbb{E}[Y_{ig}|D_{ig}=0] \\ &= \mathbb{E}[Y_{ig}(1,0)-Y_{ig}(0,0)] + \sum_{s=1}^{n_g} \mathbb{E}[Y_{ig}(1,s)-Y_{ig}(1,0)]\mathbb{P}[S_{ig}=s|D_{ig}=1] \\ &- \sum_{s=1}^{n_g} \mathbb{E}[Y_{ig}(0,s)-Y_{ig}(0,0)]\mathbb{P}[S_{ig}=s|D_{ig}=0].\end{aligned}$$

By linearity,

$$\begin{aligned}\beta_{\mathsf{D}} &= \mathbb{E}[Y_{ig}(1,0)-Y_{ig}(0,0)] + \kappa_1\sum_{s=1}^{n_g} s\mathbb{P}[S_{ig}=s|D_{ig}=1] \\ &- \kappa_0\sum_{s=1}^{n_g} s\mathbb{P}[S_{ig}=s|D_{ig}=0] \\ &= \mathbb{E}[Y_{ig}(1,0)-Y_{ig}(0,0)] + \kappa_1\mathbb{E}[S_{ig}|D_{ig}=1] - \kappa_0\mathbb{E}[S_{ig}|D_{ig}=0].\end{aligned}$$

But

$$\begin{aligned}\beta_\ell &= \beta_{\mathsf{D}} - \frac{\gamma_\ell}{n_g}(\mathbb{E}[S_{ig}|D_{ig}=1]-\mathbb{E}[S_{ig}|D_{ig}=0]) \\ &= \mathbb{E}[Y_{ig}(1,0)-Y_{ig}(0,0)] + (\kappa_1-\kappa_0)\{(1-\lambda)\mathbb{E}[S_{ig}|D_{ig}=1]+\lambda\mathbb{E}[S_{ig}|D_{ig}=0]\}\end{aligned}$$

which gives the required result. □

Proof of Theorem 3 The coefficients from Equation (3) are characterized by the minimization problem:

$$\min_{\tilde{\alpha}_\ell,\tilde{\beta}_\ell,\gamma_\ell^0,\gamma_\ell^1} \mathbb{E}\left[\left(Y_{ig} - \tilde{\alpha}_\ell - \tilde{\beta}_\ell D_{ig} - \gamma_\ell^0 \bar{D}_g^{(i)}(1-D_{ig}) - \gamma_\ell^1 \bar{D}_g^{(i)} D_{ig}\right)^2\right]$$

The objective function can be rewritten as:

$$\mathbb{E}\left[\left(Y_{ig} - \tilde{\alpha}_\ell - \gamma_\ell^0 \bar{D}_g^{(i)}\right)^2 (1 - D_{ig})\right] + \mathbb{E}\left[\left(Y_{ig} - \tilde{\alpha}_\ell - \tilde{\beta}_\ell - \gamma_\ell^1 \bar{D}_g^{(i)}\right)^2 D_{ig}\right]$$

which can be reparameterized as:

$$\mathbb{E}\left[\left(Y_{ig} - \alpha_0 - \gamma_\ell^0 \bar{D}_g^{(i)}\right)^2 (1 - D_{ig})\right] + \mathbb{E}\left[\left(Y_{ig} - \alpha_1 - \gamma_\ell^1 \bar{D}_g^{(i)}\right)^2 D_{ig}\right]$$

where $\alpha_0 = \tilde{\alpha}_\ell$ and $\alpha_1 = \tilde{\alpha}_\ell + \tilde{\beta}_\ell$. The first-order conditions for α_0 and α_1 imply:

$$\begin{aligned}
\alpha_0 &= \mathbb{E}[Y_{ig}|D_{ig} = 0] - \gamma_\ell^0 \mathbb{E}[\bar{D}_g^{(i)}|D_{ig} = 0] \\
\alpha_1 &= \mathbb{E}[Y_{ig}|D_{ig} = 1] - \gamma_\ell^1 \mathbb{E}[\bar{D}_g^{(i)}|D_{ig} = 1]
\end{aligned}$$

