**Supplement File**

*Accommodating latent XM interactions in statistical mediation analysis*

1. Derivation of effects

2. Implied covariance matrix for a single mediator model with an XM interaction

3. Derivation of the rescaled true values

4. Effect size differences for true models under summed scores and factor scores

5. Supplemental figures and tables
6. R Code to carry out the procedures

**1. Derivation of effects**

Below, we demonstrate how to derive the causal estimators using the equations from the mediation model.

**Assuming x = 1 and x\* = 0**

**Assuming x = 0.5 and x\* = -0.5**

**Mediated interaction**

**Assuming x = 1 and x\* = 0**

**Assuming x = .5 and x\* = -.5**

**2. Implied covariance of the mediation model with linear effects**

Consider the following model, where C acts as a confounder of the M-Y relation:

The implied covariance matrix of this model is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *X* | *M* | *Y* | *XM* | *C* |
| *X* |  |  |  |  |  |
| *M* |  |  |  |  |  |
| *Y* |  |  |  |  |  |
| *XM* |  |  |  |  |  |
| *C* |  |  |  |  |  |

However, for a situation in which we mean-center X and M so that and are zero, and X is randomized so that , the covariance algebra of the model reduces to:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *X* | *M* | *Y* | *XM* | *C* |
| *X* |  |  |  |  |  |
| *M* |  |  |  |  |  |
| *Y* |  |  |  |  |  |
| *XM* |  |  |  |  |  |
| *C* |  |  |  |  |  |

As such, the true covariance matrix of the model can be estimated by plugging in the true variances of the variables in the mediation model and the paths. Also, note that XM does not correlate with X, M, or C. For our case, XM only correlates with Y, which is convenient for the derivation of the expected parameter values that different models will recover.

**3. Derivation of the rescaled true values for summed scores and factor scores**

As mentioned in the text, the M estimate with summed scores and the M estimate with factor scores have different moments than the data-generating values. Therefore, for the computation of bias, the estimates in the summed score model, factor score model, corrected summed score model, and corrected factor score model must be rescaled. In essence, the rescaling entails obtaining the true covariance matrix that includes the rescaled variables, and then fit the mediation model to that covariance matrix. A convenient relation from the implied relations of the model in part 2 of the supplement is that, if X and M are centered, then XM only correlates with Y.

Suppose that the items that assess M fits a unidimensional factor model:

where Ψ is a diagonal matrix of residual variances. Then, M is part of a mediation model:

We can obtain the implied covariance matrix of X, Y, C, and the items that assess M by using the following expression:

where *I* is an identity matrix, *B* is a matrix of regression coefficients, and is a diagonal matrix containing the variances for exogenous variables and residual variances for endogenous variables in a model with X, M, Y, and C. Also, is an augmented matrix of factor loadings, where each column represents X, M, Y, and C and the rows are X, items that assess M, Y, and C. X, Y, and C are treated as single-indicator latent variables, so entries of for those variables have a value of 1 where the column match the rows. For M, the factor loadings are specified in their respective column, and there are zeros elsewhere. Finally, is an augmented diagonal matrix of residual variances for X, items that assess M, Y, and C. The entries of have the residual variances of the items that assess M, and zero in the rest of the entries.

Using this model, we can use the implied variances and covariances of the summed scores and factor scores.

* For summed score of M: **,** where **1** is a summing matrix which indicates which entries of the covariance matrix should be summed. On the rows are X, items that assess M, Y, and C, and the columns are X, summed M, Y, and C. For example, if M was assessed with three items, the summing matrix would be:



* For factor score of M: **,** where is a matrix of the Bartlett weights. The matrix is like the summing matrix, but the entries corresponding to and the items are replaced by the Bartlett weights (discussed in text).
* For the corrected summed score of M: , but the variance of the summed score for M is corrected for unreliability, so it is substituted by , where is an estimate of the reliability of the summed score.
* For the corrected factor score of M: , but the variance of the factor score for M is corrected by unreliability, so it is substituted by , where is an estimate of the reliability of the factor score.

Lastly, we need to complete , , , and with the relations among the variables and XM. As such, we manually add another row and another column with the expected relations of XM under each of these covariance matrices – zero covariances with X, M and C, a variance of XM obtained with Eq. 6 from the manuscript, and a transformed covariance of Y. In this case, the transformed covariance is determined by computing the true correlation between XM and Y and multiplying that covariance by the standard deviations of both Y and the standard deviation of the M estimate (e.g., factor score, summed score, corrected factor score, and corrected summed score).

Now that , , , and are 5x5 covariance matrices containing the relation between X, estimate M, Y, C, and XM, we can fit mediation models to , , , and , and the estimated regression coefficients will be the rescaled true values for each of these models. Note that XM is simply treated as another variable in the model.

**4. Effect size differences for true models.**

Below are approximate partial R2 for some of the paths for the single mediator model with an XM latent interaction with true M, summed scores for M, and factor scores for M. Estimates were obtained by taking the true covariance matrices between the variables (M as a summed score or M as a factor score, as shown in part 3 of the supplement) and estimating the partial correlation matrix using the cor2pcor function in the corpcor R-package. For true M, with a variance close to 1, a path of .14 accounts for 2% of variance, and a path of .39 accounts for 13% of the variance. Variables X and XM have a variance approximately to .25, so for an a-path or h-path = .14, the variables accounted for .5% of the variance, which is 2%/4. For an a-path of .39, the variables accounted for 3.6% of the variance, which is close to 13%/4. So, the effect sizes roughly match the Cohen effect sizes proposed in the text. For models with unreliable M, the effect sizes are smaller than the data-generating values as expected. Therefore, the summed score mediation model and the factor score mediation model estimate recover these attenuated effect sizes, not the data generating effect sizes.

