How to Model Mediating and Moderating Effects

Hsueh-Sheng Wu

CFDR Workshop Series

November 3, 2014

BGSU

Center for Family and Demographic Research

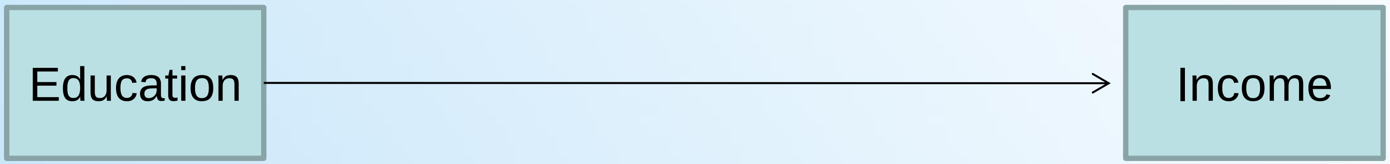
Outline

- Why study moderation and/or mediation?
- Moderation
  - Logic
  - Testing moderation in regression
  - Checklist of modeling moderation
- Mediation
  - Logic
  - Testing mediation in regression
  - Checklist of modeling mediation
- Extension: newer models
- Conclusion

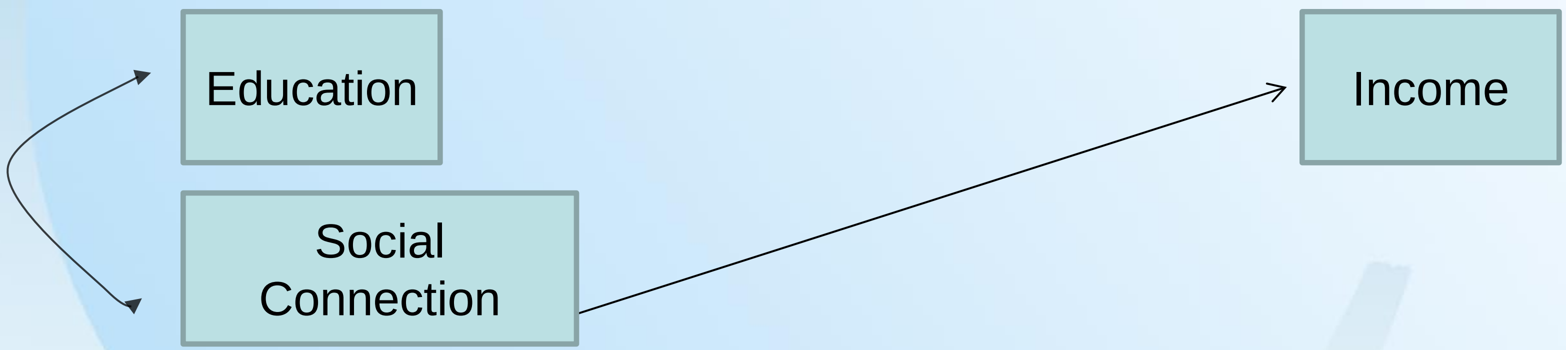
BGSU
Center for Family and Demographic Research

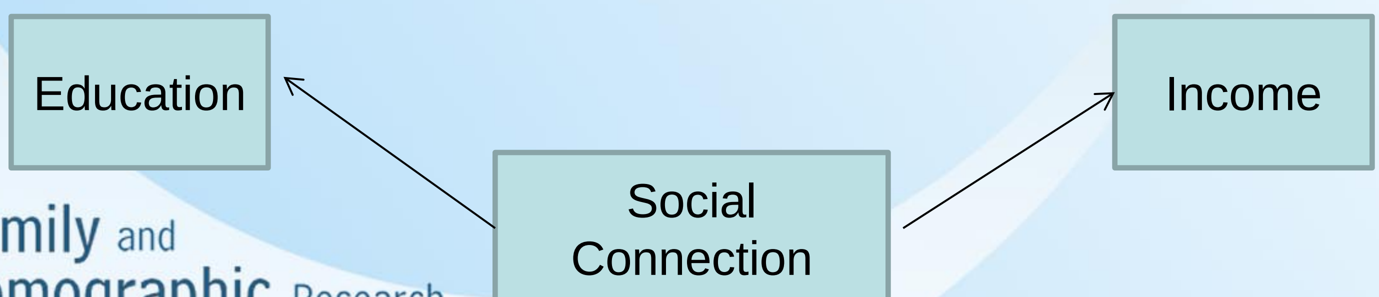
Why Study Moderation and/or Mediation

Given a significant association found between X (e.g., education) and Y (e.g., income), social science researchers can conduct three additional analyses with new variables (e.g., social connection):