and since $\tilde{\beta}_\ell = \alpha_1 - \alpha_0$,

$$\begin{aligned}
\tilde{\beta}_\ell &= \mathbb{E}[Y_{ig}|D_{ig} = 1] - \mathbb{E}[Y_{ig}|D_{ig} = 0] - \left(\gamma_\ell^1 \mathbb{E}[\bar{D}_g^{(i)}|D_{ig} = 1] - \gamma_\ell^0 \mathbb{E}[\bar{D}_g^{(i)}|D_{ig} = 0]\right) \\
&= \mathbb{E}[Y_{ig}|D_{ig} = 1] - \mathbb{E}[Y_{ig}|D_{ig} = 0] - \left(\frac{\gamma_\ell^1}{n_g} \mathbb{E}[S_{ig}|D_{ig} = 1] - \frac{\gamma_\ell^0}{n_g} \mathbb{E}[S_{ig}|D_{ig} = 0]\right).
\end{aligned}$$

On the other hand, the first-order condition for each γ_ℓ^d implies:

$$\begin{aligned}
0 &= \mathbb{E}\left[\left(Y_{ig} - \alpha_d - \gamma_\ell^d \bar{D}_g^{(i)}\right) \bar{D}_g^{(i)} \mathbb{1}(D_{ig} = d)\right] \\
&= \mathbb{E}\left[\left(Y_{ig} - \mathbb{E}[Y_{ig}|D_{ig} = d] - \gamma_\ell^d (\bar{D}_g^{(i)} - \mathbb{E}[\bar{D}_g^{(i)}|D_{ig} = d])\right) \bar{D}_g^{(i)} \mathbb{1}(D_{ig} = d)\right]
\end{aligned}$$

Thus:

$$\gamma_\ell^d = \frac{\mathbb{C}\mathrm{ov}(Y_{ig}, \bar{D}_g^{(i)}|D_{ig} = d)}{\mathbb{V}[\bar{D}_g^{(i)}|D_{ig} = d]}$$

which gives the required result by calculations shown in the proof of Theorem 2. $\square$

Proof of Corollary 2 When the true model satisfies exchangeability, potential outcomes have the form $Y_{ig}(d, s)$. From Theorem 3,

$$\begin{aligned}
\gamma_\ell^d &= \sum_{s=1}^{n_g} \omega_d(s)(\mathbb{E}[Y_{ig}|D_{ig} = d, S_{ig} = s] - \mathbb{E}[Y_{ig}|D_{ig} = d, S_{ig} = 0]) \\
&= \sum_{s=1}^{n_g} \omega_d(s) \mathbb{E}[Y_{ig}(d, s) - Y_{ig}(d, 0)].
\end{aligned}$$

By linearity, $\mathbb{E}[Y_{ig}(d,s) - Y_{ig}(d,0)] = s\kappa_d$ and thus:

$$
\begin{aligned}
\gamma_\ell^d &= \frac{\kappa_d n_g}{\mathbb{V}[S_{ig}|D_{ig}=d]} \sum_{s=1}^{n_g} s(s - \mathbb{E}[S_{ig}|D_{ig}=d])\mathbb{P}[S_{ig}|D_{ig}=d] \\
&= \kappa_d n_g \\
&= \mathbb{E}[Y_{ig}(d,n_g) - Y_{ig}(d,0)].
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\tilde{\beta}_\ell &= \beta_{\mathsf{D}} - \left(\frac{\gamma_\ell^1}{n_g}\mathbb{E}[S_{ig}|D_{ig}=1] - \frac{\gamma_\ell^0}{n_g}\mathbb{E}[S_{ig}|D_{ig}=0]\right) \\
&= \mathbb{E}[Y_{ig}(1,0) - Y_{ig}(0,0)] + \kappa_1\mathbb{E}[S_{ig}|D_{ig}=1] - \kappa_0\mathbb{E}[S_{ig}|D_{ig}=0] \\
&- \kappa_1\mathbb{E}[S_{ig}|D_{ig}=1] + \kappa_0\mathbb{E}[S_{ig}|D_{ig}=0] \\
&= \mathbb{E}[Y_{ig}(1,0) - Y_{ig}(0,0)]
\end{aligned}
$$

