Supplementary Appendix for:

Naoki Egami. "Spillover Effects in the Presence of Unobserved Networks." *Political Analysis*

A Details of Setup

We describe regularity conditions for the support of treatment exposure probabilities to ensure well-defined causal estimands.

The required regularity conditions are as follows: (1) the support of $\Pr(G_i = g, U_i = u|T_i = 1)$ is equal to the support of $\Pr(G_i = g, U_i = u|T_i = 0)$ for all i, and (2) the support of $\Pr(U_i = u|T_i = d, G_i = g^H)$ is equal to the support of $\Pr(U_i = u|T_i = d, G_i = g^L)$ for all i. We discuss them in order.

When we define the unit level direct effect, we avoid ill-defined causal effects by focusing on settings where the support of $\Pr(G_i = g, U_i = u \mid T_i = 1)$ is equal to the support of $\Pr(G_i = g, U_i = u \mid T_i = 0)$ for all i. This can be violated when the total number of treated units is small so that for some (g, u), $\Pr(G = g, U = u|T_i = 1) = 0$ and $\Pr(G = g, U = u|T_i = 0) > 0$. One extreme example is that when we use complete randomization with the total number of treated units equal to 1. In this case, whenever $T_i = 1$, $\Pr(G = g, U = u|T_i = 1) = 0$ for all (g, u), but when $T_i = 0$, $\Pr(G = g, U = u|T_i = 0) > 0$ for some (g, u). Another extreme example is that the total number of treated units is too large. For example, when we use complete randomization with the total number of treated units equal to $N - 1$. In this case, whenever $T_i = 0$, $\Pr(G = g, U = u|T_i = 0) = 0$ for all (g, u) except for $(g, u) = (1, 1)$, but when $T_i = 1$, $\Pr(G = g, U = u|T_i = 1) > 0$ for some (g, u) other than $(g, u) = (1, 1)$. It is clear that when researchers use a Bernoulli design, the support of $\Pr(G_i = g, U_i = u \mid T_i = 1)$ is equal to the support of $\Pr(G_i = g, U_i = u \mid T_i = 0)$ for all i.

When we define the unit level network-specific spillover effect, we avoid ill-defined causal effects by focusing on settings where $\Pr(U_i = u \mid T_i = d, G_i = g^H)$ and $\Pr(U_i = u \mid T_i = d, G_i = g^L)$ have the same support for all i. This requires that g^H and g^L are small enough so that the distribution over the fraction of treated neighbors in network $\mathcal{U}$ is not restricted, especially $\Pr(U_i = 0 \mid T_i = d, G_i = g^H) > 0$ and $\Pr(U_i = 0 \mid T_i = d, G_i = g^L) > 0$ for all i. Formally, $g^H, g^L \leq g_s$ where $g_s \equiv \min_i \{1 - |\mathcal{N}_i^{(\mathcal{G},\mathcal{U})}|/|\mathcal{N}_i^{\mathcal{G}}|\}$. The desired support condition can be violated when the total number of treated units is too small so that for some u, $\Pr(U = u | T_i = 1, G_i = g^H) = 0$ and $\Pr(U = u | T_i = 1, G_i = g^L) > 0$. One extreme example is that when we use complete randomization with the total number of treated units equal to $1 + g^H \times |\mathcal{N}_i^{\mathcal{G}}|$. In this case, whenever $G_i = g^H$, $\Pr(U = u | T_i = 1, G_i = g^H) = 0$ for all $u > |\mathcal{N}_i^{(\mathcal{G},\mathcal{U})}| / |\mathcal{N}_i^{\mathcal{U}}|$, but when $G_i = g^L < g^H$, $\Pr(U = u | T_i = 0, G_i = g^L) > 0$ for some $u > |\mathcal{N}_i^{(\mathcal{G},\mathcal{U})}| / |\mathcal{N}_i^{\mathcal{U}}|$. Finally, it is clear that $\Pr(U_i = u \mid T_i = d, G_i = g^H)$ and $\Pr(U_i = u \mid T_i = d, G_i = g^L)$ have the same support for all i if researchers use a Bernoulli design and $g^H, g^L \leq g_s$.

B Connection between Total Spillover Effects and Network-Specific Spillover Effects

Here, we connect the ANSE to the popular estimand in the literature. In particular, we show that the ANSE can be seen as the decomposition of the average total spillover effect (Hudgens and Halloran, 2008).