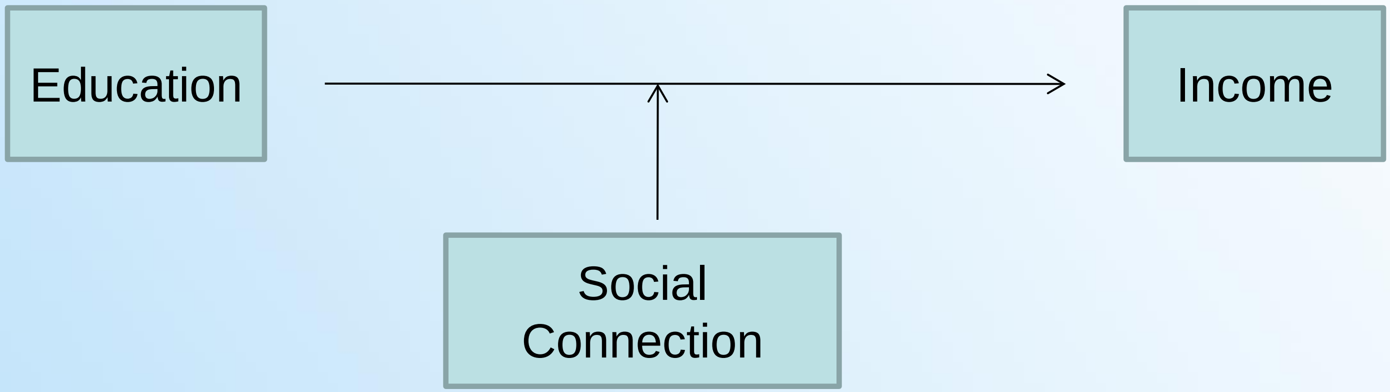
(1) Test for the spurious relation between X and Y



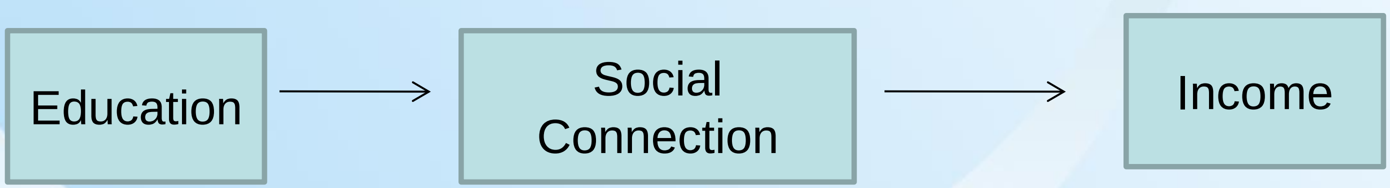


Why Study Moderation and/or Mediation?

(2) Test if Z modifies the X-Y relation



(3) Test if Z mediates the effect of X on Y



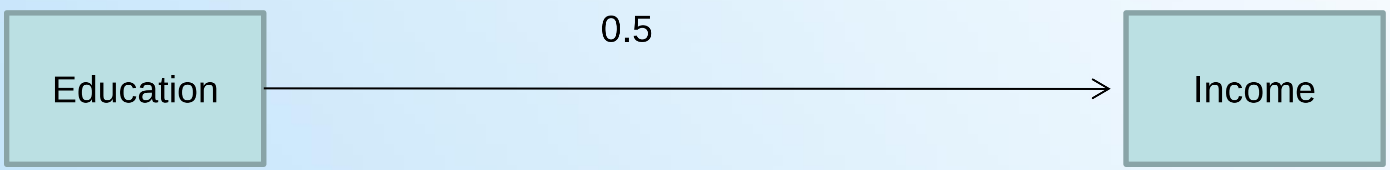
Why Study Moderation and/or Mediation?

- The basic model establishes a valid X-Y relation, and this relation is the basis for testing the spurious X-Y relation and the moderating or mediating effects of Z on the X-Y relation
- The first model describe that Z created a spurious association between X and Y.
- The second model is the moderating effect of Z, that is, the X-Y relation varies with the value of Z.
- The third model is the mediating effect of Z, that is, the effect of X on Y needed to go through Z.
- Z can play three different types of roles in the X-Y relation.

