Supplementary Materials

S1. Proofs

We collect notation from the main text and introduce new notation:

\[ m_x(y, z) = \mathbb{E}(Y_i^2 \mid D_i = 1, Z_i = z, Y_{i1} = y, X_i = x), \quad q_x(y, z) = \text{Pr}(Y_{i1} = y \mid D_i = 1, Z_i = z, X_i = x), \]
\[ m^*_x(y, z) = \mathbb{E}(Y_i^2 \mid D_i = 1, Z_i = z, Y^*_{i1} = y, X_i = x), \quad q^*_x(y, z) = \text{Pr}(Y^*_{i1} = y \mid D_i = 1, Z_i = z, X_i = x), \]
\[ p_x(y) = \text{Pr}(Y_{i1} = y \mid Y^*_{i1} = y, G_i = c, X_i = x), \quad \xi_x(y, z) = \text{pr}(Y^*_{i1} = 1 \mid Z_i = z, D_i = 1, Y_{i1} = y, X_i = x), \]
\[ r_x(z) = \frac{q_x(1, 2 - z)(1 - q_x(1, 2 - z))}{\{p_x(1) - q_x(1, 2 - z)\} \cdot [q_x(1, 2 - z) - \{1 - p_x(0)\}].} \]

We will prove our results under the following weaker version of Assumption 1.

**Assumption S1** (Strong Ignorability of Treatment Assignment).

\[ Z_i \indep D_i(z) \mid X_i, \]
\[ Z_i \indep \{Y^*_{i1}(z), Y_{i2}(z, y^*_1)\} \mid G_i = c, X_i, \]
\[ 0 < \text{Pr}(Z_i = z \mid X_i) < 1 \]

for \( z = 0, 1, 2 \) and \( y^*_1 = 0, 1 \).

S1.1. Proof of Theorem 1

We first consider the average spillover effect, by the law of total probability,

\[ \theta = \sum_x \{\mathbb{E}(Y_{i2}(1) \mid G_i = c, X_i = x) \text{pr}(X_i = x \mid G_i = c) - \mathbb{E}(Y_{i2}(0) \mid G_i = c, X_i = x) \text{pr}(X_i = x \mid G_i = c)\} \]
\[ = \sum_x \{\mathbb{E}(Y_{i2} \mid Z_i = 1, G_i = c, X_i = x) - \mathbb{E}(Y_{i2}(2) \mid G_i = c, X_i = x)\} \text{pr}(X_i = x \mid G_i = c) \]
\[ = \sum_x \{\mathbb{E}(Y_{i2} \mid Z_i = 1, D_i = 1, X_i = x) - \mathbb{E}(Y_{i2} \mid Z_i = 2, D_i = 1, X_i = x)\} \text{pr}(X_i = x \mid G_i = c) \]  \hspace{1cm} (S1)
where the second inequality follows from Assumption 3. From Assumption S1

\[ \text{pr}(X_i = x | G_i = c) = \frac{\text{pr}(G_i = c | X_i = x) \text{pr}(X_i = x)}{\text{pr}(G_i = c)} \]

\[ = \frac{\text{pr}(G_i = c | X_i = x) \text{pr}(X_i = x)}{\sum_x \text{pr}(G_i = c | X_i = x) \text{pr}(X_i = x)} \]

\[ = \frac{\text{pr}(D_i = 1 | Z_i \neq 0, X_i = x) \text{pr}(X_i = x)}{\sum_x \text{pr}(D_i = 1 | Z_i \neq 0, X_i = x) \text{pr}(X_i = x)}, \] (S2)

which simplifies to \( \text{pr}(X_i = x | D_i = 1) \) if Assumption 1 holds. Plugging (S2) into (S1), we can obtain the identification formula for \( \theta \), which becomes the same as the one in Theorem 1 if Assumption 1 holds.

