Applications of Causally Defined Direct and Indirect Effects in Mediation Analysis using SEM in Mplus

Bengt Muthén *

October 28, 2011

*I thank Tihomir Asparouhov, Michael Sobel, Hendricks Brown, David MacKinnon, and Judea Pearl for helpful advice

Abstract

This paper summarizes some of the literature on causal effects in mediation analysis. It presents causally-defined direct and indirect effects for continuous, binary, ordinal, nominal, and count variables. The expansion to non-continuous mediators and outcomes offers a broader array of causal mediation analyses than previously considered in structural equation modeling practice. A new result is the ability to handle mediation by a nominal variable. Examples with a binary outcome and a binary, ordinal or nominal mediator are given using Mplus to compute the effects. The causal effects require strong assumptions even in randomized designs, especially sequential ignorability, which is presumably often violated to some extent due to mediator-outcome confounding. To study the effects of violating this assumption, it is shown how a sensitivity analysis can be carried out. This can be used both in planning a new study and in evaluating the results of an existing study.

1 Introduction

This paper considers mediation analysis (see, e.g., Baron & Kenny, 1986; MacKinnon, 2008) as carried out in structural equation modeling (SEM; see, e.g., Goldberger & Duncan, 1973; Jöreskog and Sörbom, 1979; Bollen, 1989). Mediation analysis in SEM uses the terms direct and indirect effects. The implication that the direct and indirect effects produced by SEM are causal effects has been criticized in e.g. Holland (1988) and Sobel (2008), while generally interpreted with causal implications by others, e.g. Pearl (2010, 2011a). The challenge in using mediation for causal inference comes in interpreting the relationship between changes in the mediator and its impact on the outcome, which cannot rely on inferential support from an underlying randomized trial. SEM practitioners are left with a somewhat confusing picture of what is accomplished with mediational analysis. To exacerbate the problem, the causal inference literature is often difficult to understand for researchers using SEM. Also, key researchers disagree about the best language to use as seen in the recent debate in the journal NeuroImage (Lindquist & Sobel, 2010, 2011; Glymour, 2011; Pearl, 2011b).

As a modest attempt to help clarify part of the picture, this paper gives a summary of some of the key issues, showing relationships between SEM effect concepts and causal effect concepts in mediation analysis, and focusing on applications of mediation analyses with causally-defined direct and indirect effects produced by Mplus. The paper shows that causally-defined direct and indirect effects are not necessarily the same as effects typically presented by SEM practitioners, and in several cases provide new effects that have not been used in SEM practice. The causally-defined effects can be obtained via extended types of SEM analyses. To claim that

effects are causal, however, it is not sufficient to simply use the causally-defined effects. A set of assumptions needs to be fulfilled for the effects to be causal and the plausibility of these assumptions needs to be considered.

The paper presents causally-defined direct and indirect effects for continuous, binary, ordinal, nominal, and count variables. The expansion to non-continuous mediators and outcomes offers a broader array of causal mediation analyses than previously considered in SEM practice. A new result is the ability to handle mediation by a nominal variable. Examples with a binary outcome and a binary, ordinal and nominal mediators are given. The assumptions behind causal effects in mediation modeling are discussed and sensitivity analyses of the possible distorting effects of violations of the assumptions are exemplified. Extensions to moderated mediation and latent variable mediation are discussed. For the paper to be self-contained, an appendix gives derivations of the effects, most of which can be found in the literature. Estimation is performed by maximum-likelihood, weighted least-squares, and Bayesian analysis. The analyses can be carried out by the free demo version of Mplus at www.statmodel.com. An appendix gives the Mplus input scripts for all analyses.

2 A mediation model with treatment-mediator interaction

Consider Figure 1 which corresponds to a randomized trial with a binary treatment dummy variable x (0=control, 1=treatment), a covariate c, a continuous mediator m, and a continuous outcome y, a situation examined in detail by MacKinnon (2008). A special feature is that the treatment and

mediator interact in their influence on the outcome y. This possibility is important to the so-called MacArthur approach to mediation (Kraemer et al., 2008). As pointed out in e.g. VanderWeele and Vansteelandt (2009), the possibility of this interaction was emphasized in Judd and Kenny (1981) but not in the influential Baron and Kenny (1986) article on mediation, and is therefore often not explored. The interaction possibility is, however, stated in James and Brett (1984) and more recently in Preacher et al. (2007). The covariate c is useful in randomized studies to increase the power to detect a treatment effect. Adding an interaction between c and x, a treatment-baseline interaction effect on y can be explored; this type of moderated mediation is discussed in Section 11.1. The model of Figure 1 is used to first discuss the SEM concepts of direct and indirect effects and then the corresponding causal concepts.

[Figure 1 about here.]

3 SEM concepts of direct and indirect effects

Assuming linear relationships, Figure 1 translates into

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

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

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. SEM considers the reduced form of this model,

obtained by inserting (2) in (1),

$$y_i = \beta_0 + \beta_1\ (\gamma_0 + \gamma_1\ x_i + \gamma_2\ c_i + \epsilon_{2i}) + \beta_2\ x_i + \beta_3\ x_i\ (\gamma_0 + \gamma_1\ x_i + \gamma_2\ c_i + \epsilon_{2i}) +$$
$$\beta_4\ c_i + \epsilon_{1i}, \tag{3}$$
$$= \beta_0 + \beta_1\ \gamma_0 + \beta_1\ \gamma_1\ x_i + \beta_3\ \gamma_0\ x_i + \beta_3\ \gamma_1\ x_i^2 + \beta_2\ x_i + \beta_1\ \gamma_2\ c_i + \beta_3\ \gamma_2\ x_i\ c_i +$$
$$\beta_4\ c_i + \beta_1\ \epsilon_{2i} + \beta_3\ x_i\ \epsilon_{2i} + \epsilon_{1i}. \tag{4}$$

First, assume no treatment-mediator interaction, that is, $\beta_3 = 0$. In this case, the reduced-form expression of (4) states that the direct effect of x on y is β_2 and the indirect effect via m is $\beta_1\ \gamma_1$. In both cases, the presence of the covariate c implies that these statements are conditional on c. These are the standard formulas used in mediation modeling.

Second, let $\beta_3 \neq 0$. In this case, the definitions of the direct and indirect effect are perhaps less clear. One may consider the direct effect to be $\beta_3\ \gamma_0 + \beta_2 + \beta_3\ \gamma_2\ c$, where the first term is included because γ_0 is not part of the influence of x on m and the third term is included for the same reason. In this way, there can be a direct effect even if $\beta_2 = 0$. One may consider the indirect effect to be a sum composed of a main part $\beta_1\ \gamma_1$ and an interaction part $\beta_3\ \gamma_1$. In this way, there can be a indirect effect even if $\beta_1 = 0$.

It should be noted that the Mplus MODEL INDIRECT computations are not valid for a model such as Figure 1 due to the treatment-mediator interaction, but reports the direct effect as β_2 and the indirect effect as $\beta_1\ \gamma_1$. As shown in Section 5 the correct effects can, however, be computed via MODEL CONSTRAINT.

4 Causal inference concepts of direct and indirect effects

Causally-defined direct and indirect effects were introduced in Robins and Greenland (1992) and further elaborated in Pearl (2001) and Robins (2003). Drawing on this work, some of the more accessible treatments of direct and indirect causal effects are given in VanderWeele and Vansteelandt (2009), see also Valeri and VanderWeele (2011), and Imai et al. (2010a,b). Valeri and VanderWeele (2011) describe macros for SAS and SPSS, and Imai et al. (2010c) describe the R program mediation.

The assumptions behind the causally-defined effects are important and may often not be fulfilled in practice. VanderWeele and Vansteelandt (2009) and Imai et al. (2010b) give formal, technical statements of the assumptions using potential outcomes notation and provide proofs of identifiability. Valeri and VanderWeele (2011) use simple language to summarize these assumptions and their summary is quoted here:

> "(i) no unmeasured confounding of the treatment-outcome relationship.
>
> (ii) no unmeasured confounding of the mediator-outcome relationship.
>
> (iii) no unmeasured treatment-mediator confounding
>
> (iv) no mediator-outcome confounder affected by treatment"

Assumptions (i) and (iii) are fulfilled when X is a randomized treatment. Assumptions (i) and (ii) are sufficient for the controlled direct effect defined below. The direct and indirect effects defined below require all four assumptions (although see Pearl, 2011c, footnote 5 for exceptions). This

means that even with randomized treatment, direct and indirect effects require that assumptions (ii) and (iv) be fulfilled. Taken together, this is often referred to as the sequential ignorability assumption. Because the mediator values are not randomized within treatment groups, assumptions (ii) and (iv) may often not be fulfilled. As pointed out in VanderWeele and Vansteelandt (2009), assumptions (i)-(iii) "could potentially be satisfied, at least approximately, by collecting data on more and more confounding variables". Assumption (iv), however, "will be violated irrespective of whether data is available for all such variables." Even in randomized studies this means that the causally-defined effects are biased unless assumptions (ii) and (iv) hold, and if assumption (iv) does not hold causal effects cannot be identified. Imai (2010a, b) and VanderWeele (2010) propose sensitivity analyses to study the impact of violations of assumptions. A sensitivity analysis is illustrated in a later section for both simulated and real data.

A key concept in the causal effect literature is a counterfactual or potential outcome. Let $Y_i(x)$ denote the potential outcome that would have been observed for that subject had the treatment variable X been set at the value x, where x is 0 or 1 in the example considered here (in the following, upper-case letters denote variables and lower-case letters values of these variables). The $Y_i(x)$ outcome may not be the outcome that is observed for the subject and is therefore possibly counterfactual. The effect of treatment for a subject can be seen as $Y_i(1) - Y_i(0)$, but is clearly not identified given that a subject only experiences one of the two treatments. The average effect $E[Y(1) - Y(0)]$ is, however, identifiable if X is assigned randomly as is the case in a randomized controlled trial. Similarly, let Y(x, m) denote the potential outcome that would have been observed if the treatment for

the subject was x and the value of the mediator M was m.

Following are definitions of the total, direct, and indirect effects. The formulas are general, that is, not based on a particular model such as the linear model for continuous variables of (1) and (2). Because of this, they can be generalized to other types of variables.

The controlled direct effect is defined as

$$CDE(m) = E[Y(1, m) - Y(0, m) \mid C = c]. \tag{5}$$

where $M = m$ for a fixed value m. The first index of the first term is 1 corresponding to the treatment group and the first index of the second term is 0 corresponding to the control group.

The direct effect (often called the pure or natural direct effect) does not hold the mediator constant, but instead allows the mediator to vary over subjects in the way it would vary if the subjects were given the control condition. The direct effect is expressed as

$$DE = E[Y(1, M(0)) - Y(0, M(0)) \mid C = c] = \tag{6}$$

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

$$\times f(M \mid C = c, X = 0)\, \partial M, \tag{7}$$

where f is the density of M. A simple way to view this is to note that in (6) Y's first argument, that is x, changes values, but the second does not, implying that Y is influenced by X only directly. The expression should be read as the conditional expectation, given the covariate, of the difference between the outcome in the treatment and control group when the mediator is held constant at the values it would obtain for the control group. The

right-hand side of (7) is part of what is referred to as the Mediation Formula in Pearl (2009, 2011c).

The total indirect effect is defined as (Robins, 2003)

$$TIE = E[Y(1, M(1)) - Y(1, M(0)) \mid C = c] = \tag{8}$$

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

$$- \int_{-\infty}^{\infty} E[Y \mid C = c, X = 1, M = m] \times f(M \mid C = c, X = 0)\, \partial M. \tag{9}$$

A simple way to view this is to note that the first argument of Y in (8) does not change, but the second does, implying that Y is influenced by X due to its influence on M. The expression should be read as the conditional expectation, given the covariate, of the difference between the outcome in the treatment group when the mediator changes from values it would obtain in the treatment group to the values it would obtain in the control group. The name total indirect effect is used in Robins (2003), while Pearl (2001) and VanderWeele and Vansteelandt (2009) call it the natural indirect effect.

The total effect is (Robins, 2003)

$$TE = E[Y(1) - Y(0) \mid C = c] \tag{10}$$

$$= E[Y(1, M(1)) - Y(0, M(0)) \mid C = c]. \tag{11}$$

A simple way to view this is to note that both indices are 1 in the first term and 0 in the second term. In other words, the treatment effect on Y comes both directly and indirectly due to M. The total effect is the sum of the direct effect and the total indirect effect (Robins, 2003),

$$TE = DE + TIE. \tag{12}$$

The pure indirect effect (Robins, 2003) is defined as

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C = c] \tag{13}$$

Here, the effect of X on Y is only indirect via M. This is called the natural indirect effect in Pearl (2001) and VanderWeele and Vansteelandt (2009). The difference between TIE and PIE is shown below for the model of (1) and (2).

4.1 Applying the causal effects to the mediation model

The appendix Section 13.1 (see also the Appendix of VanderWeele & Vansteelandt, 2009) shows how the direct effect in (7) and the total indirect effect in (9), conditional on the value c, are explicated in terms of the parameters of the model of (1) and (2) by integrating over the distribution of M. The direct effect is

$$DE = \beta_2 + \beta_3\, \gamma_0 + \beta_3\, \gamma_2\, c. \tag{14}$$

This agrees with the direct effect conjectured for the reduced form of the SEM approach above, but the results are obtained via a clear definition. The total indirect effect is

$$TIE = \beta_1\, \gamma_1 + \beta_3\, \gamma_1. \tag{15}$$

This agrees with the indirect effect conjectured for the reduced form of the SEM approach above. The pure indirect effect excludes the interaction

part,

$$PIE = \beta_1 \, \gamma_1. \tag{16}$$

In summary, the SEM estimates for the mediation model of Figure 1 can be used to express the causal direct and indirect effects. The causal inference using potential outcomes clarifies how to conceptualize these effects. As will be seen in the next sections, there is not necessarily a similar agreement between effects used in SEM practice and the causal effect results when either the outcome Y or the mediator M is not continuous. In fact, the causally-defined effects to be presented have not been available in SEM software until now.

5 Monte Carlo simulation of continuous mediator, continuous outcome with treatment-mediator interaction

Monte Carlo simulations are useful for planning purposes to determine the sample size needed to recover parameter values well and to have sufficient power to detect various effects. Mplus has quite general Monte Carlo capabilities as is demonstrated in this paper; see also Muthén and Muthén (1998-2010, chapter 12). For an application of a Monte Carlo study, see Muthén and Muthén (2002).

Consider again the model of Figure 1 as explicated in (1) and (2), but simplified to not include a covariate c. Note that the interaction between

the treatment and the mediator in

$$y_i = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, m_i + \epsilon_{1i} \tag{17}$$

can be expressed via a random slope β_{1i},

$$y_i = \beta_0 + \beta_{1i}\, m_i + \beta_2\, x_i + \epsilon_{1i} \tag{18}$$

$$\beta_{1i} = \beta_1 + \beta_3\, x_i + \epsilon_i, \tag{19}$$

where the residual ϵ has not only zero mean but also zero variance. A non-zero variance can also be handled and represents heteroscedasticity in line with random coefficient regression shown in ex 3.9 in the Mplus User's Guide (Muthén & Muthén, 1998-2010). A non-zero variance is not pursued here, however. Inserting (19) in (18) gives the same as (17).

This random slope approach to create an interaction is used in the Mplus input for a Monte Carlo simulation shown in Section 14.1. 500 samples of size 400 are generated in a first step. A second step analyzes the 400 samples in a model where an interaction term $x \times m$ is created and included in the analysis model. MODEL CONSTRAINT is used to specify the causal direct and indirect effects defined in Section 4. The effects are computed by specifying NEW parameters derived from labeled model parameters. Standard errors are automatically produced using the delta method. The results are shown in Table 1 for the second step. The results for the first step are exactly the same, except for a slight difference in the standard errors using the MLR estimator instead of ML. The Mplus input gives comments to describe the quantities derived from the model parameters. The new parameters tie, pie, and de correspond to the indirect and direct effects

of (15), (16), and (14). It is seen that all parameters are well recovered and standard errors are well estimated. The last two columns show good 95% coverage and good power to reject that the parameter is zero. For a description of how to interpret the Mplus Monte Carlo output, see pp. 362-365 of the User's Guide, Muthén and Muthén (1998-2010). The setup can be used for planning purposes to study coverage and power at different sample sizes and effect sizes.

