eAPPENDICES for Hafeman DM and VanderWeele TJ, Alternative assumptions for the identification of direct and indirect effects. Epidemiology.

eAPPENDIX A. Notation and Probability Rules

The following notation will be used to derive assumptions based on these response types. For any variable V, let 1(V=v) be 1 if V=v and 0 otherwise. Define $M^T_1=1(M^T=1)$, $M^T_2=1(M^T=2)$, $M^T_{12}=1(M^T=1)+1(M^T=2)$, etc.; similarly define $Y^T_1=1(Y^T=1)$, $Y^T_2=1(Y^T=2)$, $Y^T_{12}=1(Y^T=1)+1(Y^T=2)$, $Y^T_{124}=1(Y^T=1)+1(Y^T=2)+1(Y^T=4)$, etc. so that e.g. Y^T_{124} takes the value 1 if the Y-type is 1 or 2 or 4. Note that because variables of the form M^T_b are binary we can conceive of them either as variables or as events and we will thus use $M^T_b=1$ and M^T_b interchangeably; one could also use the complement of the event, sometimes denoted by $M^{T\,C}_b$, and $M^T_b=0$ interchangeably; similar remarks hold for the Y-type variables Y^T_d. The probability that an individual is of Y-type d is thus denoted by $P(Y^T_d)$; the joint probability that an individual is a given M-type b and Y-type d is denoted by $P(M^T_b Y^T_d)$; the probability that an individual is either Y-type d or e is denoted by $P(Y^T_{de})$ which is equivalent to $P(Y^T_d+Y^T_e)$. Y-types can be made conditional on exposure status and/or M-type; for example, the probability that an individual is doomed on Y ($Y^T=1$), given X=1 and $M^T=4$ is denoted by $P(Y^T_1|X=1,M^T_4)$. Notation and probability rules used for the development of assumptions are described in Appendix A.

<u>Notation</u>

Rule (N1): $P(Y^T_a \text{ or } Y^T_b|X=x, M^T_c) = P(Y^T_a + Y^T_b|X=x, M^T_c) = P(Y^T_{ab}|X=x, M^T_c)$

e.g. $P(Y^T_1 + Y^T_{24}|X=1, M^T_4) = P(Y^T_{124}|X=1, M^T_4)$

Rule (N2): $P(M^T_a \text{ and } Y^T_b|X=x) = P(M^T_a Y^T_b|X=x)$

e.g. $P(M^T_2 \text{ and } Y^T_{26}|X=1) = P(M^T_2 Y^T_{26}|X=1)$

Probability

Rule (P1): $P(Y^T_a + Y^T_b|X=x, M^T_c) = P(Y^T_a|X=x, M^T_c) + P(Y^T_b|X=x, M^T_c)$

e.g. $P(Y^T_1 + Y^T_{24}|X=1, M^T_4) = P(Y^T_1|X=1, M^T_4) + P(Y^T_{24}|X=1, M^T_4)$

Rule (P2): $P(M^T_aY^T_b|X=x) = P(M^T_a|X=x)*P(Y^T_b|X=x, M^T_a)$

e.g. $P(M^T_2Y^T_{26}|X=1) = P(M^T_2|X=1)*P(Y^T_{26}|X=1, M^T_2)$

Rule (P3): $P(M^T_1Y^T_a + M^T_2Y^T_a + M^T_4Y^T_a) = P[(M^T_1 + M^T_2 + M^T_4)*Y^T_a] = P(Y^T_a)$

e.g. $P(M^T_{12}Y^T_{124}|X=1) + P(M^T_4Y^T_{124}|X=1) = P(Y^T_{124}|X=1)$

Rule (P4): $P(M^T_c |X=x)*[P(Y^T_a|X=1, M^T_d) – P(Y^T_b|X=0, M^T_c)] = P(M^T_c |X=x)*P(Y^T_a|X=1, M^T_d)$
$– P(M^T_c |X=x)*P(Y^T_b|X=0, M^T_c)$

e.g. $P(M^T_{24}|X=0)*[P(Y^T_{124}|X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})] = P(M^T_{24}|X=0)*P(Y^T_{124}|X=1, M^T_4)$
$- P(M^T_{24}|X=0)*P(Y^T_1|X=0, M^T_{24})$

eAPPENDIX B. Derivations of the Results in Table 5 and the Results for the New Identification Assumptions in Tables 6-9. Observable estimates are translated into potential outcomes and then response types (using the eAppendix B Table), conditional on observable populations. Assumptions necessary for these quantities to equal the true direct and indirect effects (Table 4) are derived.

eAppendix B Table Potential outcomes and response type proportions

Potential outcomes	Response type proportions
$P(M_1=1)$	$P(M^T_{12})$
$P(M_1=0)$	$P(M^T_4)$
$P(M_0=1)$	$P(M^T_1)$
$P(M_0=0)$	$P(M^T_{24})$
$P(Y_{11})$	$P(Y^T_{12468})$
$P(Y_{10})$	$P(Y^T_{124})$
$P(Y_{01})$	$P(Y^T_{126})$
$P(Y_{00})$	$P(Y^T_1)$
$P(Y_1=1) = P(Y_{1M_1}=1)$	$P(Y^T_{124} + M^T_{12}Y^T_{68})$
$P(Y_0=1) = P(Y_{0M_0}=1)$	$P(Y^T_1 + M^T_1Y^T_{26})$