First, by extending Hudgens and Halloran (2008) to settings with multiple networks, the individual average potential outcome are defined as follows.

$$\overline{Y}_i(d, g) \equiv \sum_{u \in \Delta_i^u} Y_i(d, g, u) \Pr(U_i = u \mid T_i = d, G_i = g), \tag{A1}$$

where the potential outcome of individual i is averaged over the conditional distribution of the treatment assignment $\Pr(U_i = u \mid T_i = d, G_i = g)$. Here, the individual average potential outcome represents the expected outcome of unit i when she receives the direct treatment d and the treated proportion g in network $\mathcal{G}$. Taking the difference in the two individual average potential outcomes, the *average total spillover effect* (ATSE) in network $\mathcal{G}$ is defined as follows (Halloran and Hudgens, 2016).[8]

$$\psi(g^H, g^L; d) \equiv \frac{1}{N} \sum_{i=1}^{N} \{\overline{Y}_i(d, g^H) - \overline{Y}_i(d, g^L)\}. \tag{A2}$$

This causal quantity is the *total* spillover effect of changing the treated proportion in network $\mathcal{G}$ from g^L to g^H as the following decomposition of the ATSE demonstrates.

$$\psi(g^H, g^L; d) = \tau(g^H, g^L; d) + \tag{A3}$$

[8] This quantity is called the average indirect causal effect in Hudgens and Halloran (2008). We define it as the average total spillover effect to clarify how it combines multiple network-specific spillover effects.

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\sum_{u\in\Delta_i^u}\{Y_i(d,g^H,u)-Y_i(d,g^H,u')\}\{\Pr(U_i=u\mid T_i=d,G_i=g^H)-\Pr(U_i=u\mid T_i=d,G_i=g^L)\}\right\},$$

for any $u' \in \Delta_i^u$. The first term is the ANSE in network $\mathcal{G}$ (Definition 3), which quantifies the spillover effect specific to network $\mathcal{G}$. The second term represents the spillover effect in $\mathcal{U}$, $Y_i(d, g^H, u) - Y_i(d, g^H, u')$, weighted by the change in the conditional distribution of U_i due to the change in G_i, $\Pr(U_i = u \mid T_i = d, G_i = g^H) - \Pr(U_i = u \mid T_i = d, G_i = g^L)$. This is because U_i, the treated proportion of neighbors in the other network $\mathcal{U}$, is not fixed at constant and thus, they change as G_i, the treated proportion of neighbors in $\mathcal{G}$, changes. Thus, the ATSE captures the sum of the spillover effect specific to network $\mathcal{G}$ and the spillover effect specific to $\mathcal{U}$ induced by the change in U_i associated with the change in G_i. For example, the ATSE of changing from g^L to g^H on the Facebook network captures two spillover effects together; (1) the spillover effect specific to the Facebook and (2) the spillover effect in the face-to-face network. This is because the treated proportion in the offline network U_i is associated with the change in the treated proportion in the Facebook network G_i. We discuss this issue in further details when we derive the exact bias formula in Section 3. When network $\mathcal{U}$, such as the offline network, is causally irrelevant, the ATSE is equal to the ANSE in the Facebook network, but in general, the two estimands do not coincide.

While both the ATSE and the ANSE quantify spillover effects, their substantive meanings differ. The ATSE is useful when researchers wish to know the total amount of spillover effects that result from interventions on an observed network. For instance, politicians decided to run online campaigns on Twitter and want to estimate the total amount of spillover effects they can induce by their Twitter messages. These politicians might not be interested in distinguishing whether the spillover effects arise through Twitter or through unobserved face-to-face interactions. Thus, the ATSE is of relevance when the target network is predetermined and the mechanism can be ignored.

In contrast, the ANSE is essential for disentangling different channels through which spillover effects arise. It is the main quantity of interest when researchers wish to examine the causal role of individual networks or to discover the most causally relevant network to target. For example, it is of scientific interest to distinguish how much spillover effects arise through the Twitter network or through offline communications. By estimating the ANSE, researchers can learn about the importance of online human interactions.

C Proofs

This section provides proofs for all theorems in the paper.

C.1 Proof of Theorem 1

C.1.1 ADE

First, we rewrite the estimator with the standard IPW representation.

$$
\begin{aligned}
&\widehat{\delta} \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=1\}\widetilde{w}_i Y_i - \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=0\}\widetilde{w}_i Y_i \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\sum_{(g,u)\in\Delta_i^{gu}}\Pr(G_i=g,U_i=u)\left\{\frac{\mathbf{1}\{T_i=1,G_i=g,U_i=u\}Y_i}{\Pr(T_i=1,G_i=g,U_i=u)} - \frac{\mathbf{1}\{T_i=0,G_i=g,U_i=u\}Y_i}{\Pr(T_i=0,G_i=g,U_i=u)}\right\}.
\end{aligned}
$$

Then, the theorem follows from the standard proof for the IPW estimator.

