Interactions

June 7, 2010

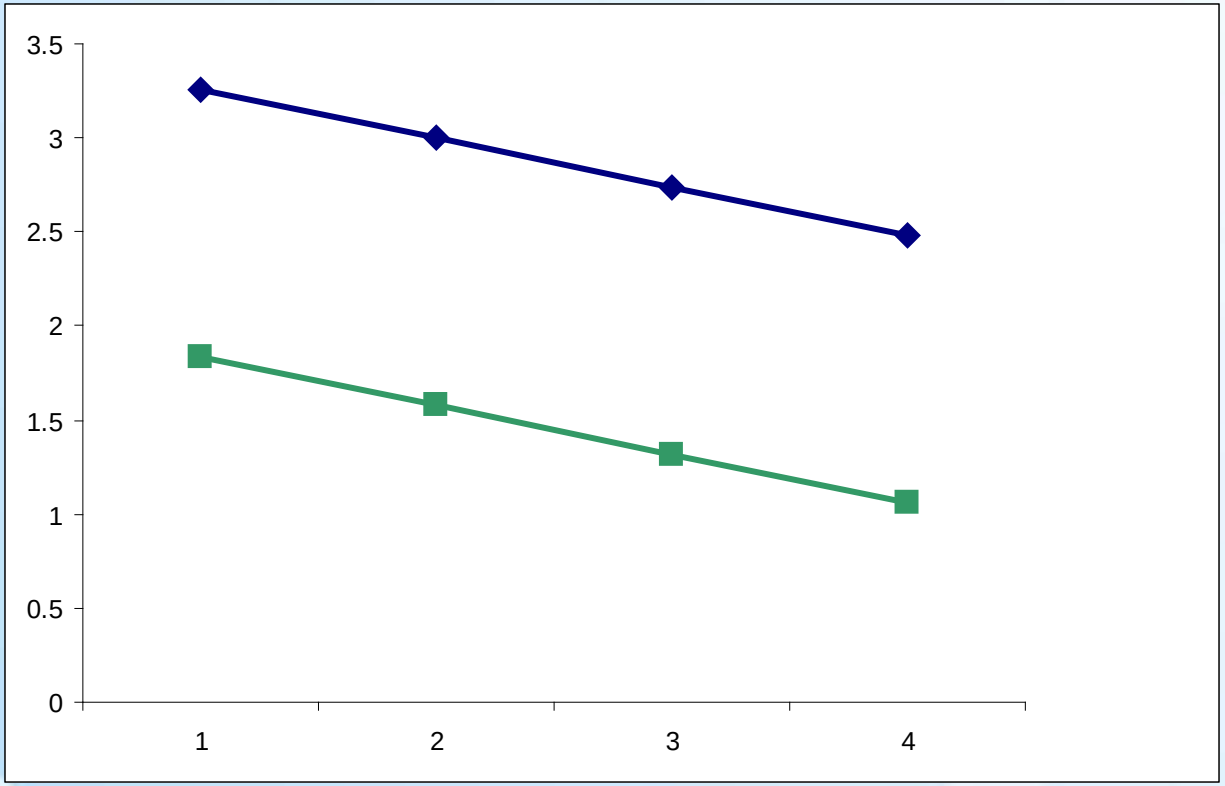
CFDR Workshops

What is an Interaction?