We then consider the average contagion and direct effects. We only need to identify \( \mathbb{E}\{Y_{i2}(z, Y_{i1}^*(z')) \mid G_i = c, X_i = x\} \) for all \( x \) and \( z, z' = 0, 1 \). By the law of total probability, we have

\[ \mathbb{E}\{Y_{i2}(z, Y_{i1}^*(z')) \mid G_i = c, X_i = x\} \]

\[ = \sum_{y_1} \mathbb{E}\{Y_{i2}(z, y_{i1}) \mid G_i = c, Y_{i1}^*(z') = y_{i1}, X_i = x\} \text{Pr}(Y_{i1}^* = y_{i1}^* | G_i = c, X_i = x) \]

\[ = \sum_{y_1} \mathbb{E}\{Y_{i2}(z, y_{i1}) \mid Z_i = z', G_i = c, Y_{i1}^*(z') = y_{i1}, X_i = x\} \text{Pr}(Y_{i1}^* = y_{i1}^* | Z_i = z', G_i = c, X_i = x) \]

\[ = \sum_{y_1} \mathbb{E}\{Y_{i2}(z, y_{i1}) \mid Z_i = z', G_i = c, X_i = x\} \text{Pr}(Y_{i1}^* = y_{i1}^* | Z_i = z', G_i = c, X_i = x) \]

\[ = \sum_{y_1} \mathbb{E}\{Y_{i2}(z, y_{i1}) \mid Z_i = z, G_i = c, X_i = x\} \text{Pr}(Y_{i1}^* = y_{i1}^* | Z_i = z', G_i = c, X_i = x) \]

\[ = \sum_{y_1} \mathbb{E}\{Y_{i2}(z, y_{i1}) \mid Z_i = z, G_i = c, Y_{i1}^*(z) = y_{i1}, X_i = x\} \text{Pr}(Y_{i1}^* = y_{i1}^* | Z_i = z', G_i = c, X_i = x) \]

\[ = \sum_{y_1} m_x^*(y, z) q_x^*(y, z'), \] (S3)

where the second and the fourth equalities follow from Assumption 1 and the third and the fifth equalities follow from Assumption 2. Therefore, we need only to identify \( m_x^*(y, z) \) and \( q_x^*(y, z') \) for \( z, y = 0, 1 \). From Assumption 1, we can identify \( q_x^*(y, 1) = q_x(y, 1) \), and from Assumption 3, we
can identify

\[ q^*_x(y, 0) = \Pr(Y^*_{i1} = 1 \mid Z_i = 2, G_i = c, X_i = x) = q_x(y, 2). \]

Finally, Assumption 4 implies \( m^*_x(y, 1) = m_x(y, 1) \), and by Assumptions 3 and 4 we have,

\[ m^*_x(y, 0) = \mathbb{E}(Y^*_{i2} \mid Z_i = 2, G_i = c, Y^*_{i1} = y^*_i, X_i = x) = m_x(y, 2). \]

By plugging the identification formulas of \( q^*_x(y, z) \) and \( m^*_x(y, z) \) for \( z = 0, 1 \) into equation (S3), we can obtain the identification formulas for the average contagion and direct effects. They become the same as the identification formulas in Theorem 1 if Assumption 1 holds.

\[ \square \]

### S1.2. Proof of Theorem 2

First, because the expression of \( \theta \) does not include \( Y_{i1} \), the identification formula does not change without Assumption 4, i.e.,

\[ \theta = \sum_x \Pr(X_i = x \mid G_i = c) \cdot \{m_x(1, 1)q_x(1, 1) + m_x(0, 1)q_x(0, 1) - m_x(1, 2)q_x(1, 2) - m_x(0, 2)q_x(0, 2)\}. \]

For \( z = 1, 2 \), by the law of total probability and Assumption 5,

\[ q_x(1, z) = p_x(1) \cdot q^*_x(1, z) + (1 - p_x(0)) \cdot \{1 - q^*_x(1, z)\}. \]

We then have

\[ q^*_x(1, z) = \frac{q_x(1, z) - (1 - p_x(0))}{p_x(1) + p_x(0) - 1}. \]

Again, by the law of total probability and Assumption 5, we have

\[ m_x(1, z) = m^*_x(1, z) \cdot \xi_x(1, z) + m^*_x(0, z) \cdot \{1 - \xi_x(1, z)\}, \]

\[ m_x(0, z) = m^*_x(1, z) \cdot \xi_x(0, z) + m^*_x(0, z) \cdot \{1 - \xi_x(0, z)\}, \]
Therefore, we have

\[ m_x^*(1, z) = \frac{(1 - \xi_x(0, z))m_x(1, z) - (1 - \xi_x(1, z))m_x(0, z)}{\xi_x(1, z) - \xi_x(0, z)}, \]

\[ m_x^*(0, z) = \frac{\xi_x(1, z)m_x(0, z) - \xi_x(0, z)m_x(1, z)}{\xi_x(1, z) - \xi_x(0, z)}. \]