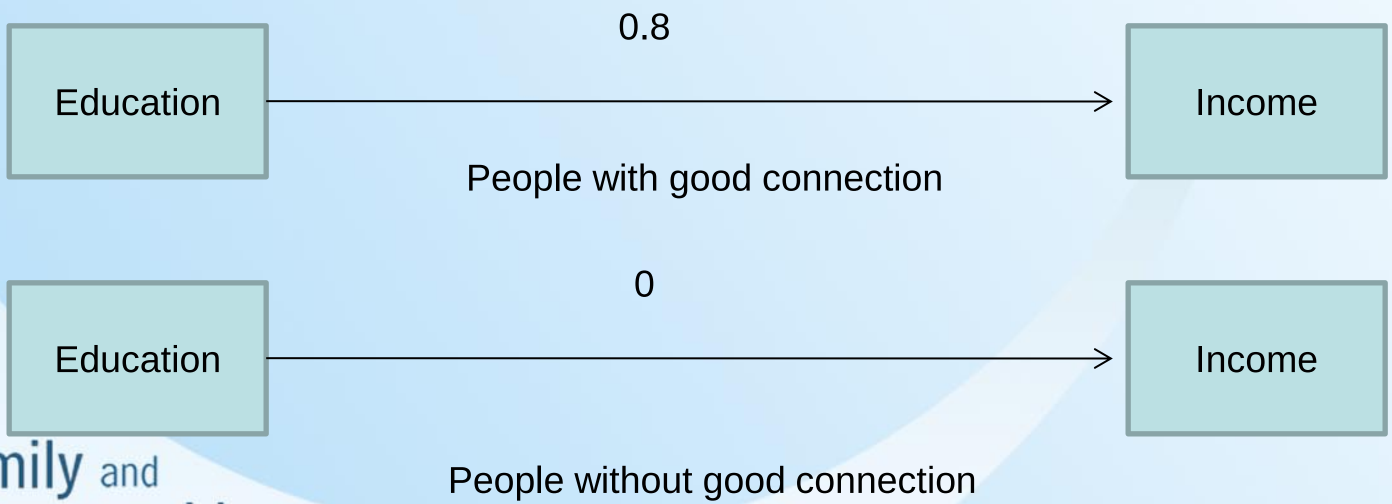
Moderation

Logic: If Z moderates the X-Y relation, the X-Y relation differs in magnitude or even sign for at least some levels of Z. For example:

(1) The X-Y relation for the whole sample

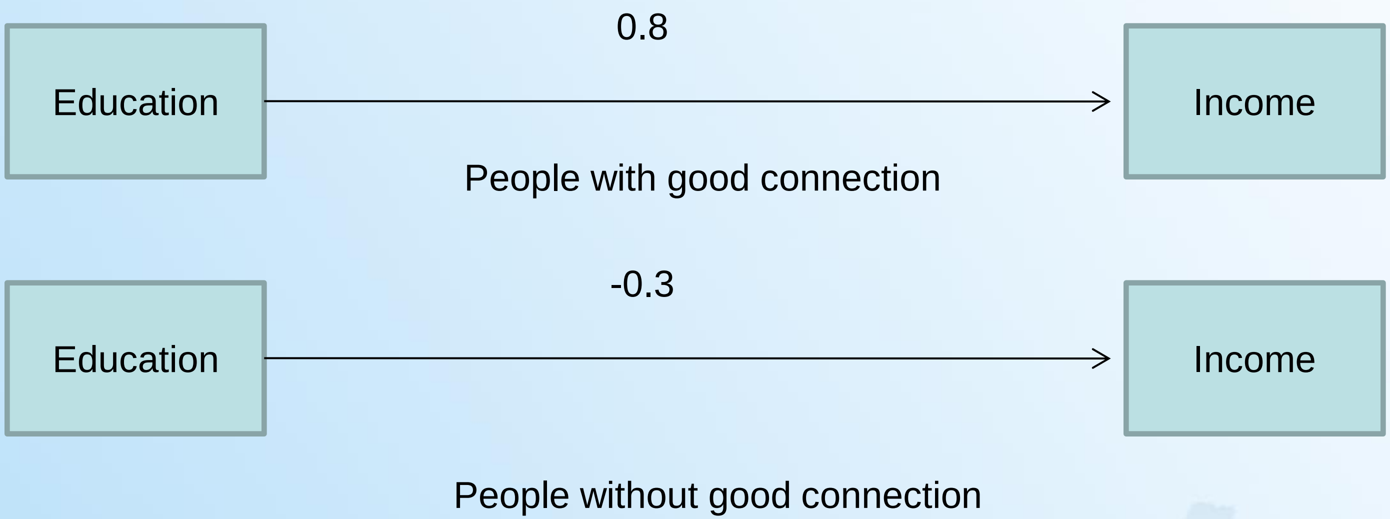


(2) The strength of X-Y relation differs for subgroups of people



Moderation

(3) The sign of the X-Y relation changes for subgroups of people



How to Model Moderation

Moderation indicates that the X-Y relation differs by the level of Z. When using multiple regression, you simply include X, Z, and an interaction term between X and Z as predictors of Y. If the regression coefficient of this interaction term is significant, it suggests that Z modifies the X-Y relation.

$$Y = b_0 + aX + bZ + cXZ + \varepsilon$$

a indicates the main effect of X on Y

b indicates the main effect of Z on Y

c indicates the interaction effect of X and Z, i.e., the effect of X on Y, given the presence of Z

BGSU

Center for Family and Demographic Research

Checklist of Modeling Moderation

- Theoretical assumption of the moderation effect
- Issues in creating the product term of X and Z
- Measurement level of Y
- Collinearity between X and Z
- Unequal variances between groups
- Measurement errors in X, Y, and Z

Theoretical Assumption of Moderation Effect

Three types of Moderating effects:

1. Enhancing effect: The presence of a moderator enhances the effect of the predictor on the outcome variable
2. Buffering effect: The presence of a moderator reduces the effect of the predictor on the outcome variable
3. Antagonistic effect: The presence of a moderator reverse the effect of the predictor on the outcome variable

BGSU

Center for Family and Demographic Research

Coding for the Interaction Term

The coding of the interaction term can be complex depending on how independent variables and moderators are measured