**True M (reliability = 1):**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| g=0, a = .14 | .005 | .000 |
| g=0, a = .39 | .036 | .000 |
| g=.2, a = .14 | .005 | .038 |
| g=.2, a = .39 | .036 | .038 |

a=.14, g=.2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .019 | .005 | .037 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .019 | .000 | .037 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .019 | .005 | .037 |
| f=.2, cp = .14, b=.39, h=-.14 | .005 | .132 | .005 | .037 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .132 | .000 | .037 |
| f=.2, cp = .14, b=.39, h=.14 | .005 | .132 | .005 | .037 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .019 | .005 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .019 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .019 | .005 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .005 | .132 | .005 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .132 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .005 | .132 | .005 | .000 |

a=.14, g=0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .019 | .005 | .039 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .019 | .000 | .038 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .019 | .005 | .039 |
| f=.2, cp = .14, b=.39, h=-.14 | .005 | .132 | .005 | .039 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .132 | .000 | .038 |
| f=.2, cp = .14, b=.39, h=.14 | .005 | .132 | .005 | .039 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .019 | .005 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .019 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .019 | .005 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .005 | .132 | .005 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .132 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .005 | .132 | .005 | .000 |

**Summed scores, six items (reliability = .687):**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| g=0, a = .14 | .003 | .000 |
| g=0, a = .39 | .026 | .000 |
| g=.2, a = .14 | .003 | .026 |
| g=.2, a = .39 | .026 | .027 |

a=.14, g=.2 (note, for true model c’-path effect was .0049, and for summed score is a bit larger than .005)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .040 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .040 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .040 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .091 | .003 | .045 |
| f=.2, cp = .14, b=.39, h=0 | .006 | .091 | .000 | .045 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .091 | .003 | .045 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .091 | .003 | .001 |
| f=0, cp = .14, b=.39, h=0 | .006 | .091 | .000 | .001 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .091 | .003 | .001 |

a=.14, g=0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .038 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .038 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .038 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .091 | .003 | .037 |
| f=.2, cp = .14, b=.39, h=0 | .006 | .091 | .000 | .037 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .091 | .003 | .037 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .091 | .003 | .000 |
| f=0, cp = .14, b=.39, h=0 | .006 | .091 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .091 | .003 | .000 |

**Summed scores, nine items (reliability = .781):**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| g=0, a = .14 | .004 | .000 |
| g=0, a = .39 | .029 | .000 |
| g=.2, a = .14 | .004 | .030 |
| g=.2, a = .39 | .029 | .030 |

a=.14, g=.2 (note, for true model c’-path effect was .0049, and for summed score is a bit larger than .005)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .040 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .040 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .040 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .104 | .004 | .042 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .103 | .000 | .042 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .104 | .004 | .042 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .104 | .004 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .103 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .104 | .004 | .000 |

a=.14, g=0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .038 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .038 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .038 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .104 | .004 | .037 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .103 | .000 | .038 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .104 | .004 | .037 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .104 | .004 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .103 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .104 | .004 | .000 |

**Factor scores, six items (reliability = .704):**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| g=0, a = .14 | .003 | .000 |
| g=0, a = .39 | .026 | .000 |
| g=.2, a = .14 | .003 | .027 |
| g=.2, a = .39 | .026 | .027 |

a=.14, g=.2 (note, for true model c’-path effect was .0049, and for summed score is a bit larger than .005)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .014 | .004 | .040 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .014 | .000 | .040 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .014 | .004 | .040 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .093 | .003 | .045 |
| f=.2, cp = .14, b=.39, h=0 | .006 | .093 | .000 | .044 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .093 | .003 | .045 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .014 | .004 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .014 | .004 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .093 | .003 | .000 |
| f=0, cp = .14, b=.39, h=0 | .006 | .093 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .093 | .003 | .000 |

a=.14, g=0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .038 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .038 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .038 |
| f=.2, cp = .14, b=.39, h=-.14 | .006 | .093 | .003 | .037 |
| f=.2, cp = .14, b=.39, h=0 | .006 | .093 | .000 | .037 |
| f=.2, cp = .14, b=.39, h=.14 | .006 | .093 | .003 | .037 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .013 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .013 | .003 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .006 | .093 | .003 | .000 |
| f=0, cp = .14, b=.39, h=0 | .006 | .093 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .006 | .093 | .003 | .000 |

**Factor scores, nine items (reliability = .802):**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| g=0, a = .14 | .004 | .000 |
| g=0, a = .39 | .030 | .000 |
| g=.2, a = .14 | .004 | .030 |
| g=.2, a = .39 | .030 | .030 |

a=.14, g=.2 (note, for true model c’-path effect was .0049, and for summed score is a bit larger than .005)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .040 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .039 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .040 |
| f=.2, cp = .14, b=.39, h=-.14 | .005 | .106 | .004 | .040 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .105 | .000 | .042 |
| f=.2, cp = .14, b=.39, h=.14 | .005 | .106 | .004 | .040 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .005 | .106 | .004 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .105 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .005 | .106 | .004 | .000 |

a=.14, g=0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| f=.2, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .038 |
| f=.2, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .038 |
| f=.2, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .038 |
| f=.2, cp = .14, b=.39, h=-.14 | .005 | .106 | .004 | .038 |
| f=.2, cp = .14, b=.39, h=0 | .005 | .105 | .000 | .037 |
| f=.2, cp = .14, b=.39, h=.14 | .005 | .106 | .004 | .038 |
| f=0, cp = .14, b=.14, h=-.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.14, h=0 | .005 | .015 | .000 | .000 |
| f=0, cp = .14, b=.14, h=.14 | .005 | .015 | .004 | .000 |
| f=0, cp = .14, b=.39, h=-.14 | .005 | .106 | .004 | .000 |
| f=0, cp = .14, b=.39, h=0 | .005 | .105 | .000 | .000 |
| f=0, cp = .14, b=.39, h=.14 | .005 | .106 | .004 | .000 |

5. Supplemental figures and tables

****

**Figure S1**. Regression trees to predict statistical power in the parameter as a function of simulation factors and the approach to handle the XM interaction. Note: Ni is sample size, a is for the a-path, b is for the b-path, h is for the h-path, c01 is the confounding effect size, where c01=0 if f = g = 0 and f = g = .20 otherwise.