$$
\begin{aligned}
&\mathbb{E}[\hat{\delta}] \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\sum_{(g,u)\in\Delta_i^{gu}}\Pr(G_i=g,U_i=u)\times \\
& \qquad \left\{\frac{\mathbb{E}[\mathbf{1}\{T_i=1,G_i=g,U_i=u\}Y_i]}{\Pr(T_i=1,G_i=g,U_i=u)} - \frac{\mathbb{E}[\mathbf{1}\{T_i=0,G_i=g,U_i=u\}Y_i]}{\Pr(T_i=0,G_i=g,U_i=u)}\right\} \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\sum_{(g,u)\in\Delta_i^{gu}}\Pr(G_i=g,U_i=u)\times \\
& \qquad \left\{\frac{\Pr(T_i=1,G_i=g,U_i=u)Y_i(1,g,u)}{\Pr(T_i=1,G_i=g,U_i=u)} - \frac{\Pr(T_i=0,G_i=g,U_i=u)Y_i(0,g,u)}{\Pr(T_i=0,G_i=g,U_i=u)}\right\} \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\sum_{(g,u)\in\Delta_i^{gu}}\Pr(G_i=g,U_i=u)\{Y_i(1,g,u)-Y_i(0,g,u)\} \\
=\ & \delta
\end{aligned}
$$

where the second equality follows from the consistency of potential outcomes. □

C.1.2 ANSE

First, we rewrite the estimator with the standard IPW representation.

$$
\begin{aligned}
&\widehat{\tau} \\
=\ & \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=d,G_i=g^H\}w_i Y_i - \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=d,G_i=g^L\}w_i Y_i
\end{aligned}
$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u} \Pr(U_i = u \mid T_i = d, G_i = g^L) \times \left\{\frac{\mathbf{1}\{T_i = d, G_i = g^H, U_i = u\}Y_i}{\Pr(T_i = d, G_i = g^H, U_i = u)} - \frac{\mathbf{1}\{T_i = d, G_i = g^L, U_i = u\}Y_i}{\Pr(T_i = d, G_i = g^L, U_i = u)}\right\},$$

Then, the theorem follows from the standard proof for the IPW estimator.

$$\begin{aligned}
& \mathbb{E}[\hat{\tau}(g, g'; d)] \\
= & \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u} \Pr(U_i = u \mid T_i = d, G_i = g^L) \times \\
& \left\{\frac{\mathbb{E}[\mathbf{1}\{T_i = d, G_i = g^H, U_i = u\}Y_i]}{\Pr(T_i = d, G_i = g^H, U_i = u)} - \frac{\mathbb{E}[\mathbf{1}\{T_i = d, G_i = g^L, U_i = u\}Y_i]}{\Pr(T_i = d, G_i = g^L, U_i = u)}\right\} \\
= & \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u} \Pr(U_i = u \mid T_i = d, G_i = g^L) \times \\
& \left\{\frac{\Pr(T_i = d, G_i = g^H, U_i = u)Y_i(d, g^H, u)}{\Pr(T_i = d, G_i = g^H, U_i = u)} - \frac{\Pr(T_i = d, G_i = g^L, U_i = u)Y_i(d, g^L, u)}{\Pr(T_i = d, G_i = g^L, U_i = u)}\right\} \\
= & \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u} \Pr(U_i = u \mid T_i = d, G_i = g^L)\{Y_i(d, g^H, u) - Y_i(d, g^L, u)\} \\
= & \tau(g^H, g^L; d),
\end{aligned}$$

which completes the proof. □

C.2 Proof of Theorem 2

The expectation of an estimator $\hat{\tau}_B(g^H, g^L; d)$ is

$$\begin{aligned}
& \mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] = \frac{1}{N}\sum_{i=1}^{N}\left\{\mathbb{E}[Y_i \mid T_i = d, G_i = g^H] - \mathbb{E}[Y_i \mid T_i = d, G_i = g^L]\right\} \\
= & \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u}\left\{Y_i(d, g^H, u)\Pr(U_i = u \mid T_i = d, G_i = g^H) - Y_i(d, g^L, u)\Pr(U_i = u \mid T_i = d, G_i = g^L)\right\}.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
& \mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] - \tau(g^H, g^L; d) \\
= & \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u}\Big\{Y_i(d, g^H, u)\{\Pr(U_i = u \mid T_i = d, G_i = g^H) - \Pr(U_i = u \mid T_i = d, G_i = g^L)\} \\
& -Y_i(d, g^L, u)\{\Pr(U_i = u \mid T_i = d, G_i = g^L) - \Pr(U_i = u \mid T_i = d, G_i = g^L)\}\Big\}
\end{aligned}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sum_{u\in\Delta_i^u}\Bigg\{\{Y_i(d,g^H,u)-Y_i(d,g^H,u')\} \times\{\Pr(U_i=u\mid T_i=d,G_i=g^H)-\Pr(U_i=u\mid T_i=d,G_i=g^L)\}\Bigg\}.$$

for any $u'\in\Delta^u$. □

C.2.1 Lemma: Bias in ADE

First, we can rewrite the estimator as the standard IPW estimator.