1. Continuous X and Z

    Original variables:

        Y: Level of happiness

        X: Income

        Z: Years of schooling

    Interaction term:

        Income*Years of schooling

BGSU

Center for Family and Demographic Research

Coding for the Interaction Term

2. Categorical causal variables and moderators

Original variables:

Y: Level of happiness

X: Marital status:

(1) Single, (2) Married, and (3) Cohabiting

Z: Gender: (1) Male and (2) Female

Dummy variables:

Married is 1 if X = 2 and = 0 otherwise

Cohabiting is 1 if X = 3 and = 0 otherwise

Female is 1 if Gender = 2 and = 0 otherwise

Two interaction terms:

Female*Married and Female*Cohabiting

Coding of the Interaction Term

3. Continuous X and Categorical Z

Original variables:

Y: Level of happiness

X: Income

Z: Marital status:

(1) Single, (2) Married, and (3) Cohabiting

Dummy variables:

Married is 1 if X = 2 and = 0 otherwise

Cohabiting is 1 if X = 3 and = 0 otherwise

Interaction terms:

Income*Married and Income*Cohabiting

Coding of the Interaction Term

4. Categorical causal variables and Continuous moderators

Original variables:

Y: Level of happiness

X: Marital status:

(1) Single, (2) Married, and (3) Cohabiting

Z: Income

Dummy variables:

Married is 1 if X = 2 and = 0 otherwise

Cohabiting is 1 if X = 3 and = 0 otherwise

Interaction terms:

Married*Income and Cohabiting*Income

BGSU

Center for Family and Demographic Research

Measurement Level of Y

- The number of response categories of Y contributes to the possibility of detecting the moderation effects.

- If X and Z both have five response categories, their interaction term has 25 different combinations. It is easier to detect the effect of the interaction term when the values of the outcome variable has a wider range than a narrow range.

Collinearity between X and Z

- When X and Z are correlated with each other, it reduces the power of detecting the moderation effects.

- In extreme cases, you won't be able to fit the model.

- You can try to reduce this problem by increasing sample size and/or centering and/or standardizing the variables

Unequal Variances between Groups

- When testing moderation, it is assumed that the variance of regression coefficients estimated from different sub-groups are the same. But, if this is not true, different sub-groups should be analyzed separately.

- Use Chow test to determine if different groups of respondents should be combined.

- You can use structural equation modeling to test moderation effects when unequal variances exist between groups.

Measurement Errors in X, Y, and Z

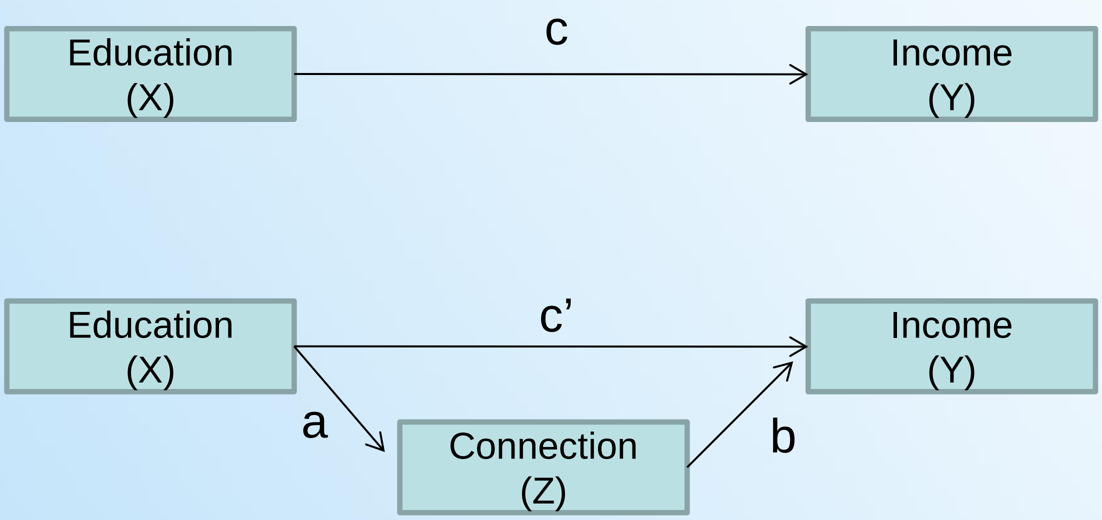
- The measurement errors in X and Z reduce the reliability of the interaction term created from X and Z.

- The measurement error in Y reduces the correlation with the predictors, lowers the overall R^2 value, and the power of the test.

BGSU
Center for Family and Demographic Research

Mediation

Logic: If Z mediates the X-Y relation, then the following conditions hold



- X predict Y
- X predict Z
- Z predicts Y
- When Y are predicted by both X and Z:
  - The regression coefficient of Z (i.e., b) should be significant
  - The regression coefficient of X differently when Z is in the regression than when Z is not (i.e., c' is different from c).