as required. □

Proof of Lemma 2 Take a constant $c \in \mathbb{R}$. Then

$$
\mathbb{P}\left[\min_{\mathbf{a}\in\mathcal{A}_n} N(\mathbf{a}) \leq c\right] \leq |\mathcal{A}_n| \max_{\mathbf{a}\in\mathcal{A}_n} \mathbb{P}[N(\mathbf{a}) \leq c].
$$

Now, for any $\delta > 0$,

$$
\begin{aligned}
\mathbb{P}[N(\mathbf{a}) \leq c] &= \mathbb{P}\left[N(\mathbf{a}) \leq c, \left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \delta\right] + \mathbb{P}\left[N(\mathbf{a}) \leq c, \left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| \leq \delta\right] \\
&\leq \mathbb{P}\left[\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \delta\right] + \mathbb{P}[N(\mathbf{a}) \leq c, G(n+1)\pi(\mathbf{a})(1-\delta) \leq N(\mathbf{a}) \leq \pi(\mathbf{a})G(n+1)(1+\delta)] \\
&\leq \mathbb{P}\left[\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \delta\right] + \mathbb{1}(G(n+1)\pi(\mathbf{a}) \leq c/(1-\delta)) \\
&\leq \mathbb{P}\left[\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \delta\right] + \mathbb{1}(G(n+1)\underline{\pi}_n \leq c/(1-\delta))
\end{aligned}
$$

which implies

$$
|\mathcal{A}_n| \max_{\mathbf{a}\in\mathcal{A}_n} \mathbb{P}[N(\mathbf{a}) \leq c] \leq |\mathcal{A}_n| \max_{\mathbf{a}\in\mathcal{A}_n} \mathbb{P}\left[\left|\frac{\hat{\pi}(\mathbf{a})}{\pi(\mathbf{a})} - 1\right| > \delta\right] + |\mathcal{A}_n|\mathbb{1}(G(n+1)\underline{\pi}_n \leq c/(1-\delta))
$$

which converges to zero under condition 4 and using Lemma A1. □

Proof of Theorem 4 All the estimators below are only defined when $\mathbb{1}(N(\mathbf{a}) > 0)$. Because under the conditions for Lemma 2 this event occurs with probability approaching one, the indicator will be omitted to simplify the notation. Let $\varepsilon_{ig}(\mathbf{a}) = Y_{ig} - \mathbb{E}[Y_{ig}|\mathbf{A}_{ig} = \mathbf{a}]$.

For the consistency part, we have that

$$\frac{\sum_g \sum_i \varepsilon_{ig}(\mathbf{a})\mathbb{1}_{ig(\mathbf{a})}}{N(\mathbf{a})} = \frac{\sum_g \sum_i (\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}| > \xi_n) - \mathbb{E}[\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}| > \xi_n)])\mathbb{1}_{ig(\mathbf{a})}}{N(\mathbf{a})} + \frac{\sum_g \sum_i (\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}| \leq \xi_n) - \mathbb{E}[\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}| \leq \xi_n)])\mathbb{1}_{ig(\mathbf{a})}}{N(\mathbf{a})}$$

for some increasing sequence of constants ξ_n whose rate will be determined along the proof. Let

$$\underline{\varepsilon}_{ig}(\mathbf{a}) = \varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}(\mathbf{a})| \leq \xi_n) - \mathbb{E}[\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}(\mathbf{a})| \leq \xi_n)]$$

and

$$\bar{\varepsilon}_{ig}(\mathbf{a}) = \varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}(\mathbf{a})| > \xi_n) - \mathbb{E}[\varepsilon_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}(\mathbf{a})| > \xi_n)]$$