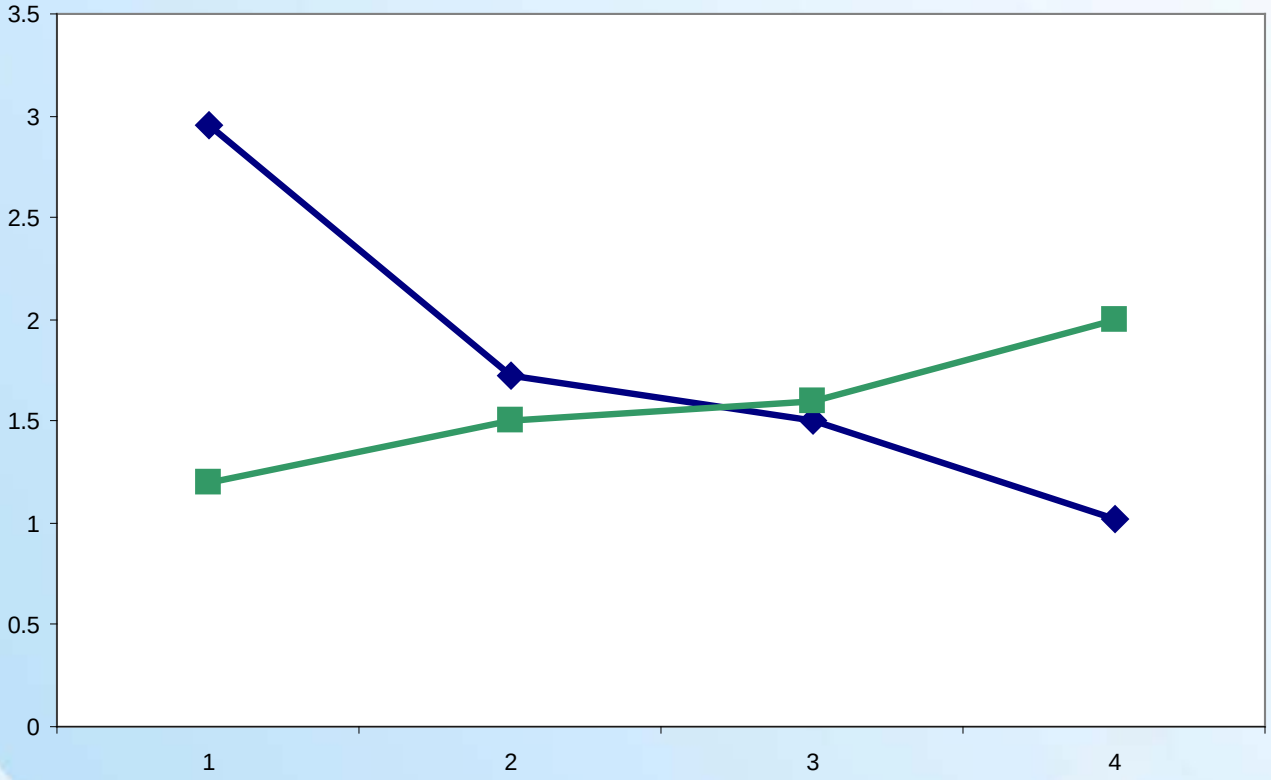
- The association between X_1 and Y varies according to levels of X_2. (Agresti and Finlay 1997)

- The association between gender and number of sexual partners varies according to levels of race
  - There is an interaction between gender and race in their effects on number of sexual partners