$$\begin{aligned}\widehat{\delta}_B &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=1\}\widetilde{w}_i^{\mathsf{B}}Y_i-\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{T_i=0\}\widetilde{w}_i^{\mathsf{B}}Y_i\\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g}\Pr(G_i=g)\left\{\frac{\mathbf{1}\{T_i=1,G_i=g\}Y_i}{\Pr(T_i=1,G_i=g)}-\frac{\mathbf{1}\{T_i=0,G_i=g\}Y_i}{\Pr(T_i=0,G_i=g)}\right\}.\end{aligned}$$

We have the following equality for any g,

$$\begin{aligned}&\mathbb{E}[\mathbf{1}\{T_i=d,G_i=g\}Y_i]\\ =\ &\mathbb{E}[\sum_{u\in\Delta_i^u(g)}\mathbf{1}\{T_i=d,G_i=g,U_i=u\}Y_i(d,g,u)]\\ =\ &\sum_{u\in\Delta_i^u(g)}\Pr(T_i=d,G_i=g,U_i=u)Y_i(d,g,u)\end{aligned}$$

where $\Delta_i^u(g)$ is the support $\{u:\Pr(U_i=u\mid G_i=g)>0\}$. Therefore, the expectation of $\widehat{\delta}_B$ is

$$\begin{aligned}&\mathbb{E}[\widehat{\delta}_B]\\ =\ &\frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g}\Pr(G_i=g)\left\{\frac{\mathbb{E}[\mathbf{1}\{T_i=1,G_i=g\}Y_i]}{\Pr(T_i=1,G_i=g)}-\frac{\mathbb{E}[\mathbf{1}\{T_i=0,G_i=g\}Y_i]}{\Pr(T_i=0,G_i=g)}\right\}\\ =\ &\frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g}\Pr(G_i=g)\Bigg\{\\ &\quad\sum_{u\in\Delta_i^u(g)}\left\{\frac{\Pr(T_i=1,G_i=g,U_i=u)Y_i(1,g,u)}{\Pr(T_i=1,G_i=g)}-\frac{\Pr(T_i=0,G_i=g,U_i=u)Y_i(0,g,u)}{\Pr(T_i=0,G_i=g)}\right\}\Bigg\}\\ =\ &\frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g}\Pr(G_i=g)\Bigg\{\\ &\quad\sum_{u\in\Delta_i^u(g)}\Big\{Y_i(1,g,u)\Pr(U_i=u\mid T_i=1,G_i=g)-Y_i(0,g,u)\Pr(U_i=u\mid T_i=0,G_i=g)\Big\}\Bigg\}.\end{aligned}$$

Then, we have

$$
\begin{aligned}
& \mathbb{E}[\hat{\delta}_B] - \delta \\
= \; & \frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g} \Pr_i(G_i = g)\Bigg\{\sum_{u\in\Delta_i^u} Y_i(1,g,u)\{\Pr_i(U_i = u \mid T_i = 1, G_i = g) - \Pr_i(U_i = u \mid G_i = g)\} \\
& \quad - \sum_{u\in\Delta_i^u} Y_i(0,g',u)\{\Pr(U_i = u \mid T_i = 0, G_i = g) - \Pr(U_i = u \mid G_i = g)\}\Bigg\} \\
= \; & \frac{1}{N}\sum_{i=1}^{N}\sum_{g\in\Delta_i^g} \Pr(G_i = g)\Bigg\{ \\
& \quad \sum_{u\in\Delta_i^u(g)} \{Y_i(1,g,u) - Y_i(1,g,u')\}\{\Pr(U_i = u \mid T_i = 1, G_i = g) - \Pr(U_i = u \mid G_i = g)\} \\
- \; & \quad \sum_{u\in\Delta_i^u(g)} \{Y_i(0,g',u) - Y_i(0,g',u')\}\{\Pr(U_i = u \mid T_i = 0, G_i = g) - \Pr(U_i = u \mid G_i = g)\}\Bigg\},
\end{aligned}
$$

which completes the proof. □

C.3 Proof of Theorem 3

Using Theorem 2, under Assumption 3,

$$
\begin{aligned}
& \mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] - \tau(g^H, g^L; d) \\
= \; & \lambda \times \frac{1}{N}\sum_{i=1}^{N}\{\mathbb{E}[U_i = u \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L]\}.
\end{aligned}
$$