Steps of Testing Mediation

1. Test if X predicts Y

$$Y = B_1 + cX + \varepsilon_1$$

2. Test if X predicts Z

$$Y = B_2 + aX + \varepsilon_2$$

3. Test if X still predicts Y when Z is in the model

$$Y = B_3 + c' X + bZ + \varepsilon_3$$

Decision Rules

- Z completely mediates the X-Y relation if all three conditions are met:
  (1) X predicts Y
  (2) X predicts Z
  (3) X no longer predicts Y, but Z does when both X and Z are used to predict Y

- Z partially mediates the X-Y relation if all three conditions are met:
  (1) X predicts Y
  (2) X predicts Z
  (3) Both X and Z predict Y, but X have a smaller regression coefficient when both X and Z are used to predict Y than when only X is used

- Z does not mediate the X-Y relation if any of
  (1) X does not predict Z
  (2) Z does not predict Y
  (3) The regression coefficient of X remain the same before and after Z is used to predict Y

BGSU Center for Family and Demographic Research

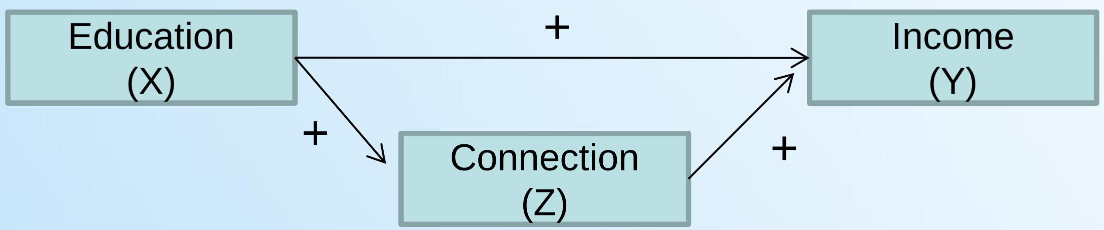
Checklist of Modeling Mediation

- Theoretical assumptions on the mediator effects
- Establish causation among X, Y and Z
- Selection of the mediator
- Significance of the mediation effect
- Omitted variables
- Collinearity between X and Z

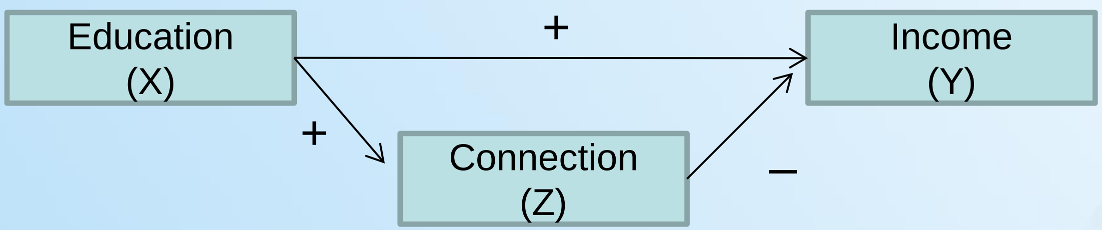
BGSU
Center for Family and Demographic Research

Theoretical Assumption on Mediation Effects

(1) Enhancer



(2) Suppresser



Establish Causation among X, Y, and Z

- The mediation study indicates that X causes Z and then Z causes Y.

- Three conditions of causation between two variables

  A association between two variables

  The association is not spurious

  The cause precedes the effect in time

- If you use cross-sectional data, you need to justify your mediation model

Education (X) → Connection (Z) → Income (Y)

Education (X) → Income (Y) → Connection (Z)

BGSU

Center for Family and Demographic Research

Selection of Mediators

- Mediator should be something changeable.

- The association of Z to X influences the possibility of detecting mediating effects. A high X-Z association implies that more variance in Z is explained by X, and there is less variance in Z to contribute to the perdition of Y.

- The association of Z to Y influences the possibility of detecting mediating effects. If the Z-Y association is slightly stronger than the X-Z association, it is easier to detect mediating effects.

- Measurement error in Z underestimates the Z-Y association and overestimates the X-Y association. Hoyle and Robinson (2003) recommended that the reliability of mediator should be at least .90

BGSU

Center for Family and Demographic Research

Significance of Mediation Effect

- Describe the proportion of the X-Y relation that is attributable to z (Shrout & Bolger, 2002)

$$ab/c$$

Where a, b, and c are unstandardized regression coefficients

- Statistical test of the mediating effect by dividing the product of paths *a* and *b* by a standard error term (Baron & Kenny, 1986)

$$Z = \frac{ab}{\sqrt{b^2 sa^2 + a^2 sb^2 + sa^2 sb^2}}$$

Where *a* and *b* are unstandardized regression coefficients and *sa* and *sb* are their standard errors