For the first term,

$$\mathbb{P}\left[\max_{\mathbf{a}\in\mathcal{A}_n}\left|\frac{\sum_g \sum_i \underline{\varepsilon}_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})}\right| > Mr_n \middle| \mathbf{A}\right] \leq |\mathcal{A}_n| \max_{\mathbf{a}\in\mathcal{A}_n} \mathbb{P}\left[\left|\frac{\sum_g \sum_i \underline{\varepsilon}_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})}\right| > Mr_n \middle| \mathbf{A}\right]$$

For the right-hand side, by Bernstein's inequality

$$\begin{aligned}\mathbb{P}\left[\left|\sum_g \sum_i \underline{\varepsilon}_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})\right| > N(\mathbf{a})Mr_n \middle| \mathbf{A}\right] &\leq 2\exp\left\{-\frac{1}{2}\frac{M^2 r_n^2 N(\mathbf{a})^2}{\sigma^2(\mathbf{a})N(\mathbf{a}) + 2\xi_n M r_n N(\mathbf{a})/3}\right\} \\ &= 2\exp\left\{-\frac{1}{2}\frac{M^2 r_n^2 N(\mathbf{a})}{\sigma^2(\mathbf{a}) + 2M\xi_n r_n/3}\right\} \\ &\leq 2\exp\left\{-\frac{1}{2}\frac{M^2 r_n^2 \min_{\mathbf{a}\in\mathcal{A}_n} N(\mathbf{a})}{\bar{\sigma}^2 + 2M\xi_n r_n/3}\right\}\end{aligned}$$

Set

$$r_n = \sqrt{\frac{\log|\mathcal{A}_n|}{G(n+1)\underline{\pi}_n}}$$

Next, use the fact that

$$\frac{\min_{\mathbf{a}\in\mathcal{A}_n} N(\mathbf{a})}{G(n+1)\underline{\pi}_n} \to_{\mathbb{P}} 1$$

which follows from Lemma A1, since

$$\mathbb{P}[\pi(\mathbf{a})(1-\varepsilon) \leq \hat{\pi}(\mathbf{a}) \leq \pi(\mathbf{a})(1+\varepsilon), \forall\, \mathbf{a}] \to 1$$

for any $\varepsilon > 0$ and

$$\begin{aligned}\mathbb{P}[\pi(\mathbf{a})(1-\varepsilon) \leq \hat{\pi}(\mathbf{a}) \leq \pi(\mathbf{a})(1+\varepsilon), \forall\, \mathbf{a}] &\leq \mathbb{P}[\underline{\pi}_n(1-\varepsilon) \leq \min_{\mathbf{a}} \hat{\pi}(\mathbf{a}) \leq \underline{\pi}_n(1+\varepsilon)] \\ &= \mathbb{P}\left[\left|\frac{\min_{\mathbf{a}} \hat{\pi}(\mathbf{a})}{\underline{\pi}_n} - 1\right| \leq \varepsilon\right].\end{aligned}$$

Then,

$$\mathbb{P}\left[\left|\sum_g \sum_i \varepsilon_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})\right| > N(\mathbf{a})Mr_n \middle| \mathbf{A}\right] \leq 2\exp\left\{-\frac{1}{2}\frac{M^2\log|\mathcal{A}_n|(1+o_{\mathbb{P}}(1))}{\bar{\sigma}^2 + 2M\xi_n r_n/3}\right\}$$

and therefore

$$\mathbb{P}\left[\max_{\mathbf{a}\in\mathcal{A}_n}\left|\frac{\sum_g\sum_i \varepsilon_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})}\right| > Mr_n \middle| \mathbf{A}\right] \leq$$
$$2\exp\left\{\log|\mathcal{A}_n|\left(1-\frac{1}{2}\frac{M^2(1+o_{\mathbb{P}}(1))}{\bar{\sigma}^2+2Mr_n\xi_n/3}\right)\right\}$$

which can be made arbitrarily small for sufficiently large M as long as $r_n\xi_n = O(1)$.

