Mediation

Total effect:

$$TE_{x,x'} = E\left[Y \mid X = x'\right] - E\left[Y \mid X = x\right]$$

$$TE_{x,x'} = \sum_z E\left[Y \mid X = x', z\right] \Pr\left(z \mid X = x'\right) - E\left[Y \mid X = x, z\right] \Pr\left(z \mid X = x\right)$$

Controlled (at $Z = z$) direct effect:

$$CDE_{x,x'} = E\left[Y \mid X = x', z\right] - E\left[Y \mid X = x, z\right]$$

Natural direct effect (weighted average of controlled direct effect):

$$NDE_{x,x'} = \sum_z \left(E\left[Y \mid X = x', z\right] - E\left[Y \mid X = x, z\right]\right) \Pr\left(z \mid X = x\right)$$

$$NDE_{x,x'} = \sum_z E[Y \mid X = x', z] \Pr\left(z \mid X = x\right) - E\left[Y \mid X = x, z\right] \Pr\left(z \mid X = x\right)$$

(Natural) indirect effect:

$$IE_{x,x'} = \sum_z E\left[Y \mid X = x, z\right]\left\{\Pr\left(z \mid X = x'\right) - \Pr\left(z \mid X = x\right)\right\}$$

$$IE_{x,x'} = \sum_z E\left[Y \mid X = x, z\right]\Pr\left(z \mid X = x'\right) - E\left[Y \mid X = x, z\right]\Pr\left(z \mid X = x\right)$$

Reverse indirect effect:

$$IE_{x',x} = \sum_z E\left[Y \mid X = x', z\right].\Pr\left(z \mid X = x\right) - E\left[Y \mid X = x', z\right]\Pr\left(z \mid X = x'\right)$$

$$TE_{x,x'} = NDE_{x,x'} - IE_{x',x}$$

$$\sum_z E\left[Y \mid X = x', z\right]\Pr\left(z \mid X = x'\right) - E\left[Y \mid X = x, z\right]\Pr\left(z \mid X = x\right)$$

$$= \sum_z E[Y \mid X = x', z]\Pr\left(z \mid X = x\right) - E\left[Y \mid X = x, z\right]\Pr\left(z \mid X = x\right)$$

$$-\left\{\sum_z E\left[Y \mid X = x', z\right]\Pr\left(z \mid X = x\right) - E\left[Y \mid X = x', z\right]\Pr\left(z \mid X = x'\right)\right\}$$

Linear model with interactions:

$$x = a_0 + \epsilon_1$$
$$z = b_0 + \beta x + \epsilon_2$$
$$y = c_0 + \alpha x + \gamma z + \delta xz + \epsilon_3$$

$$TE_{0,1} = \sum_z E\left[Y \mid X = 1, z\right] \Pr\left(z \mid X = 1\right) - E\left[Y \mid X = 0, z\right] \Pr\left(z \mid X = 0\right)$$
$$= \left\{c_0 + \alpha\left(1\right) + \gamma E\left[Z \mid X = 1\right] + \delta(1) E\left[Z \mid X = 1\right]\right\}$$
$$- \left\{c_0 + \alpha\left(0\right) + \gamma E\left[Z \mid X = 0\right] + \delta(0) E\left[Z \mid X = 0\right]\right\}$$
$$= \alpha + \gamma\left(b_0 + \beta - b_0\right) + \delta\left(b_0 + \beta - 0\right)$$
$$TE_{0,1} = \alpha + \delta b_0 + \beta\left(\gamma + \delta\right)$$

$$DE_{0,1} = \sum_{z} E[Y \mid X = x', z] \, \mathrm{Pr}\,(z \mid X = x) - E\,[Y \mid X = x, z] \, \mathrm{Pr}\,(z \mid X = x)$$

$$= \{c_0 + \alpha\,(1) + \gamma E\,[Z \mid X = 0] + \delta(1) E\,[Z \mid X = 0]\}$$

