How to Model Mediating and Moderating Effects

Hsueh-Sheng Wu

CFDR Workshop Series

Spring 2011

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):
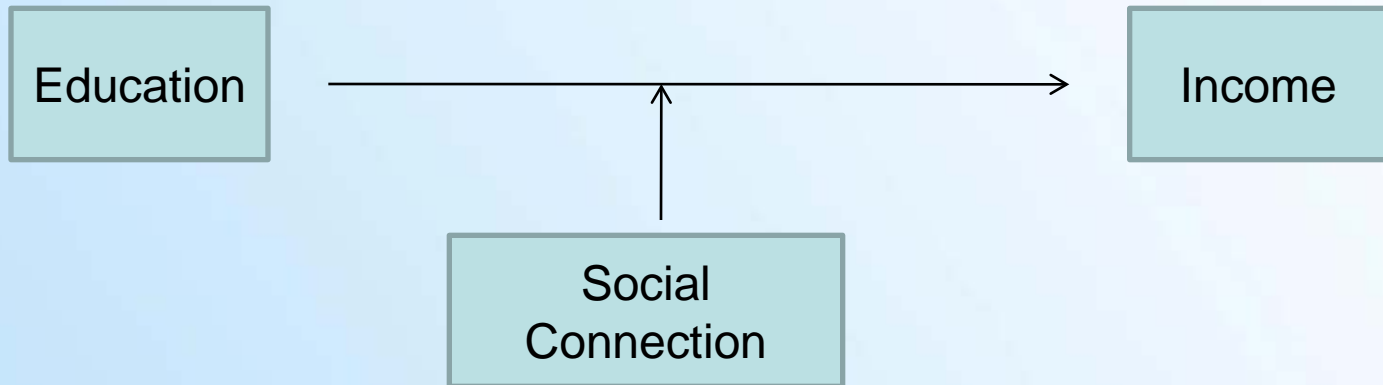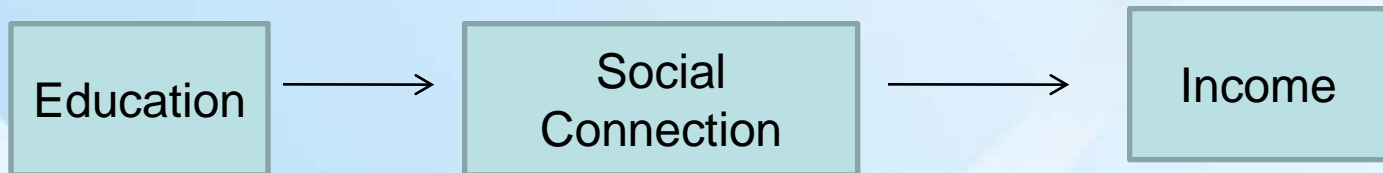