[Table 1 about here.]

Because the effects involve products of parameters, the distribution of the effect estimates may not be well approximated by a normal distribution. This is particularly the case with small sample sizes and in situations with a binary mediator and/or a binary outcome. To account for this non-normality of the effect distribution, ML estimation can use bootstrapped standard errors and bootstrap-based confidence intervals. The modification of the Mplus input is to request BOOTSTRAP=1000, say, in the ANALYSIS command, and add CINTERVAL(BOOTSTRAP) in the OUTPUT COMMAND. As an alternative, Bayesian analysis can be used, where the parameter distributions do not have to be normal. The Bayesian analysis produces posterior distributions and confidence (credibility) intervals of the effects. This is accomplished simply by specifying ESTIMATOR=BAYES in the ANALYSIS command.

6 Mediation modeling with a binary outcome and a continuous mediator

Consider next the case of Figure 1 where the outcome y is binary. This replaces (1) with a corresponding probit or logistic regression equation. In this case, the Mplus direct and indirect effects of SEM are defined for a continuous latent response variable underlying the binary outcome and therefore use the same formulas as before. This is also the approach proposed in MacKinnon et al. (2007), considering a model without the treatment-mediator interaction. The corresponding effects defined for the observed binary outcome may be less well known, but have been presented in Iamai et al. (2010a), and are restated here.

Considering a model with the treatment-mediator interaction, Vander-Weele and Vansteelandt (2010) define causal effects for the observed binary outcome. They consider logistic regression for (1) and assume that y corresponds to a rare outcome. In this case, the indirect effect can be expressed as an odds ratio that is approximately equal to

$$e^{\beta_1 \, \gamma_1 + \beta_3 \, \gamma_1}, \tag{20}$$

that is, using the same formula as in (15), but with parameters on the logit scale.

This paper considers probit regression for y in (1) without an assumption of the binary outcome being rare. Appendix Section 13.2 derives causally-defined direct and indirect effects (see also Imai et al., 2010a, Appendix F). Using the definition in (9), the causal total indirect effect is expressed as

the probability difference

$$\Phi[probit(1,1)] - \Phi[probit(1,0)], \tag{21}$$

using the standard normal distribution function Φ, and where for x, x' = 0, 1 corresponding to the control and treatment group,

$$probit(x,x') = [\beta_0 + \beta_2\ x + \beta_4\ c + (\beta_1 + \beta_3\ x)(\gamma_0 + \gamma_1\ x' + \gamma_2\ c)]/\sqrt{v(x)}, \tag{22}$$

where the variance $v(x)$ for x = 0, 1 is

$$v(x) = (\beta_1 + \beta_3\ x)^2\ \sigma_2^2 + 1. \tag{23}$$

where σ_2^2 is the residual variance for the continuous mediator m. Although not expressed in simple functions of model parameters, the quantity of (21) can be computed and corresponds to the change in the y=1 probability due to the indirect effect of the treatment (conditionally on c when that covariate is present).

The total indirect effect odds ratio for the binary y related to the binary x can be expressed as

$$\frac{\Phi[probit(1,1)]/(1 - \Phi[probit(1,1)])}{\Phi[probit(1,0)]/(1 - \Phi[probit(1,0)])}. \tag{24}$$

For any given data set, this odds ratio can be compared to that in (20) computed via logistic regression and assuming that the outcome y is rare.

Using the definition in (13), the pure indirect effect is expressed as the probability difference

$$\Phi[probit(0,1)] - \Phi[probit(0,0)]. \tag{25}$$

Using the definition in (6), the direct effect is expressed as the probability difference

$$\Phi[probit(1,0)] - \Phi[probit(0,0)]. \tag{26}$$

6.1 A closer look at the effects in a simple special case

To put the causal indirect and direct effects in perspective, consider the special case of no treatment-mediator interaction ($\beta_3 = 0$) and no covariate c. In this case the causal indirect effect $\Phi[probit(1,1)] - \Phi[probit(1,0)]$ has probit arguments

$$probit(1,1) = [\beta_0 + \beta_2 + \beta_1\ \gamma_0 + \beta_1\ \gamma_1]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}, \tag{27}$$

$$probit(1,0) = [\beta_0 + \beta_2 + \beta_1\ \gamma_0]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}. \tag{28}$$

This may be compared to a naive approach of expressing the indirect effect for the probit as the product $\beta_1\ \gamma_1$ and considering the probability difference $\Phi(a) - \Phi(b)$ with and without this indirect effect, where

$$a = [\beta_0 + \beta_1\ \gamma_0 + \beta_1\ \gamma_1]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}, \tag{29}$$

$$b = [\beta_0 + \beta_1\ \gamma_0]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}. \tag{30}$$

The difference between the causal and naive indirect effect approaches is that the direct effect slope β_2 plays a role in the former, but not in the

latter.

Noting that $\Phi(b) = \Phi[probit(0,0)]$, the causal direct effect $\Phi[probit(1,0)] - \Phi[probit(0,0)]$ has probit arguments

$$probit(1,0) = [\beta_0 + \beta_2 + \beta_1\ \gamma_0]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}, \tag{31}$$

$$probit(0,0) = [\beta_0 + \beta_1\ \gamma_0]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}. \tag{32}$$

A naive approach may instead focus on the direct effect β_2 and consider $\Phi(a') - \Phi(b')$, where

$$a' = [\beta_0 + \beta_2]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}, \tag{33}$$

$$b' = [\beta_0]/\sqrt{\beta_1^2\ \sigma_2^2 + 1}. \tag{34}$$

This leaves out the $\beta_1\ \gamma_0$ term of the causal approach.

The difference between the causal effects and the effects obtained by what is called the naive approach has been studied in Imai et al. (2010a) and Pearl (2011c). Imai et al. (2010a, Appendix E, p. 23) conducted a Monte Carlo simulation study to show the biases, while Pearl (2011c) presented graphs showing the differences.

In summary, the causal approach gives clear definitions of indirect and direct effects. Alternative, naive, approaches do not have the same causal interpretation.

6.2 Mplus computations

The direct and indirect effects can be estimated in Mplus using maximum-likelihood. Standard errors of the direct and indirect causal effects are

obtained by the delta method using the Mplus MODEL CONSTRAINT command. Bootstrapped standard errors and confidence intervals are also available, taking into account possible non-normality of the effect distributions. Furthermore, Bayesian analysis is available in order to describe the posterior distributions of the effects. Examples of Mplus analysis are shown below.

It should be noted that changing from probit to logistic regression, not assuming a rare outcome, does not lead to as simple expressions as in (21) and (26). This is because in the logistic case the integration over the mediator does not lead to an explicit form, but calls for numerical integration.

Maximum-likelihood estimation using logistic regression is also available in Mplus, where effects can be derived using approximate odds ratios under the assumption of a rare outcome.

6.3 Distributional assumption for the mediator: Latent response variable mediation

The direct and indirect effect formulas given above in the probit case assume normality for the residual ϵ_2 in the mediator regression. This may be a strong assumption and when it is violated the effects will be biased.

One type of non-normality may arise when the mediator can be viewed as an ordered categorical (ordinal) variable. In this case, the approach of Muthén (1984) may be taken where instead of the observed mediator, an underlying continuous latent response variable is viewed as the relevant mediator. In line with an ordered probit model, the observed mediator categories are determined by the latent mediator variable falling below or

exceeding thresholds as illustrated in Figure 2. Although the observed ordinal mediator m has a non-symmetric distribution with the highest frequency for $m = 0$, the latent mediator m^* can still be normal conditional on the covariates.

Figure 2 corresponds to the measurement relationship

$$m_i = \begin{cases} 0 & \text{if } m_i^* \leq \tau_1 \\ 1 & \text{if } \tau_1 < m_i^* < \tau_2 \\ 2 & \text{if } \tau_2 \leq m_i^* \end{cases}$$

where for a latent response variable y^* behind the binary outcome y

$$y_i^* = \beta_0 + \beta_1\, m_i^* + \beta_2\, x_i + \beta_3\, x_i\, m_i^* + \beta_4\, c_i + \epsilon_{1i}, \tag{35}$$

$$m_i^* = \gamma_1\, x_i + \gamma_2\, c_i + \epsilon_{2i}. \tag{36}$$

The key point is that the continuous latent response variable m^* is used not only as a dependent variable in (36) but also replaces the observed m as a predictor in (35). This implies that the probit-based direct and indirect causal effects of the previous section with a continuous mediator are still valid. This type of model can be estimated in Mplus using weighted least-squares and Bayesian analysis. An application is shown in Section 6.6.

[Figure 2 about here.]

6.4 Monte Carlo simulation with a binary outcome and a continuous mediator

To study the behavior of maximum-likelihood and Bayesian estimation with a binary outcome, a Monte Carlo study is carried out for a model like the

one in Figure 1, using $n = 200$. The same two steps are used as in the Monte Carlo study of Section 5. Data are generated using probit for the binary outcome. Appendix Section 14.2 shows the Mplus input for Step 1 and the Step 2 input for maximum-likelihood (the Bayes analysis simply changes to ESTIMATOR=BAYES, deleting LINK=PROBIT). Causal effects in terms of probabilities and odds ratios are expressed in MODEL CONSTRAINT using the formulas presented in the beginning of this section.

Table 2 and Table 3 show the results for the two estimators (the Bayes analysis uses FBITER=10000). It is seen that for both estimators all parameters, including the causal effects, are well estimated with good coverage.

[Table 2 about here.]

[Table 3 about here.]

Appendix Section 14.2 also shows the Mplus input for a Bayesian analysis of the data generated in the first replication of the simulation. This analysis produces the posterior distributions of all the parameters. Figure 3 shows the posterior for the odds ratio corresponding to the direct effect (orde) and Figure 4 shows the posterior for the odds ratio corresponding to the total indirect effect (ortie). It is seen that neither posterior is close to normally distributed. Vertical lines at the tails show the upper and lower limits of the Bayesian 95% credibility interval. Bayes has the advantage that this interval is not symmetrically placed around the mean as is the case when using the maximum-likelihood approach. In other words, as seen in the Monte Carlo simulation, maximum-likelihood and Bayes will give similar point estimates for these odds ratios but different confidence/credibility intervals.

[Figure 3 about here.]

[Figure 4 about here.]

6.5 Example 1: Aggressive behavior and juvenile court record

Data for this example are from a randomized field experiment in Baltimore public schools where a classroom-based intervention was aimed at reducing aggressive-disruptive behavior among elementary school students. Figure 5 shows the Fall baseline aggression score as agg1, observed before the intervention started. The variable agg1 is used as a covariate in the analysis to strengthen the power to detect treatment effects. The mediator variable agg5 is the aggression score in Grade 5 after the intervention ended. The outcome juvcrt is a binary variable indicating whether or not the student obtained a juvenile court record by age 18 or an adult criminal record. The analysis to be presented involves $n = 250$ boys in treatment and control classrooms with complete data. A further description of the data and related analyses is given in Muthén et al. (2002).

The juvcrt outcome is not rare, but is observed for 50% of the sample. The mediator agg5 is not normally distributed, but is quite skewed with a heavy concentration at low values. The normality assumption of Section 6, however, pertains to the mediator residual ϵ_2 and because the covariate agg1 has a distribution similar to the mediator agg5, the agg5 distribution is to some extent driven by the agg1 distribution so that the normality assumption for the residual may be a reasonable approximation. Causal effect estimates are computed using the probit approach. They

are compared with those of the logistic regression approach, mistakenly assuming that the outcome juvcrt is rare.

[Figure 5 about here.]

Appendix Section 14.3 shows the Mplus input for maximum-likelihood analysis of this model using probit and logit. The probit output is shown in Table 4. It is seen that the treatment-mediator interaction (xm) is not significant. The section New/additional parameters show the effect estimates. The causal direct effect (direct) of (26) is not significant. The causal indirect effect (indirect) of (21) is estimated as -0.064 and is significant. This is the drop in the probability of a juvenile court record due to the indirect effect of treatment. The odds ratio for the indirect effect of (24) is estimated as 0.773 which is significantly different from one ($z = (0.773 - 1)/0.092 = -2.467$). These findings can be compared with the indirect and direct effects labeled ind and dir at the top of the new parameters section, which use the regular definitions in (15) and (14), that is, considering the continuous latent response variable for the outcome as the relevant dependent variable.

[Table 4 about here.]

Using logistic regression instead, the maximum-likelihood estimate of the odds ratio under the rare outcome assumption of (20) is 0.734 and is also significantly different from one; see Table 5. This means that in the current example, the probit and logistic approaches give quite similar results despite the outcome not being rare.

[Table 5 about here.]

The Mplus input in Appendix Section 14.3 can be easily adapted to other applications. The statements in the MODEL CONSTRAINT section need not be changed if the same parameter labels are used in the MODEL command. If there is no treatment-mediator interaction in the model, the statement $beta3 = 0$ can be added in MODEL CONSTRAINT below the NEW statement. Likewise, with no covariate c for the probit analysis, $beta4 = 0$ is added. Note that in the probit analysis $beta4$ is multiplied by zero, that is, the effect is evaluated at the average of the covariate c.

6.6 Example 2: Intentions to stop smoking

MacKinnon et al. (2007) analyzed the model shown in Figure 6. There is no evidence of treatment-mediator interaction. The data are from a drug intervention program for students in Grade 6 and 7 in Kansas City. Schools were randomly assigned to treatment or control. The multilevel aspect of the data is ignored here as in MacKinnon et al. (2007). The mediator is the intention to use cigarettes in the following 2-month period, measured about six months after baseline. The outcome is cigarette use in the previous month, measured at follow-up. Cigarette use is observed for 18% of the sample. The data for $n = 864$ students are shown in Table 6.

[Figure 6 about here.]

[Table 6 about here.]

Table 6 shows that the intention mediator is not close to normally distributed in either the treatment or control group. This means that the normality assumption for the ϵ_2 residual is violated. Because of this, the data are analyzed not only using the observed mediator approach but also

the latent response variable mediator approach discussed in Section 6.3. In the former case, normality is (mistakenly) assumed for the continuous mediator given the treatment dummy variable and maximum-likelihood estimation is used. In the latter case normality is assumed for the latent response variable given the treatment dummy variable, treating the observed mediator as ordered categorical, and using weighted least-squares estimation. Appendix Section 14.4 shows the Mplus inputs. Table 7 and Table 8 show the results using probit for the observed and latent mediator approach, respectively.

For the observed mediator approach using probit, the causal direct effect odds ratio is 0.731, while the causal indirect odds ratio is 0.853. Using logistic regression (not shown), the causal indirect odds ratio is 0.843, that is, only slightly lower than the value for probit.

[Table 7 about here.]

[Table 8 about here.]

For the latent mediator approach using probit, the causal direct effect odds ratio is 0.829, while the causal indirect odds ratio is 0.796. This means that the latent mediator approach results in a stronger indirect effect and a weaker direct effect relative to the observed mediator approach. A latent mediator approach using logistic regression is not yet available in Mplus.

7 Mediation modeling with a binary mediator

When the mediator is binary, a latent mediator approach or an observed mediator approach may be used. Taking a latent mediator approach leads to the causal effect techniques described in the previous section. Taking an observed mediator approach, the causal direct and indirect approach described in Section 4 is still valid but needs to be explicated. The observed binary mediator case is interesting because SEM-based direct and indirect effects have not been developed in SEM software. Direct and indirect effects for this case have, however, been discussed in Winship and Mare (1983), although not from a causal inference perspective. Causal direct and indirect effects for the case of a binary observed mediator and a continuous outcome have been explicated in Valeri and VanderWeele (2011). This section instead focuses on the case of a binary observed mediator and a binary outcome. In VanderWeele and Vansteelandt (2009) and Valeri and VanderWeele (2011) this is studied only in the special case of logistic regression with a rare outcome. The general formulas of Section 4 can be applied without a rare outcome assumption. Pearl (2010, 2011a) explicates the effects in a general non-parametric way, without a need for probit or logistic regression, although acknowledging that in practice such parametric approaches are typically called for. The formulas are expressed here in terms of both probit and logistic regression.