From here, we focus on $\mathbb{E}[U_i \mid T_i = d, G_i = g^H]$. For notational simplicity, we use $n_G(i)$ to denote the number of neighbors in the network $\mathcal{G}$ for individual i and $n_U(i)$ is similarly defined. Also, for individual i, let $\pi_{GU}(i)$ be the fraction of the neighbors in $\mathcal{U}$ who are neighbors in $\mathcal{G}$ as well. Formally, $n_G(i) = |\mathcal{N}_i^{\mathcal{G}}|, n_U(i) = |\mathcal{N}_i^{\mathcal{U}}|$ and $\pi_{GU}(i) = |\mathcal{N}_i^{(\mathcal{G},\mathcal{U})}|/|\mathcal{N}_i^{\mathcal{U}}|$.

First, we consider Bernoulli randomization with probability p. Under this setting,

$$
\mathbb{E}[U_i \mid T_i = t, G_i = g^H] = \pi_{GU}(i) \times g^H + (1 - \pi_{GU}(i)) \times p.
$$

Therefore, we have

$$
\begin{aligned}
& \mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] - \tau(g^H, g^L; d) \\
= \; & \lambda \times \frac{1}{N}\sum_{i=1}^{N}\{\mathbb{E}[U_i = u \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L]\}.
\end{aligned}
$$

$$
\begin{aligned}
&= \lambda \times \frac{1}{N}\sum_{i=1}^{N} \pi_{GU}(i) \times (g^H - g^L) \\
&= \lambda \times \pi_{GU} \times (g^H - g^L).
\end{aligned}
$$

where the final equality follows from the definition of π_{GU}.

Next, we consider complete randomization with the number of treated units K. Under this setting,

$$
\begin{aligned}
& \mathbb{E}[U_i \mid T_i = d, G_i = g^H] \\
= \;& \frac{n_U(i) \times \pi_{GU}(i) \times g^H}{n_U(i)} + (1 - \pi_{GU}(i))\frac{K - d - n_G(i) \times g^H}{N - 1 - n_G(i)} \\
= \;& \pi_{GU}(i) \times g^H + (1 - \pi_{GU}(i))\frac{K - d - n_G(i) \times g^H}{N - 1 - n_G(i)} \\
= \;& \left\{\pi_{GU}(i) - \frac{n_G(i)}{N - 1 - n_G(i)}(1 - \pi_{GU}(i))\right\} g^H + \frac{K - d}{N - 1 - n_G(i)}(1 - \pi_{GU}(i))
\end{aligned}
$$

When N is much larger than $n_G(i)$, $n_G(i)/(N - 1 - n_G(i)) \approx 0$. Then, we have

$$
\begin{aligned}
&\mathbb{E}[U_i \mid T_i = d, G_i = g^H] \approx \pi_{GU}(i) g^H + \frac{K - d}{N - 1 - n_G(i)}(1 - \pi_{GU}(i)), \\
&\mathbb{E}[U_i \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L] \approx \pi_{GU}(i)(g^H - g^L).
\end{aligned}
$$

Therefore, when N is much larger than $n_G(i)$ for all i, we get the simplified bias formula.

$$
\begin{aligned}
& \mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] - \tau(g^H, g^L; d) \\
= \;& \lambda \times \frac{1}{N}\sum_{i=1}^{N}\{\mathbb{E}[U_i = u \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L]\}. \\
\approx \;& \lambda \times \frac{1}{N}\sum_{i=1}^{N} \pi_{GU}(i) \times (g^H - g^L) \\
= \;& \lambda \times \pi_{GU} \times (g^H - g^L).
\end{aligned}
$$

Finally, we consider a situation when N is not large enough to have the aforementioned approximation. Suppose $N \approx (C + 1)n_G(i) + 1$ for all i. Then,

$$
\begin{aligned}
&\mathbb{E}[U_i \mid T_i = d, G_i = g^H] \approx \left\{\frac{C + 1}{C}\pi_{GU}(i) - \frac{1}{C}\right\} \times g^H + \frac{K - d}{N - 1 - n_G(i)}(1 - \pi_{GU}(i)), \\
&\mathbb{E}[U_i \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L] \approx \left\{\frac{C + 1}{C}\pi_{GU}(i) - \frac{1}{C}\right\}(g^H - g^L).
\end{aligned}
$$