BGSU Center for Family and Demographic Research

Omitted Variables

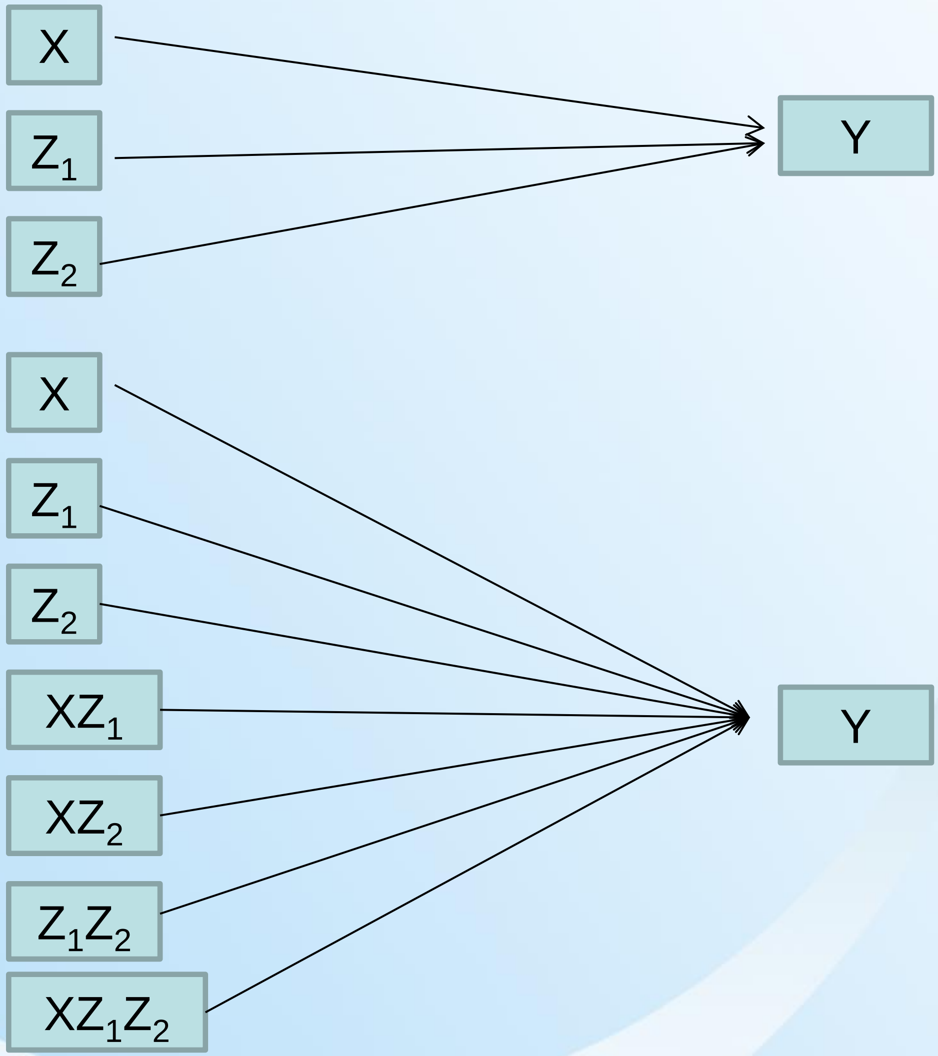
- If some omitted variables account for the X-Y or Z-Y associations, there may indeed no mediation effect.
- You need to justify that the X-Y and Z-Y relations are theoretical and empirically valid.

Collinearity between X and Z

- Because X predicts Z, there will be collinearity problem when they both in the same equation.
- In extreme cases, you won’t be able to fit the model.
- You can try to reduce this problem by increasing sample size and/or centering and/or standardizing the variables

Extension

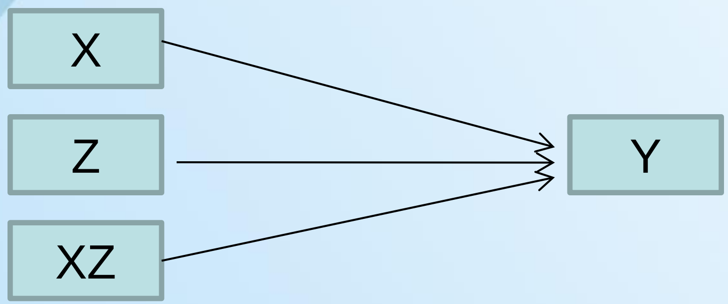
Moderation (more than one moderators)