|  |
| --- |
| Table S1. *Average raw bias of parameter estimates across conditions for each approach to handle a latent XM interaction in the single mediator model* |
|  | *Normally distributed continuous indicators* |
| model | a-path | b-path | c’-path | pnie |  pnde | tnie | tnde |  g-path |  f-path |  h-path | ah |
| summed scores | -0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| factor scores | -0.002 | 0.000 | -0.001 | 0.001 | 0.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.001 | 0.000 |
| summed scores (c) | -0.001 | 0.000 | 0.001 | 0.000 | -0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| factor scores (c) | -0.002 | 0.000 | -0.001 | 0.001 | 0.000 | 0.000 | -0.001 | 0.001 | -0.001 | 0.001 | 0.000 |
| UPI | 0.002 | 0.001 | -0.001 | 0.000 | -0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Multiple group | 0.004 | 0.002 | -0.001 | 0.000 | -0.001 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| LMS | 0.001 | 0.001 | -0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Bayesian approach | -0.020 | -0.008 | -0.014 | 0.001 | 0.000 | -0.001 | -0.001 | 0.001 | -0.001 | 0.001 | 0.000 |
|  | *Nonnormally distributed discrete indicators* |
| summed scores | -0.092 | 0.000 | 0.004 | -0.003 | 0.006 | -0.006 | 0.003 | -0.039 | 0.002 | -0.002 | -0.003 |
| factor scores | -0.005 | -0.015 | 0.005 | -0.003 | 0.007 | -0.006 | 0.003 | -0.003 | 0.002 | -0.009 | -0.003 |
| summed scores (c) | -0.092 | 0.006 | 0.001 | 0.004 | 0.006 | -0.005 | -0.004 | -0.039 | 0.001 | -0.007 | -0.010 |
| factor scores (c) | -0.005 | -0.006 | 0.003 | 0.002 | 0.007 | -0.007 | -0.002 | -0.003 | 0.001 | -0.025 | -0.009 |
| UPI | -0.001 | -0.006 | 0.002 | 0.000 | 0.004 | -0.004 | 0.000 | -0.001 | 0.001 | -0.012 | -0.004 |
| Multiple group | 0.014 | -0.013 | 0.002 | 0.001 | 0.005 | -0.004 | -0.001 | 0.002 | 0.001 | -0.013 | -0.006 |
| LMS | -0.002 | -0.006 | 0.002 | 0.000 | 0.004 | -0.004 | 0.000 | -0.002 | 0.001 | -0.012 | -0.004 |
| Bayesian approach | -0.008 | -0.002 | 0.003 | -0.001 | 0.005 | -0.005 | 0.001 | -0.005 | 0.002 | -0.013 | -0.004 |
| Note: (c) is for corrected, UPI is for unconstrained product indicator approach, LMS is for latent moderated structural equations. |

|  |
| --- |
| Table S2. *Average interval coverage of the parameters across conditions for each approach to handle a latent XM interaction in the single mediator model* |
| model | a-path | g-path | b-path | c’-path | f-path | h-path | pnie | pnde | tnie | tnde | ah: a=.14 | ah: a=.39 |
| summed scores | 0.947 | 0.944 | 0.943 | 0.946 | 0.944 | 0.945 | 0.950 | 0.947 | 0.948 | 0.946 | **0.974** | 0.947 |
| factor scores | 0.944 | 0.941 | 0.942 | 0.946 | 0.944 | 0.943 | 0.949 | 0.947 | 0.945 | 0.947 | **0.985** | 0.946 |
| summed scores (c) | 0.950 | 0.947 | 0.937 | 0.946 | 0.944 | 0.946 | 0.951 | 0.947 | 0.947 | 0.947 | **0.986** | 0.947 |
| factor scores (c) | 0.944 | 0.941 | 0.937 | 0.947 | 0.944 | 0.943 | 0.953 | 0.949 | 0.949 | 0.948 | **0.986** | 0.947 |
| UPI | 0.950 | 0.947 | 0.946 | 0.947 | 0.944 | 0.947 | 0.954 | 0.948 | 0.949 | 0.948 | **0.987** | 0.947 |
| Multiple group | 0.944 | 0.941 | 0.940 | 0.947 | 0.944 | 0.943 | 0.951 | 0.948 | 0.949 | 0.948 | **0.986** | 0.947 |
| LMS | 0.950 | 0.946 | 0.945 | 0.947 | 0.944 | 0.946 | 0.955 | 0.948 | 0.952 | 0.948 | **0.988** | 0.950 |
| Bayesian approach | 0.949 | 0.947 | 0.942 | 0.947 | 0.945 | 0.947 | 0.955 | 0.948 | 0.952 | 0.948 | **0.987** | 0.951 |
| Note: in bold are coverage values outside .925 and .975. (c) is for corrected, UPI is for unconstrained product indicator approach, LMS is for latent moderated equations, pnie is for pure natural indirect effect, pnde is for the pure natural direct effect, tnie is for the total natural indirect effect, tnde is for the total natural direct effect, and ah is the mediated interaction |

|  |
| --- |
| Table S3. Median estimates and 95% confidence intervals for the ATLAS illustration with discrete, nonnormal indicators |
| model | a-path | g-path | c’-path | b-path | h-path | f-path |
| summed scores | **1.714 [1.186, 2.246]** | -0.100 [-0.343, 0.184] | **0.191 [0.077, 0.297]** | **0.069 [0.057, 0.082]** | **-0.030 [-0.052, -0.003]** | 0.035 [-0.009, 0.094] |
| factor scores | **0.399 [0.271, 0.521]** | -0.026 [-0.082, 0.040] | **0.191 [0.077, 0.299]** | **0.297 [0.246, 0.352]** | **-0.128 [-0.221, -0.010]** | 0.036 [-0.008, 0.094] |
| summed scores (c) | **1.714 [1.186, 2.246]** | -0.100 [-0.343, 0.184] | **0.146 [0.024, 0.261]** | **0.096 [0.079, 0.114]** | -0.036 [-0.070, 0.002] | 0.038 [-0.007, 0.097] |
| factor scores (c) | **0.399 [0.271, 0.521]** | -0.026 [-0.082, 0.040] | **0.149 [0.030, 0.262]** | **0.406 [0.339, 0.483]** | -0.151 [-0.286, 0.007] | 0.038 [-0.006, 0.098] |
| UPI | **0.408 [0.275, 0.541]** | -0.026 [-0.083, 0.041] | **0.151 [0.028, 0.264]** | **0.399 [0.334, 0.473]** | **-0.173 [-0.300, -0.018**] | 0.039 [-0.006, 0.099] |
| Multiple group | **0.388 [0.265, 0.516]** | -0.031 [-0.089, 0.033] | **0.145 [0.029, 0.258]** | **0.408 [0.333, 0.488]** | -0.145 [-0.277, 0.029] | 0.033 [-0.010, 0.090] |
| LMS | **0.408 [0.275, 0.540]** | -0.024 [-0.082, 0.042] | **0.148 [0.320, 0.260]** | **0.406 [0.339, 0.478]** | **-0.216 [-0.336, -0.087]** | 0.037 [-0.007, 0.097] |
| Bayesian approach | **0.396 [0.260, 0.532]** | -0.023 [-0.082, 0.038] | **0.152 [0.028, 0.274]** | **0.407 [0.337, 0.474]** | **-0.215 [-0.349, -0.081]** | 0.038 [-0.015, 0.089] |
|  |  |  |  |  |  |  |
| model | pnie | pnde | tnie | tnde | ah |  |
| summed scores | **0.139 [0.092, 0.196]** | **0.212 [0.091, 0.327]** | **0.088 [0.051, 0.137]** | **0.162 [0.051, 0.270]** | **-0.050 [-0.101, -0.006]** |  |
| factor scores | **0.138 [0.091, 0.195]** | **0.211 [0.092, 0.328]** | **0.088 [0.051, 0.134]** | **0.163 [0.049, 0.272]** | **-0.049 [-0.101, -0.004]** |  |
| summed scores (c) | **0.191 [0.125, 0.268]** | **0.174 [0.049, 0.296]** | **0.128 [0.074, 0.200]** | 0.111 [-0.015, 0.231] | -0.060 [-0.133, 0.003] |  |
| factor scores (c) | **0.185 [0.122, 0.262]** | **0.176 [0.046, 0.297]** | **0.126 [0.073, 0.192]** | 0.115 [-0.015, 0.233] | -0.057 [-0.131, 0.003] |  |
| UPI | **0.192 [0.123, 0.270]** | **0.178 [0.057, 0.305]** | **0.121 [0.070, 0.186]** | 0.112 [-0.018, 0.231] | **-0.067 [-0.144, -0.007]** |  |
| Multiple group | **0.180 [0.117, 0.250]** | **0.171 [0.042, 0.290]** | **0.126 [0.069, 0.197]** | 0.114 [-0.013, 0.231] | -0.052 [-0.123, 0.010] |  |
| LMS | **0.202 [0.131, 0.282]** | **0.188 [0.060, 0.311]** | **0.114 [0.064, 0.170]** | 0.100 [-0.023, 0.218] | **-0.086 [-0.161, -0.030]** |  |
| Bayesian approach | **0.197 [0.126, 0.279]** | **0.190 [0.063, 0.312]** | **0.112 [0.064, 0.172]** | 0.105 [-0.027, 0.230] | **-0.084 [-0.154, -0.028]** |  |
| Note: in bold are estimates in which the 95% confidence intervals did not contain zero. (c) is for corrected, UPI is for unconstrained product indicator approach, LMS is for latent moderated equations. |