7.1 Causal effects with a binary mediator and a binary outcome

In Section 4 the direct, total indirect, and pure indirect effects are defined as

$$DE = E[Y(1, M(0)) - Y(0, M(0)) \mid C], \tag{37}$$

$$TIE = E[Y(1, M(1)) - Y(1, M(0)) \mid C], \tag{38}$$

$$PIE = E[Y(0, M(1)) - Y(0, M(0)) \mid C]. \tag{39}$$

Appendix Section 13.3 shows that with a binary mediator and a binary outcome these formulas lead to the expressions

$$DE = [F_Y(1,0) - F_Y(0,0)]\,[1 - F_M(0)] + [F_Y(1,1) - F_Y(0,1)]\,F_M(0), \tag{40}$$

$$TIE = [F_Y(1,1) - F_y(1,0)]\,[F_M(1) - F_m(0)], \tag{41}$$

$$PIE = [F_Y(0,1) - F_y(0,0)]\,[F_M(1) - F_m(0)]. \tag{42}$$

where $F_Y(x, m)$ denotes $P(Y = 1 \mid X = x, M = m)$ and $F_M(x)$ denotes $P(M = 1 \mid X = x)$, where F denotes either the standard normal or the logistic distribution function corresponding to using probit or logistic regression. These formulas agree with those of Pearl (2010, 2011a). The following sections give two examples, applying these causal effects using Mplus.

7.2 Pearl's hypothetical binary case

Pearl (2010, 2011a) provided a hypothetical example with a binary treatment X, a binary mediator M corresponding to the enzyme level in the subject's blood stream, and a binary outcome Y corresponding to being cured or not. This example was also discussed on SEMNET in September 2011 (see web reference below). Table 9 shows the design of the example.

[Table 9 about here.]

The top part of the table suggests that the percentage cured is higher in the treatment group for both enzyme levels and that the effect of treatment is higher at enzyme level 1 than enzyme level 0. There is therefore a treatment-mediator interaction in line with Figure 1, except with a binary mediator and a binary outcome. Because of the non-linear expressions of Section 7.1, however, the interaction should not be expected to take a simple linear form as in Section 4.1. An analysis is needed to clarify what role the enzyme mediator plays. While this can be done using the population values of Table 9, a Monte Carlo simulation study is carried out to also study the sampling behavior of the effects.

7.2.1 Monte Carlo simulation

Using a sample of $n = 400$, where the subjects have equal probability of being in the control and treatment groups, Mplus Monte Carlo simulations are carried out using the specifications of Table 9. Data are generated and analyzed using both logit and probit. The Mplus inputs are shown in the appendix Section 14.5, also giving the definitions of the quantities derived from the model parameters. These include ratios of direct and indirect

effects relative to the total effect as in Pearl (2010, 2011a). The effects are labeled de for direct effect, tie for total indirect effect (natural indirect effect), pie for pure indirect effect, te for total effect, with ratios dete, tiete, and piete. Furthermore, compdete refers to the direct effect complement $1 - de/te$. Note that $1 - de/te = tie/te$ because $te - de = tie$, that is $te = de + tie$.

The results for logit with maximum-likelihood estimation are shown in Table 10, the results for probit with maximum-likelihood estimation are shown in Table 11, and the results for probit with Bayesian estimation are shown in Table 12. It is seen that all causal effects are well recovered, giving good approximations to the values shown in Pearl (2010, 2011a).

The tables show a somewhat unusual situation where the y on m regression slope would be insignificant at this sample size, but the xm interaction regression slope would be significant. In terms of causal effects, the interaction effect shows up most clearly in the difference between the total indirect effect (tie) and the pure indirect effect (pie). Pearl (2011a) focuses the interpretation on the direct effect complement ($compdete = 1 - de/te$ which is the same as $tiete = tie/te$) and the pure indirect effect ratio to total effect ($piete = pie/te$), concluding (the values referred to are given in the Population column):

> "We conclude that 30.4% of all recoveries is owed to the capacity of the treatment to enhance the secretion of the enzyme, while only 7% of recoveries would be sustained by enzyme enhancement alone."

Further discussion of this example by Pearl is available at

http://www.mii.ucla.edu/causality/wp-content/uploads/2011/09/grice.pdf

[Table 10 about here.]

[Table 11 about here.]

[Table 12 about here.]

7.2.2 Example 3: N=200 data based on the Pearl example

An example that fulfills the design of Table 9 with 100 subjects in the control group and 100 in the treatment group is shown in Table 13.

[Table 13 about here.]

The Mplus input for a Bayes analysis of these data using probit is shown in Appendix Section 14.6. The results are shown in Table 14. Bayesian estimation allows for non-normal parameter distributions. As an example, the posterior distribution for the ratio of the direct effect to the total effect is shown in Figure 7.

[Table 14 about here.]

[Figure 7 about here.]

7.3 Binary mediator and continuous outcome

When the outcome is continuous instead of binary, the formulas of (40) - (42) still apply by changing $F_Y(x, m)$ to denote the expectation $E(Y \mid X = x, M = m)$. The expectation of Y is obtained for the various 0 and 1 values of x and m indicated in the three formulas.

8 Mediation modeling with a nominal mediator

Mediation modeling with a nominal mediator has apparently not been approached in the SEM literature or in the causal mediation literature. The question is how such mediation should be conceptualized. What does it mean that a nominal variable acts as a mediator? As a hypothetical example, consider an intervention aimed at reducing air pollution. An important part of the intervention is to encourage people to change from using their own car while commuting to work in favor of a van pool, bus, or light rail. The mode of transportation mediator is therefore nominal. A direct effect is also possible by the intervention also aiming to encourage other low-pollution activities.

Here again, the general formulas of Section 4 can be used. The formulas need the distribution of M conditional on X and the expectation of Y conditional on M and X, followed by integration/summation over M. The influence of X on M can be modeled by a multinomial logistic regression so that the distribution of M conditional on X is well defined. The influence of M on Y is naturally captured by different Y means for the different M categories, by different Y=1 probabilities for a binary Y, or by different rates for a count Y. Appendix Section 13.4 shows the causal effects for a continuous outcome Y. The corresponding formulas for a binary or count outcome Y follow in a straightforward way.

The joint analysis of a nominal variable as a dependent variable in one regression and as an independent variable in another regression is easily handled in Mplus by using a mixture analysis with a nominal latent class

variable that is the same as the observed nominal M. In this case, the latent class membership is known, drawing on the Mplus KNOWNCLASS feature. The Y means change over the classes as the default. An interaction between X and M is captured by letting the direct influence of X on Y vary over the latent classes. Maximum-likelihood estimation can be carried out for the two regressions and the causal effects defined in MODEL CONSTRAINT as before. The Mplus approach also allows for the nominal mediator to not be observed but latent, or partly observed, or observed with error.

8.1 Monte Carlo simulation

A Monte Carlo simulation is carried out with $n = 800$ for a 3-category mediator where the most polluting mode of transportation is the third category. The Mplus input for Step 1 and Step 2 of the simulation are shown in Section 14.7. The results are shown in Table 15, Table 16, and Table 17. It is seen that the estimation performs very well. The direct and indirect effects show good coverage. The Step 1 and Step 2 results are slightly different due to latent class being unobserved in Step 1 and observed in Step 2.

[Table 15 about here.]

[Table 16 about here.]

[Table 17 about here.]

8.2 Example 4: Hypothetical pollution data with a nominal mediator and a binary outcome

Consider the hypothetical data in Table 18 as an example of the pollution intervention with a binary outcome. The mediator category 3 corresponds to using the car and has the highest pollution percentage.

[Table 18 about here.]

The Mplus input for this analysis is shown in Appendix Section 14.8. The results are shown in Table 19 and Table 20.

[Table 19 about here.]

[Table 20 about here.]

9 Mediation modeling with a count outcome

Causal effects using a count outcome are shown in Appendix Section 13.5. A continuous mediator is considered, but as mentioned in the appendix the count variable can also be a mediator. A count outcome can also be combined with a binary or nominal mediator. To model the count variable, Mplus can handle Poisson, negative binomial, and inflation versions of those models as well as zero-truncation, hurdle modeling, and mixture (latent class) versions.

Appendix Section 14.9 shows the Mplus input for a Monte Carlo simulation study with a count outcome and a continuous mediator using maximum-likelihood estimation. The results are shown in Table 21.

[Table 21 about here.]

10 Violated assumptions and sensitivity analysis

As shown in the preceding sections, causally-derived direct and indirect effects are not necessarily the same as SEM effects, particularly with non-continuous mediators and/or outcomes. The causally-derived effects can, however, be obtained via extended types of SEM analyses using Mplus. To claim that effects are causal, however, it is not sufficient to simply use the causally-derived effects. The set of assumptions given earlier needs to be fulfilled for the effects to be causal and the plausibility of these assumptions needs to be considered in each application. One way to read Holland (1988) and Sobel (2008) is that the authors think many if not most applications are not likely to fulfill such assumptions even in randomized studies. This is also echoed in Bullock et al. (2010).

Imai et al. (2010a, b) stress the importance of sensitivity analysis as part of mediation analysis. Techniques to study sensitivity to assumptions have been proposed in Imai et al. (2010a, b) and VanderWeele (2010a). This section focuses on the critical assumption of no mediator-outcome confounding and shows how the sensitivity analysis proposed by Iamai et al. is carried out in Mplus.

Consider the violation of the no mediator-outcome confounding in the context of the simple mediation model of Figure 8. An unmeasured (latent) variable Z influences both the mediator M and the outcome Y. When Z is not included in the model, a covariance is created between the residuals in

the two equations of the regular mediation model as indicated in Figure 9. Including the residual covariance, however, makes the model not identified. An example of a mediator-outcome confounder in the aggressive behavior example of Section 6.5 is the variable poverty which may affect both the Grade 5 aggression score mediator and the juvenile court record outcome. There are presumably many such omitted variables in a typical study.

Imai et al. (2010a, b) propose a sensitivity analysis where causal effects are computed given different fixed values of the residual covariance. This is useful both in real-data analyses as well as in planning studies. As for the latter, the approach can answer questions such as how large does your sample and effects have to be for the lower confidence band on the indirect effect to not include zero when allowing for a certain degree of mediator-outcome confounding?

As a first step in understanding the Imai et al. approach, Figure 10 indicates that there is another way to estimate the mediation model. The figure shows that M and Y are regressed on X, allowing for a residual covariance, but Y is not regressed on M. To illustrate this approach, a Monte Carlo study is performed to show that the same estimates of the indirect and direct effects are obtained as when regressing M on X and regressing Y on M and X. Appendix Section 14.10.1 shows the Mplus input for generating the data using the M on X, Y on M and X model, while analyzing the data using the M on X, Y on X model of Figure 10. Table 22 shows the results, verifying that the data-generating parameters are well recovered.

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

[Table 22 about here.]

As a second step in understanding the Imai et al. approach, Appendix Section 13.6 shows how the parameters of the Figure 10 model can be used to derive indirect and direct effects under different assumptions for the residual covariance in the Figure 9 model. The coefficient β_1 of the indirect effect $\beta_1\ \gamma_1$ is obtained as

$$\beta_1 = \sigma/\sigma_2\ (\tilde{\rho} - \rho\ \sqrt{(1-\tilde{\rho}^2)/(1-\rho^2)}), \tag{43}$$

where σ and σ_2 are the standard deviations of the outcome and mediator residuals in the Figure 10 model, $\tilde{\rho}$ is the correlation between these residuals, and ρ is a sensitivity parameter representing the non-identified correlation between the residuals of the Figure 9 model. The coefficient γ_1 is obtained from the regression of M on X. Appendix Section 13.6 shows that the direct effect β_2 is obtained as

$$\beta_2 = \kappa_1 - \beta_1\ \gamma_1, \tag{44}$$

where κ_1 is obtained from the regression of Y on X.

10.1 Sensitivity analysis in a Monte Carlo study

To illustrate the sensitivity analysis, Appendix Section 14.10.2 shows the Mplus input for a Monte Carlo study that generates data according to Figure 9 with a residual correlation of 0.25. The indirect effect is 0.25 and the direct effect is 0.4. The data are analyzed by the model of Figure 10 using MODEL CONSTRAINT to derive the data-generating parameters

according to the appendix formulas while applying a fixed correlation of $\rho = 0.25$, that is, the true correlation. Table 23 shows that the indirect and direct effects (labeled ind and de) are correctly estimated with this adjustment.

[Table 23 about here.]

A sensitivity analysis is obtained by varying the fixed ρ correlation in MODEL CONSTRAINT. The above Monte Carlo study is used to illustrate this. The correct value for the indirect effect is 0.25 (marked with a horizontal broken line). The biased estimate assuming $\rho = 0$ is 0.3287, an overestimation due to ignoring the positive residual correlation. The sensitivity analysis varies the ρ values from -0.9 to $+0.9$. A graph of the indirect effect is shown in Figure 11, including a 95% confidence interval. Using $\rho = 0$, the biased estimate of 0.3287 is obtained, that is, no adjustment is made. Using the correct value of $\rho = 0.25$, the correct indirect effect value of 0.25 is obtained. For lower ρ values the effect is overestimated and for larger ρ values the effect is underestimated.

[Figure 11 about here.]

The graph provides useful information for planning new data collections. At this sample size ($n = 400$) and effect size, the lower confidence limit does not include zero until about $\rho = 0.6$. This means that a rather high degree of confounding is needed for the effect to not be detected. Also, in the range of ρ from about -0.1 to +0.4 the confidence interval covers the correct value of 0.25 for the indirect effect.

These results are obtained by maximum-likelihood estimation using regular standard errors and using symmetric confidence intervals due

to the assumption of a normal parameter estimate distribution for the indirect effect. For smaller samples it may be better to use confidence (credibility) intervals generated by Bayesian analysis, allowing for a non-normal posterior for the indirect effect, producing non-symmetric confidence intervals.

10.2 Example 5: Sensitivity analysis for head circumference at birth and mother's drinking and smoking

This example considers the effects on the baby's head circumference of mother's drinking and smoking during pregnancy (Day et al., 1994). A reduction in head circumference is frequently used as a proxy for the potential of deficient cognitive development in a child. The dependent variables in the mediation model are baby's head circumference at birth (hcirc0) and at 36 months (hcirc36). The key focus is on a binary risk factor defined by the mother's drinking and smoking during the third trimester (alcccig).

Figure 12 shows the mediation model. One may hypothesize that mothers' drinking and smoking during pregnancy affect babies' head circumference at birth, but any effect at 36 months is an indirect effect via hcirc0. That is, if head circumference is low at 36 months it is because it is low at birth. An alternative hypothesis is that a baby's head circumference at birth and 36 months are both directly affected by mother's drinking and smoking during pregnancy. That is, the growth rate in head circumference, after the baby has left the womb, is affected by mother's drinking and

smoking during pregnancy.

It should be emphasized that this is not a randomized study, so that there are many possibilities for confounding. As a minimal set, gender and ethnicity are added as covariates to be able to gauge the effects of the mother's behavior during pregnancy controlling for those variables. For example, male babies tend to have a larger head circumference at birth than female babies and males may also have a faster growth rate, hence impacting both the mediator and the outcome. The baby's gender is scored as 1 for males and 0 for females, and baby's ethnicity scored as 1 for blacks and 0 for others.

[Figure 12 about here.]