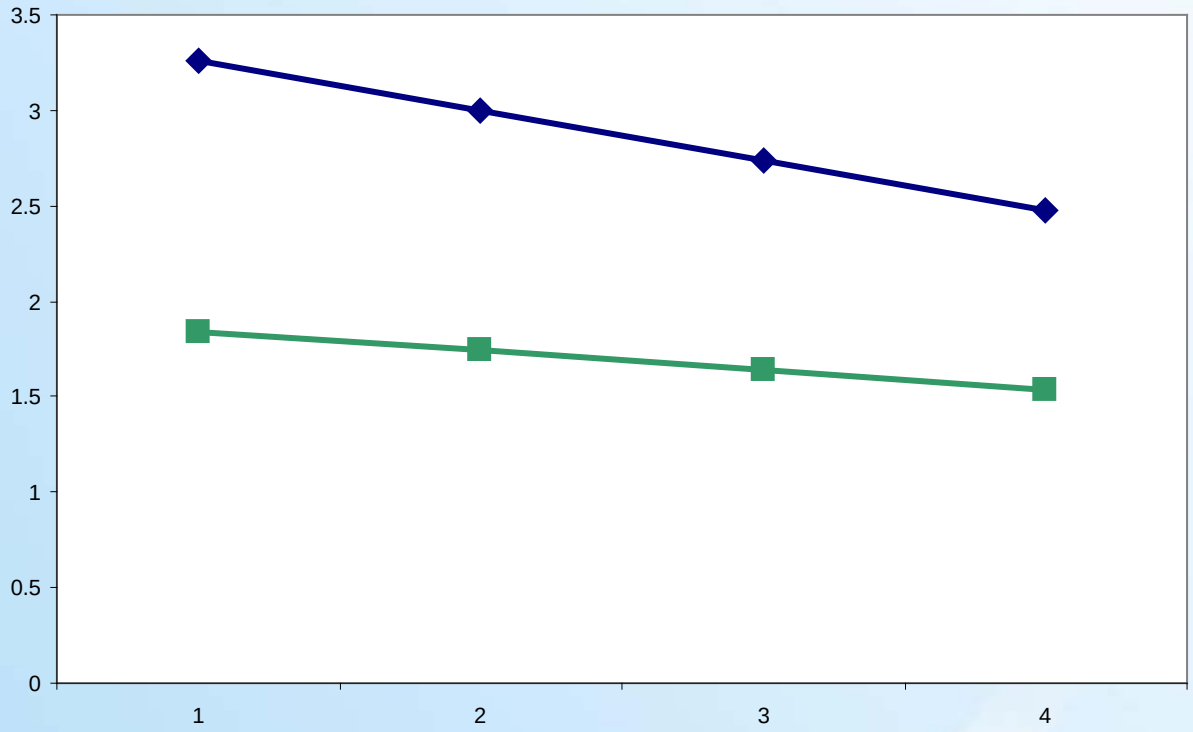
NOT and Interaction



An Interaction



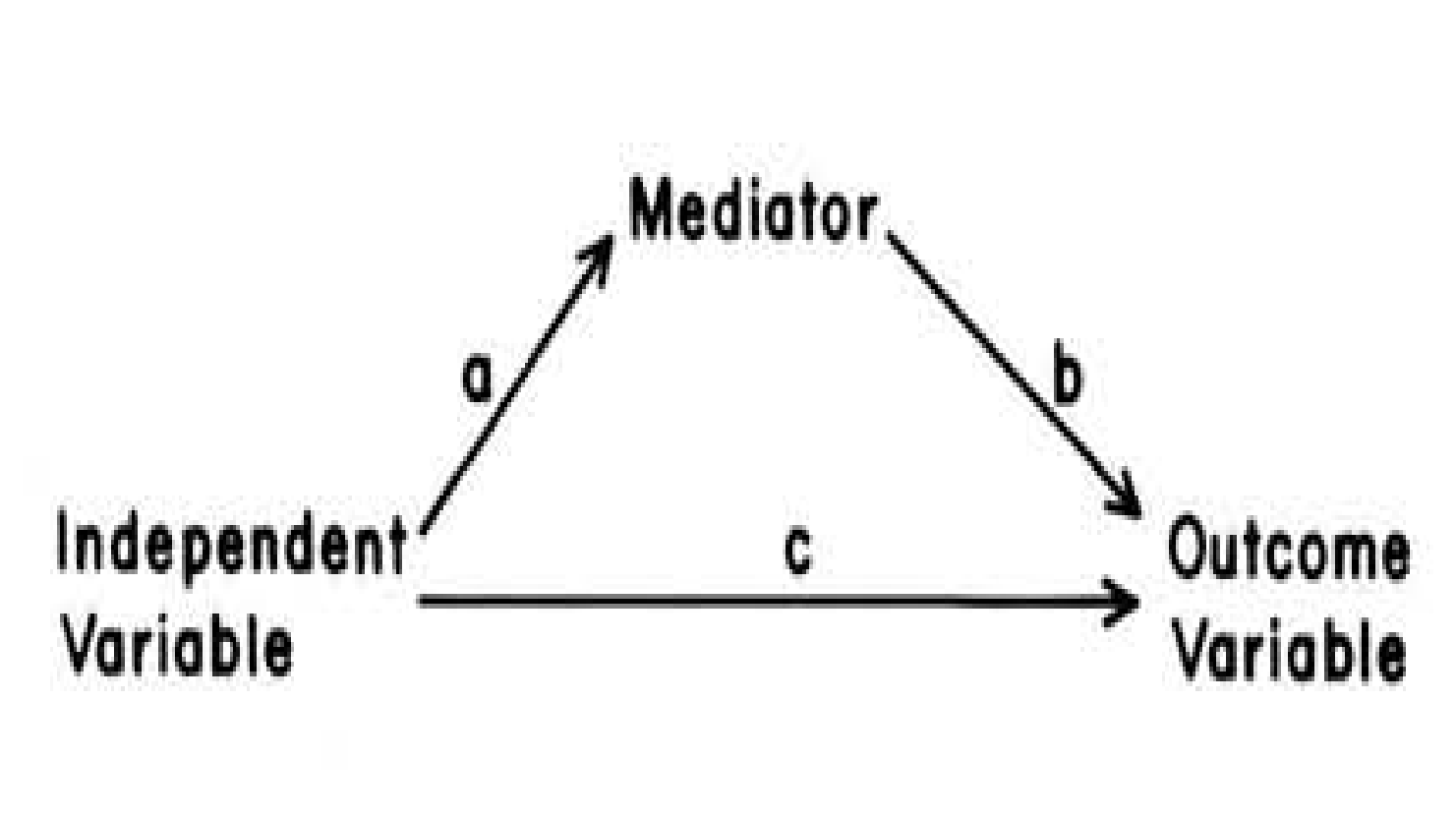
This is also an interaction



Mediating versus Moderating

- Mediating effect
  - "Impact of X on Y is *mediated* by Z if Z is partly or completely the mechanism by which X's effect on Y is realized" (DeMaris 2004, pg. 101)
    - X causes Z and Z then causes Y
    - Mediating effects explain how or why such effects occur (Baron and Kenney 1986)