|  |
| --- |
| Table S4. Summary of recommendations |
| Approach | Distributional assumptions | Pros | Cons |
| summed scores | If outcomes are nonnormal, standard errors might be underestimated. | - easily estimated in any OLS program- the method is familiar to applied researchers | - does not recover the true effects- may not be robust to nonnormal data |
| factor scores | Multivariate normality if the psychometric model is estimated with maximum likelihood | - easily estimated in any OLS program | - does not recover the true effects - one needs to determine the psychometric structure of the mediator |
| corrected summed scores | Multivariate normality if estimated with maximum likelihood | - easily estimated in any SEM program | - covariance matrix must be manually adjusted, which is cumbersome- covariance matrix is not guaranteed to be positive definite- there are many reliability estimators |
| corrected factor scores | Multivariate normality if estimated with maximum likelihood | - easily estimated in any SEM program | - covariance matrix must be manually adjusted, which is cumbersome- covariance matrix is not guaranteed to be positive definite- one needs to determine the psychometric structure of the mediator |
| multiple-group model | Multivariate normality if estimated with maximum likelihood | - easily estimated in any SEM program | - the moderator has to be categorical- one needs to determine the psychometric structure of the mediator- post-processing computations to recover the mediation paths- difficult to scale to more mediators  |
| Unconstrained product indicator approach | Multivariate normality if estimated with maximum likelihood, but products of variables are not multivariate normal | - easily estimated in any SEM program- easily scales to more latent variables | - there are many ways to build the product indicators- product of variables are not multivariate normal |
| Latent moderated structural equations | Multivariate normality | - streamlined implementation- easily scales to more latent variables | - limited software availability (mostly available in Mplus) |
| Bayesian mediation | Data matches model assumptions | - easily estimated in any program with MCMC capabilities- scalable to more latent variables | - Researchers need to be familiar with Bayesian mediation |

6. R code to conduct the procedures

## Estimating latent XM interactions in mediation analysis

Below, we provide code to conduct eight different approaches to accommodate latent XM interactions in statistical mediation analysis. First, we start by simulating a dataset to demonstrate the procedures. Next, we present four approaches that accommodate the latent XM interaction by scoring M, and then estimating an observed XM interaction. Then, we present four approaches that accommodate the XM interaction directly or indirectly using structural models. Lastly, we present some helper functions to simulate data.

Note that this supplement is mainly about model estimation. The causal effects can be calculated using the parameter estimates below and the equations that we show on the text. Standard errors for model parameters and the causal effects could be estimated using the 95% percentile bootstrap confidence intervals, although this is not shown below.

## Data simulation

Below, we simulate from a single mediator model with an XM interaction to demonstrate the code. We simulate a binary X, six continuous items that assess M, Y, and C.

# true parameters

a=.39 # a-path
b=.39 # b-path
cp=.14 # c'-path
h=.14 # h-path
f=.2 # f-path
g=.2 # g-path
n1=500 #sample size
items=6 # number of items

#variances and residual variances
varx=.25
varm=1
vary=1
varc=1

## sim data
set.seed(1000)
X0=rnorm(n1,0,1) #X variable from normal distribution
X=ifelse(X0<median(X0),0,1)-.5 #dichotomize X and center
C0=rnorm(n1,0,varc) #covariate
C=scale(C0,scale=F) #center covariate
M2=a\*X+g\*C+rnorm(n1,0,varm) #simulate M
M22=scale(M2,scale=F) #center M
Y2=cp\*X+b\*M22+f\*C+h\*X\*M22+rnorm(n1,0,vary) #simulate Y

#draw random item parameters
set.seed(2011)
k1=runif(items,.4,.7) #loadings from .4 to .7
k2=1-k1^2 #residual variances
k3=0 #intercepts
itempar=cbind(k1,k2,k3) #matrix of item parameters

#simulate item responses for M
dat2=cont\_sim(M22,itempar) #using helper function to simulate data

#dataset
data1=cbind(Y2,M22,dat2[,1:items],X,C)
colnames(data1)=c("Y","trueM",paste0('m',1:6),'X','C')
head(data1)

## Y trueM m1 m2 m3 m4 m5 m6 X
## [1,] 0.8602371 2.0115252 0.167 2.068 0.411 0.655 1.259 1.019 -0.5
## [2,] -0.6421430 -1.3486662 -1.203 -0.429 -0.171 -0.894 -0.416 -1.877 -0.5
## [3,] -0.1726501 1.0264461 0.644 0.654 1.312 1.586 1.266 0.543 0.5
## [4,] -0.3594628 0.4339889 -0.418 0.668 0.422 0.883 -0.604 1.432 0.5
## [5,] -2.3013136 -1.6338240 -1.246 -0.983 -1.573 -0.956 0.198 -0.444 -0.5
## [6,] -0.1969614 -0.2508961 -0.094 0.443 -0.168 0.142 -0.528 -0.122 -0.5
## C
## [1,] 0.72520783
## [2,] 0.42254618
## [3,] -0.06112888
## [4,] -1.91549314
## [5,] -0.63209393
## [6,] 0.04557363

## Scoring M and estimating observed XM interaction

### These procedures estimate a score for M, and then compute the XM interaction. Four procedures that use this approach are estimating summed scores for M, corrected summed scores for M, estimating a Bartlett factor score for M, and corrected factor scores for M using Croon’s correction. We discuss these procedures below.