From Theorem 1 for \( z = 0, 1, \)

\[ \tau(z) = \sum_x \left\{m_x^*(1, 2-z) - m_x^*(0, 2-z)\right\} \left\{q_x^*(1, 1) - q_x^*(1, 2)\right\} \Pr(X_i = x) \]

\[ = \sum_x \frac{m_x(1, 2-z) - m_x(0, 2-z)}{\xi_x(1, 2-z) - \xi_x(0, 2-z)} \cdot \frac{q_x(1, 1) - q_x(1, 2)}{p_x(1) + p_x(0) - 1} \cdot \Pr(X_i = x), \]

whereas

\[ \xi_x(1, 2-z) - \xi_x(0, 2-z) = \frac{p_x(1) \cdot q_x(1, 2-z)}{q_x(1, 2-z)} - \frac{(1 - p_x(1)) \cdot q_x^*(1, 2-z)}{1 - q_x(1, 2-z)} \]

\[ = \frac{\{p_x(1) - q_x(1, 2-z)\} \cdot q_x^*(1, 2-z)}{\{p_x(1) - q_x(1, 2-z)\} \cdot [q_x(1, 2-z) - \{1 - p_x(0)\}]} \cdot \frac{1}{p_x(1) + p_x(0) - 1}. \]

Therefore, we have

\[ \tau(z) = \sum_x \left[ \Pr(X_i = x \mid G_i = c) \cdot r_x(z) \{m_x(1, 2-z) - m_x(0, 2-z)\} \{q_x(1, 1) - q_x(1, 2)\} \right]. \]

For the average direct effect, we have

\[ \eta(z) = \theta - \tau(1-z) \]

\[ = \sum_x \left[ \Pr(X_i = x \mid G_i = c) \cdot \sum_{y=0}^{1} \left( m_x(y, 1)q_x(y, 1) - m_x(y, 2)q_x(y, 2) \right) \right. \]

\[ - r_x(1-z)m_x(y, 1+z) \{q_x(y, 1) - q_x(y, 2)\} \left. \right] \]

\[ = \sum_x \left[ \Pr(X_i = x \mid G_i = c) \cdot \sum_{y=0}^{1} \left( \{m_x(y, 1) - m_x(y, 2)\}q_x(y, 2-z) \right. \right. \]

\[ - \{1 - r_x(1-z)\}m_x(y, 1+z) \{q_x(y, 2) - q_x(y, 1)\} \left. \right]. \]
Plugging $S2$ into the above equations, we can obtain the identification formulas for the average contagion and direct effects. They become the same as those in Theorem 2 if Assumption 1 holds.

S1.3. Proof of Corollary 1

We only need to drive the bounds for $r_x(z)$. Because $p_x(1) + p_x(0) \geq p$ and $0 \leq p_x(1), p_x(0) \leq 1$, we can obtain the range of $(p_x(1), 1 - p_x(0))$ as,

$$1 - p_x(0) \in [0, 2 - p], \quad p_x(1) \in [1 - p_x(0) + p - 1, 1].$$

To obtain the bounds for $r_x(z)$, we first look into the range of the following function:

$$f_a(x_1, x_2) = (a - x_1)(a - x_2), \quad a \in [0, 1], \quad x_1 \in [0, 2 - p], \quad x_2 \in [x_1 + p - 1, 1].$$

We enumerate all the cases for different relative magnitude among $2 - p$, $p - 1$ and $a$.

Case 1: $p \geq 3/2$

(a) When $a < 2 - p$, we have $x_2 < 1 - p \geq a$. $f_a(x_1, x_2)$ reaches its maximum $\{a - (2 - p)\}(a - 1)$ at $(x_1, x_2) = (2 - p, 1)$, and reaches its minimum $a(a - 1)$ at $(x_1, x_2) = (0, 1)$. Thus, we have $f_a(x_1, x_2) \in [a(a - 1), \{a - (2 - p)\}(a - 1)]$.

(b) When $a = 2 - p$, we have $x_2 < 1 - p \geq a$. $f_a(x_1, x_2)$ reaches its maximum 0 at $(x_1, x_2) = (2 - p, 1)$, and reaches its minimum $a(a - 1)$ at $(x_1, x_2) = (0, 1)$. Thus, we have $f_a(x_1, x_2) \in [a(a - 1), 0]$.