<u>CDE (m=0)</u>
$=P(Y=1|X=1, M=0) – P(Y=1|X=0, M=0)$
$=P(Y_{10}=1|X=1, M_1=0) – P(Y_{00}=1|X=0, M_0=0)$
$= P(Y^T_{124}| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})$
$= P(Y^T_{24}| X=1, M^T_4) + P(Y^T_1| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})$

Assumption #1: $P(Y^T_1| X=1, M^T_4) = P(Y^T_1|X=0, M^T_{24})$
$= P(Y^T_{24}| X=1, M^T_4)$

Assumption #2: $P(Y^T_{24}| X=1, M^T_{12}) = P(Y^T_{24}|X=1, M^T_4) = P(Y^T_{24}|X=1)$
$= P(Y^T_{24}| X=1)$

No-Confounding Assumption: $P(Y^T_{24}|X=1) = P(Y^T_{24}|X=0)$
$= P(Y^T_{24})$

<u>CDE (m=1)</u>
$=P(Y=1|X=1, M=1) – P(Y=1|X=0, M=1)$
$=P(Y_{11}=1|X=1, M_1=1) – P(Y_{01}=1|X=0, M_0=1)$
$= P(Y^T_{12468}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)$
$= P(Y^T_{48}| X=1, M^T_{12}) + P(Y^T_{126}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)$

Assumption #1: $P(Y^T_{126}| X=1, M^T_{12}) = P(Y^T_{126}|X=0, M^T_1)$
$= P(Y^T_{48}| X=1, M^T_{12})$

Assumption #2: $P(Y^T_{48}| X=1, M^T_{12}) = P(Y^T_{48}|X=1, M^T_4) = P(Y^T_{48}|X=1)$
$= P(Y^T_{48}| X=1)$

No-Confounding Assumption: $P(Y^T_{24}|X=1) = P(Y^T_{24}|X=0)$
$= P(Y^T_{48})$

<u>PDE</u>
$= P(M=1|X=0)*[P(Y=1|X=1, M=1) + P(M=0|X=0)*P(Y=1|X=1,M=0) – P(Y=1|X=0)$
$= P(M_0=1|X=0)*[P(Y_{11}=1|X=1, M_1=1) + P(M_0=0|X=0)*P(Y_{10}=1|X=1,M_1=0) – P(Y_1=1|X=0)$
$=P(M^T_1|X=0)* P(Y^T_{12468}|X=1, M^T_{12}) + P(M^T_{24}|X=0)*P(Y^T_{124}|X=1, M^T_4) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$=P(M^T_1|X=0)* P(Y^T_{68}|X=1, M^T_{12}) + P(M^T_1|X=0)* P(Y^T_{124}|X=1, M^T_{12}) + P(M^T_{24}|X=0)*P(Y^T_{124}|X=1, M^T_4) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$

Assumption #1: $P(Y^T_{124}|X=1, M^T_{12}) = P(Y^T_{124}|X=1, M^T_4) = P(Y^T_{124}|X=1)$
$=P(M^T_1|X=0)* P(Y^T_{68}|X=1, M^T_{12}) + P(M^T_1|X=0)* P(Y^T_{124}|X=1) + P(M^T_{24}|X=0)*P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$=P(M^T_1|X=0)* P(Y^T_{68}|X=1, M^T_{12}) + P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$

Assumption #2: $P(Y^T_{68}|X=1, M^T_{12}) = P(Y^T_{68}|X=0, M^T_1)$
$=P(M^T_1|X=0)* P(Y^T_{68}|X=0, M^T_1) + P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$=P(M^T_1Y^T_{68}|X=0) + P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$

No-Confounding Assumption: $P(M^T_1Y^T_{68}|X=0) = P(M^T_1Y^T_{68}|X=1)$
$=P(M^T_1Y^T_{68}|X=1) + P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)$

No-Confounding Assumption: $P(Y^T_1 + M^T_1Y^T_{26}|X=1) = P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$=P(M^T_1Y^T_{68}|X=1) + P(Y^T_{124}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=1)$
$=P(Y^T_{24} + M^T_1Y^T_8–M^T_1Y^T_2|X=1)$
$=P(Y^T_4 + M^T_{24}Y^T_2 + M^T_1Y^T_8|X=1)$

No-Confounding Assumption: $=P(Y^T_4 + M^T_{24}Y^T_2 + M^T_1Y^T_8|X=1) = P(Y^T_4 + M^T_{24}Y^T_2 + M^T_1Y^T_8)$
$=P(Y^T_4 + M^T_{24}Y^T_2 + M^T_1Y^T_8)$

<u>TDE</u>
$= P(Y=1|X=1) - P(M=1|X=1)*P(Y=1|X=0,M=1) – P(M=0|X=1)P(Y=1|X=0, M=0)$
$= P(Y_1=1|X=1) - P(M_1=1|X=1)*P(Y_{01}=1|X=0,M_0=1) – P(M_1=0|X=1)P(Y_{00}=1|X=0, M_0=0)$
$= P(Y^T_1 + Y^T_{24} + M^T_{12}Y^T_{68} |X=1) - P(M^T_{12}|X=1)* P(Y^T_{126}|X=0, M^T_1) - P(M^T_4|X=1)*P(Y^T_1|X=0, M^T_{24})$