### Summed scores for M

For this procedure, we first estimate a summed score of the items that assess M for each respondent, and then we estimate an observed XM interaction. Parameter estimates for the single mediator model are below.

ms = rowSums(data1[,paste0('m',1:6)]) #summed score for M
data\_ms=data.frame(cbind(data1,ms)) #make dataset

lm(Y~X+ms+C+X\*ms,data=data\_ms) #fit models

##
## Call:
## lm(formula = Y ~ X + ms + C + X \* ms, data = data\_ms)
##
## Coefficients:
## (Intercept) X ms C X:ms
## -0.01516 0.13478 0.08936 0.27991 0.03418

lm(ms~X+C,data=data\_ms) #fit models

##
## Call:
## lm(formula = ms ~ X + C, data = data\_ms)
##
## Coefficients:
## (Intercept) X C
## 0.01391 1.37726 0.75552

### Corrected summed scores for M

For this procedure, we first estimate a summed score of the items that assess M for each respondent, along with a reliability estimate for those item responses (e.g., coefficient alpha). Then, we take the covariance matrix between X, summed score M, Y, C, and the XM interaction and correct the variances of summed score M and XM for unreliability. Lastly, we fit the single mediator model with an XM interaction to the corrected covariance matrix using any structural equation modeling program (e.g., lavaan in R). Parameter estimates for the single mediator model are below. Note that this code relies on the previous section to estimate summed scores of M.

library(lavaan)
library(psych)

#obtain reliability estimate of summed score
p1=alpha(data\_ms[,paste0('m',1:6)])$total$raw\_alpha

## Number of categories should be increased in order to count frequencies.

Xms=data\_ms[,'X']\*data\_ms[,'ms'] #estimate observed xm interaction
q1=cov(cbind(data\_ms[,c('X','ms','Y','C')],Xms)) #get covariance matrix
q1

## X ms Y C Xms
## X 0.250501002 0.35057315 0.06727314 0.007369311 0.003483968
## ms 0.350573146 14.56318258 1.57171652 0.801072087 -0.034309621
## Y 0.067273138 1.57171652 1.35465141 0.368340963 0.140056425
## C 0.007369311 0.80107209 0.36834096 1.046854320 0.080016102
## Xms 0.003483968 -0.03430962 0.14005643 0.080016102 3.518188372

#adjust covariance matrix
q1[2,2]=q1[2,2]\*p1
q1[5,5]=q1[1,1]\*q1[2,2]-q1[1,2]^2

#fit model
mod1='
ms~X+C
Y~X+ms+C+Xms'

msc=cfa(mod1,sample.cov=q1,sample.nobs=500)
summary(msc)

## lavaan 0.6-7 ended normally after 27 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 8
##
## Number of observations 500
##
## Model Test User Model:
##
## Test statistic 0.228
## Degrees of freedom 1
## P-value (Chi-square) 0.633
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Regressions:
## Estimate Std.Err z-value P(>|z|)
## ms ~
## X 1.377 0.269 5.126 0.000
## C 0.756 0.131 5.749 0.000
## Y ~
## X 0.074 0.094 0.789 0.430
## ms 0.133 0.015 8.785 0.000
## C 0.245 0.046 5.317 0.000
## Xms 0.052 0.029 1.756 0.079
##
## Variances:
## Estimate Std.Err z-value P(>|z|)
## .ms 9.020 0.571 15.811 0.000
## .Y 1.040 0.066 15.811 0.000

### Factor scores for M

For this procedure, we first estimate a factor score of the items that assess M for each respondent, and then we estimate an observed XM interaction. The factor scores can be estimated in lavaan by fitting a unidimensional factor model to XM and extracting the Bartlett factor scores. Parameter estimates for the single mediator model are below.

library(lavaan)

#get scores
m1=paste0('m=~',paste0('m',1:6,collapse="+")) #write model
f1=cfa(m1,data=data1[,paste0('m',1:6)],std.lv=T) #fit unidimensional cfa
k2=lavPredict(f1,method='bartlett',fsm=T) #obtain scores
colnames(k2)='k2'
tsmp=attr(k2,'fsm')[[1]] #save bartlett weights
tsmp

## m1 m2 m3 m4 m5 m6
## m 0.2801805 0.260266 0.350783 0.2248121 0.2826818 0.4319148

#fit model
data\_fa = data.frame(cbind(data1,k2))
lm(Y~X+k2+C+X\*k2,data=data\_fa) #fit models

##
## Call:
## lm(formula = Y ~ X + k2 + C + X \* k2, data = data\_fa)
##
## Coefficients:
## (Intercept) X k2 C X:k2
## -0.01276 0.12929 0.28878 0.27921 0.09521

lm(k2~X+C,data=data\_fa) #fit models

##
## Call:
## lm(formula = k2 ~ X + C, data = data\_fa)
##
## Coefficients:
## (Intercept) X C
## 4.378e-17 4.469e-01 2.364e-01

### Corrected factor scores for M with Croon’s correction

For this procedure, we first estimate a factor score of the items that assess M for each respondent, along with a reliability estimate for those item responses (e.g., 1/variance(M)). Then, we take the covariance matrix between X, factor score M, Y, C, and the XM interaction and correct the variances of factor score M and XM for unreliability, which largely resembles Croon’s correction. Lastly, we fit the single mediator model with an XM interaction to the corrected covariance matrix using any structural equation modeling program (e.g., lavaan in R). Parameter estimates for the single mediator model are below. Note that this code relies on the previous section.

library(lavaan)

#obtain reliability estimate of factor score

items=6
rel1=1/(sum(coef(f1)[(items+1):(items\*2)]\*tsmp^2)+1)
Xk2=data\_fa[,'X']\*data\_fa[,'k2'] #estimate observed xm interaction
q2=cov(cbind(data\_fa[,c('X','k2','Y','C')],Xk2)) #get covariance matrix
q2 #covariance matrix