Mediating versus Moderating



Mediating versus Moderating

- Moderating effect
  - Causal relationship between two variables changes as a *function* of the moderating variable
  - "The impact of X on Y is moderated by or conditioned on the level of Z" (DeMaris 2004, pg. 104)
  - Moderating effects are interactions

Mediating versus Moderating

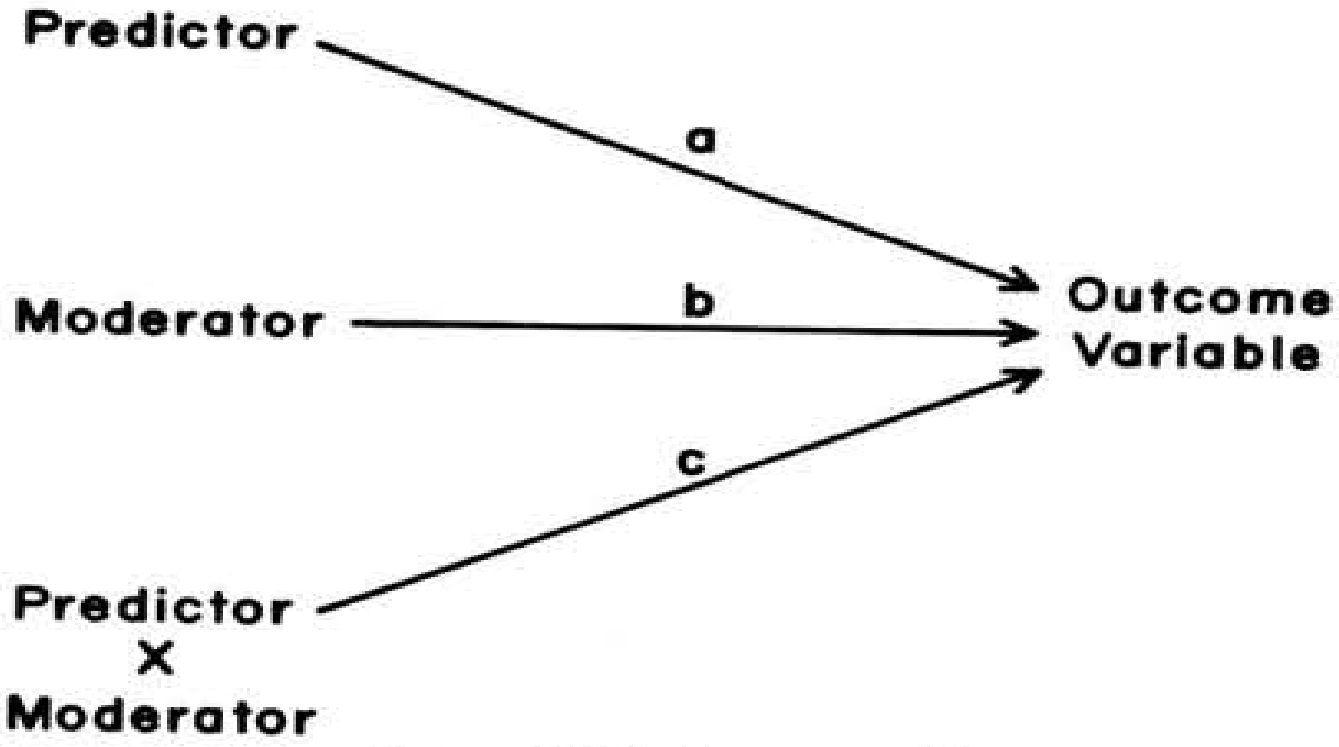


Figure 1. Moderator model.