(c) When $2 - p < a < p - 1$, we have $x_2 < 2 - p < a < p - 1 \leq x_1$. $f_a(x_1, x_2)$ reaches its minimum $a(a - 1)$ at $(x_1, x_2) = (0, 1)$. If $a \leq 1/2$, then $f_a(x_1, x_2)$ reaches its maximum $\{a - (2 - p)\}(a - 1)$ at $(x_1, x_2) = (2 - p, 1)$ and if $a > 1/2$, then $f_a(x_1, x_2)$ reaches its maximum $\{a - (p - 1)\}a$ at $(x_1, x_2) = (0, p - 1)$.

(d) When $a = p - 1$, we have $x_1 \leq 2 - p \leq a$. $f_a(x_1, x_2)$ reaches its maximum 0 at $(x_1, x_2) = (0, p - 1)$, and reaches its minimum $a(a - 1)$ at $(x_1, x_2) = (0, 1)$. Thus, we have $f_a(x_1, x_2) \in [a(a - 1), 0]$. 


(e) When \( a > p - 1 \), we have \( x_1 \leq 2 - p \leq a \). \( f_a(x_1, x_2) \) reaches its maximum \( \{ a - (p - 1) \} a \) at \((x_1, x_2) = (0, p - 1)\), and reaches its minimum \( a(a - 1) \) at \((x_1, x_2) = (0, 1)\). Thus, we have \( f_a(x_1, x_2) \in [a(a - 1), \{ a - (p - 1) \} a] \).

**Case 2: \( p < 3/2 \)**

(a) When \( a \leq p - 1 \), we have \( a \leq x_2 \). \( f_a(x_1, x_2) \) reaches its maximum \( \{ a - (2 - p) \} (a - 1) \) at \((x_1, x_2) = (2 - p, 1)\), and reaches its minimum \( a(a - 1) \) at \((x_1, x_2) = (0, 1)\). Thus, we have \( f_a(x_1, x_2) \in [a(a - 1), \{ a - (2 - p) \} (a - 1)] \).

(b) When \( p - 1 < a < 2 - p \), \( f_a(x_1, x_2) \) reaches its minimum \( a(a - 1) \) at \((x_1, x_2) = (0, 1)\). If \( a \leq 1/2 \), then \( f_a(x_1, x_2) \) reaches its maximum \( \{ a - (2 - p) \} (a - 1) \) at \((x_1, x_2) = (2 - p, 1)\) and if \( a > 1/2 \), then \( f_a(x_1, x_2) \) reaches its maximum \( \{ a - (p - 1) \} a \) at \((x_1, x_2) = (0, p - 1)\).

(c) When \( a \geq 2 - p \), we have \( a \geq x_1 \). \( f_a(x_1, x_2) \) reaches its maximum \( \{ a - (p - 1) \} a \) at \((x_1, x_2) = (0, p - 1)\), and reaches its minimum \( a(a - 1) \) at \((x_1, x_2) = (0, 1)\). Thus, we have \( f_a(x_1, x_2) \in [a(a - 1), \{ a - (p - 1) \} a] \).

To obtain the results in Corollary 1, we examine the case when \( 2 - p < q_x(1, 2 - z) < p - 1 \). From the bounds for \( f_{q_x(1, 2 - z)}(p_x(1), 1 - p_x(0)) \), we have \( r_x(z) \in [1, u_x(z)] \) where

\[
u_x(z) = \left\{ \begin{array}{ll}
\frac{q_x(1, 2 - z)}{q_x(1, 2 - z) - (2 - p)} & , \\
\frac{1 - q_x(1, 2 - z)}{(p - 1) - q_x(1, 2 - z)} & , 
\end{array} \right.
\]

According to Theorem 2,

\[
\tau(z) = \sum_x \Pr(X_i = x \mid D_i = 1) \cdot r_x(z)Q_x(z),
\]

where \( Q_x(z) = \{ m_x(1, 2 - z) - m_x(0, 2 - z) \} \{ q_x(1, 1) - q_x(1, 2) \} \). Therefore, under \( 2 - p < \min_x \{ q_x(1, 2 - z) \} \leq \max_x \{ q_x(1, 2 - z) \} < p - 1 \), the upper bound of \( \tau(z) \) is

\[
\sum_x \Pr(X_i = x \mid D_i = 1) \cdot [I\{ Q_x(z) \geq 0 \} Q_x(z) u_x(z) + I\{ Q_x(z) < 0 \} Q_x(z)],
\text{ (S4)}
\]
and the upper bound of $\tau(z)$ is

$$
\sum_x \Pr(X_i = x \mid D_i = 1) \cdot [I\{Q_x(z) \geq 0\}Q_x(z) + I\{Q_x(z) < 0\}Q_x(z)] u_x(z).
$$

(S5)

When Assumption $S1$ holds instead of Assumption $1$, we can replace $\Pr(X_i = x \mid D_i = 1)$ with (S2) in (S4) and (S5) to obtain the bounds. For other cases with different relative magnitude among $p - 1, 2 - p$ and $q_x(1, 2 - z)$, we can obtain bounds for $\tau(z)$ using a similar technique.