Assumption #1: $P(Y^T_1|X=1) = P(Y^T_1|X=0, M^T_1) = P(Y^T_1|X=0, M^T_{24})$
$= P(Y^T_{24} + M^T_{12}Y^T_{68} |X=1) - P(M^T_{12}|X=1)* P(Y^T_{26}|X=0, M^T_1)$

Assumption #2: $P(Y^T_{26}|X=0, M^T_1) = P(Y^T_{26}|X=1, M^T_{12})$
$= P(Y^T_{24} + M^T_{12}Y^T_{68} - M^T_{12}Y^T_{26}|X=1)$
$= P(Y^T_4 + M^T_4Y^T_2 + M^T_{12}Y^T_8|X=1)$

No-Confounding Assumption: $P(Y^T_4 + M^T_4Y^T_2 + M^T_{12}Y^T_8|X=1) = P(Y^T_4 + M^T_4Y^T_2 + M^T_{12}Y^T_8)$
$P(Y^T_4 + M^T_4Y^T_2 + M^T_{12}Y^T_8)$

<u>CIE (m=0)</u>
$= [P(Y=1|X=1) – P(Y=1|X=0)] – [P(Y=1|X=1, M=0) – P(Y=1|X=0, M=0)]$

$= [P(Y_1=1|X=1) – P(Y_0=1|X=0)] – [P(Y_{10}=1|X=1, M_1=0) – P(Y_{00}=1|X=0, M_0=0)]$
$= [P(Y^T_{142} + M^T_{12}Y^T_{68}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)] - [P(Y^T_{124}| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})]$
$= [P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) + P(Y^T_1 + M^T_1Y^T_{26}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)] - [P(Y^T_{124}| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})]$

No-Confounding Assumption: $P(Y^T_1 + M^T_1Y^T_{26}|X=1) = P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - [P(Y^T_{124}| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})]$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - [P(Y^T_{24}| X=1, M^T_4) + P(Y^T_1| X=1, M^T_4) – P(Y^T_1|X=0, M^T_{24})]$
Assumption #1: $P(Y^T_1| X=1, M^T_4) = P(Y^T_1|X=0, M^T_{24})$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - P(Y^T_{24}| X=1, M^T_4)$

Assumption #2: $P(Y^T_{24}| X=1, M^T_{12}) = P(Y^T_{24}|X=1, M^T_4) = P(Y^T_{24}|X=1)$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - P(Y^T_{24}| X=1)$

$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8 – Y^T_{24} |X=1)$
$= P(M^T_2Y^T_6 + M^T_{12}Y^T_8 – M^T_1Y^T_2 |X=1)$

No-Confounding Assumption: $P(M^T_2Y^T_6 + M^T_{12}Y^T_8 – M^T_1Y^T_2 |X=1) = P(M^T_2Y^T_6 + M^T_{12}Y^T_8 – M^T_1Y^T_2)$
$= P(M^T_2Y^T_6 + M^T_{12}Y^T_8 – M^T_1Y^T_2)$

<u>CIE (m=1)</u>
$=[P(Y=1|X=1) – P(Y=1|X=0)] – [P(Y=1|X=1, M=1) – P(Y=1|X=0, M=1)]$
$=[P(Y_1=1|X=1) – P(Y_0=1|X=0)] – [P(Y_{11}=1|X=1, M_1=1) – P(Y_{01}=1|X=0, M_0=1)]$
$= [P(Y^T_{124} + M^T_{12}Y^T_{68}|X=1) – P(Y^T_1 + M^T_1Y^T_{62}|X=0)] - [P(Y^T_{12468}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)]$
$= [P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) + P(Y^T_1 + M^T_1Y^T_{26}|X=1) – P(Y^T_1 + M^T_1Y^T_{26}|X=0)] - [P(Y^T_{12468}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)]$

No-Confounding Assumption: $P(Y^T_1 + M^T_1Y^T_{26}|X=1) = P(Y^T_1 + M^T_1Y^T_{26}|X=0)$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - [P(Y^T_{12468}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)]$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - [P(Y^T_{48}| X=1, M^T_{12}) + P(Y^T_{126}| X=1, M^T_{12}) – P(Y^T_{126}|X=0, M^T_1)]$

Assumption #1: $P(Y^T_{126}| X=1, M^T_{12}) = P(Y^T_{126}|X=0, M^T_1)$
$= [P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - P(Y^T_{48}| X=1, M^T_{12})$

Assumption #2: $P(Y^T_{48}| X=1, M^T_{12}) = P(Y^T_{48}| X=1, M^T_4) = P(Y^T_{48}|X=1)$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8|X=1) - P(Y^T_{48}| X=1)$
$= P(Y^T_4 + M^T_2Y^T_6 + M^T_{24}Y^T_2 + M^T_{12}Y^T_8 – Y^T_{48} |X=1)$
$= P(M^T_2Y^T_6 + M^T_{24}Y^T_2 – M^T_4Y^T_8 |X=1)$

No-Confounding Assumption: $P(M^T_2Y^T_6 + M^T_{24}Y^T_2 – M^T_4Y^T_8 |X=1) = P(M^T_2Y^T_6 + M^T_{24}Y^T_2 – M^T_4Y^T_8)$
$= P(M^T_2Y^T_6 + M^T_{24}Y^T_2 – M^T_4Y^T_8)$

