13 Appendix

13.1 Causal effects with continuous mediator and continuous outcome

Consider the model of Section 3,

$$y_i = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, m_i + \beta_4\, c_i + \epsilon_{1i}, \tag{49}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \epsilon_{2i}, \tag{50}$$

where the residuals ϵ_1 and ϵ_2 are assumed normally distributed with zero means, variances σ_1^2, σ_2^2, and uncorrelated with each other and with the predictors in their equations. The definitions for the direct, total indirect, pure indirect, and total effects are

$$DE = E[Y(1, M(0)) - Y(0, M(0)) \mid C], \tag{51}$$

$$TIE = E[Y(1, M(1)) - Y(1, M(0)) \mid C], \tag{52}$$

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C], \tag{53}$$

$$TE = E[Y(1, M(1)) - Y(0, M(0)) \mid C]. \tag{54}$$

The general expression used in these differences is $E[Y(x, M(x')) \mid C = c]$, where x, x' equal 0 or 1. Because the expectation is not conditioned on m, it is obtained by integrating out the mediator M,

$$E[Y(x, M(x')) \mid C = c] = \int_{-\infty}^{+\infty} (\beta_0 + \beta_1\, m + \beta_2\, x + \beta_3\, x\, m + \beta_4\, c) \tag{55}$$

$$\times\, f(m; \gamma_0 + \gamma_1\, x' + \gamma_2\, c, \sigma_2^2)\, \partial M, \tag{56}$$

$$= \beta_0 + \beta_2\, x + \beta_4\, c \tag{57}$$

$$+\, \beta_1\, E[M \mid X = x', C = c] + \beta_3\, x\, E[M \mid X = x', C = c], \tag{58}$$

where the last step is obtained by the fact that $\int\, Z \times f(Z; E[Z], V(Z))\, \partial Z = E[Z]$. Using this general expression, the four effects are obtained by alternating the 0 and 1 values of x and x'. Straightforward simplifications of the differences give

$$DE = E[Y(1, M(0)) - Y(0, M(0)) \mid C = c] = \beta_2 + \beta_3\, \gamma_0 + \beta_3\, \gamma_2\, c, \tag{59}$$

$$TIE = E[Y(1, M(1)) - Y(1, M(0)) \mid C = c] = \beta_1\, \gamma_1 + \beta_3\, \gamma_1, \tag{60}$$

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C = c] = \beta_1\, \gamma_1, \tag{61}$$

$$TE = E[Y(1, M(1)) - Y(0, M(0)) \mid C = c] = \beta_2 + \beta_1\, \gamma_1 + \beta_3\, \gamma_0 + \beta_3\, \gamma_1 + \beta_3\, \gamma_2\, c. \tag{62}$$

This agrees with the results in the Appendix of VanderWeele and Vansteelandt (2009), setting $a = 1$ and $a^* = 0$ using their notation for x.

13.2 Causal effects with a continuous mediator and a binary outcome

Consider a mediation model for a binary outcome y and a continuous mediator m. Assume a probit link for the binary outcome,

$$probit(y_i) = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, m_i + \beta_4\, c_i, \tag{63}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \epsilon_{2i}, \tag{64}$$

where the residual ϵ_2 is assumed normally distributed as before and where $probit(y_i)$ is defined via

$$P(Y_i = 1 \mid m, x, c) = \int_{-\infty}^{probit(y_i)} f(z; 0, 1)\, \partial z = \Phi[probit(y_i)], \tag{65}$$

where $f(z; 0, 1)$ is a standard normal density with mean zero and variance one and Φ is a standard normal distribution function. Equivalently, (63) can be expressed with a continuous latent response variable y_i^* as the dependent variable, adding a residual with variance one.

As in section 13.1 consider the general expression $E[Y(x, M(x'))\ |C = c]$. Integrating over M,

$$E[Y(x, M(x')) \mid C = c] = \int_{-\infty}^{\infty} E[Y \mid C = c, X = x, M = m] \times f(M \mid C = c, X = x')\, \partial M \tag{66}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{probit(y)} f(z; 0, 1) \partial z \times f(M; \gamma_0 + \gamma_1\, x' + \gamma_2\, c, \sigma_2^2)\, \partial M \tag{67}$$

$$= \int_{-\infty}^{probit(x, x')} f(z; 0, 1)\, \partial z, \tag{68}$$

where the last equality can be derived by a variable transformation and a change of the order of integration as in Muthén (1979; p. 810, Appendix Theorem) with

$$probit(x,x') = [\beta_0 + \beta_2\, x + \beta_4\, c + (\beta_1 + \beta_3\, x)(\gamma_0 + \gamma_1\, x' + \gamma_2\, c)]/\sqrt{v(x)}, \quad (69)$$

where the variance $v(x)$ is

$$v(x) = (\beta_1 + \beta_3\, x)^2\, \sigma_2^2 + 1. \quad (70)$$

where σ_2^2 is the residual variance for the continuous mediator m.