Therefore, the bias can be written as,

$$
\mathbb{E}[\hat{\tau}_B(g^H, g^L; d)] - \tau(g^H, g^L; d)
$$

$$
\begin{aligned}
&= \lambda \times \frac{1}{N}\sum_{i=1}^{N}\{\mathbb{E}[U_i = u \mid T_i = d, G_i = g^H] - \mathbb{E}[U_i = u \mid T_i = d, G_i = g^L]\}. \\
&\approx \lambda \times \{\frac{C+1}{C} \times \frac{1}{N}\sum_{i=1}^{N}\pi_{GU}(i) - \frac{1}{C}\} \times (g^H - g^L). \\
&= \lambda \times \{\frac{C+1}{C}\pi_{GU} - \frac{1}{C}\} \times (g^H - g^L). \qquad \text{(A4)}
\end{aligned}
$$

□

C.4 Proof of Theorem 4

First, we set the following notations. We define the support Δ_s^u to be the support Δ_i^u for all i with $S_i = s$. We drop subscript s whenever it is obvious from contexts. For $\bar{g} \in \{g^H, g^L\}$,

$$
\begin{aligned}
r_{\bar{g}}(u) &\equiv \frac{1}{N}\sum_{i:S_i=s} Y_i(d, \bar{g}, u) \\
v_{g^H}(\bar{g}) &\equiv \frac{\sum_{u\in\Delta_s^u}\{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\}\ \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\max_u r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)} \\
v_{g^L}(\bar{g}) &\equiv \frac{\sum_{u\in\Delta_s^u}\{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\}\Pr(U_i = u \mid T_i = d, G_i = g^L)}{\max_u r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)} \\
\Gamma(\bar{g}) &\equiv \frac{v_{g^H}(\bar{g})}{v_{g^L}(\bar{g})} = \frac{\sum_{u\in\Delta_s^u}\{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\}\Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u\in\Delta_s^u}\{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\}\Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
MR^{obs}(g^H, g^L; s) &\equiv \frac{\sum_{u\in\Delta_s^u} r_{g^H}(u)\Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u\in\Delta_s^u} r_{g^L}(u)\Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
MR_{\bar{g}}^{true}(g^H, g^L; s) &\equiv \frac{\sum_{u\in\Delta_s^u} r_{g^H}(u)\Pr(U_i = u \mid T_i = d, G_i = \bar{g})}{\sum_{u\in\Delta_s^u} r_{g^L}(u)\Pr(U_i = u \mid T_i = d, G_i = \bar{g})}
\end{aligned}
$$

where $0 \leq v_{g^H}(\bar{g}), v_{g^L}(\bar{g}) \leq 1$ because of non-negative outcomes.

Lemma 1 For (g^H, g^L),

$$
\begin{aligned}
\frac{MR^{obs}(g^H, g^L; s)}{MR_{g^L}^{true}(g^H, g^L; s)} &\leq B \qquad & \frac{MR^{obs}(g^H, g^L; s)}{MR_{g^H}^{true}(g^H, g^L; s)} &\leq B, \\
\frac{MR^{obs}(g^L, g^H; s)}{MR_{g^L}^{true}(g^L, g^H; s)} &\leq B \qquad & \frac{MR^{obs}(g^L, g^H; s)}{MR_{g^H}^{true}(g^L, g^H; s)} &\leq B.
\end{aligned}
$$

Proof This proof closely follows Ding and VanderWeele (2016). The key difference is that we study bias due to an unmeasured relevant network in the presence of interference in multiple networks in contrary to bias due to an unmeasured confounder in observational studies without interference (Ding and VanderWeele, 2016).

For $\bar{g} \in \{g^H, g^L\}$ and s,

$$\begin{aligned}
\Gamma(\bar{g}) &= \frac{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
&= \frac{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \frac{\Pr(U_i = u \mid T_i = d, G_i = g^H)}{\Pr(U_i = u \mid T_i = d, G_i = g^L)} \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \ \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
&\leq \mathrm{RR}_{GU}
\end{aligned}$$

Also, for $\bar{g} \in \{g^H, g^L\}$ and s,

$$\begin{aligned}
\frac{1}{\Gamma(\bar{g})} &= \frac{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^H)} \\
&= \frac{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \frac{\Pr(U_i = u \mid T_i = d, G_i = g^L)}{\Pr(U_i = u \mid T_i = d, G_i = g^H)} \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} \{r_{\bar{g}}(u) - \min_u r_{\bar{g}}(u)\} \ \Pr(U_i = u \mid T_i = d, G_i = g^H)} \\
&\leq \mathrm{RR}_{GU}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\frac{MR^{obs}(g^H, g^L; s)}{MR^{true}_{g^L}(g^H, g^L; s)} \\
= &\frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \times \frac{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} r_{g^H}(U) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
= &\frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
= &\frac{\{\max_u r_{g^H}(u) - \min_u r_{g^H}(u)\} v_{g^H}(g^H) + \min_u r_{g^H}(u)}{\{\max_u r_{g^H}(u) - \min_u r_{g^H}(u)\} \frac{v_{g^H}(g^H)}{\Gamma(g^H)} + \min_u r_{g^H}(u)}
\end{aligned}$$