S2. Computation

In this section, we provide the details of the EM algorithms for the the proposed sensitivity analyses. We will give the algorithms under Assumption $1$. Note that when sensitivity parameter is zero, we obtain the point estimates under Assumptions $1-4$. Recall that the following model is fit to the units with $D_i = 1$ ($G_i = c$).

$$
Y_{i1}(z) = I(\tilde{Y}_{i1}(z) > 0) \quad \text{where} \quad \tilde{Y}_{i1}(z) = g(z, X_i) + \epsilon_{i1},
$$

$$
Y_{i2}(z, y_{1i}) = I(\tilde{Y}_{i2}(z, y_{1i}) > 0) \quad \text{where} \quad \tilde{Y}_{i2}(z, y_{1i}) = f(z, y_{1i}, X_i) + \epsilon_{i2},
$$

$$
\begin{pmatrix}
\epsilon_{i1} \\
\epsilon_{i2}
\end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\},
$$

where $g(\cdot)$ and $f(\cdot)$ have linear forms

$$
g(z, x) = \alpha_0 + \alpha_Z z + x \alpha_X + z x \alpha_{ZX},
$$

$$
f(z, y_{1i}, x) = \beta_0 + \beta_Z z + \beta_Y y_{1i} + \beta_{YZ} z y_{1i} + x \beta_X + z x \beta_{ZX} + y_{1i} x \beta_{XY}.
$$

Define $W_{i1} = (1, Z_i, X_i, Z_i X_i)^T$, $W_{i2} = (1, Z_i, Y_{i1}^*, Z_i Y_{i1}^*, X_i, Z_i X_i, Y_{i1}^* X_i)^T$, $\alpha = (\alpha_0, \alpha_Z, \alpha_X, \alpha_{ZX})^T$, and $\beta = (\beta_0, \beta_Z, \beta_Y, \beta_{YZ}, \beta_X, \beta_{ZX}, \beta_{XY})^T$. Because $\tilde{Y}_{i1}(z) = \tilde{Y}_{i1}$, $Y_{i1}^*(z) = Y_{i1}^*$ and $Y_{i1}(z) = Y_{i1}$ if $Z_i = z$. Similarly, $\tilde{Y}_{i2}(z, y_{1i}) = \tilde{Y}_{i2}$ and $Y_{i2}(z, y_{1i}) = Y_{i2}$ if $Z_i = z$ and $Y_{i1}^* = y_{1i}^*$. Thus, we can rewrite
our model using the observed data:

\[ Y_{i1}^* = I(\tilde{Y}_{i1} > 0) \quad \text{where} \quad \tilde{Y}_{i1} = W_{i1}^T \alpha + \epsilon_{i1}, \]

\[ Y_{i2} = I(\tilde{Y}_{i2} > 0) \quad \text{where} \quad \tilde{Y}_{i2} = W_{i2}^T \beta + \epsilon_{i2}, \]

\[
\begin{pmatrix}
\epsilon_{i1} \\
\epsilon_{i2}
\end{pmatrix}
\sim N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right\}.
\]

S2.1. Sensitivity Analysis for Unobserved Confounding

We present the EM algorithm for the sensitivity analysis regarding unobserved confounding. We write our model as,

\[ Y_{i1}^* = I(\tilde{Y}_{i1} > 0) \quad \text{where} \quad \tilde{Y}_{i1} = W_{i1}^T \alpha + \epsilon_{i1}, \]

\[ Y_{i2} = I(\tilde{Y}_{i2} > 0) \quad \text{where} \quad \tilde{Y}_{i2} = W_{i2}^T \beta + \epsilon_{i2}, \]

\[
\begin{pmatrix}
\epsilon_{i1} \\
\epsilon_{i2}
\end{pmatrix}
\sim N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right\}.
\]