(1) Test for the spurious relation between X and Y

3
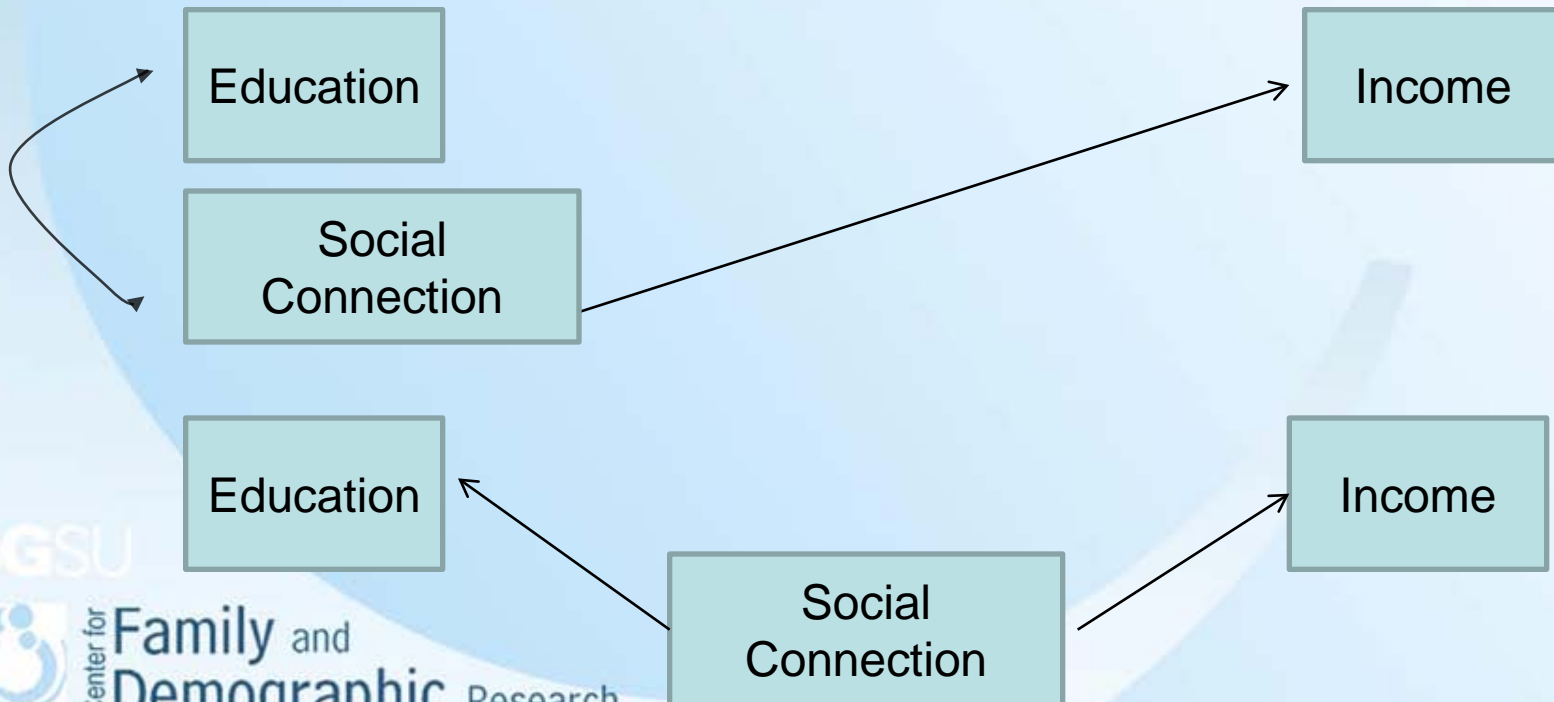
Why Study Moderation and/or Mediation?

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

Education		Income

Social Connection

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

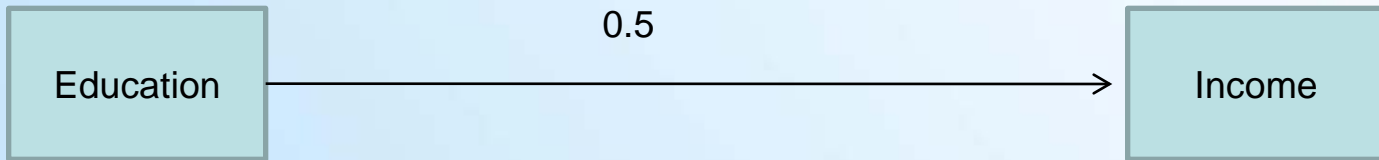| Education	Social Connection	Income

Why Study Moderation and/or Mediation?

- The first model establishes a valid X-Y relation, and this relation is the basis for testing moderation or mediation.

- The second model is the moderation model, examining under what conditions, the X-Y relation varies.

- The third model is the mediation model, examining why the X-Y relation occurs.

- The same Z variable can be included in any of these models, so you should use theory to guide your research.