$$- \{c_0 + \alpha\,(0) + \gamma E\,[Z \mid X = 0] + \delta(0) E\,[Z \mid X = 0]\}$$

$$= \alpha + \delta\,(b_0 - 0)$$

$$DE_{0,1} = \alpha + b_0 \delta$$

$$IE_{0,1} = \sum_{z} E\,[Y \mid X = x, z] \,.\, \mathrm{Pr}\,(z \mid X = x') - E\,[Y \mid X = x, z] \, \mathrm{Pr}\,(z \mid X = x)$$

$$= \{c_0 + \alpha\,(0) + \gamma E\,[Z \mid X = 1] + \delta(0) E\,[Z \mid X = 1]\}$$

$$- \{c_0 + \alpha\,(0) + \gamma E\,[Z \mid X = 0] + \delta(0) E\,[Z \mid X = 0]\}$$

$$IE_{0,1} = \gamma \beta$$

$$IE_{1,0} = \sum_{z} E\left[Y \mid X = x', z\right] . \Pr\left(z \mid X = x\right) - E\left[Y \mid X = x', z\right] \Pr\left(z \mid X = x'\right)$$

$$= \left\{c_0 + \alpha\left(1\right) + \gamma E\left[Z \mid X = 0\right] + \delta(1) E\left[Z \mid X = 0\right]\right\}$$

$$- \left\{c_0 + \alpha\left(1\right) + \gamma E\left[Z \mid X = 1\right] + \delta(1) E\left[Z \mid X = 1\right]\right\}$$

$$= -\gamma\beta - \delta\beta$$

$$IE_{1,0} = -\beta\left(\delta + \gamma\right)$$

$$TE_{0,1} = DE_{0,1} - IE_{1,0} = \alpha + b_0\delta - \left\{-\beta\left(\delta + \gamma\right)\right\}$$

$$TE_{0,1} = \alpha + \delta b_0 + \beta\left(\gamma + \delta\right)$$

Binary example:

$$E\left[Y \mid X=0, Z=0\right]=0.2$$
$$E\left[Y \mid X=0, Z=1\right]=0.3$$
$$E\left[Y \mid X=1, Z=0\right]=0.4$$
$$E\left[Y \mid X=1, Z=1\right]=0.8$$
$$E\left[Z \mid X=0\right]=0.4$$
$$E\left[Z \mid X=1\right]=0.75$$

$$TE_{0,1}=\sum_{z} E\left[Y \mid X=1, z\right]\operatorname{Pr}\left(z \mid X=1\right)-E\left[Y \mid X=0, z\right]\operatorname{Pr}\left(z \mid X=0\right)$$
$$=\{0.4\,(1-0.75)+0.8\,(0.75)\}-\{0.2\,(1-0.4)+0.3\,(0.4)\}=0.46$$

$$DE_{0,1} = \sum_z E[Y \mid X = x', z] \Pr(z \mid X = x) - E[Y \mid X = x, z] \Pr(z \mid X = x)$$

$$= \{0.4(1 - 0.4) + 0.8(0.4)\} - \{0.2(1 - 0.4) + 0.3(0.4)\} = 0.32$$

$$IE_{0,1} = \sum_z E[Y \mid X = x, z] . \Pr(z \mid X = x') - E[Y \mid X = x, z] \Pr(z \mid X = x)$$

$$= \{0.2(1 - 0.75) + 0.3(0.75)\} - \{0.2(1 - 0.4) + 0.3(0.4)\} = 0.035$$

$$IE_{1,0} = \sum_z E[Y \mid X = x', z] . \Pr(z \mid X = x) - E[Y \mid X = x', z] \Pr(z \mid X = x')$$

$$= \{0.4(1 - 0.4) + 0.8(0.4)\} - \{0.4(1 - 0.75) + 08.(0.75)\} = -0.14$$

$$TE_{0,1} = DE_{0,1} - IE_{1,0}$$

$$0.46 = 0.32 - (-0.14)$$