The complete-data log-likelihood function is given by,

\[
\log L_c(\alpha, \beta) = \sum_{i=1}^{N} I\{D_i = 1\} \cdot \left[ -\frac{(\tilde{Y}_{i1} - W_{i1}^T \alpha)^2}{2(1 - \rho^2)} + \frac{\rho(\tilde{Y}_{i1} - W_{i1}^T \alpha)(\tilde{Y}_{i2} - W_{i2}^T \beta)}{1 - \rho^2} - \frac{(\tilde{Y}_{i2} - W_{i2}^T \beta)^2}{2(1 - \rho^2)} \right] \\
\cdot I\{\tilde{Y}_{i1}(Y_{i1} - 0.5) > 0\} \cdot I\{\tilde{Y}_{i2}(Y_{i2} - 0.5) > 0\} + \text{constant}.
\]

Let \( O_i \) be the observed data for unit \( i \), i.e., \( O_i = (Y_{i1}, Y_{i2}, Z_i, D_i = 1, X_i)^T \), and let \( \xi^{(k)} \) be the estimate of \( \xi \) after the \( k \)-th iteration. In the E-step, we need to compute:

\[
E(\tilde{Y}_{i1} \mid O_i, \alpha^{(k)}, \beta^{(k)}), \quad E(\tilde{Y}_{i2} \mid O_i, \alpha^{(k)}, \beta^{(k)}).
\]

Because \( (\tilde{Y}_{i1}, \tilde{Y}_{i2})^T \mid O_i, \alpha^{(k)}, \beta^{(k)} \) follows a truncated bivariate Normal distribution with mean \((W_{i1}^T \alpha, W_{i2}^T \beta)\) and covariance matrix \( \Sigma_2 \), we use R package \texttt{tmvtnorm} to compute them.
In the M-step, we need to update the parameters based on

\[
\tilde{Y}_{i1} = W_{i1}^{\top} \alpha + \epsilon_{i1}, \quad \tilde{Y}_{i2} = W_{i2}^{\top} \beta + \epsilon_{i2}.
\]

Because we know the covariance matrix of the error terms \((\epsilon_{i1}, \epsilon_{i2})\), we can transform the two regression equations to

\[
\Sigma_2^{-1/2} \begin{pmatrix} \tilde{Y}_{i1} \\ \tilde{Y}_{i2} \end{pmatrix} = \Sigma_2^{-1/2} \begin{pmatrix} W_{i1} & 0 \\ 0 & W_{i2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Sigma_2^{-1/2} \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \end{pmatrix}.
\]

Then, we can use ordinary least squares regression to update the parameters.

After obtaining the maximum likelihood estimates of the parameters, we can then write,

\[
P\{Y_{i2}(z, Y_{i1}^*(z')) = 1 \mid G_i = c\} = \sum_x P\{Y_{i2}(z, Y_{i1}^*(z')) = 1 \mid X_i = x, D_i = 1\} \text{pr}(X_i = x \mid D_i = 1)
\]

\[
= \sum_x [P\{Y_{i2}(z, 1) = 1, Y_{i1}^*(z') = 1 \mid X_i = x, D_i = 1\} + P\{Y_{i2}(z, 0) = 1, Y_{i1}^*(z') = 0 \mid X_i = x, D_i = 1\}]
\]

\[
\cdot \text{pr}(X_i = x \mid D_i = 1)
\]

\[
= \sum_x [P\{\epsilon_{i2} > -f(z, 1, x), \epsilon_{i1} > -g(z, x) \mid X_i = x, D_i = 1\}
\]

\[
+ P\{\epsilon_{i2} > -f(z, 0, x), \epsilon_{i1} \leq -g(z, x) \mid X_i = x, D_i = 1\}] \text{pr}(X_i = x \mid D_i = 1).
\]

We calculate the terms above using the cumulative distribution function of bivariate Normal distributions. Then, based on \(P\{Y_{i2}(z, Y_{i1}^*(z')) = 1 \mid G_i = c\}\), we compute the estimated average contagion and direct effects.

