Interpreting (Even Tricky) Regression Coefficients
Dummy Coding Binary Predictors: Main Effects and Interactions

Karen Grace-Martin

Dummy Coding Binary Predictors: Main Effects and Interactions

1. The Basic Difference Between Dummy and Effect Coding

2. Dummy Coding Binary Variables

3. Interactions Between a Dummy-Coded Binary and a Numerical Predictor

4. Interpreting Coefficients with an Interaction: Binary Predictor is Key Variable

5. Interpreting Coefficients with an Interaction: Numerical Predictor is Key Variable

The Basic Difference Between Dummy and Effect Coding

Dummy and Effect Coding

$$E(Y|X) = \beta_0 + \beta_1 X$$

How do we code categories of X?

Dummy and Effect Coding

$$E(Y \mid X) = \beta_0 + \beta_1 X$$

Effect Coding
X = 1 for first group
X = -1 for second group

Dummy Coding
X = 1 for comparison group
X = 0 for reference group

Dummy and Effect Coding

Effect Coding:

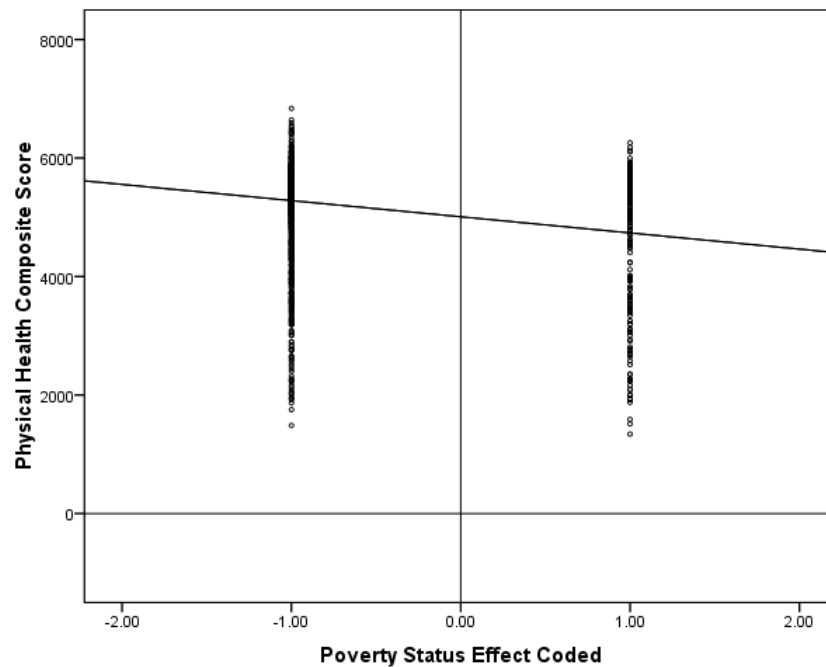Compare group means to the Grand Mean

Dummy Coding:

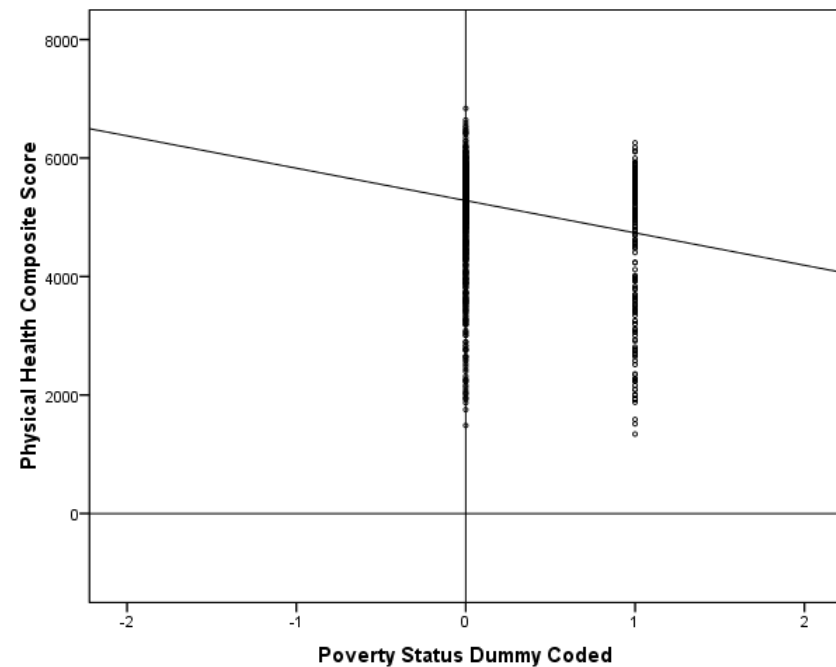Compare other group means to the mean of one group (reference, base group)
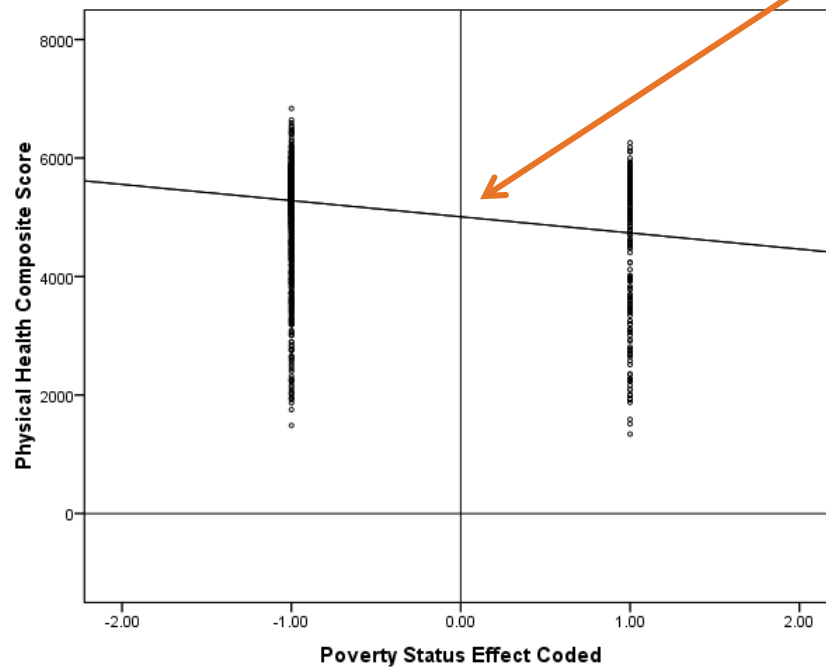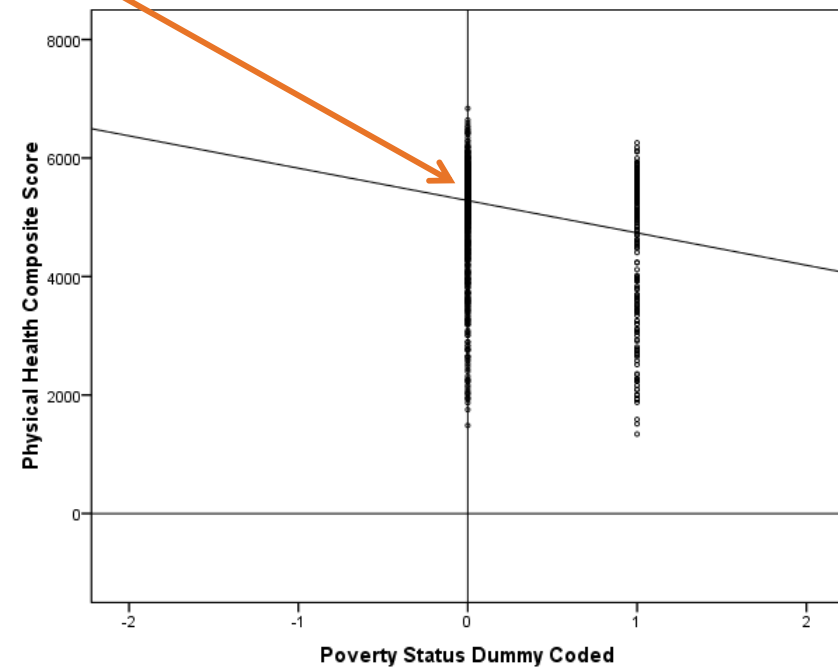
Dummy and Effect Coding