For the second term, by Markov's inequality

$$\begin{aligned}\mathbb{P}\left[\left|\sum_g\sum_i \bar{\varepsilon}_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})\right| > N(\mathbf{a})Mr_n\middle|\mathbf{A}\right] &\leq \frac{\mathbb{E}[\varepsilon^2_{ig}(\mathbf{a})\mathbb{1}(|\varepsilon_{ig}(\mathbf{a})|>\xi_n)]N(\mathbf{a})}{M^2r_n^2N(\mathbf{a})^2} \\ &\leq \frac{b^{2+\delta}}{M^2\xi_n^\delta}\frac{1}{r_n^2\min_{\mathbf{a}\in\mathcal{A}_n}N(\mathbf{a})} \\ &= \frac{b^{2+\delta}}{M^2}\frac{1}{\xi_n^\delta\log|\mathcal{A}_n|(1+o_{\mathbb{P}}(1))}\end{aligned}$$

so that

$$\mathbb{P}\left[\max_{\mathbf{a}\in\mathcal{A}_n}\left|\sum_g\sum_i \bar{\varepsilon}_{ig}(\mathbf{a})\mathbb{1}_{ig}(\mathbf{a})\right| > N(\mathbf{a})Mr_n\middle|\mathbf{A}\right] \leq \frac{b^{2+\delta}}{M^2}\frac{|\mathcal{A}_n|}{\xi_n^\delta\log|\mathcal{A}_n|(1+o_{\mathbb{P}}(1))}$$

Finally, set $\xi_n = r_n^{-1}$. Then, the above term can be made arbitrarily small for M sufficiently large, as long as

$$\frac{|\mathcal{A}_n|}{\log|\mathcal{A}_n|}\left(\frac{\log|\mathcal{A}_n|}{G(n+1)\underline{\pi}_n}\right)^{\delta/2} = O(1)$$

Setting $\delta = 2$, this condition reduces to:

$$\frac{|\mathcal{A}_n|}{G(n+1)\underline{\pi}_n} = O(1)$$

Therefore,

$$\max_{\mathbf{a}\in\mathcal{A}_n}|\hat{\mu}(\mathbf{a})-\mu(\mathbf{a})| = O_{\mathbb{P}}\left(\sqrt{\frac{\log|\mathcal{A}_n|}{G(n+1)\underline{\pi}_n}}\right)$$

The proof for the standard error estimator uses the same reasoning after replacing $\varepsilon_{ig}(\mathbf{a})$ by $\hat{\varepsilon}^2_{ig}(\mathbf{a})$ and using consistency of $\hat{\mu}(\mathbf{a})$.

For the second part, we want to bound

$$\Delta = \max_{\mathbf{a}\in\mathcal{A}_\mathbf{n}}\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[\frac{\hat{\mu}(\mathbf{a})-\mu(\mathbf{a})}{\sqrt{\mathbb{V}[\hat{\mu}(\mathbf{a})|\mathbf{A}]}}\leq x\right]-\Phi(x)\right|$$

$$\begin{aligned}
\Delta &= \max_{\mathbf{a}\in\mathcal{A}_{\mathbf{n}}} \sup_{x\in\mathbb{R}} \left| \mathbb{P}\left[\frac{\hat{\mu}(\mathbf{a}) - \mu(\mathbf{a})}{\sqrt{\mathbb{V}[\hat{\mu}(\mathbf{a})|\mathbf{A}]}} \leq x \right] - \Phi(x) \right| \\
&= \max_{\mathbf{a}\in\mathcal{A}_{\mathbf{n}}} \sup_{x\in\mathbb{R}} \left| \mathbb{E}\left\{ \mathbb{P}\left[\frac{\hat{\mu}(\mathbf{a}) - \mu(\mathbf{a})}{\sqrt{\mathbb{V}[\hat{\mu}(\mathbf{a})|\mathbf{A}]}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right\} \right| \\
&\leq \mathbb{E}\left\{ \max_{\mathbf{a}\in\mathcal{A}_{\mathbf{n}}} \sup_{x\in\mathbb{R}} \left| \mathbb{P}\left[\frac{\hat{\mu}(\mathbf{a}) - \mu(\mathbf{a})}{\sqrt{\mathbb{V}[\hat{\mu}(\mathbf{a})|\mathbf{A}]}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right| \right\}
\end{aligned}$$