**S2.2. Sensitivity Analysis for Additive Measurement Error**

We assume

\[
Y_{i1}(z) = I\{\tilde{Y}_{i1}(z) + \zeta_i > 0\} \quad \text{and} \quad \begin{pmatrix} \zeta_i \\ \epsilon_{i2} \end{pmatrix} \overset{i.i.d.}{\sim} N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right\} = \begin{pmatrix} \sigma^2 & \rho_e \sigma \\ \rho_e \sigma & 1 \end{pmatrix},
\]

where \(\sigma^2\) and \(\rho_e\) are pre-specified.
Therefore, we can write the model as:

\[ Y_{i1}^* = I(Y_{i1}' - \zeta_i > 0), \quad Y_{i1} = I(Y_{i1}' > 0) \quad \text{where} \quad Y_{i1}' = W_{i1}^\top \alpha + \epsilon_{i1}, \]

\[ Y_{i2} = I(\tilde{Y}_{i2} > 0) \quad \text{where} \quad \tilde{Y}_{i2} = W_{i2}^\top \beta + \epsilon_{i2}, \]

\[ \zeta_i \sim N(0, \sigma^2), \quad \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 1 + \sigma^2 & \rho_c \sigma \\ \rho_c \sigma & 1 \end{pmatrix} \right\}. \tag{S7} \]

Treating \( Y_{i1}', Y_{i1}^* \) and \( \tilde{Y}_{i2} \) as missing data, we can write the complete-data log-likelihood for the units with \( D_i = 1 \) as

\[
\log L_c(\xi) = \sum_{i=1}^N I(D_i = 1) \cdot \left[ \frac{1}{1 - \rho_e^2} \left\{ \frac{(Y_{i1}' - W_{i1}^\top \alpha)^2}{2(1 + \sigma^2)} - \rho_e (Y_{i1}' - W_{i1}^\top \alpha)(\tilde{Y}_{i2} - W_{i2}^\top \beta) \right\} \sqrt{1 + \sigma^2} + \frac{(\tilde{Y}_{i2} - W_{i2}^\top \beta)^2}{2} \right] + h(Y_{i1}', Y_{i1}^*) \cdot I\{Y_{i1}'(Y_{i1}') - 0.5 > 0\}I\{\tilde{Y}_{i2}(\tilde{Y}_{i2}) - 0.5 > 0\} + \text{constant},
\]

where \( \rho_e' = \rho_e / \sqrt{1 + \sigma^2} \), and \( h(Y_{i2}', Y_{i2}^*) \) is the likelihood that corresponds to \( \text{pr}(Y_{i2}' | Y_{i2}) \) which does not affect our parameter estimation.

We use the EM algorithm to obtain the MLEs of \( \alpha \) and \( \beta \). We ignore \( D_i = 1 \) in the following derivation. In the M-step, we update the parameters conditionally. In particular, we update \( \alpha \) conditional on \( \beta \):

\[
\alpha^{(k+1)} = \left\{ \sum_{i=1}^N E(W_{i1}W_{i1}^\top | O_i, \xi^{(k)}) \right\}^{-1} \left\{ \sum_{i=1}^N E \left( W_{i1}\tilde{Y}_{i1}' - \rho_e \sqrt{1 + \sigma^2} W_{i1}(\tilde{Y}_{i2} - W_{i2}^\top \beta^{(k)}) | O_i, \xi^{(k)} \right) \right\}.
\]

and update \( \beta \) conditional on \( \alpha \):

\[
\beta^{(k+1)} = \left\{ \sum_{i=1}^N E(W_{i2}W_{i2}^\top | O_i, \xi^{(k)}) \right\}^{-1} \left\{ \sum_{i=1}^N E \left( W_{i2}\tilde{Y}_{i2}' - \rho_e \sigma W_{i2}(\tilde{Y}_{i1} - W_{i1}^\top \alpha^{(k+1)}) | O_i, \xi^{(k)} \right) \right\}.
\]
Therefore, we need to calculate the following conditional expectations in the E-step:

\[
E(Y_{i1}^{*2} \mid O_i, \xi^{(k)}) = E(Y_{i1}^{*} \mid O_i, \xi^{(k)}), \quad E(Y_{i1}^{*} Y_{i1}' \mid O_i, \xi^{(k)})
\]

\[
E(Y_{i1}' \mid O_i, \xi^{(k)}), \quad E(Y_{i2}' Y_{i1}^{*} \mid O_i, \xi^{(k)}), \quad E(Y_{i2}' \mid O_i, \xi^{(k)}).
\]