## X k2 Y C Xk2
## X 2.505010e-01 0.113682170 0.06727314 0.007369311 -1.086375e-19
## k2 1.136822e-01 1.415835963 0.49317068 0.250773339 -4.270334e-03
## Y 6.727314e-02 0.493170675 1.35465141 0.368340963 3.909160e-02
## C 7.369311e-03 0.250773339 0.36834096 1.046854320 2.812305e-02
## Xk2 -1.086375e-19 -0.004270334 0.03909160 0.028123053 3.410612e-01

#adjust covariance matrix
q1[2,2]=q1[2,2]\*rel1
q1[5,5]=q1[1,1]\*q1[2,2]-q1[1,2]^2

#fit model
mod2='
k2~X+C
Y~X+k2+C+Xk2'

mfa=cfa(mod2,sample.cov=q2,sample.nobs=500)
summary(mfa)

## lavaan 0.6-7 ended normally after 21 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of free parameters 8
##
## Number of observations 500
##
## Model Test User Model:
##
## Test statistic 0.134
## Degrees of freedom 1
## P-value (Chi-square) 0.714
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Regressions:
## Estimate Std.Err z-value P(>|z|)
## k2 ~
## X 0.447 0.102 4.376 0.000
## C 0.236 0.050 4.733 0.000
## Y ~
## X 0.129 0.095 1.356 0.175
## k2 0.289 0.041 7.045 0.000
## C 0.279 0.047 5.960 0.000
## Xk2 0.095 0.080 1.186 0.236
##
## Variances:
## Estimate Std.Err z-value P(>|z|)
## .k2 1.303 0.082 15.811 0.000
## .Y 1.095 0.069 15.811 0.000

## Structural models to accommodate a latent XM interaction M

### These procedures use structural models to accommodate the latent XM interaction directly or indirectly into the model. Four procedures that use this approach are the multiple-group approach, the unconstrained product indicator approach, latent moderated structural (LMS) equations, and Bayesian mediation. We discuss these procedures below.

###

### Multiple-group model

Given that the X variable is binary, we could estimate a multiple-group model where we study the relation between (1) latent M and C and (2) the relations between Y with latent M and C per group. Note that we assume measurement invariance of the latent variable model for M across levels of X. Using algebra, we can determine the estimates of the a-path, b-path per group, c’-path, and the h-path. Parameter estimates for the single mediator model are below.

library(lavaan)

#Multiple group model
m3=paste(paste0('m=~',paste0('m',1:6,collapse="+")),'\nm~C',
 '\nY~m+C')
t3=data1[order(data1[,'X'],decreasing=F),] #reorder data
f4=cfa(m3,data=t3,
 std.lv=T,group='X',group.equal=c('loadings','intercepts'),
 group.partial=c(y~1)) #fit mediation model

f5=parameterEstimates(f4)
f5

## lhs op rhs block group label est se z pvalue ci.lower ci.upper
## 1 m =~ m1 1 1 .p1. 0.490 0.053 9.253 0.000 0.387 0.594