In line with the Imai et al. sensitivity approach, hcirc36 is regressed on alccig, gender, and ethnicity and hcirc0 is regressed on alccig, gender, and ethnicity. A first analysis uses a residual correlation ρ fixed at zero, that is, carrying out a regular mediation analysis equivalent to that of Figure 12. The Mplus input is given in Appendix Section 14.10.3. Table 24 shows the results. In the section New/additional parameters this gives a significant indirect effect of -0.162 and an insignificant direct effect of 0.084, both in standard deviation units for hcirc36. In terms of the parameters of the original model of Figure 12, the estimate for β_1 is significant at 0.444, and the estimate of γ_1 is found at the top of the table under the regression of hcirc0 on alccig, namely -0.366. The β_1 γ_1 product is the reported indirect effect. The results indicate that mother's drinking and smoking are detrimental to the child's head circumference at birth, having an indirect effect also three years later, but having no direct effect. A sensitivity analysis is, however, needed to study effects

of potential omitted mediator-outcome confounders. There are presumably many omitted variables influencing head circumference at both birth and 36 months. It is likely that these omitted variables create a positive correlation between the residuals of the mediator and the outcome.

[Table 24 about here.]

Figure 13 shows the results of the sensitivity analysis. The data for the graph are produced by a series of analyses using the Mplus input in Appendix Section 14.10.3, varying the ρ value of MODEL CONSTRAINT. The figure shows that if the residual correlation ρ is less than about 0.4, the negative indirect effect is still bounded away from zero. A residual correlation as large as 0.4 or larger might, however, be considered quite possible in this application. If so, the detrimental indirect causal effect of mother's drinking and smoking may not be convincingly demonstrated in this case.

[Figure 13 about here.]

The direct effect also changes as a function of the residual correlation ρ (see (44). Figure 14 shows that the direct effect is not significantly different from zero in the range of ρ from -0.3 to 0.75. Assuming that the residual correlation falls somewhere in this wide range, a direct effect is not detected.

[Figure 14 about here.]

11 General mediation modeling

The basic mediation models discussed so far are simple versions of what is often seen in practice. This section lists a few of the generalizations and

outlines how the causally-defined effects come into play in these models.

11.1 Moderated mediation

The need to study moderated mediation frequently arises in applications. Figure 1 of Section 2 is an example where the binary treatment variable X moderates the influence of the mediator M on the outcome Y. An example of moderation of the regression of M on X and the regression of Y on X is shown in Figure 15, where the observed covariate Z is a moderator. Using the aggressive behavior example of Section 6.5, the Grade 1 Fall aggression score may serve as a moderator in that initially more aggressive boys are somewhat more likely to benefit from the intervention. This is often referred to as treatment-baseline interaction.

[Figure 15 about here.]

Figure 15 corresponds to the model

$$y_i = \beta_0 + \beta_1\, m_i + \beta_2\, x_i + \beta_3\, x_i\, z_i + \epsilon_{1i}, \tag{45}$$

$$m_i = \gamma_0 + \gamma_1\, x_i + \gamma_2\, z_i + \gamma_3\, x_i\, z_i + \epsilon_{2i}, \tag{46}$$

Applying the Appendix Section 13.1 formulas, it follows that the direct and total indirect effects are

$$DE = \beta_2 + \beta_3\, z, \tag{47}$$

$$TIE = \beta_1\, (\gamma_1 + \gamma_3\, z). \tag{48}$$

The effects can then be evaluated at different z values of interest.

For a binary moderator, multiple-group SEM gives a flexible approach. Again using the aggressive behavior example, females have less of an effect of the intervention than boys. The multiple-group approach can estimate the same parameters as in (45) and (46), leading to the same effect definitions, but also allows further flexibility such as group-varying residual variances.

11.2 Mediation analysis with latent variables

In a more general setting, latent variables may often play the roles of mediators and outcomes. The latent variables may represent continuous latent response variables, continuous factors, or categorical latent class variables.

11.2.1 Latent response variables: Latent versus observed binary and ordinal mediators and outcomes

In the smoking example of Section 6.6, the analyses compared treating the mediator as an observed variable versus a latent response variable, or response tendency, m^* behind an ordered categorical (ordinal) observed variable. Similarly, a binary mediator can be treated as either the observed binary variable or as the latent continuous response variable. The substantively relevant mediator may be the response tendency or the actual manifestation. This same line of thinking applies to the outcome. For example, the causal effects for an ordinal outcome can be expressed by the causal formulas in terms of the expectation of this observed categorical variable, where an intervention attempts to increase or decrease the probabilities of certain observed categories. Or, the substantively relevant outcome may be the response tendency, where the observed categories are

merely crude categorizations of this tendency. The choice decides if the causal effects for continuous or categorical variables should be used.

11.2.2 Factors

Figure 16 shows an example of factors measured by multiple indicators. In this case, the causally-defined effects pertain to the continuous latent mediator fm and the continuous latent outcome fy, that is, the usual formulas for continuous variables apply. Adding moderated mediation implies modeling with interactions involving latent variables, which is available in Mplus using maximum-likelihood estimation.

[Figure 16 about here.]

11.2.3 Latent class variables

Figure 17 shows an example where the mediator is a latent class variable measured by multiple indicators. The multiple indicators may correspond to repeated measures with random effects (i and s) as in growth mixture modeling (Muthén & Asparouhov, 2009). In these cases, the mediator is nominal and the formulas of Section 8 apply. This involves mixture analysis, which is available in Mplus using maximum-likelihood or Bayesian estimation.

[Figure 17 about here.]

11.3 Multilevel mediation

Causal inference in multilevel settings presents further challenges for mediational modeling and is beyond the scope of this paper. Additional

assumptions are needed for causally-defined effects. Key references include Hong and Raudenbush (2006) and VanderWeele (2010b).

12 Conclusions

This paper summarizes some of the literature on causal effects in mediation analysis. Applications are shown where the effects are estimated using Mplus. This broadens mediation analysis as currently carried out in SEM practice, where causal effects have been considered only in the case of continuous mediators and outcomes. In this paper, causal effects are computed also for mediators and outcomes that are binary, ordinal, nominal, or count variables. The causal effects require strong assumptions even in randomized designs, especially sequential ignorability, which is presumably often violated to some extent due to mediator-outcome confounding. To study the effects of violating this assumption, it is shown how a sensitivity analysis developed by Imai et al. (2010a,b) can be carried out using Mplus. This can be used both in planning a new study and in evaluating the results of an existing study.

Reports on SEM analyses often use language to interpret their findings which implies that the effects found are causal. The causal effects literature indicates how difficult it can be for such claims to be correct. It is likely that more often only approximations to causal findings are obtained. In this sense, SEM mediation analysis perhaps serves more as a useful exploratory tool rather than a confirmatory causal analysis device, as is sometimes claimed.

Ongoing research on the mediation topic focuses on the Achilles heel of

the analysis, namely that the mediator is not randomized. To avoid this, new designs are explored, such as parallel designs, encouragement designs, and crossover designs; see, e.g., Bullock et al. (2010) and Imai et al (2011). These designs, however, come with their own challenges and assumptions and much further research is needed.

References

[1] Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 11731182.

[2] Bollen, K. A. (1989). Structural equation models with latent variables. New York: John Wiley.

[3] Bullock, J.G., Green, D.P. & Ha, S.E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). Journal of Personality and Social Psychology, 98, 550-558.

[4] Day, N.L., Richardson, G.A., Geva, D. & Robles, N. (1994). Alcohol, marijuana, and tobacco: The effects of prenatal exposure on offspring growth and morphology at age six. Alcoholism: Clinical and Experimental Research 18, 786794.

[5] Glymour, C. (2011). Counterfactuals, graphicalc ausal models and potential outcomes: Response to Lindquist and Sobel. NeuroImage, In Press.

[6] Goldberger, A.S. & Duncan, O.D. (1973). Structural equation models in the social sciences. New York: Seminar Press.

[7] Holland, P. W. (1988). Causal inference, path analysis and recursive structural equation models (with discussion). In: Sociological Methodology 1988, Ed C.C. Clogg, American Sociological Association.

[8] Hong, G. & Raudenbush, S.W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel

observational data. Journal of the American Statistical Association, 101, 901-910.

[9] Imai, K., Keele, L., & Tingley, D. (2010a). A general approach to causal mediation analysis. Psychological Methods, 15, 309-334.

[10] Imai, K., Keele, L., & Yamamoto, Y. (2010b). Identification, inference and sensitivity analysis for causal mediation effects. Statistical Science, 25, 51-71.

[11] Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2010c). Advances in Social Science Research Using R (ed. H. D. Vinod), chapter Causal Mediation Analysis Using R, pages 129-154. Lecture Notes in Statistics. Springer, New York.

[12] Imai. K., Tingley, D., & Yamamoto, T. (2011). Experimental designs for identifying causal mechanisms. Forthcoming in Journal of the Royal Statistical Society, Series A (with discussions).

[13] James, L.R. & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. Journal of Applied Psychology, 69, 307-321.

[14] Jöreskog, K.G. & Sörbom, D. (1979). Advances in factor analysis and structural equation models. Cambridge, MA: Abt Books.

[15] Judd, C.M. & Kenny, D.A. (1981). Process analysis: estimating mediation in treatment evaluations. Evaluation Review, 5:602-619.

[16] Kraemer, H.K., Kiernan, M., Essex, M. & Kupfer, D.J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. Health Psychology, 27, S101-S108.

[17] Lindquist, M.A. & Sobel M.E. (2010). Graphical models, potential outcomes and causal inference: Comment on Ramsey, Spirtes and Glymour. NeuroImage, 57, 334-336.

[18] Lindquist, M.A. & Sobel M.E. (2011). Cloak and DAG: A response to the comments on our comments. Forthcoming in NeuroImage.

[19] MacKinnon, D.P., Lockwood, C.M., Brown, C.H., Wang, W., & Hoffman, J.M. (2007). The intermediate endpoint effect in logistic and probit regression. Clinical Trials, 4, 499-513.

[20] MacKinnon D.P. (2008). An introduction to statistical mediation analysis. New York, NY: Lawrence Erlbaum Associates.

[21] Muthén, B. (1979). A structural probit model with latent variables. Journal of the American Statistical Association, 74, 807-811.

[22] Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika, 49, 115-132.

[23] Muthén, B. & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G. & Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.

[24] Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P., Kellam, S., Carlin, J., & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. Biostatistics, 3, 459-475.

[25] Muthén, L.K. & Muthén, B.O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. Structural Equation

Modeling, 4, 599-620.

[26] Muthén, B. & Muthén, L. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén. Available at www.statmodel.com.

[27] Pearl, J. (2001). Direct and indirect effects. In Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence. San Francisco: Morgan Kaufmann, 411420.

[28] Pearl, J. (2009). Causality: Models, reasoning, and inference. Second edition. New York: Cambridge University Press.

[29] Pearl, J. (2010). The Foundations of Causal inference. In Sociological Methodology 2010, Ed. T. Liao., Wiley.

[30] Pearl, J. (2011a). The causal mediation formula - A guide to the assessment of pathways and mechanisms. Forthcoming in Prevention Science.

[31] Pearl, J. (2011b). Graphical models, potential outcomes and causal inference: Comment on Lindquist and Sobel. NeuroImage, In Press.

[32] Pearl, J. (2011c). The mediation formula: A guide to the assessment of causal pathways in nonlinear models. To appear in C. Berzuini, P. Dawid, and L. Bernadinelli (Eds.), Causality: Statistical Perspectives and Applications.

[33] Preacher, K.J., Rucker, D.D. & Hayes, A.F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. Multivariate Behavioral Research, 42, 185-227.

[34] Robins, J.M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In Highly Structured

Stochastic Systems, Eds. P. Green, N.L. Hjort, & S. Richardson, Oxford University Press, New York, 7081.

[35] Robins, J.M., and Greenland, S. (1992). Identifiability and exchangeability of direct and indirect effects. Epidemiology 3, 143-155.

[36] Sobel, M. (2008). Identification of causal parameters in randomized studies with mediating variables. Journal of Educational and Behavioral Statistics, 33, 230-251.

[37] Valeri, L. & VanderWeele, T.J. (2011). Extending the Baron and Kenny analysis to allow for exposure-mediator interactions: SAS and SPSS macros. Submitted to Psychological Methods.

[38] VanderWeele, T.J. (2010a). Bias formulas for sensitivity analysis for direct and indirect effects. Epidemiology, 21, 540-551.

[39] VanderWeele, T.J. (2010b). Direct and indirect effects for neighborhood-based clustered and longitudinal data. Sociological Methods & Research, 38, 515-544/

[40] VanderWeele T.J. & Vansteelandt S. (2009). Conceptual issues concerning mediation, interventions and composition. Statistics and Its Interface, 2, 457-468.

[41] VanderWeele T.J. & Vansteelandt S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. Am J Epidemiol. 2010; 172(12):13391348.

[42] Winship, C. & Mare, R.D. (1983). Structural equations and path analysis for discrete data. American Journal of Sociology, 89, 54-110.

List of Figures

Figure 1: A mediation model with treatment-mediator interaction. The filled circle represents an interaction term consisting of the variables connected to it without arrow heads, in this case x and m.

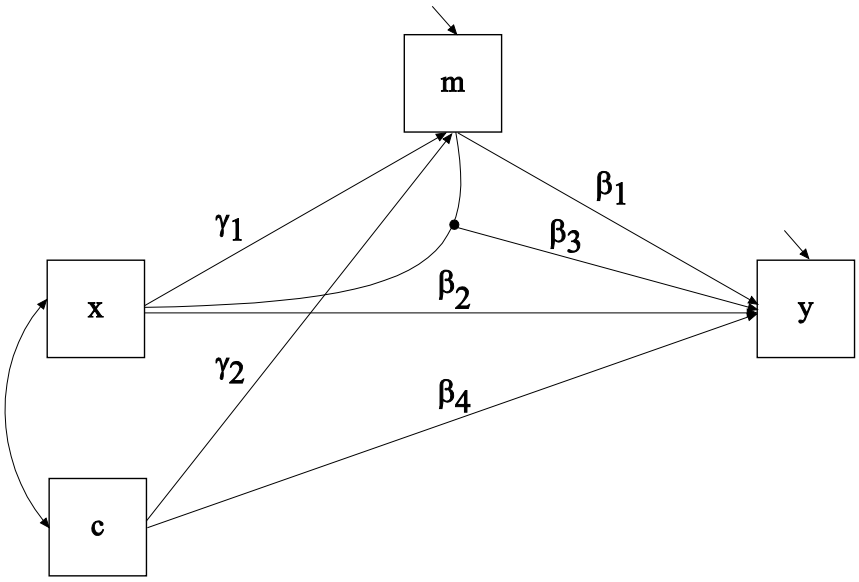


Figure 2: Latent response variable m^* behind a three-category ordinal variable m