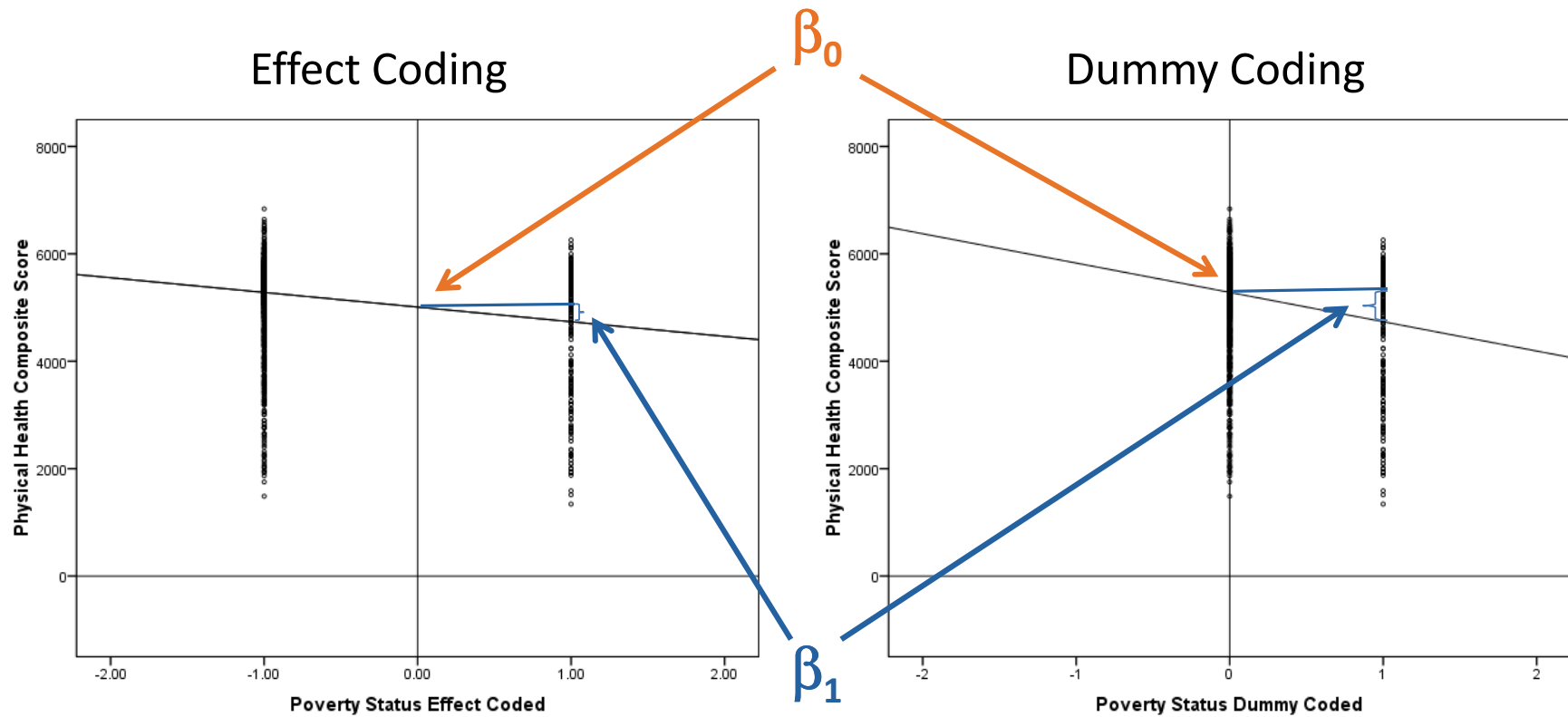# Dummy and Effect Coding

β_0

Effect Coding · Dummy Coding

Dummy and Effect Coding

Dummy Coding Binary Variables

Dummy Coding Binary Variables

	Original Code	Value Label	Dummy Code	N
Poverty Status	2.00	Poverty	1	266
	3.00	Not in Poverty	0	1434

THE ANALYSIS FACTOR

Dummy Coding Binary Variables

One binary predictor: Poverty Status

ANOVA

Dependent Variable: PCS

Source	Type III Sum of Squares	df	Mean Square	F	p
Model	66877513.406[a]	1	66877513.406	102.343	.000
Poverty Status	66877513.406	1	66877513.406	102.343	.000
Error	1109582872.150	1698	653464.589		
Total	1176460385.555	1699			

Dummy Coding Binary Variables

One binary predictor: Poverty Status

Marginal Means

Dependent Variable: PCS

Poverty Status	Mean	se
Poverty	4736.812	49.564
Not in Poverty	5282.757	21.347

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-545.945*	53.966	.000

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5282.757	21.347	247.471	.000
Poverty	-545.945	53.966	-10.116	.000
Not In Poverty	0[a]	.	.	.

a. This parameter is set to zero because it is redundant.