13.2.1 Indirect effect

Consider the causally-defined indirect effect of the binary treatment variable x scored 0 for controls and 1 for treatment,

$$\begin{aligned}
&E[Y(1,M(1)) - Y(1,M(0)) \mid C] = \\
&= \int_{-\infty}^{\infty} E[Y \mid C=c, X=1, M=m] \times f(M \mid C=c, X=1)\, \partial M \\
&- \int_{-\infty}^{\infty} E[Y \mid C=c, X=1, M=m] \times f(M \mid C=c, X=0)\, \partial M. \quad (71)
\end{aligned}$$

This corresponds to the total indirect effect of (8).

For the first term in (71), x = x'=1. For the second term, x=1, x'=0, so that the $\gamma_1\, x'$ term of (69) is zero, thereby blocking the SEM, reduced-form indirect effect for the probit. The two terms of (71) can therefore be expressed as

$$\Phi[probit(1,1)] - \Phi[probit(1,0)], \quad (72)$$

where again the first probit index refers to x and the second to x' with values 0 and 1 for the control and treatment group inserted in (69).

Similarly, the pure indirect effect of (13) can be expressed as

$$\Phi[probit(0,1)] - \Phi[probit(0,0)], \tag{73}$$

These results agree with those presented in Imai et al. (2010a, Appendix F).

13.2.2 Direct effect

The causally-defined direct effect is

$$E[Y(1, M(0)) - Y(0, M(0)) \mid C] = \tag{74}$$
$$= \int_{-\infty}^{\infty} \{E[Y \mid C = c, X = 1, M = m] - E[Y \mid C = c, X = 0, M = m]\}$$
$$\times f(M \mid C = c, X = 0)\ \partial M. \tag{75}$$

It follows from the above derivations of the indirect effect that the direct effect can be expressed as

$$\Phi[probit(1,0)] - \Phi[probit(0,0)]. \tag{76}$$

13.3 Causal effects for binary mediator and binary outcome

In Section 4 the direct effect is defined as

$$E[Y(1, M(0)) - Y(0, M(0)) \mid C] = \tag{77}$$

$$= \int_{-\infty}^{\infty} \{E[Y \mid C = c, X = 1, M = m] - E[Y \mid C = c, X = 0, M = m]\}$$

$$\times f(M \mid C = c, X = 0)\, \partial M \tag{78}$$

and the total indirect effect as

$$E[Y(1, M(1)) - Y(1, M(0)) \mid C] = \tag{79}$$

$$= \int_{-\infty}^{\infty} E[Y \mid C = c, X = 1, M = m] \times f(M \mid C = c, X = 1)\, \partial M$$

$$- \int_{-\infty}^{\infty} E[Y \mid C = c, X = 1, M = m] \times f(M \mid C = c, X = 0)\, \partial M. \tag{80}$$

With a binary mediator M, the integral is replaced by a sum over the two values of M and the density f is replaced by the probability of M=0, 1. Note also that for a binary variable Y, $E[Y] = P(Y = 1)$. Let $F_Y(x, m)$ denote $P(Y = 1 \mid X = x, M = m)$ and let $F_M(x)$ denote $P(M = 1 \mid X = x)$, where F denotes either the standard normal or the logistic distribution function corresponding to using probit or logistic regression. It follows that the direct effect is

$$E[Y(1, M(0)) - Y(0, M(0)) \mid C] = \tag{81}$$

$$[F_Y(1, 0) - F_Y(0, 0)]\, [1 - F_M(0)] + [F_Y(1, 1) - F_Y(0, 1)]\, F_M(0) \tag{82}$$