From Lemma A.1 in Ding and VanderWeele (2016), when $\Gamma(g^H) > 1$, $\frac{\mathrm{MR}^{obs}(g^H, g^L; s)}{\mathrm{MR}^{true}_{g^L}(g^H, g^L; s)}$ is increasing in $v_{g^H}(g^H)$. Therefore, it takes the maximum value when $v_{g^H}(g^H) = 1$.

$$\begin{aligned}
\frac{\mathrm{MR}^{obs}(g^H, g^L; s)}{\mathrm{MR}^{true}_{g^L}(g^H, g^L; s)} &\leq \frac{\Gamma(g^H) \times \mathrm{MR}_{UY}(g^H, s)}{\Gamma(g^H) + \mathrm{MR}_{UY}(g^H, s) - 1} \\
&\leq \frac{\mathrm{RR}_{GU} \times \mathrm{MR}_{UY}}{\mathrm{RR}_{GU} + \mathrm{MR}_{UY} - 1}
\end{aligned}$$

where the second inequality comes from Lemma A.2 in Ding and VanderWeele (2016) and $\Gamma(g^H) \leq \mathrm{RR}_{GU}, \mathrm{MR}_{UY} = \max_{g,s} \mathrm{MR}_{UY}(g, s)$.

From Lemma A.1 in Ding and VanderWeele (2016), when $\Gamma(g^H) \leq 1$, $\frac{\mathrm{MR}^{obs}(g^H, g^L; s)}{\mathrm{MR}^{true}_{g^L}(g^H, g^L; s)}$ is non-increasing in $v_{g^H}(g^H)$. Therefore, it takes the maximum value at $v_{g^H}(g^H) = 0$.

$$\frac{\mathrm{MR}^{obs}(g^H, g^L; s)}{\mathrm{MR}^{true}_{g^L}(g^H, g^L; s)} \leq 1 \leq \frac{\mathrm{RR}_{GU} \times \mathrm{MR}_{UY}}{\mathrm{RR}_{GU} + \mathrm{MR}_{UY} - 1}$$

where the second inequality comes from Lemma A.2 in Ding and VanderWeele (2016) and $\text{RR}_{GU} \geq 1, \text{MR}_{UY} \geq 1$.

Hence, we obtain the desired result.

$$\frac{\text{MR}^{obs}(g^H, g^L; s)}{\text{MR}^{true}_{g^L}(g^H, g^L; s)} \leq \frac{\text{RR}_{GU} \times \text{MR}_{UY}}{\text{RR}_{GU} + \text{MR}_{UY} - 1}.$$

Similar derivations apply to the other three inequalities. □

Proof of the theorem. For notational simplicity, we use the following representation.

$$\begin{aligned}
m(d, g; s) &\equiv \frac{1}{N} \sum_{i:S_i=s} \mathbb{E}[Y_i \mid T_i = d, G_i = g] \\
&= \frac{1}{N} \sum_{i:S_i=s} \frac{\sum_{u \in \Delta_s^u} \Pr(T_i = d, G_i = g, U_i = u) Y_i(d, g, u)}{\Pr(T_i = d, G_i = g)} \\
&= \frac{1}{N} \sum_{i:S_i=s} \sum_{u \in \Delta_s^u} \Pr(U_i = u \mid T_i = d, G_i = g) Y_i(d, g, u) \\
&= \sum_{u \in \Delta_s^u} \left\{ \frac{1}{N} \sum_{i:S_i=s} Y_i(d, g, u) \right\} \Pr(U_i = u \mid T_i = d, G_i = g) \\
&= \sum_{u \in \Delta_s^u} r_g(u) \Pr(U_i = u \mid T_i = d, G_i = g).
\end{aligned}$$

We want to show that, for g^H, g^L,

$$\frac{m(d, g^H; s)}{B} - B \times m(d, g^L; s) \ \leq \ \frac{1}{N} \sum_{i:S_i=s} \tau_i(g^H, g^L; d) \ \leq B \times m(d, g^H; s) - \frac{m(d, g^L; s)}{B}.$$

Because this implies the desired result.