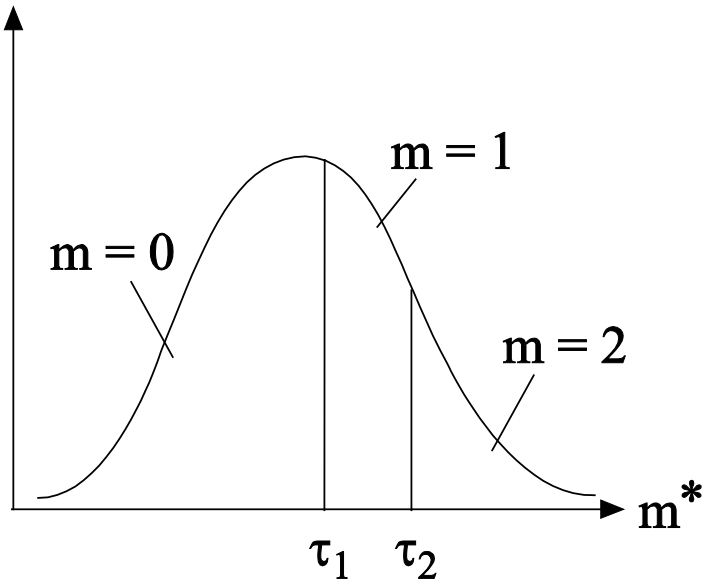


Figure 3: Bayes posterior distribution for the direct effect odds ratio

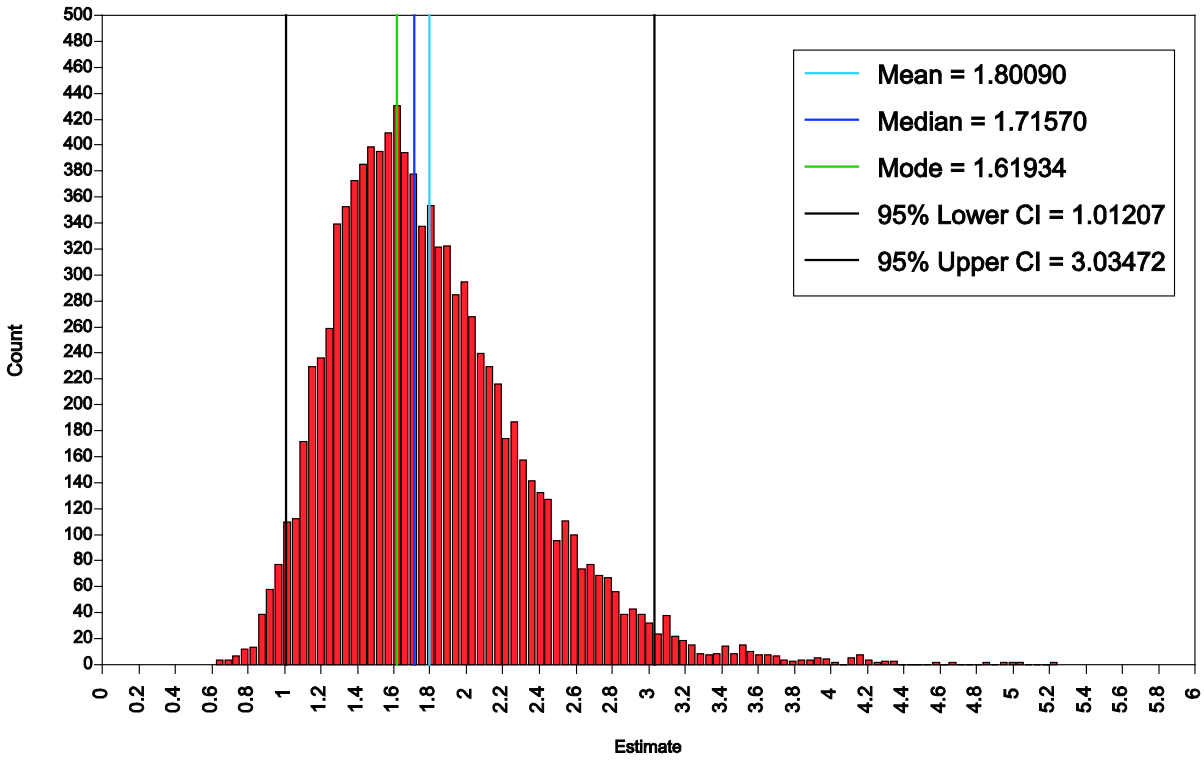


Figure 4: Bayes posterior distribution for the total indirect effect odds ratio

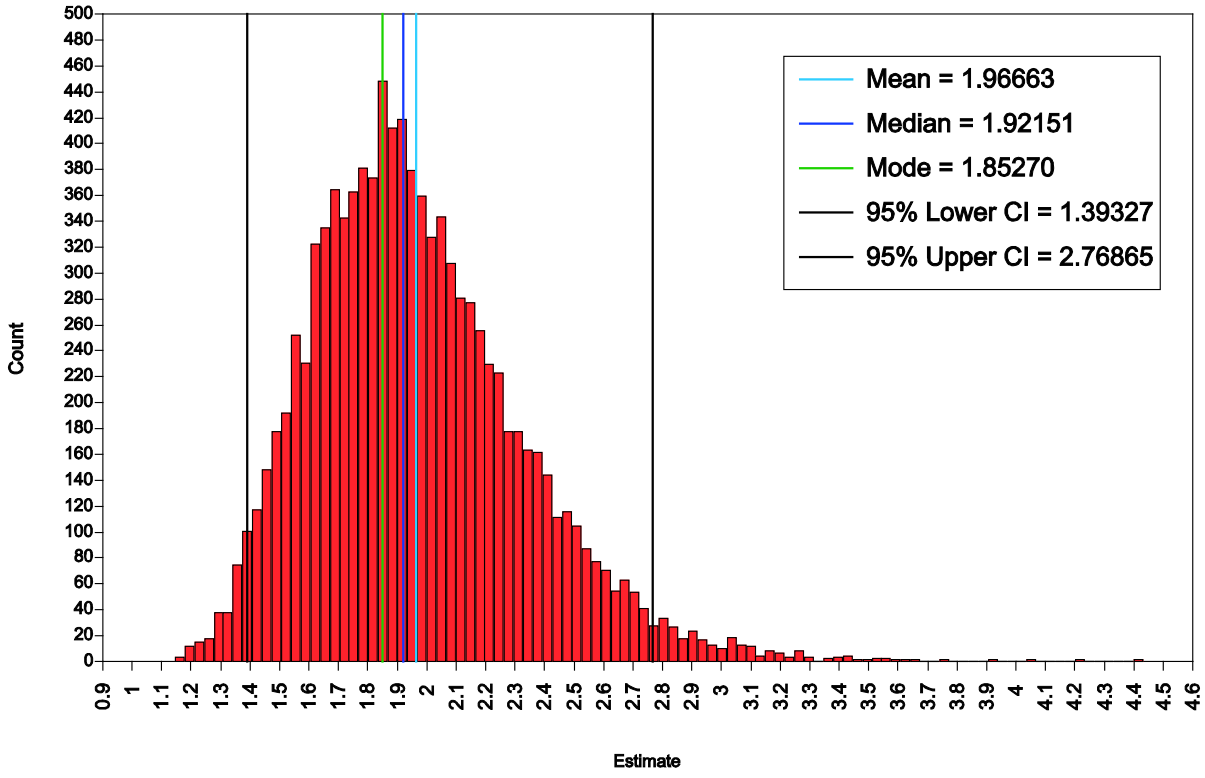


Figure 5: A mediation model for aggressive behavior and juvenile court outcome

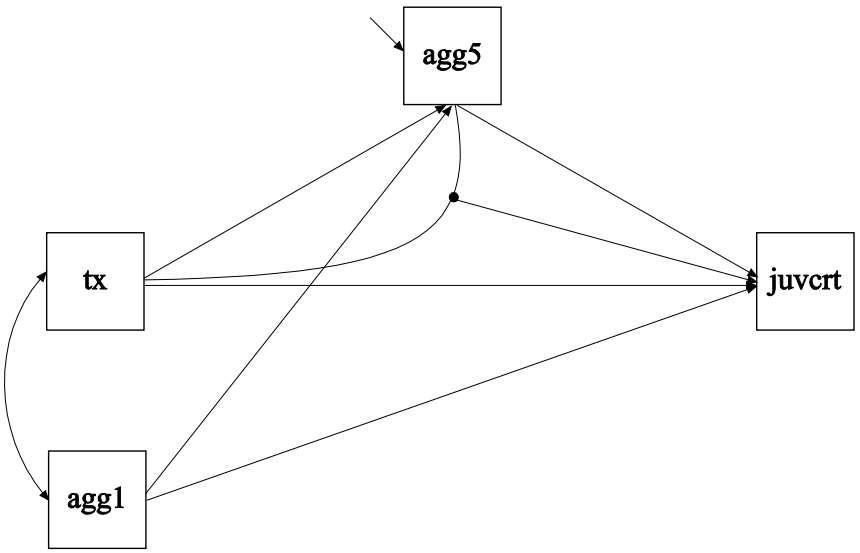


Figure 6: A mediation model for intentions to stop smoking

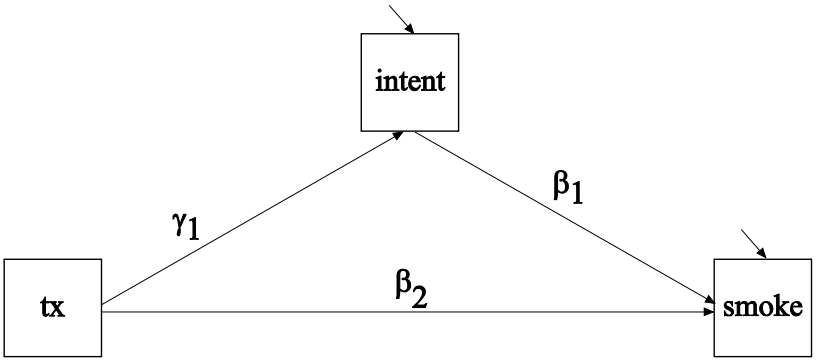


Figure 7: Bayes posterior distribution for the ratio of the direct effect to the total effect for n=200 data based on Pearl

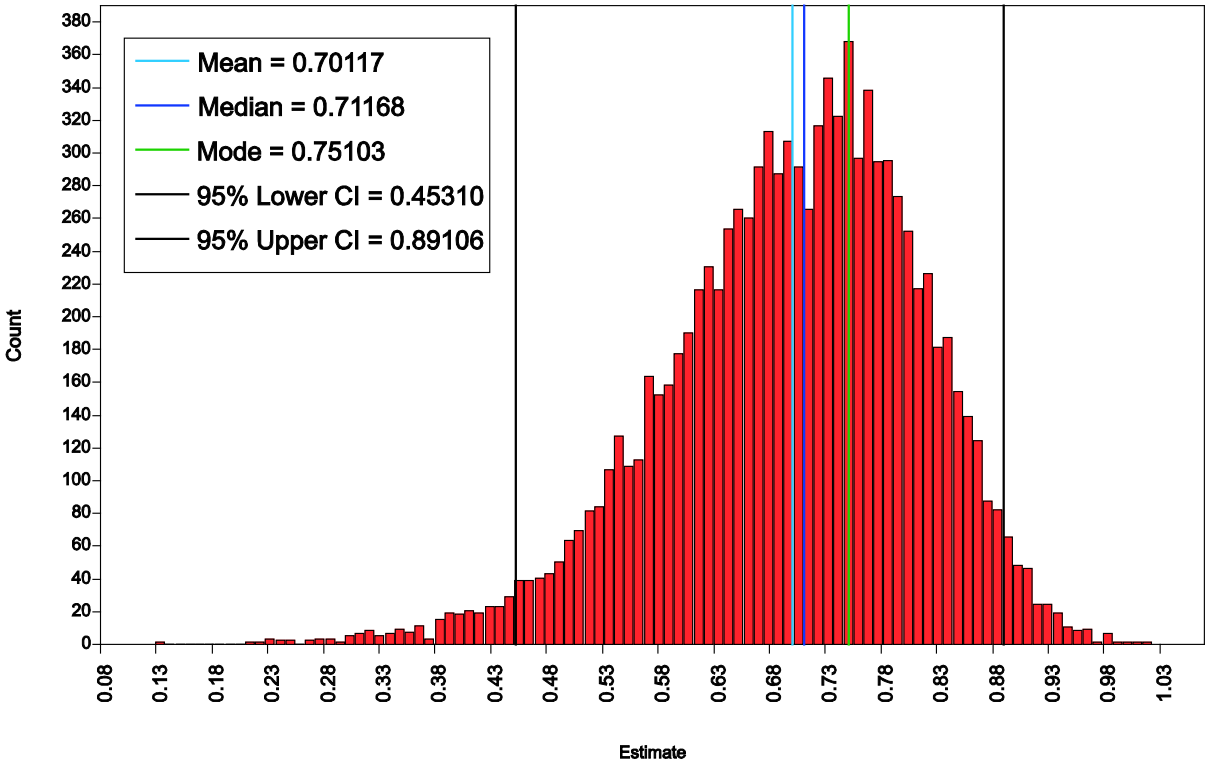


Figure 8: Mediator-outcome confounding 1

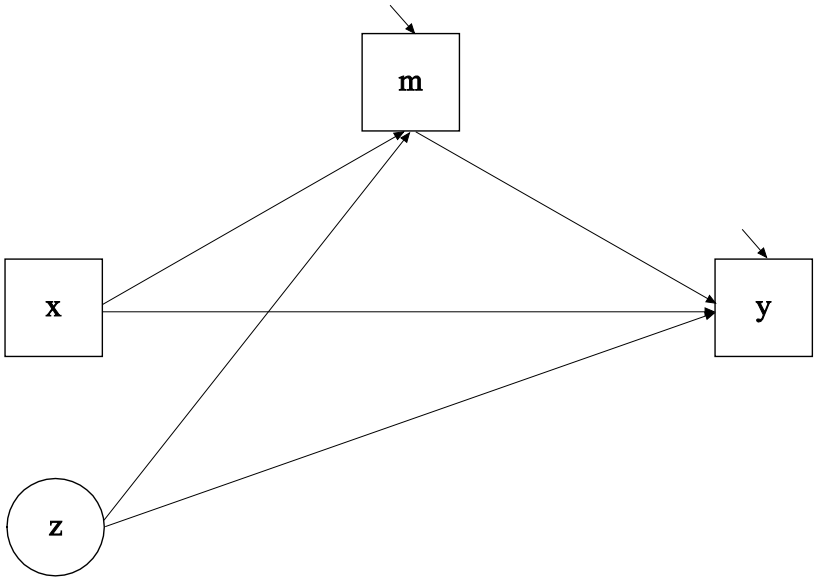


Figure 9: Mediator-outcome confounding 2

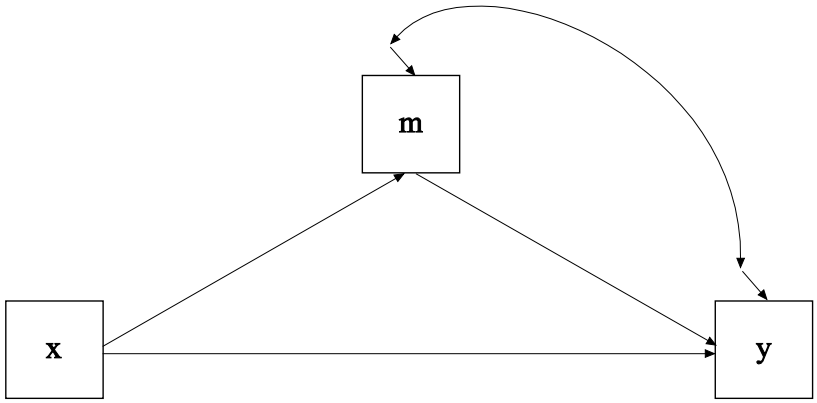


Figure 10: Mediator-outcome confounding 3

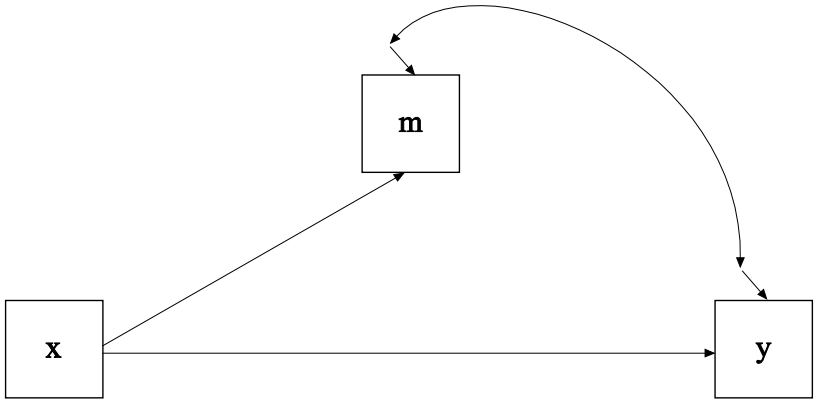


Figure 11: Indirect effect based on sensitivity analysis with ρ varying from -0.9 to +0.9 and true residual correlation 0.25