and the total indirect effect is

$$E[Y(1, M(1)) - Y(1, M(0)) \mid C] = \tag{83}$$

$$F_Y(1, 0)\ [1 - F_M(1)] + F_Y(1, 1)\ F_M(1) \tag{84}$$

$$-\ F_Y(1, 0)\ [1 - F_M(0)] - F_Y(1, 1)\ F_M(0) \tag{85}$$

$$= F_Y(1, 0)\ [F_M(0) - F_M(1)] + F_Y(1, 1)\ [F_M(1) - F_M(0)] \tag{86}$$

$$= [F_Y(1, 1) - F_y(1, 0)]\ [F_M(1) - F_M(0)]. \tag{87}$$

The pure indirect effect is similarly derived as

$$E[Y(0, M(1)) - Y(0, M(0)) \mid C] = \tag{88}$$

$$= [F_Y(0, 1) - F_Y(0, 0)]\ [F_M(1) - F_M(0)]. \tag{89}$$

The direct effect and pure indirect effect expressions agree with those in Pearl (2010, 2011a) as given in connection with the artificial example.

Odds ratio expressions can be obtained as well. For the direct effect, it follows that the numerator and denominator probabilities of the odds ratio are

$$E[Y(1, M(0))] = F_Y(1, 0)\ (1 - F_M(0)) + F_Y(1, 1)\ F_M(0), \tag{90}$$

$$E[Y(0, M(0))] = F_Y(0, 0)\ (1 - F_M(0)) + F_Y(0, 1)\ F_M(0). \tag{91}$$

For the total indirect effect, it follows that the numerator and denominator probabilities of the odds ratio are

$$E[Y(1, M(1)] = F_Y(1, 0)\ (1 - F_M(1)) + F_Y(1, 1)\ F_M(1), \tag{92}$$

$$E[Y(1, M(0))] = F_Y(1, 0)\ (1 - F_M(0)) + F_Y(1, 1)\ F_M(0). \tag{93}$$

13.4 Causal effects for a nominal mediator

As in Section 13.2.1, consider

$$E[Y(x, M(x')) \mid C = c] = \sum_{s=1}^{S} E[Y \mid C = c, X = x, M = m] \times P(M \mid C = s, X = x'), \tag{94}$$

where Y is a continuous outcome, and P is the multinomial logistic regression formula for the probabilities of the nominal mediator M,

$$P(M \mid C = s, X = x') = e^{\gamma_{0s} + \gamma_{1s}\ x' + \gamma_{2s}\ c} / \sum_{d=1}^{D} e^{\gamma_{0d} + \gamma_{1d}\ x' + \gamma_{2d}\ c}, \tag{95}$$

where the γ coefficients are all zero for the last category. This formula can be applied as before to the direct, total indirect, pure indirect, and total effects defined as

$$E[Y(1, M(0)) - Y(0, M(0)) \mid C], \tag{96}$$

$$E[Y(1, M(1)) - Y(1, M(0)) \mid C], \tag{97}$$

$$E[Y(0, M(1)) - Y(0, M(0)) \mid C], \tag{98}$$

$$E[Y(1, M(1)) - Y(0, M(0)) \mid C]. \tag{99}$$

13.5 Causal effects for a count outcome

As in Section 13.2.1, consider

$$E[Y(x, M(x'))] = \int_{-\infty}^{\infty} E[Y \mid C = c, X = x, M = m] \times f(M \mid C = c, X = x')\,\partial M. \tag{100}$$

For a count outcome Y the log rate is modeled, so that the rate (mean) is

$$E[Y \mid C = c, X = x, M = m] = e^{\beta_0 + \beta_1\, m + \beta_2\, x + \beta_3\, x\, m + \beta_4\, c}. \tag{101}$$

Letting $\mu_M = \gamma_0 + \gamma_1\, x' + \gamma_2\, c$ denote the mean of the normal density f, it follows that

$$E[Y(x, M(x'))] = e^{\beta_0 + \beta_2\, x + \beta_4\, c} \int e^{\beta_1\, M + \beta_3\, x\, M} f(M; \mu_M, \sigma^2)\,\partial M, \tag{102}$$

$$= e^{\beta_0 + \beta_2\, x + \beta_4\, c}\, e^{(b^2 - 1)\, \mu_M^2 / 2\sigma^2} \int f(M; b\, \mu_M, \sigma^2)\,\partial M, \tag{103}$$

where the integral over the normal density f is one, and where

$$b = (2\sigma^2(\beta_1 + \beta_3\, x)/\mu_M + 2)/2. \tag{104}$$

The expressions for the direct and indirect effects then follow as before.