How is an interaction modeled?

- Basic multiple regression - “main effects” model
  - $E(Y) = b_0 + b_1X + b_2Z + \ldots b_kX_k + e$
- Interaction model
  - $E(Y) = b_0 + b_1X + b_2Z + gXZ$

Is there an Interaction Effect?

- Do a nested F test using the original model and the model containing the interaction terms.
- Remember $F=\frac{(R^2_A-R^2_B)/\Delta df}{(1-R^2_A)/(n-K-1)}$
- If the test is significant you can conclude that there is an interaction.

How big is the effect of the interaction?

- Eta squared
  - $Eta^2 = \frac{SS(AxB)}{SS(T)}$

Sum of squares for the interaction term

Sum of squares total

  - Multiply Eta^2 by 100 to get the percentage of variance in the dependent variable accounted for by the interaction effect

Interpreting the Main Effect

- When an interaction is included in the model, the main effects are interpreted as the effect of x_1 when x_2 is 0
  - Example: Your two independent variables are age and years of school. Your main effect would be the effect of age on income when schooling is 0 or the effect of schooling on income when age is 0. (Allison 1999)
- When an interaction is in the model, you typically do not worry about the significance of the main effects (Allison 1999, pg. 168).

Centering Continuous Variables

- **What is centering and how does it help me?**
  - Centering a variable is to subtract it's mean from it, therefore the centered variable (Xc) will have a mean of zero
  - Advantages of centering
    - Easier to interpret
    - Reduces multicollinearity
      - This has been called into question (Echambadi & Hess 2007)

Interpreting the Interaction

- Calculate the slope of Y on X1 at different values of X2 (Jaccard, Turrisi, & Wan 1990)
- Values should be guided by theory, but without theory choose low (one standard deviation below the mean X2 score), medium (the mean), and high (one standard deviation above the mean) values

Interpreting the Interaction

- Factor the equation to isolate common multipliers. This finds the impact of X or its partial slope (DeMaris 2004: 104).
  - Interaction Model:
    - $E(Y) = b_0 + b_1X + b_2Z + gXZ$
  - Partial Slope: $E(Y) = b_0 + b_2Z + (b_1 + gZ)X$
  - $(b_1 + gZ)$ is the partial slope of X
    - Partial slope has a unit-impact interpretation, similar to a non-interaction model

Interpreting interactions in Logistic models

- Similar to OLS but the interpretation is in log odds of the event.

What if I want to know if there is a significant relationship between other variables?

- Simply change the omitted variable in the main effect and interaction term
  - $E(Y) = b_0 + b_1X + b_2Z + gXZ$

Interactions versus Chow Test

- Interactions test whether impact of one variable depends on the level of another variable.
- They DO NOT examine if there are differences between groups.
- Use a Chow test to determine if two groups (females/males, single/marrieds) have significantly different parameters from each other.

Chow Test

- Run the full model with your variable of interest in the model
- Next run two models: one for males and one for females (Note: take gender out of the model)
- Do an F-test
  - $F=\frac{[SSE_c-(SSE_1+SSE_2)]/p}{(SSE_1+SSE_2)/(n-2p)}$

*If the F-test is sig than the models are different for the two groups

Chow Test for Logistic

- $X^{2=}$-2 ln Lc- [(-2 ln L_1) + (-2 ln L_2)]

Table 1

	Model 1		
	B		SE
Intercept	2.74	***	0.411
Female	-0.094		0.567
Black	3.98	***	0.806
Hispanic	2.86	*	1.434
Other Race	5.97	***	1.413
BlackXFemale	-3.23	***	1.097
HispanicXFemale	-2.5		2.001
OtherRaceXFemale	-4.53	*	1.966

- The effect -.094 is the effect of female when all the race variables = 0 or white.
- The effect – 3.23 is the additional effect of being female when the respondent is black.
- What is the effect of a Hispanic male? Female?

	Model 1	
	B	SE
Intercept	-5.56 ***	1.08
Female	2.16	1.45
Age	0.29 ***	0.05
AgeXFemale	-0.012 †	0.07

- The effect 2.16 is the effect of female when age equals 0
- What is the effect of an 18 year old female?
- What is the effect of an 18 year male?