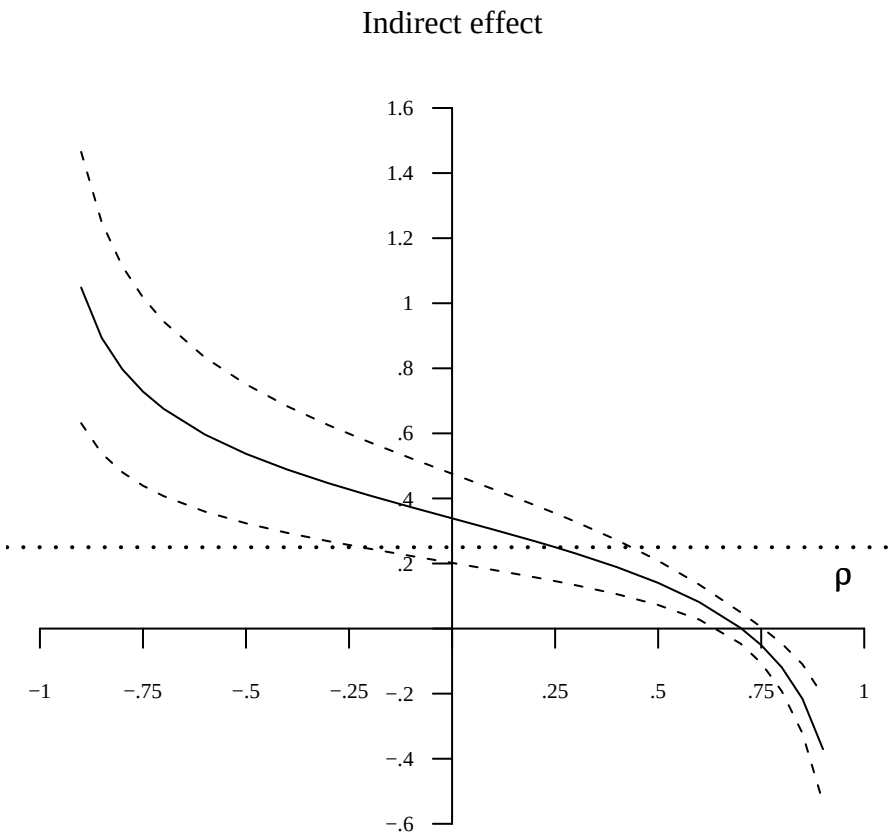


Figure 12: Mediation model for mother's drinking and smoking related to child's head circumference

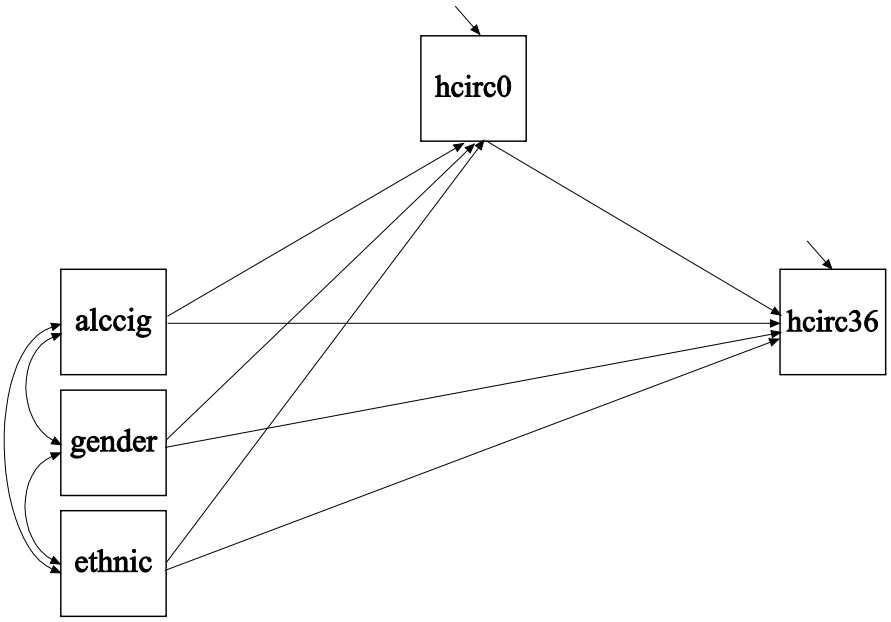


Figure 13: Sensitivity analysis for indirect effect of head circumference example

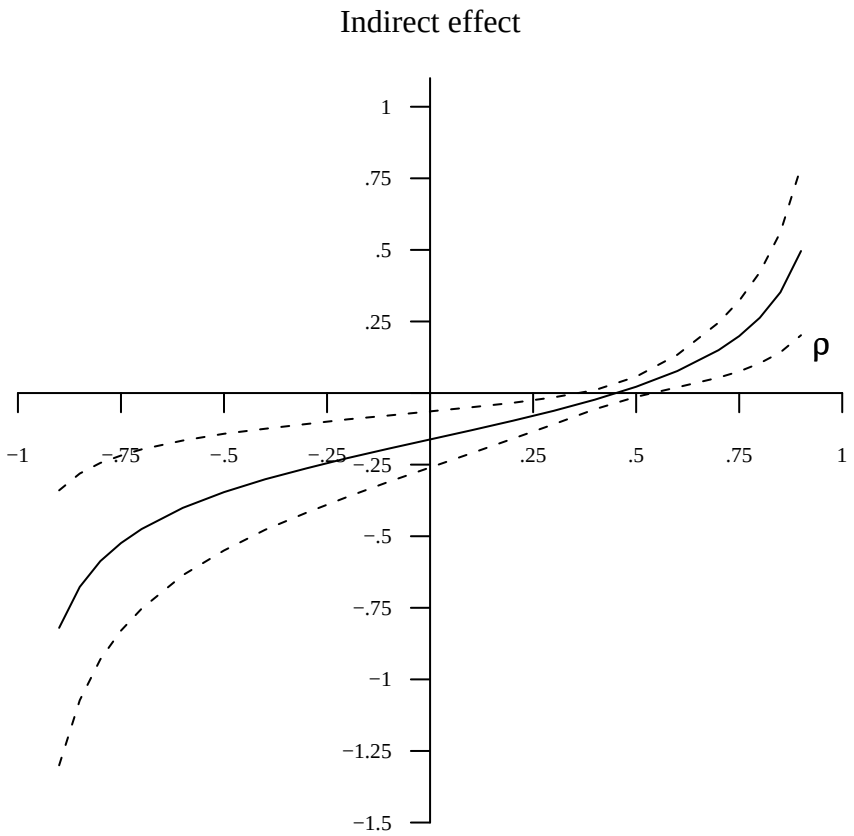


Figure 14: Sensitivity analysis for direct effect of head circumference example

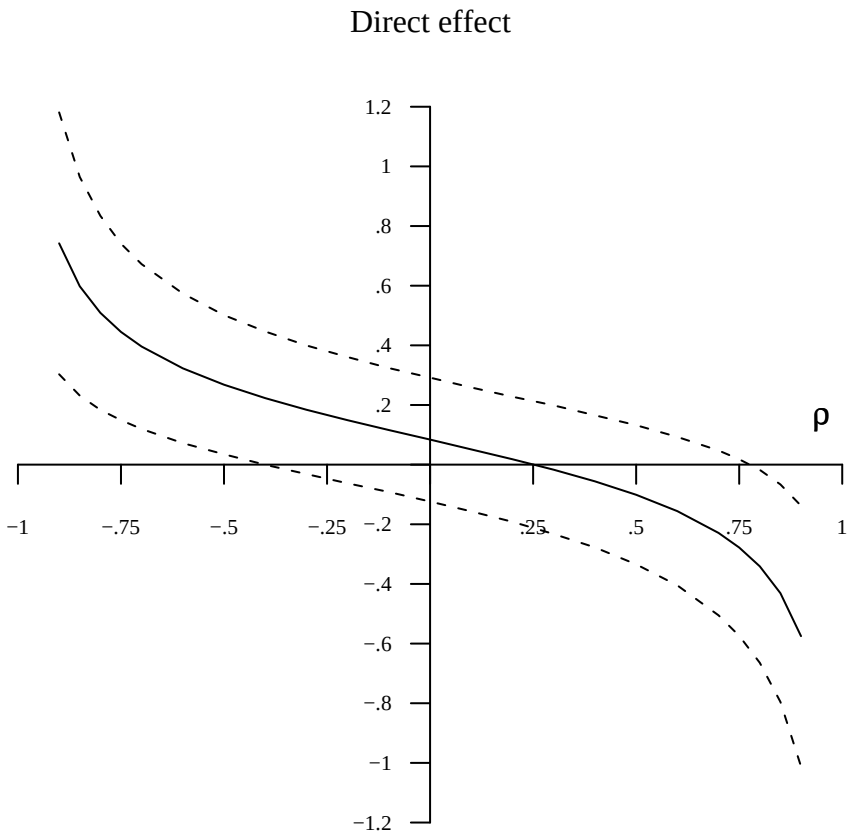


Figure 15: Z moderating the effect of X on M and Y

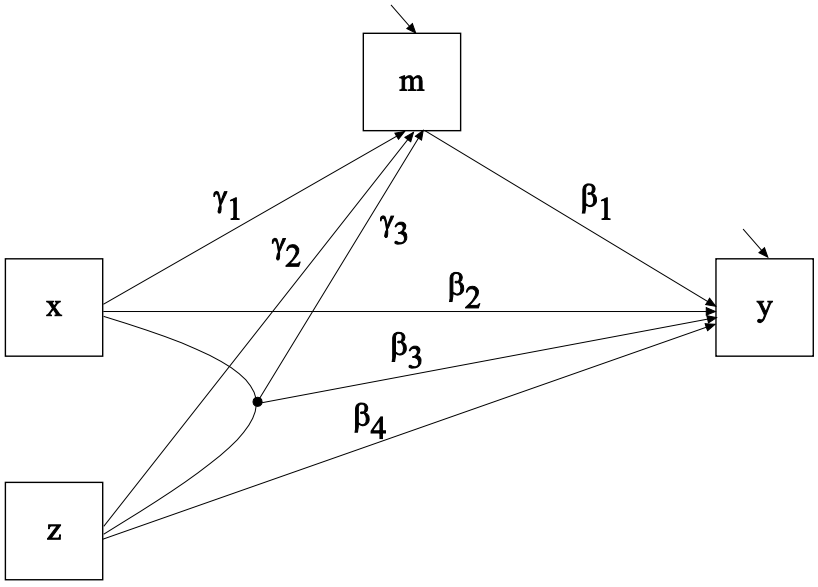


Figure 16: Continuous latent factors as mediator and outcome

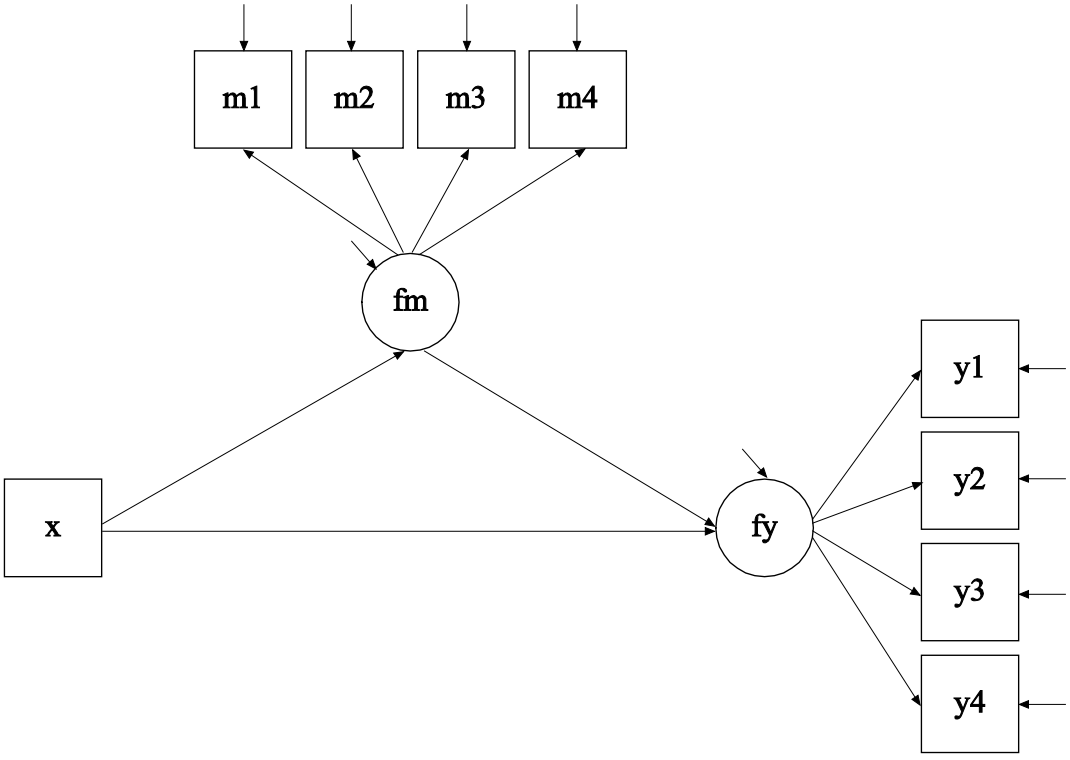
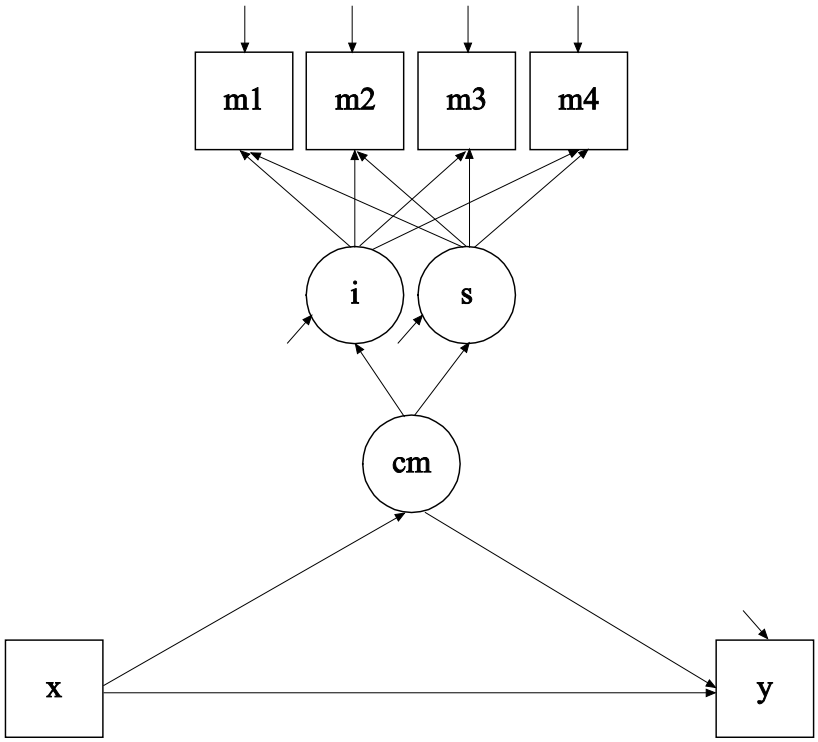


Figure 17: Latent class variable as mediator



List of Tables

Table 1: Output for continuous mediator, continuous outcome with treatment-mediator interaction, Step 2

	Population	Estimates Average	Std. Dev.	S.E. Average	M.S.E.	95% Cover	% Sig Coeff
y ON							
x	0.400	0.4011	0.1784	0.1761	0.0318	0.950	0.616
xm	0.000	0.2006	0.0716	0.0711	0.0051	0.958	0.780
m	0.500	0.5006	0.0493	0.0501	0.0024	0.964	1.000
m ON							
x	0.500	0.5015	0.0981	0.0997	0.0096	0.940	0.998
Intercepts							
y	1.000	0.9984	0.1107	0.1122	0.0122	0.954	1.000
m	2.000	2.0032	0.0683	0.0705	0.0047	0.962	1.000
Residual variances							
y	0.500	0.4974	0.0372	0.0352	0.0014	0.936	1.000
m	1.000	0.9933	0.0667	0.0702	0.0045	0.960	1.000
New/additional parameters							
tie	0.350	0.3518	0.0748	0.0745	0.0056	0.932	0.998
pie	0.250	0.2509	0.0544	0.0561	0.0029	0.950	0.998
de	0.800	0.8027	0.0802	0.0766	0.0064	0.936	1.000