<u>PIE</u>
$= [P(M=1|X=1) – P(M=1|X=0)]*[P(Y=1|X=0,M=1) – P(Y=1|X=0,M=0)]$
$= [P(M_1=1|X=1) – P(M_0=1|X=0)]*[P(Y_{01}=1|X=0,M_0=1) – P(Y_{00}=1|X=0,M_0=0)]$
$= [P(M^T_{12}|X=1) – P(M^T_1|X=0)]*[P(Y^T_{126}|X=0,M^T_1) – P(Y^T_1|X=0,M^T_{24})]$

$= [P(M^T_2|X=1) + P(M^T_1|X=1) - P(M^T_1|X=0)]*[P(Y^T_{126}|X=0,M^T_1) - P(Y^T_1|X=0,M^T_{24})]$

No-Confounding Assumption: $P(M^T_1|X=1) = P(M^T_1|X=0)$
$=[P(M^T_2|X=1)*[P(Y^T_{126}|X=0,M^T_1) - P(Y^T_1|X=0,M^T_{24})]$
$= P(M^T_2|X=1)*[P(Y^T_{26}|X=0,M^T_1) + P(Y^T_1|X=0,M^T_1) - P(Y^T_1|X=0,M^T_{24})]$

Assumption #1: $P(Y^T_1|X=0,M^T_1) = P(Y^T_1|X=0,M^T_{24})$
$= P(M^T_2|X=1)*P(Y^T_{26}|X=0,M^T_1)$

Assumption #2: $P(Y^T_{26}|X=0,M^T_1)=P(Y^T_{26}|X=1,M^T_2)$
$= P(M^T_2|X=1)*P(Y^T_{26}|X=1,M^T_2)$
$= P(M^T_2Y^T_{26}|X=1)$

No-Confounding Assumption: $P(M^T_2Y^T_{26}|X=1) = P(M^T_2Y^T_{26})$
$= P(M^T_2Y^T_{26})$

Note: Assumption #2 for the PIE is identical to Assumption #2 for the TDE, assuming no confounding of the exposure-disease relationship.

PIE Assumption #2: $P(Y^T_{26}|X=0,M^T_1)=P(Y^T_{26}|X=1,M^T_2)$
TDE Assumption #2: $P(Y^T_{26}|X=0, M^T_1) = P(Y^T_{26}|X=1, M^T_{12})$

If (1) $P(Y^T_{26}|X=0, M^T_1) = P(Y^T_{26}|X=1, M^T_2)$ <PIE Assumption #2> and (2) exposed and unexposed fully exchangeable, then the following equality must hold:

$P(Y^T_{26}|X=0, M^T_1) = P(Y^T_{26}|X=1, M^T_{12})$ <TDE Assumption #2>

Proof:

$P(Y^T_{26}|X=1, M^T_{12})$ is a weighted average of $P(Y^T_{26}|X=1, M^T_1)$ and $P(Y^T_{26}|X=1, M^T_2)$

$P(Y^T_{26}|X=1, M^T_{12}) = P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{26}|X=1, M^T_1) +$
$P(M^T_2|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{26}|X=1, M^T_2)$

Based on (1), $P(Y^T_{26}|X=1, M^T_2) = P(Y^T_{26}|X=0, M^T_1)$
Based on (2), $P(Y^T_{26}|X=1, M^T_1) = P(Y^T_{26}|X=0, M^T_1)$

$P(Y^T_{26}|X=1, M^T_{12}) = P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)* P(Y^T_{26}|X=0, M^T_1) +$
$P(M^T_2|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{26}|X=0, M^T_1)$

$P(Y^T_{26}|X=1, M^T_{12}) = P(Y^T_{26}|X=0, M^T_1)$

<u>TIE</u>
$= [P(M=1|X=1) - P(M=1|X=0)]*[P(Y=1|X=1,M=1) - P(Y=1|X=1,M=0)]$
$= [P(M_1=1|X=1) - P(M_0=1|X=0)]*[P(Y_{11}=1|X=1,M_1=1) - P(Y_{10}=1|X=1,M_1=0)]$
$= [P(M^T_{12}|X=1) - P(M^T_1|X=0)]*[P(Y^T_{12468}|X=1,M^T_{12}) - P(Y^T_{124}|X=1,M^T_4)]$
$= [P(M^T_2|X=1) + P(M^T_1|X=1) - P(M^T_1|X=0)]* [P(Y^T_{12468}|X=1,M^T_{12}) - P(Y^T_{124}|X=1,M^T_4)]$

No-Confounding Assumption: $P(M^T_1|X=1) = P(M^T_1|X=0)$
$= P(M^T_2|X=1)* [P(Y^T_{12468}|X=1,M^T_{12}) - P(Y^T_{124}|X=1,M^T_4)]$
$= P(M^T_2|X=1)* [P(Y^T_{68}|X=1,M^T_{12}) + P(Y^T_{124}|X=1,M^T_{12}) - P(Y^T_{124}|X=1,M^T_4)]$

Assumption #1: $P(Y^T_{124}|X=1,M^T_{12}) = P(Y^T_{124}|X=1,M^T_4)$
$= P(M^T_2|X=1)*P(Y^T_{68}|X=1,M^T_{12})$

Assumption #2: $P(Y^T_{68}|X=1,M^T_{12})=P(Y^T_{68}|X=1,M^T_2)$
$= P(M^T_2|X=1)*P(Y^T_{68}|X=1,M^T_2)$
$= P(M^T_2Y^T_{68}|X=1)$

No-Confounding Assumption: $P(M^T_2Y^T_{68}|X=1) = P(M^T_2Y^T_{68})$
$= P(M^T_2Y^T_{68})$

Note: Assumption #2 for the PDE is identical to Assumption #2 for the TIE.