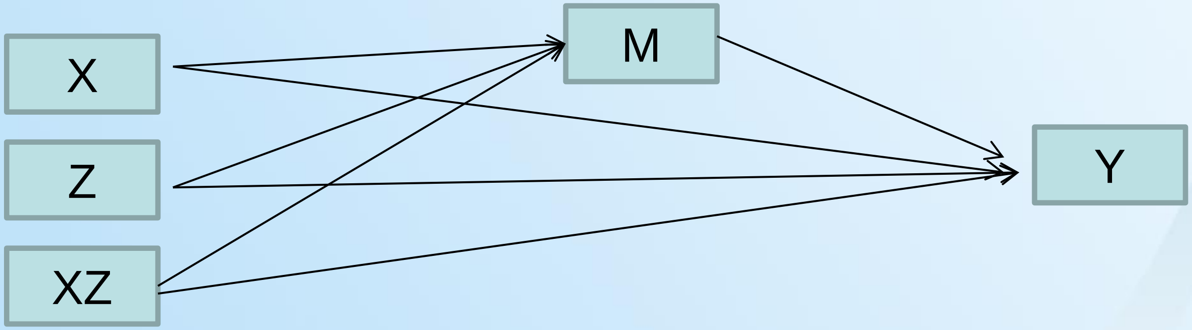


BGSU Center for Family and Demographic Research

Extension

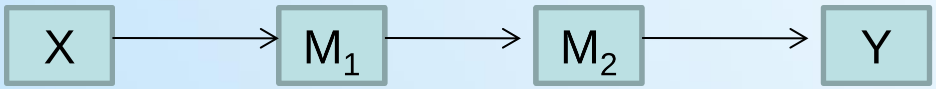
Mediated moderation

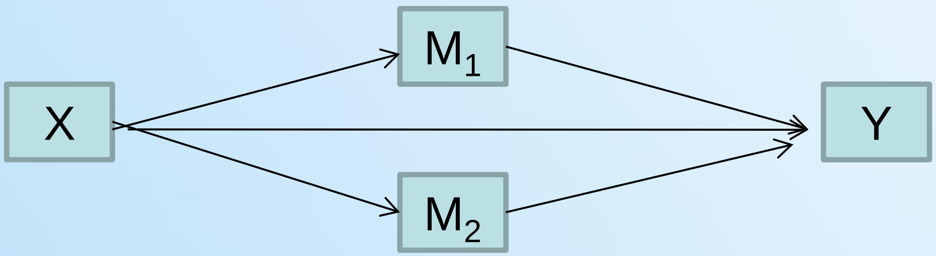




Extension

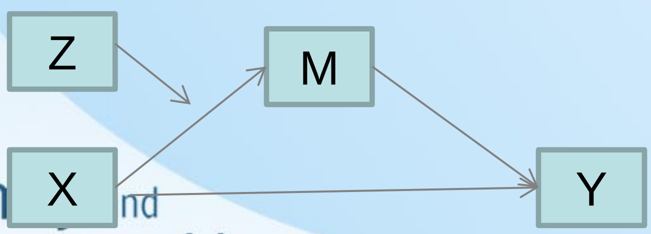
- Mediation
  - More than two mediators





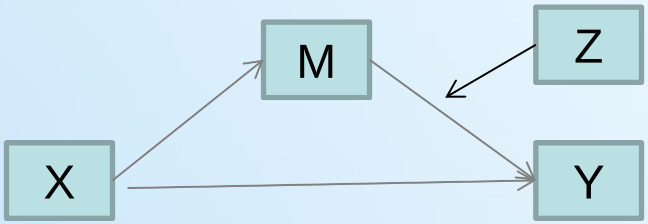
- Moderated mediator (Edwards and Lambert, 2007)

  (1) first stage moderation model

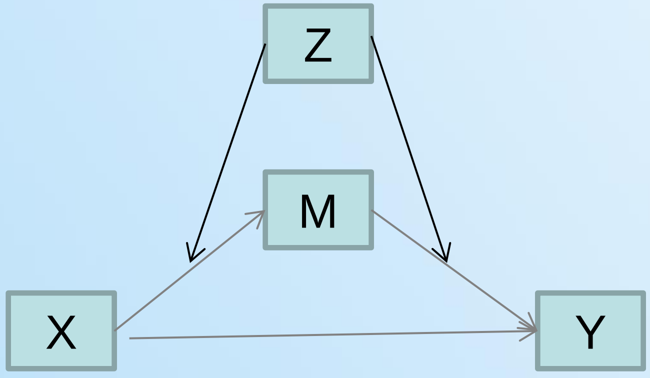


Extension

(2) Second Stage Moderation Model



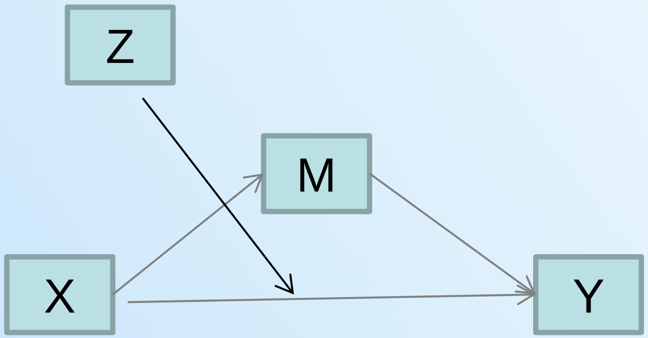
(3) First and Second Stage Moderation Model