## 2 m =~ m2 1 1 .p2. 0.463 0.053 8.750 0.000 0.359 0.567

## 3 m =~ m3 1 1 .p3. 0.599 0.057 10.509 0.000 0.488 0.711

## 4 m =~ m4 1 1 .p4. 0.392 0.049 7.933 0.000 0.295 0.489

## 5 m =~ m5 1 1 .p5. 0.508 0.054 9.459 0.000 0.403 0.614

## 6 m =~ m6 1 1 .p6. 0.602 0.054 11.245 0.000 0.497 0.707

## 7 m ~ C 1 1 0.217 0.081 2.685 0.007 0.059 0.375

## 8 Y ~ m 1 1 0.344 0.074 4.617 0.000 0.198 0.489

## 9 Y ~ C 1 1 0.224 0.067 3.353 0.001 0.093 0.355

## 10 m1 ~~ m1 1 1 0.718 0.073 9.828 0.000 0.575 0.861

## 11 m2 ~~ m2 1 1 0.829 0.082 10.159 0.000 0.669 0.989

## 12 m3 ~~ m3 1 1 0.744 0.081 9.231 0.000 0.586 0.902

## 13 m4 ~~ m4 1 1 0.738 0.071 10.354 0.000 0.598 0.878

## 14 m5 ~~ m5 1 1 0.765 0.078 9.831 0.000 0.613 0.918

## 15 m6 ~~ m6 1 1 0.573 0.067 8.613 0.000 0.443 0.704

## 16 Y ~~ Y 1 1 0.908 0.086 10.567 0.000 0.739 1.076

## 17 m ~~ m 1 1 1.000 0.000 NA NA 1.000 1.000

## 18 C ~~ C 1 1 0.913 0.000 NA NA 0.913 0.913

## 19 m1 ~1 1 1 .p19. -0.078 0.053 -1.474 0.140 -0.181 0.026

## 20 m2 ~1 1 1 .p20. -0.065 0.052 -1.231 0.218 -0.167 0.038

## 21 m3 ~1 1 1 .p21. -0.192 0.058 -3.317 0.001 -0.305 -0.078

## 22 m4 ~1 1 1 .p22. -0.077 0.048 -1.598 0.110 -0.172 0.018

## 23 m5 ~1 1 1 .p23. -0.105 0.054 -1.959 0.050 -0.210 0.000

## 24 m6 ~1 1 1 .p24. -0.179 0.055 -3.275 0.001 -0.286 -0.072

## 25 Y ~1 1 1 -0.132 0.064 -2.058 0.040 -0.257 -0.006

## 26 C ~1 1 1 -0.015 0.000 NA NA -0.015 -0.015

## 27 m ~1 1 1 0.000 0.000 NA NA 0.000 0.000

## 28 m =~ m1 2 2 .p1. 0.490 0.053 9.253 0.000 0.387 0.594

## 29 m =~ m2 2 2 .p2. 0.463 0.053 8.750 0.000 0.359 0.567

## 30 m =~ m3 2 2 .p3. 0.599 0.057 10.509 0.000 0.488 0.711

## 31 m =~ m4 2 2 .p4. 0.392 0.049 7.933 0.000 0.295 0.489

## 32 m =~ m5 2 2 .p5. 0.508 0.054 9.459 0.000 0.403 0.614

## 33 m =~ m6 2 2 .p6. 0.602 0.054 11.245 0.000 0.497 0.707

## 34 m ~ C 2 2 0.267 0.071 3.774 0.000 0.128 0.406

## 35 Y ~ m 2 2 0.472 0.097 4.885 0.000 0.283 0.662

## 36 Y ~ C 2 2 0.264 0.070 3.755 0.000 0.126 0.401

## 37 m1 ~~ m1 2 2 0.838 0.083 10.084 0.000 0.675 1.001

## 38 m2 ~~ m2 2 2 0.737 0.073 10.052 0.000 0.593 0.881

## 39 m3 ~~ m3 2 2 0.747 0.080 9.304 0.000 0.590 0.904

## 40 m4 ~~ m4 2 2 0.749 0.072 10.396 0.000 0.608 0.890

## 41 m5 ~~ m5 2 2 0.750 0.076 9.847 0.000 0.601 0.899

## 42 m6 ~~ m6 2 2 0.586 0.067 8.756 0.000 0.455 0.718

## 43 Y ~~ Y 2 2 1.174 0.114 10.293 0.000 0.950 1.397

## 44 m ~~ m 2 2 0.934 0.175 5.338 0.000 0.591 1.277

## 45 C ~~ C 2 2 1.176 0.000 NA NA 1.176 1.176

## 46 m1 ~1 2 2 .p19. -0.078 0.053 -1.474 0.140 -0.181 0.026

## 47 m2 ~1 2 2 .p20. -0.065 0.052 -1.231 0.218 -0.167 0.038

## 48 m3 ~1 2 2 .p21. -0.192 0.058 -3.317 0.001 -0.305 -0.078

## 49 m4 ~1 2 2 .p22. -0.077 0.048 -1.598 0.110 -0.172 0.018

## 50 m5 ~1 2 2 .p23. -0.105 0.054 -1.959 0.050 -0.210 0.000

## 51 m6 ~1 2 2 .p24. -0.179 0.055 -3.275 0.001 -0.286 -0.072

## 52 Y ~1 2 2 -0.094 0.091 -1.036 0.300 -0.272 0.084

## 53 C ~1 2 2 0.015 0.000 NA NA 0.015 0.015

*##* 54 m ~1 2 2 0.467 0.111 4.220 0.000 0.250 0.684

#then, parameter estimates for the single mediator model can be derived

#a-path is the intercept for group = 0.5
## lhs op rhs block group label est se z pvalue ci.lower ci.upper
## 54 m ~1 2 2 0.467 0.111 4.220 0.000 0.250 0.684#b-path per group are:
## lhs op rhs block group label est se z pvalue ci.lower ci.upper
## 8 Y ~ m 1 1 0.344 0.074 4.617 0.000 0.198 0.489
## 35 Y ~ m 2 2 0.472 0.097 4.885 0.000 0.283 0.662

#c-path at M=0 is i\_y(g=2) – i\_y(g=1) -.5(i\_m; g=2)\*h
## 25 Y ~1 1 1 -0.132 0.064 -2.058 0.040 -0.257 -0.006

## 52 Y ~1 2 2 -0.094 0.091 -1.036 0.300 -0.272 0.084

## 54 m ~1 2 2 0.467 0.111 4.220 0.000 0.250 0.684

## 8 Y ~ m 1 1 0.344 0.074 4.617 0.000 0.198 0.489
## 35 Y ~ m 2 2 0.472 0.097 4.885 0.000 0.283 0.662

#h-path is the difference between the b-paths per group

## lhs op rhs block group label est se z pvalue ci.lower ci.upper
## 8 Y ~ m 1 1 0.343 0.074 4.615 0.000 0.198 0.489
## 35 Y ~ m 2 2 0.486 0.090 5.420 0.000 0.310 0.662

### Unconstrained product indicator (UPI) approach

For our implementation of the UPI approach, we create three parcels (i.e., three sums) of the item responses that estimate M, and multiplied those parcels by X. Those products served as indicators of a latent XM term for the single mediator model. To identify the latent XM term, its mean and variance are constrained to values shown below. Parameter estimates for the single mediator model are below

#create product indicators with parcels
xms1=(data1[,'m1']+data1[,'m2'])\*data1[,'X']
xms2=(data1[,'m3']+data1[,'m4'])\*data1[,'X']
xms3=(data1[,'m5']+data1[,'m6'])\*data1[,'X']

data\_upi=cbind(data1,xms1,xms2,xms3)

upi1='

fm~a\*fx+f\*C
Y~fx+fm+fxm+C

fx=~X
fm=~NA\*m1+m2+m3+m4+m5+m6
fxm=~NA\*xms1+xms2+xms3
fx~~vax\*fx
C~~vc\*C
fm~~1\*fm

fxm~~vxm\*fxm
fxm~covxm\*1

vxm == (a\*a\*vax+f\*f\*vc+1)\*vax-(vax\*a)^2
covxm == vax\*a
X~~0\*X

'
f3=cfa(upi1,data=data\_upi,meanstructure=T)

upiscore=coef(f3)[1:6]

upiscore

## a f Y~fx Y~fm Y~fxm Y~C
## 0.47517181 0.24993257 0.06824222 0.40025092 0.13499441 0.24752006

### Latent moderated structural (LMS) equations

The LMS approach is available in Mplus. Here, the latent XM interaction is specified with the XWITH function. For our purposes we conduct the analysis in R using the R-package MplusAutomation, which calls Mplus from R to run the Mplus input file we create below, and then brings in the parameter estimates into R. Parameter estimates for the single mediator model are below

library(MplusAutomation)

write.table(data1,file='mplus\_data.dat',row.names=F,col.names=F)
mplustemplate=c(
 'TITLE: single mediator model with XM interaction;',
 'DATA: FILE = mplus\_data.dat;',
 'VARIABLE: NAMES = y truem m1-m6 x c;',
 'usevariables are y m1-m6 x c;',

 'ANALYSIS: TYPE = random;',
 'estimator=ML; algorithm=integration;',
 'MODEL:',

 'f1 on x (a)',
 'c;',
 'g1 | f1 XWITH x;',
 'y on f1 (b)',
 'x (c)',
 'c',
 'g1 (h);',
 'f1 by m1\* m2-m6;',
 'f1@1;',

 'OUTPUT: TECH9;'
)

writeLines(mplustemplate,'mp.inp')
runModels('mp.inp')
lms0=readModels(paste0('mp','.out'))$parameters[[1]]
lms0