SAS Code

SAS - [SAS_code regressions 1-29-09.sas *]

File Edit View Tools Run Solutions Window Help

```
libname a "R:\CFDR\CFDR\Kristy\Interaction Workshop";run;
data a.workshop; set a.workshop;
whitexfemale=white*female;
blackxfemale=black*female;
hispxfemale=hisp*female;
otherrxfemale=otherr*female;run;

data a.workshop; set a.workshop;
agexfemale=rage4*female;run;
data a.workshop; set a.workshop;
proc reg data=a.workshop;
model csex=female black hisp otherr blackxfemale hispxfemale otherrxfemale;run;
proc reg data=a.workshop;
model csex= female age agexfemale;run;
proc logistic data=a.workshop descending;
model dv =female black hisp otherr blackxfemale hispxfemale otherrxfemale;run;
proc logistic data=a.workshop descending;
model dv= female age agexfemale;run;
```

Don't forget the descending in logistic regression

SPSS Code


```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT csex
  /METHOD=ENTER female black hisp otherr blackxfemale hispxfemale
  otherrxfemale  .

REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT csex
  /METHOD=ENTER female age agexfemale.

LOGISTIC REGRESSION VARIABLES  dv
  /METHOD = ENTER female age agexfemale
  /CONTRAST (female)=Indicator
  /PRINT = SUMMARY
  /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

LOGISTIC REGRESSION VARIABLES  dv
  /METHOD = ENTER female black hisp otherr blackxfemale hispxfemale
  /CONTRAST (female)=Indicator
  /PRINT = SUMMARY
  /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .
```

STATA Code

Dependent variable

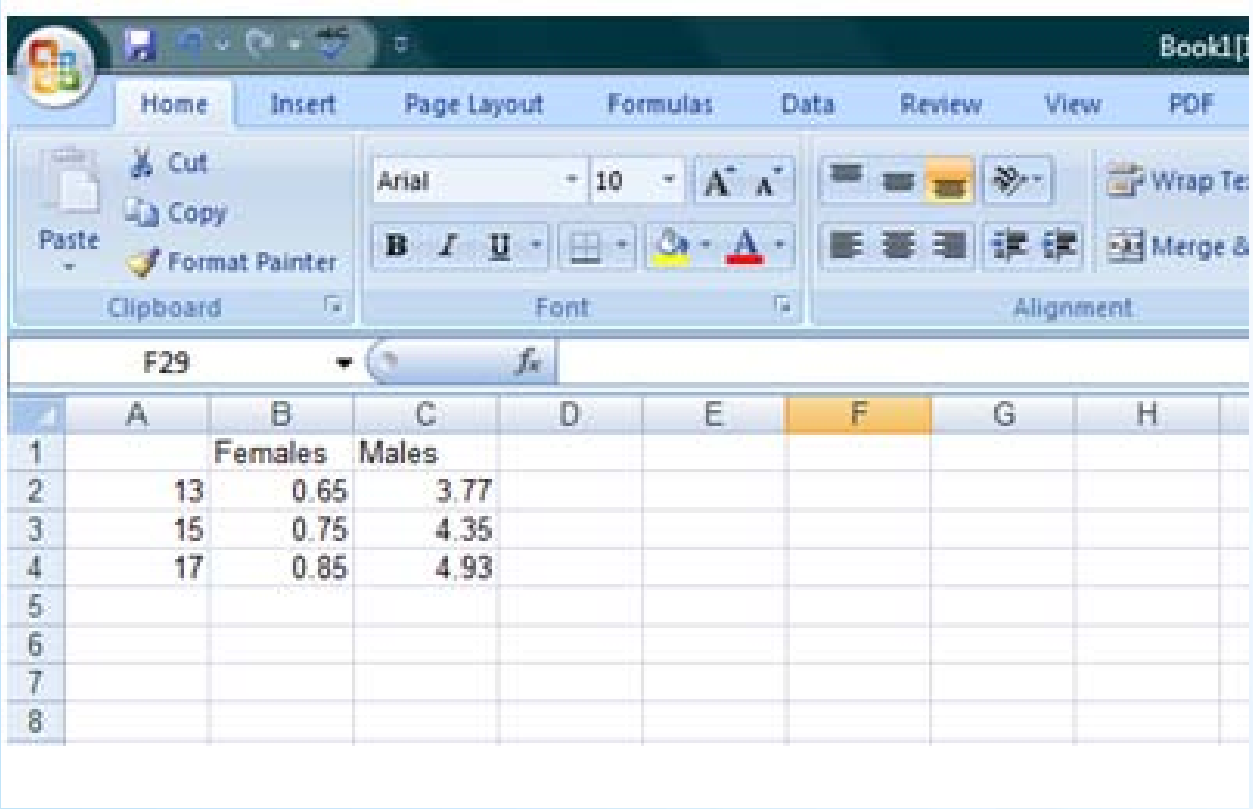
- regress csex female black hisp otherr blackxfemale hispxfemale otherrxfemale
- regress csex female age agexfemale
- logistic dv female black hisp otherr blackxfemale hispxfemale otherrxfemale
- logistic dv female age agexfemale