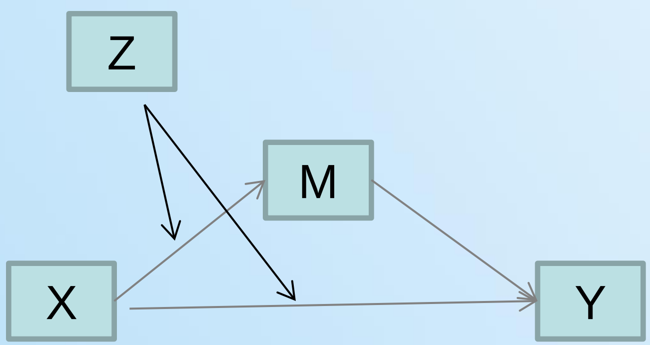
BGSU
Center for Family and Demographic Research

Extension

(4) Direct effect moderation model



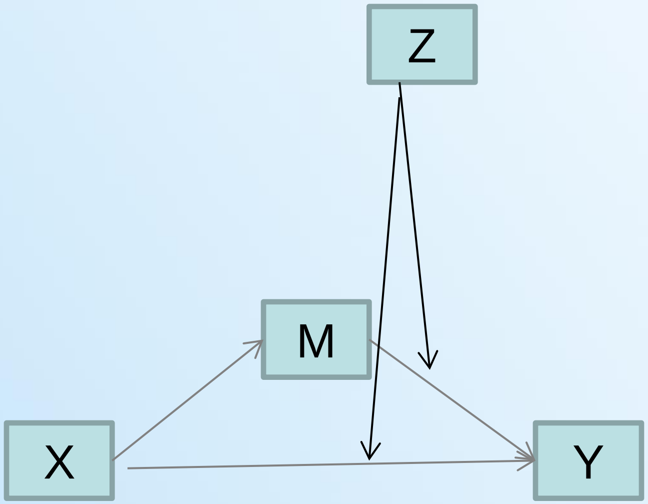
(5) Direct Effect and First Stage Moderation Model



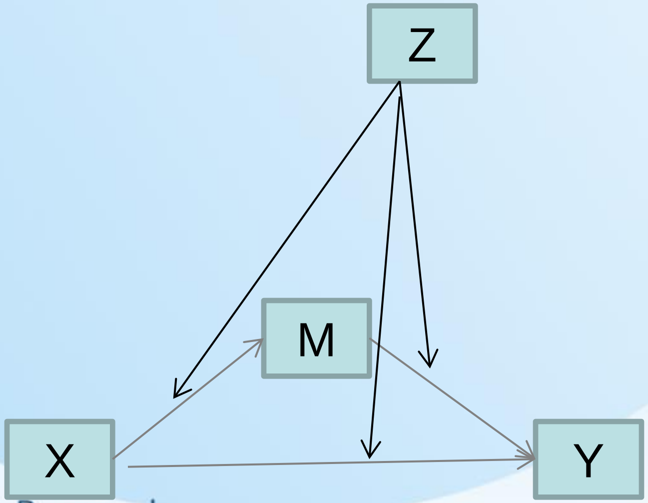
BGSU
Center for Family and Demographic Research

Extension

(6) Direct Effect and Second Stage Moderation Model



(7)Total Effect Moderation Model



Conclusion

- Moderation examines under what conditions, the X-Y relation varies, while mediation examines why the X-Y relation occurs.
- You should use theory to guide the examination of moderation and mediation because the same variables can play the role of mediator or moderator.
- Moderation and mediation can be examined simultaneously in mediated moderation and moderated mediation.
- Moderation and mediation can be examined by multiple regression. However, some complex moderation and mediation models may need to be examined by structural equation modeling.
- After examining moderators or mediators in your analyses, you probably want to focus on the theoretical links of these variables when writing your papers.

BGSU Center for Family and Demographic Research

References

Journal articles

Baron, R.M. & Kenny, D.A.(1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical consideration. Journal of Personality and Psychology, 51, 1173-1182.

Edwards, J.R. & Lambert, L.S. (2007) Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. Psychological Methods, 12, 1-22.

Frazier, P.A., Tix, A,P., and Barron, K.E. (2004) Testing Moderator and mediator effects in counseling psychology, Journal of Counseling Psychology, 51, 115-134.

Hoyle, R.H. & Robinson, J.I. (2003) Mediated and moderated effects in social psychological research: Measurement, design, and analysis issues. In C. Sansone, C. Morf, & A.T. Panter (Eds.), Handbook of methods in social psychology. Thousand Oaks, CA: Sage.

MacKinnon, D.A. ; Lockwood, C.M.; Hoffman, J.M.; West, S.G. and Sheets, V. (2002) A Comparison of Methods to Test Mediation and Other Intervening Variable Effects. Psychological Methods,7,1-35

Zhang, Z. , Zyphur, M. and Preacher, K.J. (2009) Testing multilevel mediation using hierarchical liner model. Organizational Research Methods, 12, 695-719.

Morgan-Lopez, A.A.& MacKinnon, D.P. (2006) Demonstration and evaluation of a method for assessing mediated moderation. Behavior Research Methods, 38, 77-87.

Shrout, P.E., & Bolger, N. (2002) Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods, 7, 422-445.

Website

For Moderation
http://www.davidakenny.net/cm/moderation.htm#LI

For Mediation
http://www.davidakenny.net/cm/mediate.htm