BGSU

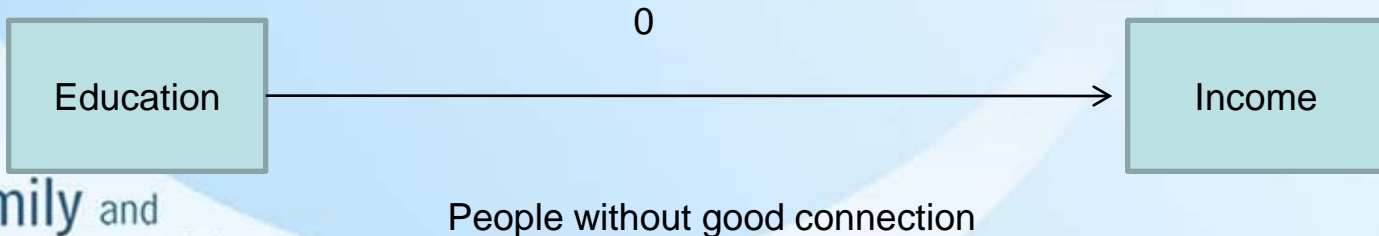Center for Family and Demographic Research

Moderation

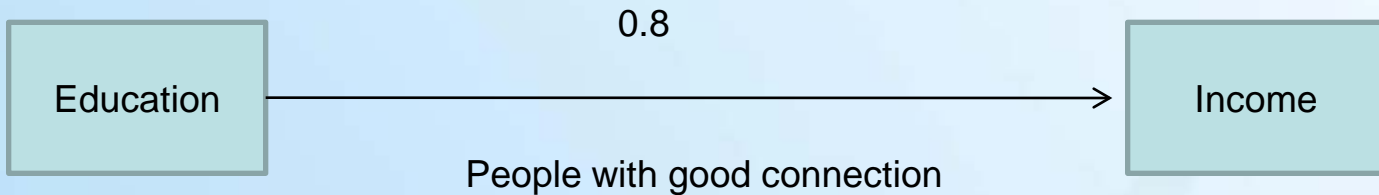Logic: If Z moderates the X-Y relation, the X-Y relation differs in magnitude or even sign for at least some levels of Z. For example:

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

| Education | 0.5 → | Income |

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

| Education | 0.8 → | Income |

People with good connection

| Education | 0 → | Income |

People without good connection

6

Moderation

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

0.8

| Education | → | Income |

People with good connection

-0.3

| Education | → | Income |

People without good connection
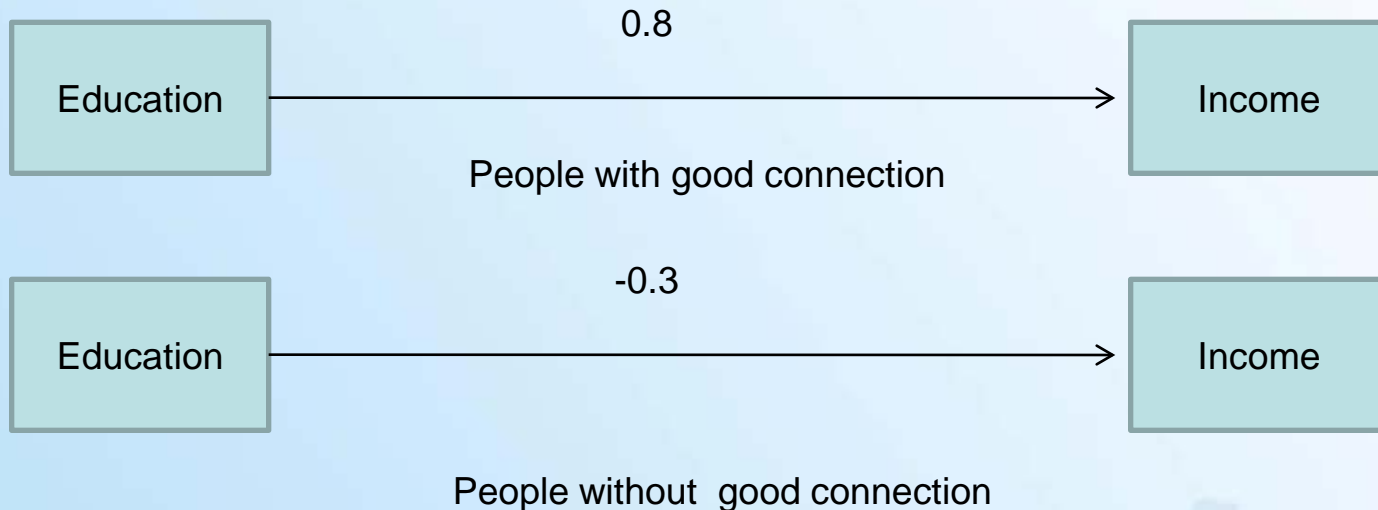
How to Model Moderation