## paramHeader param est se est\_se pval
## 1 F1.BY M1 0.479 0.048 9.933 0.000
## 2 F1.BY M2 0.456 0.048 9.587 0.000
## 3 F1.BY M3 0.592 0.049 12.087 0.000
## 4 F1.BY M4 0.388 0.045 8.549 0.000
## 5 F1.BY M5 0.501 0.048 10.506 0.000
## 6 F1.BY M6 0.587 0.046 12.852 0.000
## 7 F1.ON X 0.475 0.111 4.289 0.000
## 8 F1.ON C 0.250 0.054 4.612 0.000
## 9 Y.ON F1 0.406 0.057 7.080 0.000
## 10 Y.ON G1 0.174 0.107 1.630 0.103
## 11 Y.ON X 0.066 0.099 0.668 0.504
## 12 Y.ON C 0.244 0.049 5.018 0.000
## 13 Intercepts Y -0.023 0.051 -0.451 0.652
## 14 Intercepts M1 0.032 0.045 0.721 0.471
## 15 Intercepts M2 0.044 0.045 0.982 0.326
## 16 Intercepts M3 -0.051 0.047 -1.100 0.271
## 17 Intercepts M4 0.014 0.042 0.338 0.736
## 18 Intercepts M5 0.014 0.045 0.301 0.763
## 19 Intercepts M6 -0.038 0.043 -0.885 0.376
## 20 Residual.Variances Y 1.035 0.071 14.591 0.000
## 21 Residual.Variances M1 0.781 0.056 13.836 0.000
## 22 Residual.Variances M2 0.782 0.056 14.039 0.000
## 23 Residual.Variances M3 0.743 0.059 12.624 0.000
## 24 Residual.Variances M4 0.741 0.051 14.461 0.000
## 25 Residual.Variances M5 0.756 0.056 13.579 0.000
## 26 Residual.Variances M6 0.586 0.050 11.808 0.000
## 27 Residual.Variances F1 1.000 0.000 999.000 999.000

### Bayesian mediation

We used JAGS to estimate the single mediator model with a latent XM interaction. In this case, Gibbs sampling estimates the model, in which an intermediate step consists of estimating the latent score of M. As such, the score of M is multiplied times X and also regressed on Y to estimate the XM interaction. Note that users must download the JAGS software to conduct this analysis. In this case, the R2jags package communicates with JAGS to send the input and receive the output of the analysis. Parameter estimates for the single mediator model and 95% credible intervals are below.

library(R2jags)

## Loading required package: rjags

## Loading required package: coda

## Linked to JAGS 4.3.0

## Loaded modules: basemod,bugs

##
## Attaching package: 'R2jags'

## The following object is masked from 'package:coda':
##
## traceplot

bcafa=function() {

 ######################################################################
 # Specify the IRT model for the observables
 ######################################################################
 for (i in 1:n){
 for(j in 1:J){
 mu[i,j] <- tau[j] + ksi[i]\*lambda[j]
 m[i,j] ~ dnorm(mu[i,j], inv.psi[j])
 }
 }

 ######################################################################
 # Specify the (prior) distribution for the latent variables + pars
 ######################################################################
 for(j in 1:J){

 inv.psi[j] ~ dgamma(.5, .5) # Precisions for observables
 psi[j] <- 1/inv.psi[j] # Variances for observables
 lambda[j] ~ dnorm(1,.5)
 }
 tau[1] <- 0
 for(j in 2:J){
 tau[j] ~ dnorm(0,.5) }

 ######################################################################
 # Structural pars
 ######################################################################
 int1~dnorm(0,.001)
 int2~dnorm(0,.001)
 apath~dnorm(0,.001)
 bpath~dnorm(0,.001)
 cpath~dnorm(0,.001)
 gpath~dnorm(0,.001)
 fpath~dnorm(0,.001)
 hpath~dnorm(0,.001)
 tau.e ~ dgamma(1,1)

 for (i in 1:n){
 kappa[i]<-int1+apath\*x[i]+gpath\*c[i]
 ksi[i] ~ dnorm(kappa[i],1) # distribution for the latent variables
 }

 for(i in 1:n){
 y.prime[i] <- int2+bpath\*ksi[i]+cpath\*x[i]+fpath\*c[i]+hpath\*ksi[i]\*x[i]
 y[i]~dnorm(y.prime[i],tau.e)
 }

} # closes the model

jagsdat=list(m=data1[,paste0('m',1:6)],n=nrow(data1),J=6,y=data1[,'Y'],x=c(data1[,'X']),c=data1[,'C'])
bpar <- c("tau", "lambda",'psi','int1','int2','apath','bpath','cpath',
 'gpath','fpath','hpath','tau.e')

jagcfa <- jags(data = jagsdat,
 parameters.to.save = bpar, n.chains=3,
 n.iter = 5000, n.burnin = 1000, model.file = bcafa)

## module glm loaded

## Compiling model graph
## Resolving undeclared variables
## Allocating nodes
## Graph information:
## Observed stochastic nodes: 3500
## Unobserved stochastic nodes: 526
## Total graph size: 14042
##
## Initializing model

a1=round(jagcfa$BUGSoutput$summary[,c(1,6,7)],4)
a1

 mean 2.5% 97.5%

## apath 0.4701 0.2398 0.6869

## bpath 0.4092 0.2930 0.5225

## cpath 0.0531 -0.1533 0.2506

## deviance 9024.2039 8953.2043 9095.2146

## fpath 0.2435 0.1477 0.3370

## gpath 0.2489 0.1446 0.3587

## hpath 0.1759 -0.0391 0.3953

## int1 0.0704 -0.1166 0.2618

## int2 -0.0508 -0.1662 0.0682

## lambda[1] 0.4773 0.3845 0.5741

## lambda[2] 0.4609 0.3669 0.5548

## lambda[3] 0.5954 0.5006 0.6936

## lambda[4] 0.3912 0.3025 0.4811

## lambda[5] 0.5047 0.4108 0.5976

## lambda[6] 0.5913 0.4999 0.6828

## psi[1] 0.7878 0.6789 0.9072

## psi[2] 0.7891 0.6869 0.9069

## psi[3] 0.7508 0.6375 0.8672

## psi[4] 0.7487 0.6506 0.8575

## psi[5] 0.7638 0.6617 0.8779

## psi[6] 0.5945 0.5018 0.7008

## tau[1] 0.0000 0.0000 0.0000

## tau[2] 0.0119 -0.0962 0.1193

## tau[3] -0.0924 -0.2137 0.0297

## tau[4] -0.0140 -0.1133 0.0820

## tau[5] -0.0221 -0.1340 0.0902

## tau[6] -0.0782 -0.1987 0.0349

## tau.e 0.9586 0.8329 1.0929