Both the Poisson and negative binomial model for counts have the rate (mean) as given above and therefore have the same effect formulas. Zero-inflated models need to take into account that the mean is the rate multiplied by $(1 - \pi)$, where π is the probability of being in the zero class.

The count variable can also be a mediator in which case the integral is replaced by a sum over the possible counts and using the probabilities determined by the Poisson distribution. Using m to predict y, m may be treated as continuous.

13.6 Imai et al. sensitivity analysis

The following derivation explicates that of Imai (2010b, Appendix D), with the trivial extension of including a covariate C. Consider again the model of Section 3, but for simplicity without the treatment-mediator interaction,

$$y_i = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_4\, c_i + \epsilon_{1i}, \tag{105}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \epsilon_{2i}, \tag{106}$$

where the residuals ϵ_1 and ϵ_2 are assumed normally distributed with zero means and variances σ_1^2, σ_2^2. To reflect omitted mediator-outcome confounders, let the residuals have correlation ρ, that is, covariance $\rho\, \sigma_1\, \sigma_2$. The reduced-form is

$$y_i = \beta_0 + \beta_1\, (\gamma_0 + \gamma_1\, x_i + \gamma_2\, c_i + \epsilon_{2i}) + \beta_2\, x_i + \beta_4\, c_i + \epsilon_{1i}, \tag{107}$$

$$= \beta_0 + \beta_1\, \gamma_0 + \beta_1\, \gamma_1\, x_i + \beta_2\, x_i + \beta_1\, \gamma_2\, c_i + \beta_4\, c_i + \beta_1\, \epsilon_{2i} + \epsilon_{1i}. \tag{108}$$

Consider the regression

$$y_i = \kappa_0 + \kappa_1\, x_i + \kappa_2\, c_i + \epsilon_i, \tag{109}$$

where from the reduced-form

$$\kappa_0 = \beta_0 + \beta_1\ \gamma_0, \tag{110}$$

$$\kappa_1 = \beta_2 + \beta_1\ \gamma_1, \tag{111}$$

$$\kappa_2 = \beta_4 + \beta_1\ \gamma_2, \tag{112}$$

$$\epsilon_i = \beta_1\ \epsilon_{2i} + \epsilon_{1i}. \tag{113}$$

Consider what can be identified using the two regressions of (106) and (109), first in the case of $\rho = 0$ and then for $\rho \neq 0$. Let $\tilde{\rho}$ denote the correlation between the residuals in these two regressions, ϵ_2 and ϵ, so that the residual covariance is $\tilde{\rho}\ \sigma_2\ \sigma$, where σ^2 is the variance of ϵ. Note that the residual covariance can be expressed as

$$Cov(\epsilon_2, \epsilon) = \tilde{\rho}\ \sigma_2\ \sigma = \beta_1\ \sigma_2^2 + \rho\ \sigma_1\ \sigma_2, \tag{114}$$

and the variance of the residual σ^2 as

$$\sigma^2 = \beta_1^2\ \sigma_2 + \sigma_1 + 2\ \beta_1\ \rho\ \sigma_1\ \sigma_2. \tag{115}$$

13.6.1 The case of $\rho = 0$

From (114) it follows that β_1 is identified as

$$\beta_1 = \tilde{\rho}\ \sigma_2\ \sigma/\sigma_2^2 = \tilde{\rho}\ \sigma/\sigma_2. \tag{116}$$

Together with γ_1 being identified from (106), this identifies the usual indirect effect $\beta_1\ \gamma_1$.

13.6.2 The case of $\rho \neq 0$

From (114) it is possible to express σ_1 as

$$\sigma_1 = (\tilde{\rho}\,\sigma - \beta_1\,\sigma_2)/\rho. \tag{117}$$

Inserting this in (115) gives a quadratic equation with solution

$$\beta_1 = \sigma/\sigma_2\,(\tilde{\rho} - \rho\,\sqrt{(1-\tilde{\rho}^2)/(1-\rho^2)}). \tag{118}$$

Together with γ_1 being identified from (106), this identifies the usual indirect effect $\beta_1\,\gamma_1$, assuming a fixed, known value of ρ. In addition, it is seen that the remaining parameters β_2, β_0, σ_1 are identified. For example, the direct effect β_2 is obtained via (111) as

$$\beta_2 = \kappa_1 - \beta_1\,\gamma_1, \tag{119}$$

where β_1 is given in (118), κ_1 is obtained from the outcome regression (109), and γ_1 is obtained from the mediator regression in (106).