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 effect of X when Z is zero
b indicates the effect of Z when X is zero
c indicates how much the effect of X changes as Z changes one unit

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

BGSU

Center for Family and Demographic Research

Theoretical Assumption of Moderation Effect

	X-Y relation	**Z-Y relation**	**X*Z –Y relation**
Enhancing Effect	Positive (or negative)	Positive (or negative)	Positive (or negative)
Buffering Effect	Positive	Negative	Negative
Antagonistic Effect	Positive	Positive	Negative

10

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

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

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.

BGSU

Center for Family and Demographic Research

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.

Mediation

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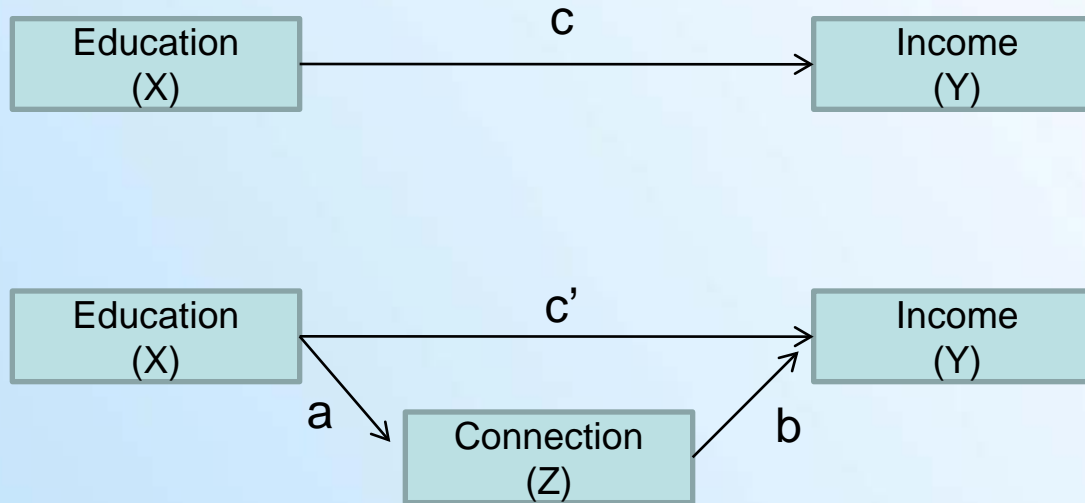
- X predict Y
- X predict Z
- Z predicts Y
- When Y are predicted by both X and Z:
    - The regression coefficient of Z (i.e., b) should be significant
    - The regression coefficient of X differently when Z is in the regression than when Z is not (i.e., c' is different from c).

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

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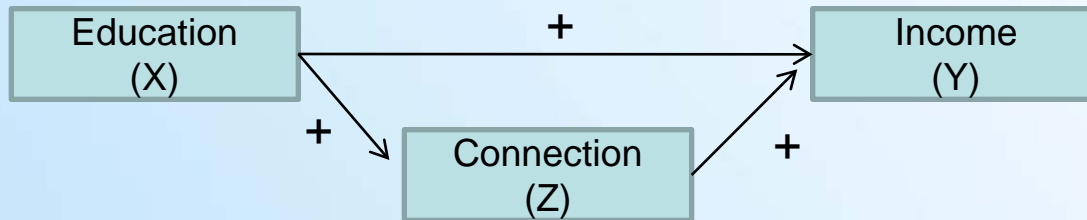
Decision Rules