Table 2: Output for Monte Carlo simulation with a binary outcome and a continuous mediator, n = 200, Step 2, ML

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
y ON							
x	0.300	0.2740	0.2796	0.2770	0.0787	0.952	0.194
m	0.700	0.7138	0.1848	0.1799	0.0343	0.956	0.990
xm	0.200	0.2370	0.2865	0.2842	0.0833	0.954	0.110
m ON							
x	0.500	0.4894	0.1207	0.1223	0.0146	0.942	0.972
Intercepts							
m	0.500	0.5044	0.0863	0.0861	0.0074	0.970	1.000
Thresholds							
y$1	0.500	0.5058	0.1670	0.1672	0.0279	0.952	0.880
Residual variances							
m	0.750	0.7465	0.0808	0.0746	0.0065	0.920	1.000
New/additional parameters							
ind	0.450	0.4661	0.1621	0.1600	0.0265	0.948	0.950
dir	0.450	0.3935	0.2140	0.2122	0.0458	0.952	0.462
arg11	0.700	0.7134	0.1858	0.1819	0.0346	0.962	0.992
arg10	0.250	0.2473	0.1846	0.1807	0.0340	0.950	0.304
arg01	0.200	0.2037	0.1788	0.1699	0.0319	0.942	0.218
arg00	-0.150	-0.1462	0.1546	0.1489	0.0239	0.948	0.188
v1	1.607	1.7107	0.3486	0.3263	0.1319	0.948	1.000
v0	1.367	1.4057	0.2238	0.1998	0.0515	0.942	1.000
probit11	0.552	0.5484	0.1317	0.1327	0.0173	0.952	0.992
probit10	0.197	0.1948	0.1468	0.1442	0.0215	0.946	0.260
probit01	0.171	0.1678	0.1437	0.1383	0.0206	0.942	0.240
probit00	-0.128	-0.1244	0.1315	0.1256	0.0173	0.952	0.190
tie	0.131	0.1303	0.0391	0.0388	0.0015	0.946	0.958
de	0.129	0.1255	0.0689	0.0676	0.0047	0.952	0.450
pie	0.119	0.1151	0.0358	0.0632	0.0013	0.936	0.950

Table 3: Output for Monte Carlo simulation with a binary outcome and a continuous mediator, n = 200, Step 2, Bayes

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
y ON							
x	0.300	0.2677	0.2760	0.2762	0.0771	0.954	0.188
m	0.700	0.7126	0.1830	0.1812	0.0336	0.950	0.990
xm	0.200	0.2513	0.2841	0.2869	0.0832	0.958	0.128
m ON							
x	0.500	0.4897	0.1207	0.1240	0.0147	0.946	0.968
Intercepts							
m	0.500	0.5044	0.0863	0.0875	0.0075	0.972	1.000
Thresholds							
y$1	0.500	0.5062	0.1655	0.1656	0.0274	0.950	0.886
Residual variances							
m	0.750	0.7650	0.0828	0.0777	0.0071	0.926	1.000
New/additional parameters							
ind	0.450	0.4616	0.1629	0.1664	0.0266	0.956	0.966
dir	0.400	0.3961	0.2133	0.2134	0.0454	0.956	0.452
arg11	0.700	0.7204	0.1879	0.1851	0.0357	0.946	0.992
arg10	0.250	0.2510	0.1851	0.1818	0.0342	0.944	0.296
arg01	0.200	0.2012	0.1785	0.1714	0.0318	0.940	0.234
arg00	-0.150	-0.1460	0.1544	0.1500	0.0238	0.946	0.194
v1	1.607	1.7456	0.3648	0.3644	0.1519	0.940	1.000
v0	1.367	1.4134	0.2252	0.2157	0.0527	0.954	1.000
probit11	0.552	0.5465	0.1305	0.1312	0.0170	0.952	0.992
probit10	0.197	0.1949	0.1451	0.1421	0.0210	0.946	0.296
probit01	0.171	0.1645	0.1427	0.1376	0.0204	0.940	0.234
probit00	-0.128	-0.1234	0.1305	0.1250	0.0170	0.946	0.194
tie	0.131	0.1266	0.0385	0.0387	0.0015	0.950	0.966
de	0.129	0.1245	0.0673	0.0665	0.0045	0.954	0.468
pie	0.119	0.1106	0.0352	0.0363	0.0013	0.956	0.960
ortie	1.779	1.7858	0.3050	0.3211	0.0929	0.944	1.000
orde	1.681	1.7321	0.4935	0.5272	0.2457	0.956	1.000
orpie	1.614	1.5914	0.2379	0.2553	0.0570	0.958	1.000

Table 4: Output for aggressive behavior and juvenile court record using probit

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
juvcrt ON				
tx	0.003	0.192	0.013	0.990
agg5	0.451	0.103	4.374	0.000
xm	0.263	0.231	1.140	0.254
agg1	-0.003	0.096	-0.036	0.972
agg5 ON				
tx	-0.267	0.115	-2.325	0.020
agg1	0.462	0.060	7.730	0.000
Intercepts				
agg5	0.074	0.070	1.054	0.292
Thresholds				
juvcrt$1	-0.035	0.097	-0.364	0.716
Residual variances				
agg5	0.787	0.074	10.706	0.000
New/additional parameters				
ind	-0.191	0.096	-1.983	0.047
dir	0.022	0.197	0.111	0.911
arg11	-0.100	0.174	-0.576	0.565
arg10	0.090	0.176	0.514	0.607
arg00	0.069	0.102	0.672	0.502
v1	1.401	0.247	5.664	0.000
v0	1.160	0.076	15.310	0.000
probit11	-0.085	0.147	-0.574	0.566
probit10	0.076	0.147	0.521	0.602
probit00	0.064	0.095	0.673	0.501
indirect	-0.064	0.030	-2.158	0.031
direct	0.005	0.067	0.076	0.940
orind	0.773	0.092	8.371	0.000
ordir	1.021	0.275	3.714	0.000

Table 5: Output for aggressive behavior and juvenile court record using logit

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
juvcrt ON				
tx	0.002	0.316	0.006	0.995
agg5	0.726	0.171	4.237	0.000
xm	0.431	0.393	1.096	0.273
agg1	0.000	0.159	-0.002	0.998
agg5 ON				
tx	-0.267	0.115	-2.325	0.020
agg1	0.462	0.060	7.730	0.000
Intercepts				
agg5	0.074	0.070	1.054	0.292
Thresholds				
juvcrt$1	-0.059	0.160	-0.366	0.714
Residual variances				
agg5	0.787	0.074	10.706	0.000
New/additional parameters				
ind	-0.309	0.158	-1.957	0.050
dir	0.034	0.325	0.103	0.918
oddsrat	0.734	0.116	6.334	0.000

Table 6: Intentions to stop smoking data (Source: MacKinnon et al., 2007, Clinical Trials, 4, p. 510)

		Cigarette use		
	Intention	No Use	Use	Total
Ctrl	4 (Yes)	9	20	29
	3 (Probably)	14	20	34
	2 (Don't think so)	36	13	49
	1 (No)	229	30	259
Tx	4 (Yes)	9	19	28
	3 (Probably)	15	11	26
	2 (Don't think so)	43	11	54
	1 (No)	353	32	385

Table 7: Output for intentions to stop smoking using probit with the mediator treated as an observed continuous variable using ML

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
ciguse ON				
tx	-0.203	0.109	-1.867	0.062
intent	0.538	0.048	11.227	0.000
intent ON				
tx	-0.186	0.070	-2.664	0.008
Intercepts				
intent	0.106	0.056	1.906	0.057
Thresholds				
ciguse$1	0.912	0.080	11.432	0.000
Residual variances				
intent	0.990	0.069	14.291	0.000
New/additional parameters				
ind	-0.100	0.038	-2.602	0.009
dir	-0.203	0.109	-1.867	0.062
arg11	-1.158	0.079	-14.579	0.000
arg10	-1.058	0.081	-13.072	0.000
arg00	-0.855	0.085	-10.105	0.000
v1	1.287	0.055	23.545	0.000
v0	1.287	0.055	23.545	0.000
probit11	-1.021	0.072	-14.240	0.000
probit10	-0.933	0.075	-12.514	0.000
probit00	-0.754	0.076	-9.947	0.000
indirect	-0.022	0.009	-2.548	0.011
direct	-0.050	0.027	-1.853	0.064
orind	0.853	0.051	16.587	0.000
ordir	0.731	0.123	5.941	0.000

Table 8: Output for intentions to stop smoking using probit with the mediator treated as a latent continuous variable using WLSMV

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
ciguse ON				
tx	-0.131	0.093	-1.409	0.159
intent	0.631	0.042	15.114	0.000
intent ON				
tx	-0.246	0.089	-2.756	0.006
Thresholds				
ciguse$1	0.760	0.072	10.496	0.000
intent$1	0.525	0.067	7.849	0.000
intent$2	0.970	0.071	13.581	0.000
intent$3	1.378	0.082	16.721	0.000
New/additional parameters				
ind	-0.155	0.057	-2.711	0.007
dir	-0.131	0.093	-1.409	0.159
arg11	-1.045	0.069	-15.102	0.000
arg10	-0.890	0.078	-11.443	0.000
arg00	-0.760	0.072	-10.496	0.000
v1	1.398	0.053	26.557	0.000
v0	1.398	0.053	26.557	0.000
probit11	-0.884	0.062	-14.195	0.000
probit10	-0.753	0.070	-10.727	0.000
probit00	-0.643	0.063	-10.189	0.000
indirect	-0.037	0.014	-2.645	0.008
direct	-0.035	0.024	-1.410	0.158
orind	0.796	0.066	12.037	0.000
ordir	0.829	0.111	7.454	0.000

Table 9: Pearl's hypothetical binary case (Source: Pearl, 2010, 2011)

Treatment X	Enzyme M	Percentage Cured Y = 1
1	1	$F_Y(1,1) = 80\%$
1	0	$F_Y(1,0) = 40\%$
0	1	$F_Y(0,1) = 30\%$
0	0	$F_Y(0,0) = 20\%$

Treatment X	Percentage M = 1
0	$F_M(0) = 40\%$
1	$F_M(1) = 75\%$

Table 10: Output for Pearl's hypothetical binary case using logit with ML, Step 2

	Estimates			S.E.	M.S.E.	95%	% Sig
	Population	Average	Std. Dev.	Average		Cover	Coeff
m ON							
x	1.504	1.5144	0.2193	0.2191	0.0481	0.964	1.000
y ON							
x	0.981	1.0020	0.3741	0.3745	0.1401	0.958	0.774
m	0.539	0.5405	0.3446	0.3399	0.1185	0.952	0.340
xm	1.253	1.2701	0.4816	0.4953	0.2318	0.968	0.750
Thresholds							
y$1	1.386	1.4085	0.2366	0.2315	0.0564	0.962	1.000
m$1	0.405	0.4136	0.1423	0.1449	0.0203	0.948	0.822
New/additional parameters							
fm0	0.400	0.3986	0.0338	0.0346	0.0011	0.940	1.000
fm1	0.750	0.7490	0.0318	0.0306	0.0010	0.938	1.000
fy00	0.200	0.1991	0.0363	0.0362	0.0013	0.950	1.000
fy10	0.400	0.4018	0.0692	0.0690	0.0048	0.954	1.000
fy01	0.300	0.2981	0.0489	0.0510	0.0024	0.954	1.000
fy11	0.800	0.8009	0.0312	0.0325	0.0010	0.956	1.000
de	0.320	0.3222	0.0543	0.0539	0.0030	0.944	1.000
pie	0.035	0.0348	0.0229	0.0227	0.0005	0.950	0.296
tie	0.140	0.1399	0.0329	0.0329	0.0011	0.940	1.000
te	0.460	0.4621	0.0435	0.0442	0.0019	0.950	1.000
iete	0.070	0.0761	0.0501	0.0505	0.0025	0.962	0.272
dete	0.696	0.6945	0.0778	0.0762	0.0060	0.938	1.000
compdete	0.304	0.3055	0.0778	0.0762	0.0060	0.938	1.000
tiete	0.304	0.3055	0.0778	0.0762	0.0060	0.938	1.000

Table 11: Output for Pearl's hypothetical binary case using probit with ML, Step 2

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
m ON							
x	0.929	0.9341	0.1321	0.1321	0.0174	0.962	1.000
y ON							
x	0.586	0.5973	0.2242	0.2244	0.0503	0.958	0.766
m	0.315	0.3148	0.2008	0.1990	0.0402	0.952	0.336
xm	0.779	0.7866	0.2857	0.2943	0.0815	0.968	0.794
Thresholds							
y$1	0.840	0.8506	0.1339	0.1315	0.0180	0.956	1.000
m$1	0.254	0.2588	0.0883	0.0899	0.0078	0.0946	0.824
New/additional parameters							
de	0.320	0.3216	0.0543	0.0539	0.0029	0.946	1.000
tie	0.140	0.1399	0.0329	0.0329	0.0011	0.938	1.000
pie	0.035	0.0347	0.0229	0.0227	0.0005	0.950	0.294
te	0.460	0.4615	0.0434	0.0442	0.0019	0.950	1.000
tiete	0.304	0.3060	0.0780	0.0764	0.0061	0.942	1.000
piete	0.070	0.0758	0.0501	0.0506	0.0025	0.964	0.272
dete	0.696	0.6940	0.0780	0.0764	0.0061	0.942	1.000
compdete	0.304	0.3060	0.0780	0.0764	0.0061	0.942	1.000
pfm0	0.400	0.3983	0.0339	0.0345	0.0011	0.940	1.000
pfm1	0.750	0.7492	0.0319	0.0306	0.0010	0.938	1.000
pfy00	0.200	0.1996	0.0364	0.0363	0.0013	0.950	1.000
pfy10	0.400	0.4017	0.0692	0.0690	0.0048	0.954	1.000
pfy01	0.300	0.2980	0.0488	0.0511	0.0024	0.956	1.000
pfy11	0.800	0.8003	0.0312	0.0326	0.0010	0.956	1.000

Table 12: Output for Pearl's hypothetical binary case using probit with Bayes, Step 2

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
m ON							
x	0.929	0.9334	0.1318	0.1310	0.0174	0.958	1.000
y ON							
x	0.586	0.5963	0.2204	0.2241	0.0486	0.958	0.772
m	0.315	0.3110	0.1976	0.1993	0.0390	0.954	0.330
xm	0.779	0.7916	0.2792	0.2919	0.0780	0.970	0.808
Thresholds							
y$1	0.840	0.8481	0.1320	0.1308	0.0175	0.952	1.000
m$1	0.254	0.2581	0.0881	0.0894	0.0078	0.946	0.824
New/additional parameters							
de	0.320	0.3208	0.0536	0.0537	0.0029	0.956	1.000
tie	0.140	0.1371	0.0323	0.0324	0.0011	0.946	1.000
pie	0.035	0.0334	0.0221	0.0227	0.0005	0.958	0.330
te	0.460	0.4598	0.0431	0.0441	0.0019	0.958	1.000
tiete	0.304	0.3027	0.0773	0.0770	0.0060	0.946	0.330
piete	0.070	0.0735	0.0488	0.0518	0.0024	0.956	1.000
dete	0.696	0.6972	0.0773	0.0770	0.0060	0.946	1.000
compdete	0.304	0.3027	0.0773	0.0770	0.0060	0.946	1.000
orde	4.030	4.2200	1.0343	1.1117	1.1036	0.950	1.000
ortie	1.833	1.8375	0.2559	0.2614	0.0654	0.954	1.000
pfm0	0.500	0.3986	0.0338	0.0342	0.0114	0.176	1.000
pfm1	0.500	0.7492	0.0319	0.0303	0.0631	0.000	1.000
pfy00	0.500	0.2002	0.0360	0.0361	0.0912	0.000	1.000
pfy10	0.500	0.4021	0.0688	0.0681	0.0143	0.712	1.000
pfy01	0.500	0.2974	0.0485	0.0507	0.0434	0.034	1.000
pfy11	0.500	0.8008	0.0312	0.0319	0.0915	0.000	1.000
numde	0.500	0.5614	0.0461	0.0453	0.0059	0.730	1.000
dende	0.500	0.2400	0.0288	0.0299	0.0684	0.000	1.000
numtie	0.500	0.7008	0.0319	0.0320	0.0413	0.000	1.000
dentie	0.500	0.5614	0.0461	0.0453	0.0059	0.730	1.000