Then,

$$\left| \mathbb{P}\left[\frac{\hat{\mu}(\mathbf{a}) - \mu(\mathbf{a})}{\sqrt{\mathbb{V}[\hat{\mu}(\mathbf{a})|\mathbf{A}]}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right| = \left| \mathbb{P}\left[\frac{\sum_g \sum_i \varepsilon_{ig} \mathbb{1}_{ig}(\mathbf{a})}{\sigma(\mathbf{a})\sqrt{N(\mathbf{a})}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right|$$

By the Berry-Esseen bound,

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left[\frac{\sum_g \sum_i \varepsilon_{ig} \mathbb{1}_{ig}(\mathbf{a})}{\sigma(\mathbf{a})\sqrt{N(\mathbf{a})}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right| \leq \frac{Cb^3}{\underline{\sigma}^3} \cdot \frac{1}{\sqrt{N(\mathbf{a})}}$$

But

$$\frac{1}{N(\mathbf{a})} = O_{\mathbb{P}}\left(\frac{1}{G(n+1)\pi(\mathbf{a})} \right)$$

Therefore,

$$\max_{\mathbf{a}\in\mathcal{A}_n} \sup_{x\in\mathbb{R}} \left| \mathbb{P}\left[\frac{\sum_g \sum_i \varepsilon_{ig} \mathbb{1}_{ig}(\mathbf{a})}{\sigma(\mathbf{a})\sqrt{N(\mathbf{a})}} \leq x \middle| \mathbf{A} \right] - \Phi(x) \right| \leq \frac{Cb^3}{\underline{\sigma}^3} \cdot O_{\mathbb{P}}\left(\frac{1}{\sqrt{G(n+1)\underline{\pi}_n}} \right)$$

and the result follows. □

Proof of Theorem 5 We want to bound:

$$\Delta^*(\mathbf{a}) = \sup_x \left| \mathbb{P}^*\left[\frac{\hat{\mu}^*(\mathbf{a}) - \hat{\mu}(\mathbf{a})}{\sqrt{\mathbb{V}^*[\hat{\mu}(\mathbf{a})]}} \leq x \right] - \Phi(x) \right|$$

uniformly over $\mathbf{a}$, where

$$\hat{\mu}^*(\mathbf{a}) = \frac{\sum_g \sum_i Y_{ig}^* \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})}$$

if the denominator is non-zero, and zero otherwise, and where

$$Y_{ig}^* \mathbb{1}_{ig}(\mathbf{a}) = (\bar{Y}(\mathbf{a}) + (Y_{ig} - \bar{Y}(\mathbf{a}))w_{ig})\mathbb{1}_{ig}(\mathbf{a}) = (\bar{Y}(\mathbf{a}) + \hat{\varepsilon}_{ig} w_{ig})\mathbb{1}_{ig}(\mathbf{a})$$

Then, if $N(\mathbf{a}) > 0$,

$$\mathbb{E}^*[\hat{\mu}^*(\mathbf{a})] = \hat{\mu}(\mathbf{a})$$
$$\mathbb{V}^*[\hat{\mu}^*(\mathbf{a})] = \frac{\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})^2}$$

The centered and scaled statistic is given by:

$$\frac{\sum_g \sum_i \hat{\varepsilon}_{ig} \mathbb{1}_{ig}(\mathbf{a}) w_{ig}}{\sqrt{\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a})}}$$