- 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
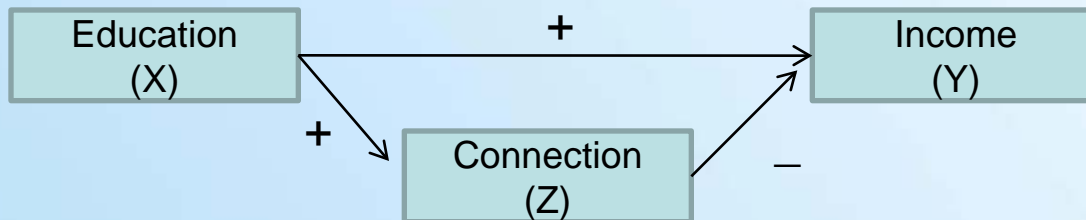
Checklist of Modeling Mediation

- 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

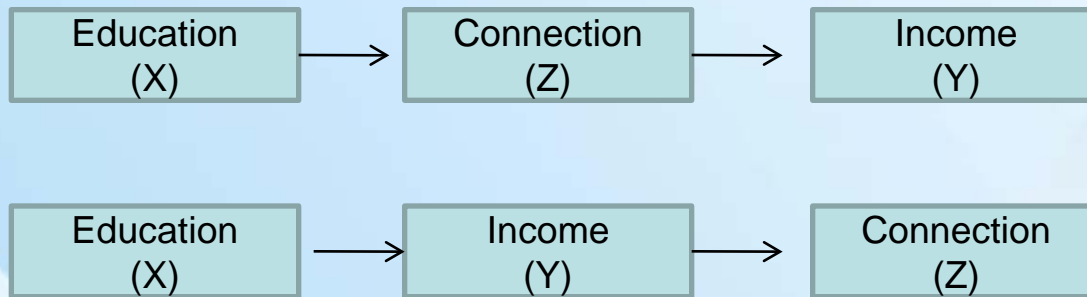# Theoretical Assumption on Mediation Effects

(1) Enhancer

```
┌─────────────┐              +              ┌─────────────┐
│  Education  │ ──────────────────────────→ │   Income    │
│     (X)     │                             │     (Y)     │
└─────────────┘                             └─────────────┘
        │  +                          +  ↗
        ↓                          
      ┌─────────────┐
      │  Connection │
      │     (Z)     │
      └─────────────┘
```

(2) Suppresser

```
┌─────────────┐              +              ┌─────────────┐
│  Education  │ ──────────────────────────→ │   Income    │
│     (X)     │                             │     (Y)     │
└─────────────┘                             └─────────────┘
        │  +                          _  ↗
        ↓                          
      ┌─────────────┐
      │  Connection │
      │     (Z)     │
      └─────────────┘
```

23

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) |

24

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

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)

$$ab \big/ c$$

  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

26

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.