Dependent variable

Stata Code

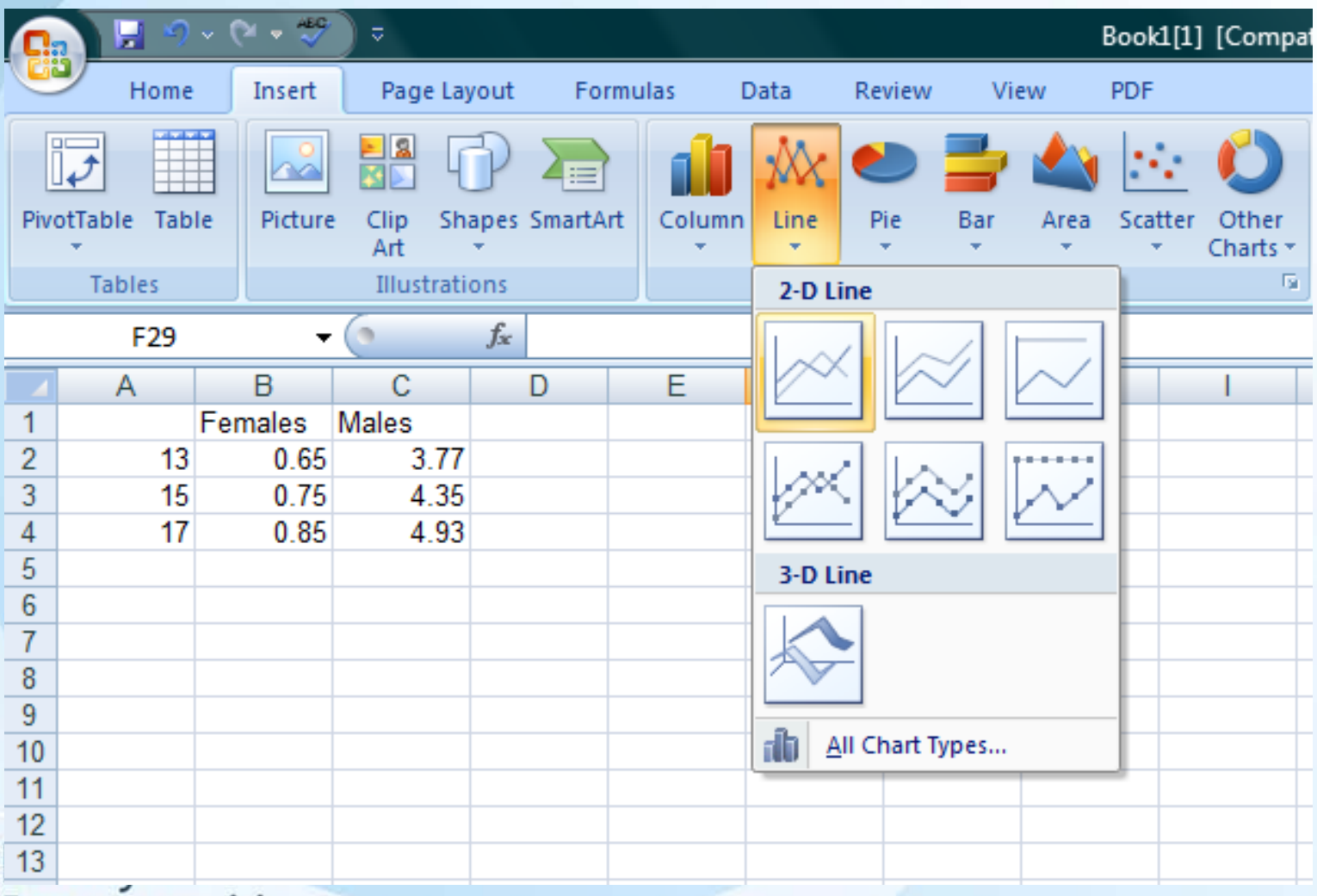
- Xi
- xi: regress csex i.female*i.religion
- xi: regress csex i.female*mtrust
- -creates dummy variables