TIE Assumption #2: $P(Y^T_{68}|X=1,M^T_{12})=P(Y^T_{68}|X=1,M^T_2)$
PDE Assumption #2: $P(Y^T_{68}|X=1, M^T_{12}) = P(Y^T_{68}|X=0, M^T_1)$

If $P(Y^T_{68}|X=1,M^T_{12})=P(Y^T_{68}|X=1,M^T_2)$ <TIE Assumption #2>, then the following equality must hold:

$P(Y^T_{68}|X=1, M^T_{12}) = P(Y^T_{68}|X=0, M^T_1)$ <PDE Assumption #2>

Proof:

$P(Y^T_{68}|X=1, M^T_{12})$ is a weighted average of $P(Y^T_{68}|X=1, M^T_1)$ and $P(Y^T_{68}|X=1, M^T_2)$

$P(Y^T_{68}|X=1, M^T_{12}) = P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_1) +$
$P(M^T_2|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_2)$

Based on TIE Assumption #2, $P(Y^T_{68}|X=1,M^T_{12})=P(Y^T_{68}|X=1,M^T_2)$

$P(Y^T_{68}|X=1, M^T_{12}) = P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_1) +$
$P(M^T_2|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_{12})$

$P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_{12}) = P(M^T_1|X=1)/P(M^T_1+M^T_2|X=1)*P(Y^T_{68}|X=1, M^T_1)$
$P(Y^T_{68}|X=1, M^T_{12}) = P(Y^T_{68}|X=1, M^T_1)$

By randomization, $P(Y^T_{68}|X=1, M^T_1) = P(Y^T_{68}|X=0, M^T_1)$
$P(Y^T_{68}|X=1, M^T_{12}) = P(Y^T_{68}|X=0, M^T_1)$ <PDE Assumption #2>

eAPPENDIX C. This appendix translates previously derived assumptions into our notation of potential outcomes and response types.

CDE(M=0)/CIE(M=0): *Robins and Greenland 1992*

Assumption #1: $R_{00} = R_{01B}$
R_{00} = risk of the outcome in unexposed, mediator-negative population [$P(Y_{00}=1|X=0,M=0)$]
R_{01B}= risk of the outcome in unexposed, mediator-positive population, if the mediator had been prevented [$P(Y_{00}=1|X=0,M=1)$]

This assumption can be stated in terms of potential outcomes and response types:
$P(Y_{00}=1|X=0,M=0) = P(Y_{00}=1|X=0,M=1)$
$P(Y^T_1|X=0, M^T_{24}) = P(Y^T_1|X=0, M^T_1)$

Assumption #2: $R_{10} = R_{11B}$
R_{10} = risk of the outcome in exposed, mediator-negative population [$P(Y_{10}=1|X=1,M=0)$]
R_{11B}= risk of the outcome in exposed, mediator-positive population, if the mediator had been prevented [$P(Y_{10}=1|X=1,M=1)$]

This assumption can be stated in terms of potential outcomes and response types:
$P(Y_{10}=1|X=1,M=0) = P(Y_{10}=1|X=1,M=1)$
$P(Y^T_{124}|X=1, M^T_4) = P(Y^T_{124}|X=1, M^T_{12})$

CDE(M=1)/CIE(M=1): *Kaufman et al. 2004*

Assumption #1: $R_{00|SET[Z=1]} = R_{01}$
R_{01} = risk of the outcome in the unexposed, mediator-positive population [$P(Y_{01}=1|X=0,M=1)$]
$R_{00|SET[Z=1]}$ = risk of the outcome in the unexposed, mediator-negative population, if they were given the mediator [$P(Y_{01}=1|X=0,M=0)$]

This assumption can be stated in terms of potential outcomes and response types:
$P(Y_{01}=1|X=0, M=1) = P(Y_{01}=1|X=0, M=0)$
$P(Y^T_{126}|X=0, M^T_1) = P(Y^T_{126}|X=0, M^T_{24})$

Assumption #2: $R_{10|SET[Z=1]} = R_{11}$
R_{11} = risk of the outcome in the exposed, mediator-positive population [$P(Y_{11}=1|X=1,M=1)$]
$R_{10|SET[Z=1]}$ = risk of the outcome in the exposed, mediator-negative population, if they were given the mediator [$P(Y_{11}=1|X=1,M=0)$]

This assumption can be stated in terms of potential outcomes and response types:
$P(Y_{11}=1|X=1, M=1) = P(Y_{11}=1|X=1, M=0)$
$P(Y^T_{12468}|X=1, M^T_{12}) = P(Y^T_{12468}|X=1, M^T_4)$

PDE/TIE: *Pearl 2001*

Assumption #1: $P(Y_{1m}=1|C)$ is identifiable
If there are no measured confounders and the outcome is dichotomous, the assumption is simplified to the following: $P(Y_{1m}=1)$ is identifiable. $P(Y_{1m}=1)$ is identifiable if the risk for the entire population [$P(Y_{1m}=1)$] is estimated by the observable subset [$P(Y_{1m}=1|X=1,M=m)$]. This yields:

$P(Y_{1m}=1|X=1,M=m) = P(Y_{1m}=1)$

Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{11}=1|X=1,M=1) = P(Y_{11}=1)$ AND
$P(Y_{10}=1|X=1,M=0) = P(Y_{10}=1)$

These can be stated in terms of response types:
$P(Y^T_{12468}|X=1, M^T_{12}) = P(Y^T_{12468})$ AND
$P(Y^T_{124}|X=1, M^T_{4}) = P(Y^T_{124})$

<u>Assumption #2:</u> $Y_{1m} \perp\!\!\!\perp M_0|C$
This assumption states that the risk of disease in the exposed individuals with a specified mediator status (Y_{1m}) is independent of the mediator in the absence of exposure (M_0), conditional on a set of measured confounders delineated by C. If there are no measured confounders, the assumption is simplified to the following: $Y_{1m} \perp\!\!\!\perp M_0$. This yields:

$P(Y_{1m}=1|M_0=1) = P(Y_{1m}=1|M_0=0)$

Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{11}=1|M_0=1) = P(Y_{11}=1|M_0=0)$ AND
$P(Y_{10}=1|M_0=1) = P(Y_{10}=1|M_0=0)$

These can be stated in terms of response types:
$P(Y^T_{12468}|M^T_{1}) = P(Y^T_{12468}|M^T_{24})$ AND
$P(Y^T_{124}|M^T_{1}) = P(Y^T_{124}|M^T_{24})$

PDE/TIE: *Petersen et al. 2006*

<u>Assumption #1:</u> $M \perp\!\!\!\perp Y_{xm}|X,C$
If there are no measured confounders, the assumption is simplified to the following: $M \perp\!\!\!\perp Y_{xm}|X$. The outcome that would be observed given a particular exposure and mediator must be independent of the observed mediator status. Given no unmeasured confounders of the exposure-disease relationship, this assumption is equivalent to Assumption #1 by Pearl (2001). This assumption can thus be stated in terms of potential outcomes:

$P(Y_{1m}=1|X=1,M=m) = P(Y_{1m}=1)$

Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{11}=1|X=1,M=1) = P(Y_{11}=1)$ AND
$P(Y_{10}=1|X=1,M=0) = P(Y_{10}=1)$

These can be stated in terms of response types:
$P(Y^T_{12468}|X=1, M^T_{12}) = P(Y^T_{12468})$ AND
$P(Y^T_{124}|X=1, M^T_{4}) = P(Y^T_{124})$

<u>Assumption #2:</u> $E(Y_{1m} - Y_{0m}|M_0=m,C) = E(Y_{1m} - Y_{0m}|C)$
If there are no measured confounders and the outcome is dichotomous, this assumption is simplified to the following: $P(Y_{1m} - Y_{0m}|M_0=m) = P(Y_{1m} - Y_{0m})$. This assumption implies that the

direct effect of X (on Y) (either in the presence or absence of the mediator) does not depend on the value of the mediator in the absence of exposure.

$P(Y_{1m} - Y_{0m}=1|M_0=1) = P(Y_{1m} - Y_{0m}=1|M_0=0)$

Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{11} - Y_{01}=1|M_0=1) = P(Y_{11} - Y_{01}=1|M_0=0)$ AND
$P(Y_{10} - Y_{00}=1|M_0=1) = P(Y_{10} - Y_{00}=1|M_0=0)$

These can be stated in terms of response types:
$P(Y^T_{48}|M^T_1) = P(Y^T_{48}|M^T_{24})$
$P(Y^T_{24}|M^T_1) = P(Y^T_{24}|M^T_{24})$

PIE/TDE: *Pearl 2001*

<u>Assumption #1:</u> $E(Y_{0m}|C)$ is identifiable
If there are no measured confounders and the outcome is dichotomous, the assumption is simplified to the following: $P(Y_{0m}=1)$ is identifiable. $P(Y_{0m}=1)$ is identifiable if the risk for the entire population $[P(Y_{0m}=1)]$ is estimated by the observable subset $[P(Y_{0m}=1|X=0,M=m)]$.

$P(Y_{0m}=1|X=0,M=m) = P(Y_{0m}=1)$

Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{01}=1|X=0,M=1) = P(Y_{01}=1)$ AND
$P(Y_{00}=1|X=0,M=0) = P(Y_{00}=1)$

These can be stated in terms of response types:
$P(Y^T_{126}|X=0, M^T_1) = P(Y^T_{126})$
$P(Y^T_1|X=0, M^T_{24}) = P(Y^T_1)$

<u>Assumption #2:</u> $Y_{0m} \perp\!\!\!\perp M_1|C$
If there are no measured confounders, the assumption is simplified to the following: $Y_{0m} \perp\!\!\!\perp M_1$. This assumption implies that risk of disease in the unexposed individuals with a specified mediator status (M=0 or M=1) is independent of the value of the mediator in the presence of exposure. This yields:

$P(Y_{0m}=1|M_1=1) = P(Y_{0m}=1|M_1=0)$

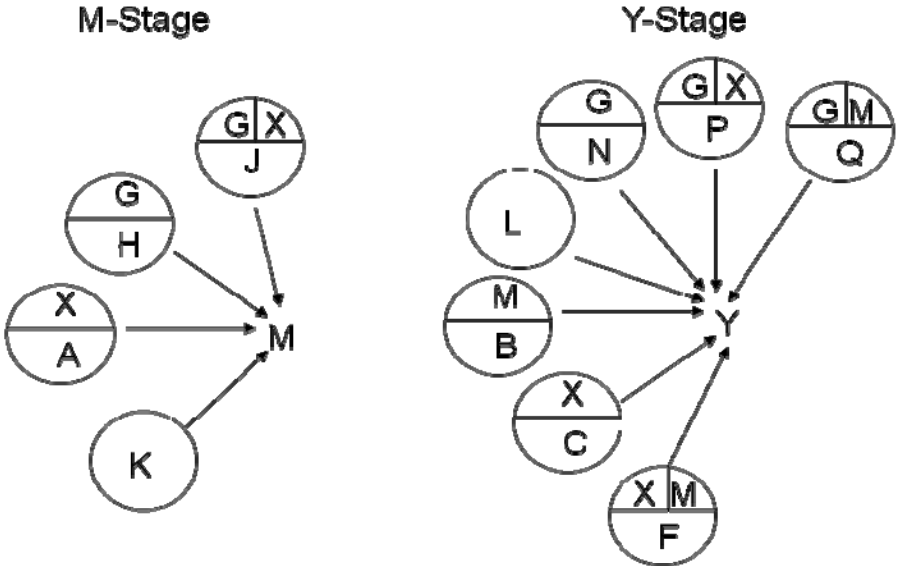
Because the mediator is dichotomous (M=1 or M=0), this assumption can be re-stated as:
$P(Y_{01}=1|M_1=1) = P(Y_{01}=1|M_1=0)$ AND
$P(Y_{00}=1|M_1=1) = P(Y_{00}=1|M_1=0)$

These can be stated in terms of response types:
$P(Y^T_{126}|M^T_{12}) = P(Y^T_{126}|M^T_4)$
$P(Y^T_1|M^T_{12}) = P(Y^T_1|M^T_4)$

eAPPENDIX D. This appendix provides support for the assertion that certain confounders will violate previous, but not current, assumptions. Specifically, in the absence of a true direct effect of exposure, a confounder that does not interact with exposure to cause the mediator will not bias the estimate of CDE(m=0).

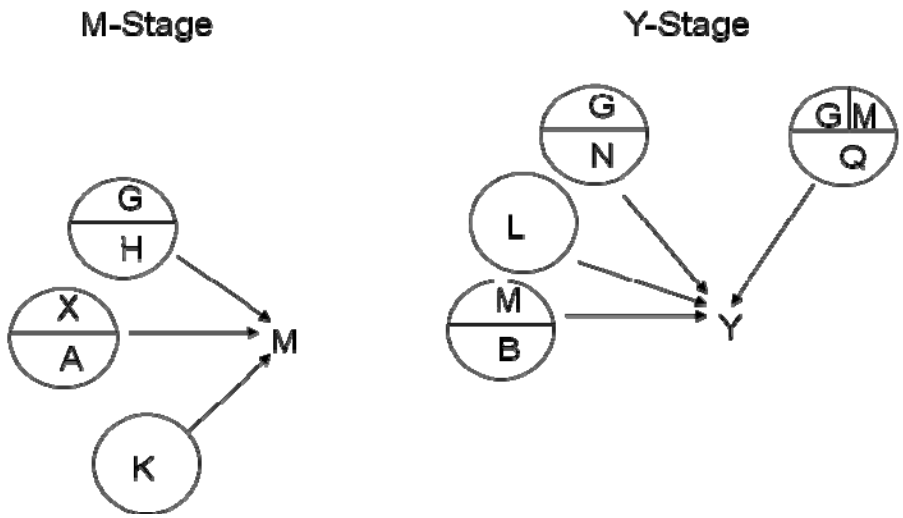
Appendix Figure 1 shows a minimal sufficient cause (MSC) model for mediation, with common cause (G) of the mediator (M) and outcome (Y). Based on the principles of the MSC model, an outcome will occur if at least one sufficient cause is completed; that is, if every component cause within a sufficient cause is present. This MSC model for mediation is a two-stage model, reflecting that an indirect effect involves 2 steps. In the M-stage, the exposure (X) causes the mediator (M) in the presence of other factors (A). In the Y-stage, the mediator (M) causes the outcome (Y) in the presence of other factors [B, (G and Q), or (X and F)]. There is also a direct effect of exposure (X) on the outcome (Y), in the presence of other factors [C, (G and P), or (M and F)]. Finally, the confounder (G) is a common cause of the mediator and outcome. Note that this model allows for two-way interaction between exposure (X) and confounder (G) to cause the mediator and outcome; exposure (X) and mediator (M) to cause the outcome; and mediator (M) and confounder (G) to cause the outcome. For more details, see reference 7.

eAppendix D Figure 1



Appendix Figure 2 is a simplified MSC model, given the following stipulations (as described in the text): (1) no interaction between exposure (X) and confounder (G) to cause the mediator and (2) no direct effect of exposure (X) on outcome (Y). We assume the mechanisms for M and for Y are independent i.e. that the background causes A, H, K, B, L, N, Q are independent. Under these conditions, previous assumptions indicate that G will bias our estimate of the CDE(m=0), thus leading us to incorrectly conclude that exposure has a direct effect on outcome. Current assumptions, on the other hand, indicate that G will not bias assessment of the direct effect. To prove this, we define each assumption (both current and previous) in terms of common causes. We then simplify.

eAppendix D Figure 2



To distinguish between random variables and probabilities, we denote random variables with bold font. For example, **G** refers to the random variable, while G refers to the proportion with the random variable [i.e. G=P(**G**=1)]. [X **V** Y] indicates X or Y, which implies Boolean addition: [X **V** Y] = X+Y-XY. G_{M^T} is short-hand for the probability: P(**G**=1|M^T), where M^T is one or more M-types (M^T_1, M^T_2, and/or M^T_4). Note that for the following calculations we assume that all component causes are independent, unless one causes the other or they share a common cause.