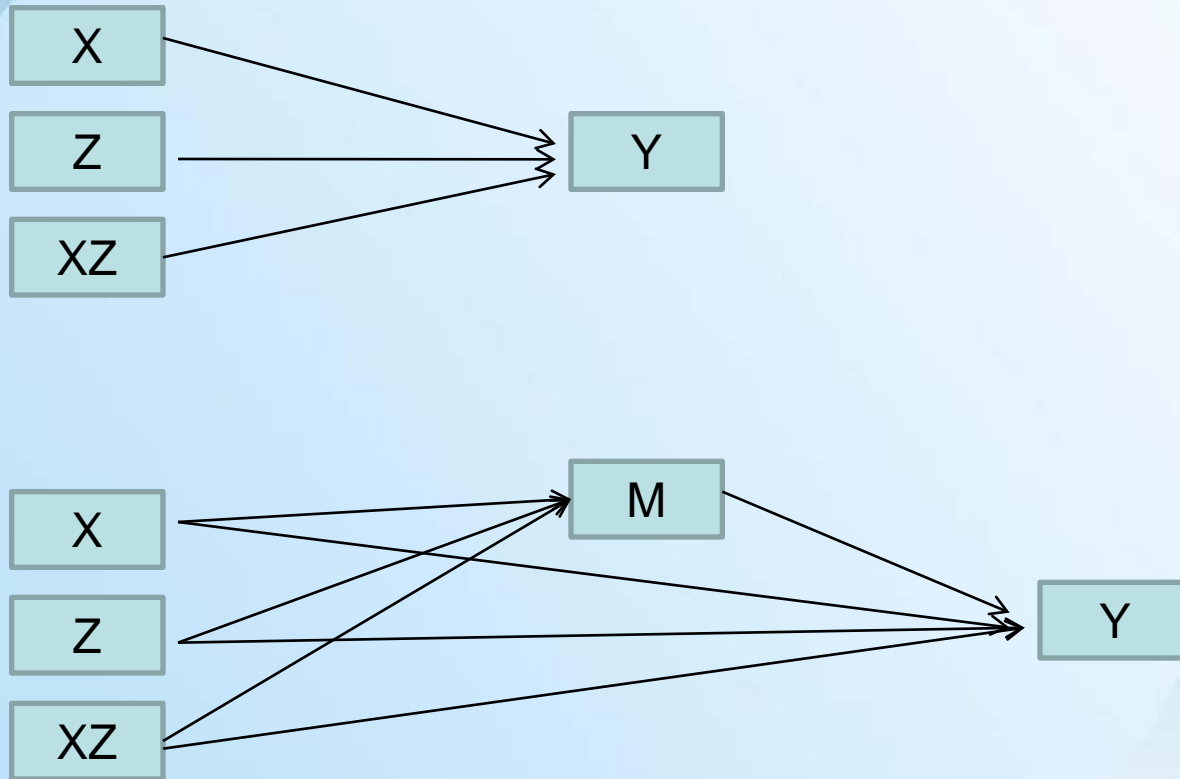- You can try to reduce this problem by increasing sample size and/or centering and/or standardizing the variables

Extension

Moderation (more than one moderators)
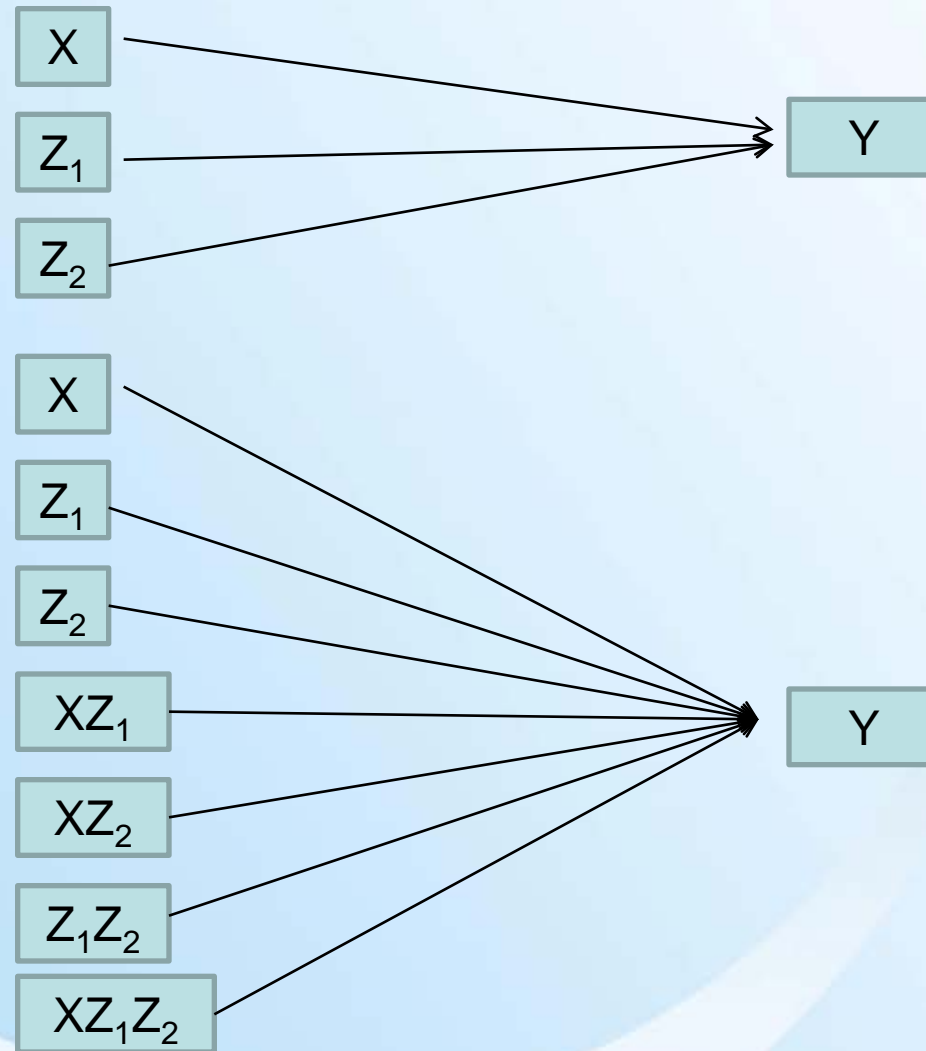
Extension

