Proofs for Identification of Direct and Indirect Effects

Identification of Controlled Direct Effects: The average controlled direct effect $\mathbb{E}[Y_{am} - Y_{a^*m}]$ is identifiable if for some set of variables C

$$Y_{am} \coprod A|C \text{ (no unmeasured confounding for the A-Y relationship)} \tag{1}$$

and

$$Y_{am} \coprod M|A, C \text{ (no unmeausured confouning for the M-Y relationship)} \tag{2}$$

If (1) and (2) hold then the average controlled direct effect is given by:

$$\mathbb{E}[Y_{am} - Y_{a^*m}] = \sum_c \{\mathbb{E}[Y|A = a, M = m, C = c] - \mathbb{E}[Y|A = a^*, M = m, C = c]\}P(C = c).$$

Proof: $\mathbb{E}[Y_{am}] - \mathbb{E}[Y_{a^*m}] = \sum_c \{\mathbb{E}[Y_{am}|C = c] - \mathbb{E}[Y_{a^*m}|C = c]\}P(C = c)$ by iterated expectations
$= \sum_c \{\mathbb{E}[Y_{am}|A = a, C = c] - \mathbb{E}[Y_{a^*m}|A = a^*, C = c]\}P(C = c)$ by (1)
$= \sum_c \{\mathbb{E}[Y_{am}|A = a, M = m, C = c] - \mathbb{E}[Y_{a^*m}|A = a^*, M = m, C = c]\}P(C = c)$ by (2)
$= \sum_c \{\mathbb{E}[Y|A = a, M = m, C = c] - \mathbb{E}[Y|A = a^*, M = m, C = c]\}P(C = c)$ by consistency.

Identification of Controlled Direct Effects Under Time Dependent Confounding: The average controlled direct effect $\mathbb{E}[Y_{am} - Y_{a^*m}]$ is identifiable if for some set of variables C and L

$$Y_{am} \coprod A|C \text{ (no unmeasured confounding for the A-Y relationship)} \tag{1}$$

and

$$Y_{am} \coprod M|A, C, L \text{ (no unmeausured confouning for the M-Y relationship)} \tag{2}$$

If (1) and (2) hold then the average controlled direct effect is given by:

$$\begin{aligned}\mathbb{E}[Y_{am} - Y_{a^*m}] &= \sum_c \sum_l \mathbb{E}[Y|A = a, M = m, C = c, L = l]P(L = l|A = a, C = c)P(C = c) \\ &\quad - \sum_c \sum_l \mathbb{E}[Y|A = a^*, M = m, C = c, L = l]P(L = l|A = a^*, C = c)P(C = c)\end{aligned}$$

Proof: $\mathbb{E}[Y_{am}] - \mathbb{E}[Y_{a^*m}] = \sum_c \{\mathbb{E}[Y_{am}|C = c] - \mathbb{E}[Y_{a^*m}|C = c]\}P(C = c)$ by iterated expectations
$= \sum_c \{\mathbb{E}[Y_{am}|A = a, C = c] - \mathbb{E}[Y_{a^*m}|A = a^*, C = c]\}P(C = c)$ by (1)
$= \sum_c \sum_l \mathbb{E}[Y_{am}|A = a, C = c, L = l]P(L = l|A = a, C = c)P(C = c) - \sum_c \sum_l \mathbb{E}[Y_{a^*m}|A = a^*, C = c, L = l]P(L = l|A = a^*, C = c)P(C = c)$ by iterated expectations
$= \sum_c \sum_l \mathbb{E}[Y_{am}|A = a, M = m, C = c, L = l]P(L = l|A = a, C = c)P(C = c) - \sum_c \sum_l \mathbb{E}[Y_{a^*m}|A = a^*, M = m, C = c, L = l]P(L = l|A = a^*, C = c)P(C = c)$ by (2)
$= \sum_c \sum_l \mathbb{E}[Y|A = a, M = m, C = c, L = l]P(L = l|A = a, C = c)P(C = c) - \sum_c \sum_l \mathbb{E}[Y|A = a^*, M = m, C = c, L = l]P(L = l|A = a^*, C = c)P(C = c)$ by consistency.

Identification of Natural Direct and Indirect Effects: Average natural direct and indirect effects are identified if

$$Y_{am} \coprod A|C \text{ (no unmeasured confounding for the A-Y relationship)} \quad (1)$$

$$Y_{am} \coprod M|A, C \text{ (no unmeausured confouning for the M-Y relationship)} \quad (2)$$

$$M_a \coprod A|C \text{ (no unmeasured confounding for the A-M relationship)} \quad (3)$$

and

$$Y_{am} \coprod M_{a^*}|C \text{ (no M-Y confounders which are effects of A).} \quad (4)$$

If (1)-(4) hold then the average natural direct effect is given by

$$\mathbb{E}[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}] = \sum\nolimits_{c,m} \{\mathbb{E}[Y|a, m, c] - \mathbb{E}[Y|a^*, m, c]\}P(m|a^*, c)P(c)$$

and the average natural indirect effect is given by

$$\mathbb{E}[Y_{aM_a} - Y_{aM_{a^*}}] = \sum\nolimits_{c,m} \mathbb{E}[Y|a, m, c]\{P(m|a, c) - P(m|a^*, c)\}P(c)$$

Proof: $\mathbb{E}[Y_{aM_{a^*}}] = \sum_c \mathbb{E}[Y_{aM_{a^*}}|C = c]P(C = c)$ by iterated expectations
$= \sum_{c,m} \mathbb{E}[Y_{am}|C = c, M_{a^*} = m]P(M_{a^*} = m|C = c)P(C = c)$ by iterated expectations
$= \sum_{c,m} \mathbb{E}[Y_{am}|C = c]P(M_{a^*} = m|A = a^*, C = c)P(C = c)$ by (4) & (3)
$= \sum_{c,m} \mathbb{E}[Y_{am}|A = a, C = c]P(M = m|A = a^*, C = c)P(C = c)$ by (1) and consistency
$= \sum_{c,m} \mathbb{E}[Y_{am}|A = a, M = m, C = c]P(M = m|A = a^*, C = c)P(C = c)$ by (2)
$= \sum_{c,m} \mathbb{E}[Y|A = a, M = m, C = c]P(M = m|A = a^*, C = c)P(C = c)$ by consistency.
If we apply this result and replace a with a^* we get $\mathbb{E}[Y_{a^*M_{a^*}}] = \sum_{c,m} \mathbb{E}[Y|a^*, m, c]P(m|a^*, c)P(c)$ and from this it follows that the average natural direct effect is given by:

$$\mathbb{E}[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}] = \sum\nolimits_{c,m} \{\mathbb{E}[Y|a, m, c] - \mathbb{E}[Y|a^*, m, c]\}P(m|a^*, c)P(c).$$

If we apply this result and replace a^* with a we get $\mathbb{E}[Y_{aM_a}] = \sum_{c,m} \mathbb{E}[Y|a, m, c]P(m|a, c)P(c)$ and from this it follows that the average natural indirect effect is given by:$\mathbb{E}[Y_{aM_a} - Y_{aM_{a^*}}] = \sum_{c,m} \mathbb{E}[Y|a, m, c]\{P(m|a, c) - P(m|a^*, c)\}P(c)$.