By Berry-Esseen,

$$\sup_x \left| \mathbb{P}^* \left[\frac{\sum_g \sum_i \hat{\varepsilon}_{ig} \mathbb{1}_{ig}(\mathbf{a}) w_{ig}}{\sqrt{\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a})}} \le x \right] - \Phi(x) \right| \le C \frac{\sum_g \sum_i |\hat{\varepsilon}_{ig}|^3 \mathbb{1}_{ig}(\mathbf{a}) / N(\mathbf{a})}{\left(\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a}) / N(\mathbf{a}) \right)^{3/2}} \cdot \frac{1}{\sqrt{N(\mathbf{a})}}$$

We also have that

$$\begin{aligned}
\frac{\sum_g \sum_i |\hat{\varepsilon}_{ig}|^3 \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})} &\le \frac{\sum_g \sum_i |Y_{ig} - \mu(\mathbf{a})|^3 \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})} + |\bar{Y}(\mathbf{a}) - \mu(\mathbf{a})|^3 + O_{\mathbb{P}}(N(\mathbf{a})^{-2}) \\
&= \mathbb{E}[|Y_{ig} - \mu(\mathbf{a})|^3] + O_{\mathbb{P}}(N(\mathbf{a})^{-1})
\end{aligned}$$

and

$$\begin{aligned}
\frac{\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})} &= \frac{\sum_g \sum_i (Y_{ig} - \mu(\mathbf{a}))^2 \mathbb{1}_{ig}(\mathbf{a})}{N(\mathbf{a})} + (\bar{Y}(\mathbf{a}) - \mu(\mathbf{a}))^2 \\
&= \sigma^2(\mathbf{a}) + O_{\mathbb{P}}(N(\mathbf{a})^{-1})
\end{aligned}$$

Then,

$$\begin{aligned}
\Delta^*(\mathbf{a}) &\le \sup_x \left| \mathbb{P}^* \left[\frac{\sum_g \sum_i \hat{\varepsilon}_{ig} \mathbb{1}_{ig}(\mathbf{a}) w_{ig}}{\sqrt{\sum_g \sum_i \hat{\varepsilon}_{ig}^2 \mathbb{1}_{ig}(\mathbf{a})}} \le x \right] - \Phi(x) \right| \mathbb{1}(N(\mathbf{a}) > 0) + 2\mathbb{1}(N(\mathbf{a}) = 0) \\
&= C \frac{\mathbb{E}[|Y_{ig} - \mu(\mathbf{a})|^3] + O_{\mathbb{P}}(N(\mathbf{a})^{-1})}{[\sigma^2(\mathbf{a}) + O_{\mathbb{P}}(N(\mathbf{a})^{-1}))]^{3/2}} \cdot \frac{\mathbb{1}(N(\mathbf{a}) > 0)}{\sqrt{N(\mathbf{a})}} + 2\mathbb{1}(N(\mathbf{a}) = 0)
\end{aligned}$$

and the result follows from the facts that $\mathbb{P}[\min_{\mathbf{a}} N(\mathbf{a}) = 0] \to 0$ and by Lemma A1. $\square$

References

Angrist, J. D. (2014): "The perils of peer effects," *Labour Economics*, 30, 98–108.

Davezies, L., X. D'Haultfoeuille, and D. Fougère (2009): "Identification of peer effects using group size variation," *Econometrics Journal*, 12(3), 397–413.

Kline, B., and E. Tamer (2019): "Econometric Analysis of Models with Social Interactions," in *The Econometric Analysis of Network Data*, ed. by B. S. Graham, and A. De Paula. Elsevier.

Manski, C. F. (2013a): "Comment," *Journal of Business & Economic Statistics*, 31(3), 273–275.

——— (2013b): "Identification of treatment response with social interactions," *The Econometrics Journal*, 16(1), S1–S23.