Mediated moderation
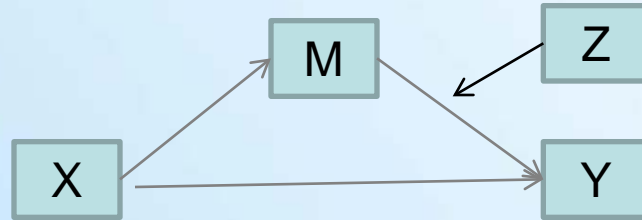
Extension

- Mediation
  - More than two mediators

$$X \longrightarrow M_1 \longrightarrow M_2 \longrightarrow Y$$

$$
\begin{array}{c}
M_1 \\
X \qquad\qquad\qquad Y \\
M_2
\end{array}
$$
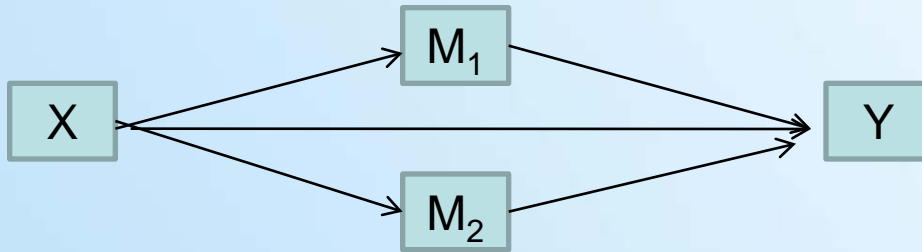
- Moderated mediator (Edwards and Lambert, 2007)

(1) first stage moderation model

$$
\begin{array}{c}
Z \longrightarrow M \\
X \longrightarrow Y
\end{array}
$$
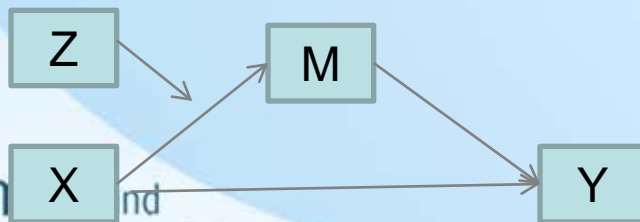
Extension