Step #1: Define conditional probabilities of G given M-type defined subsets. We use Bayes' theorem to calculate these probabilities:
P(G=g|M-type) = P(M-type|G=g)*P(G=g)/P(M-type).

G_{M^T1}= P(**G**=1| M^T_1)
= P(M^T_1|**G**=1)*P(**G**=1)/P(M^T_1)
= (H **V** K)G/(GH **V** K)

G_{M^T2}= P(**G**=1| M^T_2)
= P(M^T_2|**G**=1)*P(**G**=1)/P(M^T_2)
= (1-K)(1-H)AG/[(1-G)(1-K)A+G(1-K)(1-H)A]
= (1-H)G/[(1-G)+G(1-H)]
= (1-H)G/(1-GH)

G_{M^T4}= P(**G**=1| M^T_4)
= P(M^T_4|**G**=1)*P(**G**=1)/P(M^T_4)
= (1-K)(1-A)(1-H)G/[(1-G)(1-K)(1-A) + G(1-H)(1-K)(1-A)]
= (1-H)G/[(1-G) + G(1-H)]
= (1-H)G/(1-GH)

G_{M^T12}= P(**G**=1| M^T_{12})
= P(M^T_{12}|**G**=1)*P(**G**=1)/P(M^T_{12})

$= (K \mathbf{V} A \mathbf{V} H)G/[G(K \mathbf{V} A \mathbf{V} H) + (1-G)(K \mathbf{V} A)]$

$G_{M^T24} = P(\mathbf{G}=1 | M^T_{24})$
$= P(M^T_{24}|\mathbf{G}=1)*P(\mathbf{G}=1)/P(M^T_{24})$
$=(1-K)(1-H)G/[(1-K)(1-GH)]$
$=(1-H)G/(1-GH)$

2. Define relevant response types according to component causes.

Y^T_1: L or (G and N)
$P(Y^T_1) = L+(1-L)NG_{M^T}$

Y^T_{24}: In the absence of a direct effect of X on Y, $P(Y^T_{24}) = 0$

Y^T_{124}: In the absence of a direct effect of X on Y, $P(Y^T_{124}) = P(Y^T_1)$

<u>Current assumptions</u>

Assumption #1: $P(Y^T_1|X=0, M^T_{24})=P(Y^T_1|X=1, M^T_4)$

This assumption is not violated by confounder G. Proof:
$P(Y^T_1|X=0,M^T_{24}) = L+(1-L)NG_{M^T24} = L + (1-L)NG(1-H)/(1-GH)$
$P(Y^T_1|X=1, M^T_4) = L+(1-L)NG_{M^T4} = L + (1-L)NG(1-H)/(1-GH)$
$(1-L)NG(1-H)/(1-GH)= (1-L)NG(1-H)/(1-GH)$

Assumption #2: $P(Y^T_{24}|X=1) = P(Y^T_{24}|X=1, M^T_4)$

This assumption is not violated by confounder G. Proof:
$P(Y^T_{24}|X=1) = P(Y^T_{24}|X=1, M^T_4) = 0$

<u>Previous assumptions</u>

Assumption #1: $P(Y^T_1|X=0, M^T_1)=P(Y^T_1|X=0, M^T_{24})$

This assumption is violated by confounder G.Proof:
$P(Y^T_1|X=0, M^T_1)= L+(1-L)NG_{M^T1} = L + (1-L)N(H \mathbf{V} K)G/(GH \mathbf{V} K)$
$P(Y^T_1|X=0, M^T_{24})= L+(1-L)NG_{M^T24}= L + (1-L)N(1-H)G/(1-GH)$
$L + (1-L)N(H \mathbf{V} K)G/(GH \mathbf{V} K) \neq L + (1-L)N(1-H)G/(1-GH)$

Assumption #2: $P(Y^T_{124}| X=1, M^T_4) =P(Y^T_{124}| X=1, M^T_{12})$

This assumption is violated by confounder G.Proof:
$P(Y^T_{124}| X=1, M^T_4) = L+(1-L)NG_{M^T4} = L + (1-L)N(1-H)G/(1-GH)$
$P(Y^T_{124}| X=1, M^T_{12}) = L+(1-L)NG_{M^T12}= L + (1-L)N(K \mathbf{V} A \mathbf{V} H)G/[G(K \mathbf{V} A \mathbf{V} H) + (1-G)(K \mathbf{V} A)]$
$L + (1-L)N(1-H)G/(1-GH) \neq L + (1-L)N(K \mathbf{V} A \mathbf{V} H)G/[G(K \mathbf{V} A \mathbf{V} H) + (1-G)(K \mathbf{V} A)]$