Table 13: Pearl data n=200

X	M	Y		Total
		Not Cured	Cured	
Ctrl	Enzyme Absent	48	12	60
	Enzyme Present	28	12	40
Tx	Enzyme Absent	15	10	25
	Enzyme Present	15	60	75

Table 14: Output for n=200 data based on the Pearl example

Parameter	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%
m ON					
x	0.960	0.187	0.000	0.598	1.325
y ON					
x	0.596	0.314	0.031	-0.030	1.212
m	0.328	0.259	0.103	-0.179	0.843
xm	0.757	0.406	0.031	-0.030	1.553
Thresholds					
y$1	0.709	0.170	0.000	0.378	1.051
m$1	0.232	0.122	0.028	-0.005	0.469
New/additional parameters					
de	0.322	0.077	0.000	0.168	0.470
tie	0.131	0.046	0.000	0.051	0.231
pie	0.038	0.033	0.103	-0.021	0.111
te	0.456	0.061	0.000	0.332	0.569
tiete	0.288	0.113	0.000	0.109	0.547
piete	0.084	0.076	0.103	-0.047	0.258
dete	0.712	0.113	0.000	0.453	0.891
compdete	0.288	0.113	0.000	0.109	0.547
pfm0	0.408	0.047	0.000	0.319	0.502
pfm1	0.767	0.043	0.000	0.674	0.844
pfy00	0.239	0.052	0.000	0.147	0.353
pfy10	0.455	0.102	0.000	0.259	0.658
pfy01	0.351	0.071	0.000	0.221	0.497
pfy11	0.834	0.044	0.000	0.735	0.906
numde	0.609	0.066	0.000	0.477	0.736
dende	0.285	0.043	0.000	0.208	0.376
orde	3.908	1.508	0.000	2.024	7.852
numind	0.744	0.045	0.000	0.649	0.825
denind	0.609	0.066	0.000	0.477	0.736
orind	1.841	0.398	0.000	1.290	2.831

Table 15: Output for Monte Carlo simulation of a nominal mediator and a continuous outcome, Step 1

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
Latent class 1							
y ON							
x	-0.500	-0.4884	0.2647	0.2461	0.0701	0.936	0.546
Intercepts							
y	-2.000	-2.0254	0.2186	0.2050	0.0483	0.946	0.998
Residual variances							
y	0.750	0.7420	0.0776	0.0739	0.0061	0.920	1.000
Latent class 2							
y ON							
x	-0.300	-0.3037	0.3664	0.3472	0.1340	0.934	0.180
Intercepts							
y	0.000	0.0107	0.2900	0.2651	0.0840	0.918	0.082
Residual variances							
y	0.750	0.7420	0.0776	0.0739	0.0061	0.920	1.000
Latent class 3							
y ON							
x	-0.200	-0.2000	0.1675	0.1609	0.0280	0.938	0.254
Intercepts							
y	2.000	2.0155	0.1260	0.1173	0.0161	0.938	1.000
Residual variances							
y	0.750	0.7420	0.0776	0.739	0.0061	0.920	1.000
Categorical latent variables							
c#1 ON							
x	0.700	0.7059	0.4183	0.3374	0.1746	0.950	0.526
c#2 ON							
x	0.300	0.2761	0.3466	0.3321	0.1205	0.944	0.134
Intercepts							
c#1	-1.000	-1.0041	0.3520	0.3067	0.1237	0.956	0.900
c#2	-0.500	-0.4559	0.2599	0.2513	0.0694	0.956	0.512

Table 16: Output for Monte Carlo simulation of a nominal mediator and a continuous outcome, Step 2, part 1

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
Latent class 1							
y ON							
x	-0.500	-0.5045	0.1332	0.1285	0.0177	0.944	0.972
Intercepts							
y	-2.000	-2.0007	0.1011	0.1001	0.0102	0.958	1.000
Residual variances							
y	0.750	0.7465	0.0360	0.0373	0.0013	0.954	1.000
Latent class 2							
y ON							
x	-0.300	-0.2976	0.1125	0.1093	0.0126	0.942	0.772
Intercepts							
y	0.000	0.0021	0.0799	0.0780	0.0064	0.944	0.056
Thresholds							
Residual variances							
y	0.750	0.7465	0.0360	0.0373	0.0013	0.954	1.000
Latent class 3							
y ON							
x	-0.200	-0.1948	0.0917	0.0923	0.0084	0.954	0.554
Intercepts							
y	2.000	2.0002	0.0629	0.0609	0.0039	0.936	1.000
Residual variances							
y	0.750	0.7465	0.0360	0.0373	0.0013	0.954	1.000

Table 17: Output for Monte Carlo simulation of a nominal mediator and a continuous outcome, Step 2, part 2

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
Categorical latent variables							
c#1 ON							
x	0.700	0.6916	0.1667	0.1832	0.0278	0.966	0.982
c#2 ON							
x	0.300	0.2982	0.1693	0.1656	0.0286	0.946	0.426
Intercepts							
c#1	-1.000	-0.9920	0.1233	0.1357	0.0152	0.962	1.000
c#2	-0.500	-0.4950	0.1142	0.1146	0.0130	0.966	0.998
New/additional parameters							
denom0	1.974	1.9872	0.0950	0.0989	0.0092	0.964	1.000
denom1	2.559	2.5729	0.1614	0.1617	0.0262	0.966	1.000
p10	0.186	0.1877	0.0178	0.0195	0.0003	0.970	1.000
p11	0.289	0.2892	0.0216	0.0226	0.0005	0.964	1.000
p20	0.307	0.3080	0.0230	0.0231	0.0005	0.954	1.000
p21	0.320	0.3207	0.0233	0.0233	0.0005	0.960	1.000
p30	0.507	0.5044	0.0240	0.0250	0.0006	0.968	1.000
p31	0.391	0.3902	0.0241	0.0244	0.0006	0.962	1.000
term11	-0.116	-0.1148	0.0936	0.0981	0.0088	0.960	0.214
term10	0.354	0.3494	0.0944	0.0940	0.0089	0.952	0.956
term01	0.203	0.2028	0.0906	0.0934	0.0082	0.956	0.592
term00	0.640	0.6340	0.0850	0.0882	0.0072	0.960	1.000
de	-0.287	-0.2846	0.0640	0.0627	0.0041	0.928	0.992
tie	-0.470	-0.4642	0.1114	0.1213	0.0124	0.958	0.974
total	-0.757	-0.7488	0.1196	0.1319	0.0143	0.980	1.000
pie	-0.438	-0.4312	0.1040	0.1131	0.0108	0.966	0.972

Table 18: Hypothetical pollution data with a nominal mediator and a binary outcome

X	M	Y			Total
		0	1	%	
Ctrl	1	30	30	50	60
	2	20	60	75	80
	3	20	80	70	100
Tx	1	50	30	38	80
	2	40	60	60	100
	3	20	40	68	60

Table 19: Output for hypothetical pollution data with a nominal mediator and a binary outcome, part 1

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-value
Latent class 1				
y ON				
x	-0.511	0.346	-1.475	0.140
Thresholds				
y$1	0.000	0.258	0.000	1.000
Latent class 2				
y ON				
x	-0.693	0.329	-2.106	0.035
Thresholds				
y$1	-1.099	0.258	-4.255	0.000
Latent class 3				
y ON				
x	-0.693	0.371	-1.869	0.062
Thresholds				
y$1	-1.386	0.250	-5.545	0.000

Table 20: Output for hypothetical pollution data with a nominal mediator and a binary outcome, part 2

Parameter	Estimates	S.E.	Est./S.E.	Two-Tailed P-value
Categorical latent variables				
c#1 ON				
x	0.799	0.236	3.379	0.001
c#2 ON				
x	0.734	0.222	3.310	0.001
Intercepts				
c#1	-0.511	0.163	-3.128	0.002
c#2	-0.223	0.150	-1.488	0.137
New/additional parameters				
denom0	2.400	0.183	13.093	0.000
denom1	4.000	0.447	8.944	0.000
p10	0.250	0.028	8.944	0.000
p11	0.333	0.030	10.954	0.000
p20	0.333	0.030	10.954	0.000
p21	0.417	0.032	13.093	0.000
p30	0.417	0.032	13.093	0.029
p31	0.250	0.028	8.944	0.000
term11	0.542	0.032	16.842	0.000
term10	0.572	0.034	16.855	0.000
term01	0.679	0.032	21.077	0.000
term00	0.708	0.029	24.142	0.000
de	-0.137	0.043	-3.145	0.002
tie	-0.030	0.016	-1.860	0.063
total	-0.167	0.044	-3.828	0.000
pie	-0.029	0.015	-1.965	0.049
orde	0.549	0.106	5.199	0.000
ortie	0.886	0.058	15.306	0.000
orpie	0.872	0.060	14.517	0.000

Table 21: Output for mediation modeling with a count outcome, Step 2

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
y ON							
x	0.300	0.3042	0.1743	0.1691	0.0303	0.936	0.432
m	0.400	0.4051	0.1042	0.1036	0.0109	0.946	0.964
xm	0.200	0.2004	0.1258	0.1251	0.0158	0.952	0.394
m ON							
x	0.500	0.5016	0.0852	0.0863	0.0072	0.954	1.000
Intercepts							
m	0.500	0.4999	0.0612	0.0611	0.0037	0.948	1.000
u	-0.700	-0.7123	0.1226	0.1213	0.0152	0.956	1.000
Residual variances							
m	0.750	0.7431	0.0490	0.0525	0.0024	0.960	1.000
New/additional parameters							
ind	0.450	0.3036	0.0608	0.0632	0.0251	0.374	1.000
dir	0.400	0.4047	0.1323	0.1308	0.0175	0.942	0.860
ey1	0.670	0.6693	0.0759	0.0783	0.0057	0.952	1.000
ey0	0.497	0.4942	0.0600	0.0595	0.0036	0.956	1.000
mum1	1.000	1.0015	0.0639	0.0609	0.0041	0.936	1.000
mum0	0.500	0.4999	0.0612	0.0611	0.0037	0.948	1.000
ay1	0.900	0.8955	0.1111	0.1216	0.0123	0.960	1.000
ay0	0.600	0.6011	0.1571	0.1597	0.0246	0.956	0.958
bym11	1.450	1.4509	0.0628	0.0671	0.0039	0.962	1.000
bym10	1.900	1.9130	0.1582	0.1695	0.0251	0.964	1.000
bym01	1.300	1.3012	0.0807	0.0823	0.0065	0.956	1.000
bym00	1.600	1.6113	0.1834	0.1810	0.0337	0.950	1.000
eym11	2.086	2.1154	0.2193	0.2299	0.0489	0.956	1.000
eym10	1.545	1.5575	0.1154	0.1199	0.0134	0.960	1.000
eym01	1.584	1.6165	0.2251	0.2244	0.0516	0.952	1.000
eym00	1.297	1.3108	0.1120	0.1148	0.0127	0.946	1.000
tie	0.336	0.3668	0.0756	0.0773	0.0066	0.956	1.000
de	0.392	0.3930	0.0945	0.0948	0.0089	0.944	0.988
total	0.754	0.7598	0.1166	0.1156	0.0136	0.952	1.000
pie	0.143	0.1462	0.0508	0.0505	0.0026	0.942	0.942

Table 22: Output for Monte Carlo simulation, analyzing by M and Y regressed on X only

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
y ON							
x	0.000	0.6545	0.0877	0.0866	0.4360	0.000	1.000
m ON							
x	0.500	0.4995	0.1033	0.0998	0.0107	0.952	1.000
y WITH							
m	0.500	0.4978	0.0512	0.0498	0.0026	0.942	1.000
Intercepts							
y	0.000	2.0014	0.0637	0.0611	4.0098	0.000	1.000
m	2.000	2.0000	0.0751	0.0705	0.0056	0.942	1.000
Residual variances							
y	0.750	0.7486	0.0515	0.0529	0.0027	0.952	1.000
m	1.000	0.9956	0.0714	0.0704	0.0051	0.952	1.000
New/additional parameters							
rhocurl	0.577	0.5760	0.0345	0.0334	0.0012	0.938	1.000
beta1	0.500	0.5000	0.0366	0.0354	0.0013	0.928	1.000
beta2	0.400	0.4049	0.0730	0.0729	0.0053	0.938	1.000
beta0	1.000	1.0014	0.0897	0.0867	0.0080	0.940	1.000
sig1	0.500	0.4984	0.0338	0.0352	0.0011	0.950	1.000
ind	0.250	0.2495	0.0543	0.0531	0.0029	0.952	1.000
de	0.400	0.4049	0.0730	0.0729	0.0053	0.938	1.000

Table 23: Output for generating data with true residual correlation 0.25 and analyzing data with Imai's ρ fixed at the true value 0.25

	Estimates			S.E.	M.S.E.	95%	% Sig
Parameter	Population	Average	Std. Dev.	Average		Cover	Coeff
y ON							
x	0.000	0.6551	0.0975	0.0962	0.4386	0.000	1.000
m ON							
x	0.500	0.5007	0.1033	0.0998	0.0107	0.956	1.000
y WITH							
m	0.854	0.6743	0.0597	0.0586	0.0357	0.170	1.000
Intercepts							
y	0.000	2.0016	0.0708	0.0679	4.0112	0.000	1.000
m	2.000	2.0003	0.0752	0.0705	0.0056	0.938	1.000
Residual variances							
y	1.104	0.9251	0.0637	0.0654	0.0359	0.232	1.000
m	1.000	0.9957	0.0714	0.0704	0.0051	0.958	1.000
New/additional parameters							
rho	0.250	0.2500	0.0000	0.0000	0.0000	0.000	1.000
rhocurl	0.812	0.7021	0.0262	0.0253	0.0129	0.002	1.000
beta1	0.500	0.5001	0.0366	0.0354	0.0013	0.928	1.000
beta2	0.400	0.4049	0.0732	0.0729	0.0054	0.938	1.000
beta0	1.000	1.0011	0.0892	0.0867	0.0079	0.944	1.000
sig1	0.707	0.3528	0.0121	0.0125	0.1257	0.000	1.000
ind	0.250	0.2502	0.0544	0.0532	0.0030	0.952	1.000
de	0.400	0.4049	0.0732	0.0729	0.0054	0.938	1.000

Table 24: Output for head circumference analysis using the Imai et al. sensitivity approach with $\rho = 0$

Parameter	Estimate	S.E.	Est./S.E.	Two-Tailed P-value
hcirc36 ON				
alccig	-0.079	0.115	-0.684	0.494
gender	0.697	0.082	8.467	0.000
eth	0.090	0.083	1.093	0.274
hcirc0 ON				
alccig	-0.366	0.108	-3.384	0.001
gender	0.345	0.079	4.363	0.000
eth	0.368	0.079	4.641	0.000
hcirc36 WITH				
hcirc0	0.408	0.044	9.304	0.000
Intercepts				
hcirc0	-0.301	0.071	-4.264	0.000
hcirc36	-0.400	0.073	-5.477	0.000
Residual variances				
hcirc0	0.919	0.054	17.108	0.000
hcirc36	0.878	0.056	15.797	0.000
New/additional parameters				
rho	0.000	0.000	0.000	1.000
rhocurl	0.454	0.036	12.566	0.000
beta1	0.444	0.040	11.074	0.000
beta2	0.084	0.106	0.790	0.429
beta0	-0.266	0.067	-3.983	0.000
sig1	0.000	0.000	0.000	1.000
indirect	-0.162	0.050	-3.239	0.001
direct	0.084	0.106	0.790	0.429