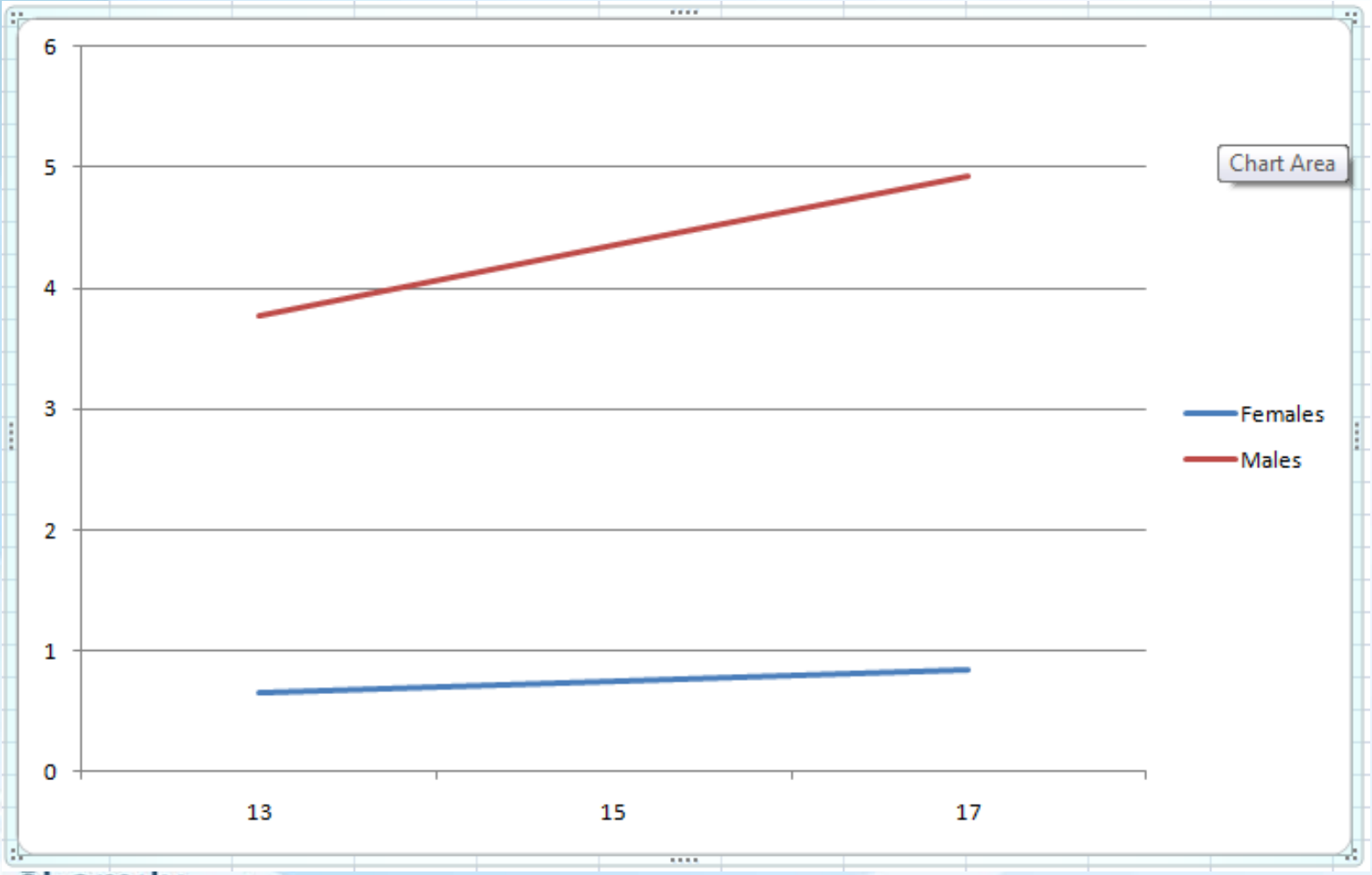
Graphing Interactions



	A	B	C	D	E	F	G	H
1		Females	Males					
2	13	0.65	3.77					
3	15	0.75	4.35					
4	17	0.85	4.93					
5								
6								
7								
8								

Graphing Interactions




6
5
4
3
2
1
0
13
15
17
Chart Area
Females
Males

Questions?