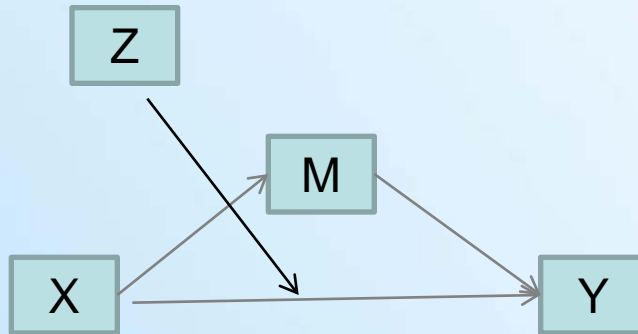## (2) Second Stage Moderation Model

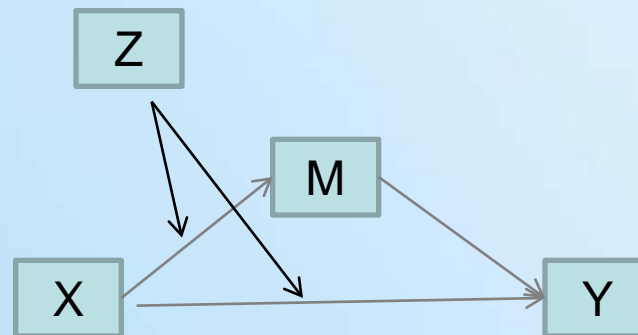## (3) First and Second Stage Moderation Model
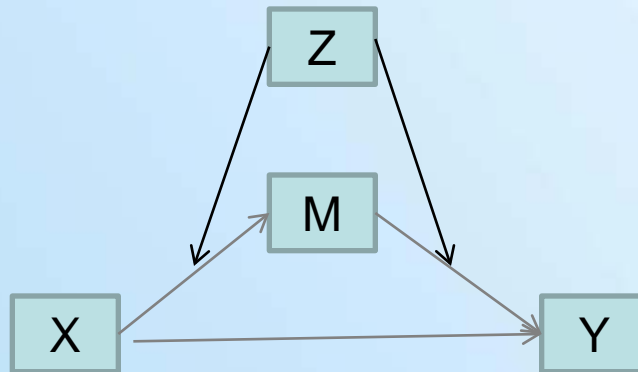
Extension
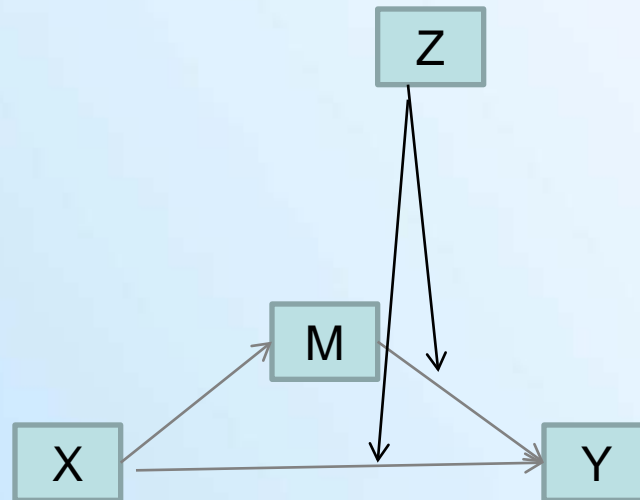
(4) Direct effect moderation model
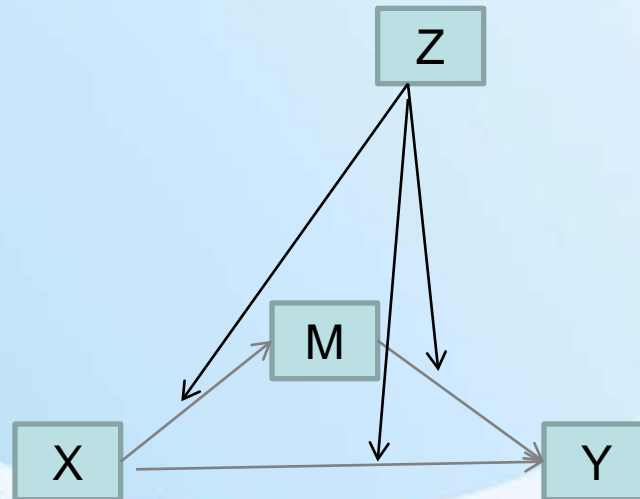

(5) Direct Effect and First Stage Moderation Model

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.

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, 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.

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

36