We calculate them separately. First, given \(Y_{i1}'^{*}, Y_{i1} \) and \(Y_{i2}, (Y_{i1}' - \zeta_i, Y_{i1}', \tilde{Y}_{i2})\) follows a trivariate truncated Normal distribution:

\[
\begin{pmatrix}
  Y_{i1}' - \zeta_i \\
  Y_{i1}' \\
  \tilde{Y}_{i2}
\end{pmatrix} \sim TN_3\left(\begin{pmatrix} W_{i1} \alpha \\ W_{i1} \alpha \\ W_{i2} \beta \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 + \sigma^2 & \rho_\sigma \\ 0 & \rho_\sigma & 1 \end{pmatrix}\right)
\]

where the truncated intervals are given by \(Y_{i1}'^{*} = I(Y_{i1}' - \zeta_i > 0), Y_{i1} = I(Y_{i1}' > 0)\) and \(Y_{i2} = I(\tilde{Y}_{i2} > 0)\). By Bayes Theorem, we have

\[
E(Y_{i1}^{*2} \mid O_i, \xi^{(k)}) = \frac{\text{pr}(Y_{i1}'^{*} = 1 \mid Z_i, D_i = 1, X_i, \xi^{(k)})}{\text{pr}(Y_{i1}'^{*} = 1, Y_{i1}, Y_{i2} \mid Z_i, D_i = 1, X_i, \xi^{(k)})} \cdot \frac{\text{pr}(Y_{i1}^{*} = 1, Y_{i1}, Y_{i2} \mid O_i, \xi^{(k)})}{\text{pr}(Y_{i1}^{*} = 1, Y_{i1}, Y_{i2} \mid O_i, \xi^{(k)})},
\]

where \(\text{pr}(Y_{i1}'^{*}, Y_{i1}, Y_{i2} \mid Z_i, D_i = 1, X_i, \xi^{(k)})\) can be calculated from R package \texttt{tmvtnorm}, which contains functions for calculating cumulative distribution functions and expectations for multivariate truncated Normal distributions. Using this package, we can then calculate

\[
E(Y_{i1}' Y_{i1}' \mid O_i, \xi^{(k)}) = E(Y_{i1}' \mid Y_{i1}'^{*} = 1, O_i, \xi^{(k)}) \cdot \text{pr}(Y_{i1}'^{*} = 1 \mid O_i, \xi^{(k)}),
\]

\[
E(Y_{i1}' \mid O_i, \xi^{(k)}) = E(Y_{i1}' Y_{i1}' \mid O_i, \xi^{(k)}) + E(Y_{i1}' (1 - Y_{i1}'^{*}) \mid O_i, \xi^{(k)}),
\]

\[
E(Y_{i2}' Y_{i1}^{*} \mid O_i, \xi^{(k)}) = E(Y_{i2}' \mid Y_{i1}' = 1, O_i, \xi^{(k)}) \cdot \text{pr}(Y_{i1}'^{*} = 1 \mid O_i, \xi^{(k)}),
\]

\[
E(Y_{i2}' \mid O_i, \xi^{(k)}) = E(Y_{i2}' Y_{i1}^{*} \mid O_i, \xi^{(k)}) + E(Y_{i2}' (1 - Y_{i1}'^{*}) \mid O_i, \xi^{(k)}).
\]

Based on these conditional expectations, we can update the parameters using (S8) and (S9).
After obtaining the maximum likelihood estimates of the parameters, we then write,

\[ P\{Y_i^2(z, Y_{i1}^*(z')) = 1 \mid G_i = c\} = \sum_x P\{Y_i^2(z, Y_{i1}^*(z')) = 1 \mid X_i = x, D_i = 1\} \text{pr}(X_i = x \mid D_i = 1) \]

\[ = \sum_x [P\{Y_i^2(z, 1) = 1, Y_{i1}^*(z') = 1 \mid X_i = x, D_i = 1\} + P\{Y_i^2(z, 0) = 1, Y_{i1}^*(z') = 0 \mid X_i = x, D_i = 1\}] \cdot \text{pr}(X_i = x \mid D_i = 1) \]

\[ = \sum_x [P\{\epsilon_i > -f(z, 1, x), \epsilon_{i1} > -g(z, x) \mid X_i = x, D_i = 1\} + P\{\epsilon_i > -f(z, 0, x), \epsilon_{i1} \leq -g(z, x) \mid X_i = x, D_i = 1\}] \text{pr}(X_i = x \mid D_i = 1). \]

We calculate the terms above using the cumulative distribution function of bivariate Normal distributions. Then, based on \( P\{Y_i^2(z, Y_{i1}^*(z')) = 1 \mid G_i = c\} \), we compute the estimated average contagion and direct effects.