$$\begin{aligned}
& \frac{m(d, g^H; s)}{B} - B \times m(d, g^L; s) \ \leq \ \frac{1}{N} \sum_{i:S_i=s} \tau_i(g^H, g^L; d) \ \leq B \times m(d, g^H; s) - \frac{m(d, g^L; s)}{B} \\
\Leftrightarrow \ & \begin{cases} \sum_{s \in \mathcal{S}} \left\{ \frac{m(d,g;s)}{B} - B \times m(d, g^L; s) \right\} \ \leq \ \sum_{s \in \mathcal{S}} \left\{ \frac{1}{N} \sum_{i:S_i=s} \tau_i(g^H, g^L; d) \right\} \\ \sum_{s \in \mathcal{S}} \left\{ \frac{1}{N} \sum_{i:S_i=s} \tau_i(g^H, g^L; d) \right\} \ \leq \ \sum_{s \in \mathcal{S}} \left\{ B \times m(d, g^H; s) - \frac{m(d,g^L;s)}{B} \right\} \end{cases} \\
\Leftrightarrow \ & \frac{\mathbb{E}[\hat{m}(d, g^H)]}{B} - B \times \mathbb{E}[\hat{m}(d, g^L)] \ \leq \ \tau(g^H, g^L; d) \ \leq B \times \mathbb{E}[\hat{m}(d, g^H)] - \frac{\mathbb{E}[\hat{m}(d, g^L)]}{B}.
\end{aligned}$$

First, using Lemma 1,

$$\begin{aligned}
& \frac{m(d, g^H; s)}{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
= \ & \frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^H)}{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \times \frac{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \\
&= \frac{MR^{obs}(g^H, g^L; s)}{MR^{true}_{g^L}(g^H, g^L; s)} \leq B
\end{aligned}
$$

where the final equality follows from the lemma. Therefore,

$$
\frac{m(d, g^H; s)}{B} \ \leq \ \sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L). \tag{A5}
$$

Also, since $B \geq 1$,

$$
\sum_{u \in \Delta_s^u} r_{g^L}(U) \Pr(U_i = u \mid T_i = d, G_i = g^L) \ = m(d, g^L; s) \ \leq \ B \times m(d, g^L; s). \tag{A6}
$$

Finally, taking equations (A5) and (A6) together,

$$
\begin{aligned}
& \frac{m(d, g^H; s)}{B} - B \times m(d, g^L; s) \\
\leq \ & \sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L) - \sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \sum_{u \in \Delta_s^u} \{r_{g^H}(u) - r_{g^L}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \frac{1}{N} \sum_{i: S_i = s} \sum_{u \in \Delta_s^u} \{Y_i(d, g^H, u) - Y_i(d, g^L, u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \frac{1}{N} \sum_{i: S_i = s} \tau_i(g^H, g^L; d).
\end{aligned}
$$

Similarly, we want to prove

$$
\frac{1}{N} \sum_{i: S_i = s} \tau_i(g^H, g^L; d) \leq B \times m(d, g^H; s) - \frac{m(d, g^L; s)}{B}.
$$

First, since $B \geq 1$,

$$
\frac{m(d, g^L; s)}{B} \ \leq \ m(d, g^L; s) \ = \ \sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L). \tag{A7}
$$

Then, using Lemma 1,

$$
\begin{aligned}
& \frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{m(d, g^H; s)} \\
= \ & \frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^H)}
\end{aligned}
$$

$$= \frac{\sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)} \times \frac{\sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L)}{\sum_{u \in \Delta_s^u} r_{g^H}(U) \Pr(U_i = u \mid T_i = d, G_i = g^H)}$$
$$= \frac{MR^{obs}(g^L, g^H; s)}{MR^{true}_{g^L}(g^L, g^H; s)} \leq B.$$

Therefore, we have

$$\sum_{u \in \Delta_s^u} r_{g^H}(U) \Pr(U_i = u \mid T_i = d, G_i = g^L) \ \leq \ B \times m(d, g^H; s). \tag{A8}$$

Finally, taking equations (A7) and (A8) together,

$$\begin{aligned}
& B \times m(d, g^H; s) - \frac{m(d, g^L; s)}{B} \\
\geq \ & \sum_{u \in \Delta_s^u} r_{g^H}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L) - \sum_{u \in \Delta_s^u} r_{g^L}(u) \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \sum_{u \in \Delta_s^u} \{r_{g^H}(u) - r_{g^L}(u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \frac{1}{N} \sum_{i: S_i = s} \sum_{u \in \Delta_s^u} \{Y_i(d, g^H, u) - Y_i(d, g^L, u)\} \Pr(U_i = u \mid T_i = d, G_i = g^L) \\
= \ & \frac{1}{N} \sum_{i: S_i = s} \tau_i(g^H, g^L; d).
\end{aligned}$$

Hence we have

$$\frac{m(d, g^H; s)}{B} - B \times m(d, g^L; s) \ \leq \ \frac{1}{N} \sum_{i: S_i = s} \tau_i(g^H, g^L; d) \ \leq B \times m(d, g^H; s) - \frac{m(d, g^L; s)}{B},$$

which completes the proof. □