Dummy Coding Binary Variables

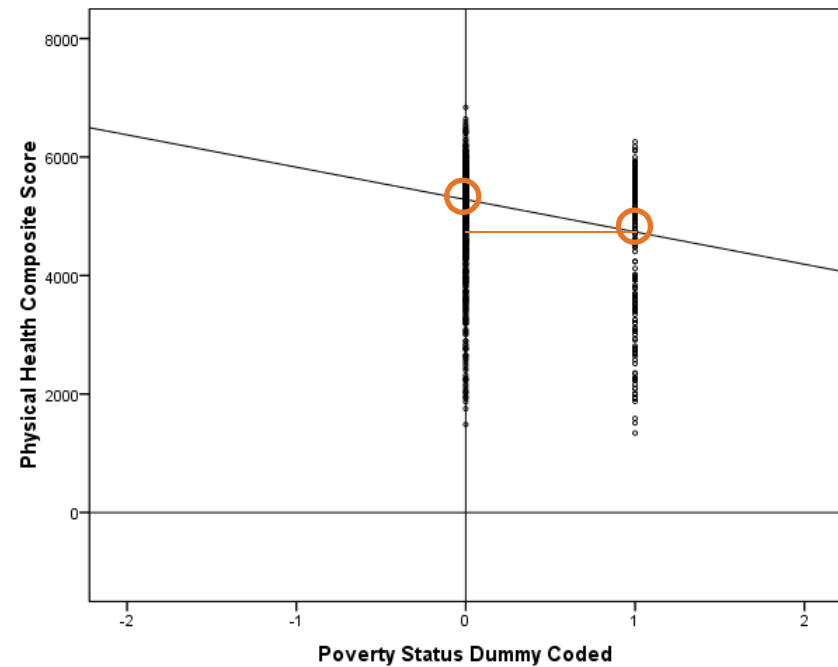## One binary predictor: Poverty Status

Marginal Means

Dependent Variable: PCS

Poverty Status	Mean	se
Poverty	4736.812	49.564
Not in Poverty	5282.757	21.347

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5282.757	21.347	247.471	.000
Poverty	-545.945	53.966	-10.116	.000
Not In Poverty	0[a]	.	.	.

a. This parameter is set to zero because it is redundant.

Dummy Coding Binary Variables

One binary predictor: Poverty Status
One continuous predictor: Depression

ANOVA

Dependent Variable: PCS

Source	Type III Sum of Squares	df	Mean Square	F	p
Model	162260943.378[a]	2	81130471.689	135.751	.000
Poverty Status	25730163.034	1	25730163.034	43.053	.000
Depression Centered	95383429.973	1	95383429.973	159.599	.000
Error	1014199442.177	1697	597642.571		
Total	1176460385.555	1699			

Dummy Coding Binary Variables

Regression Coefficients

Dependent Variable: PCS

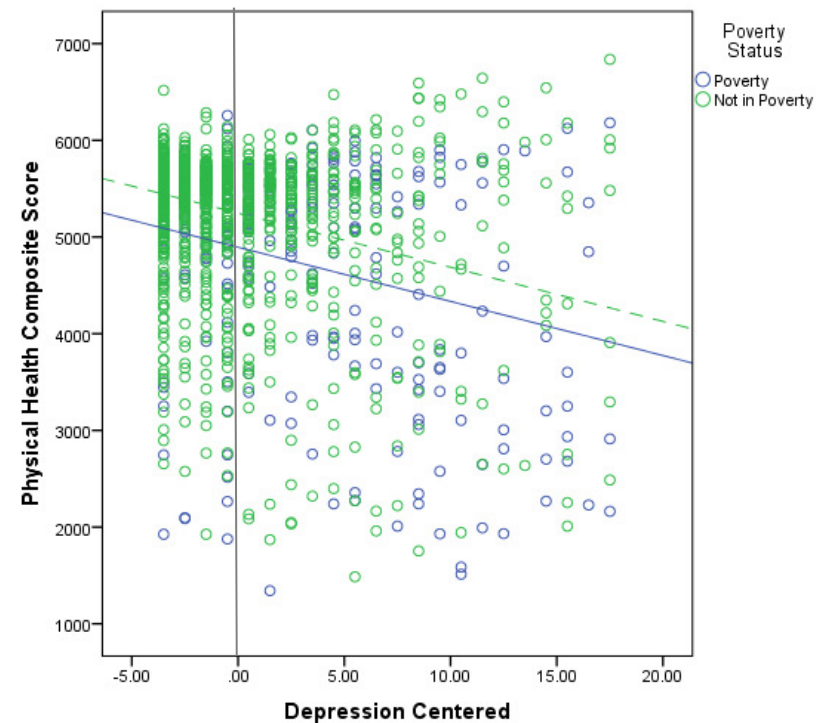| Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000
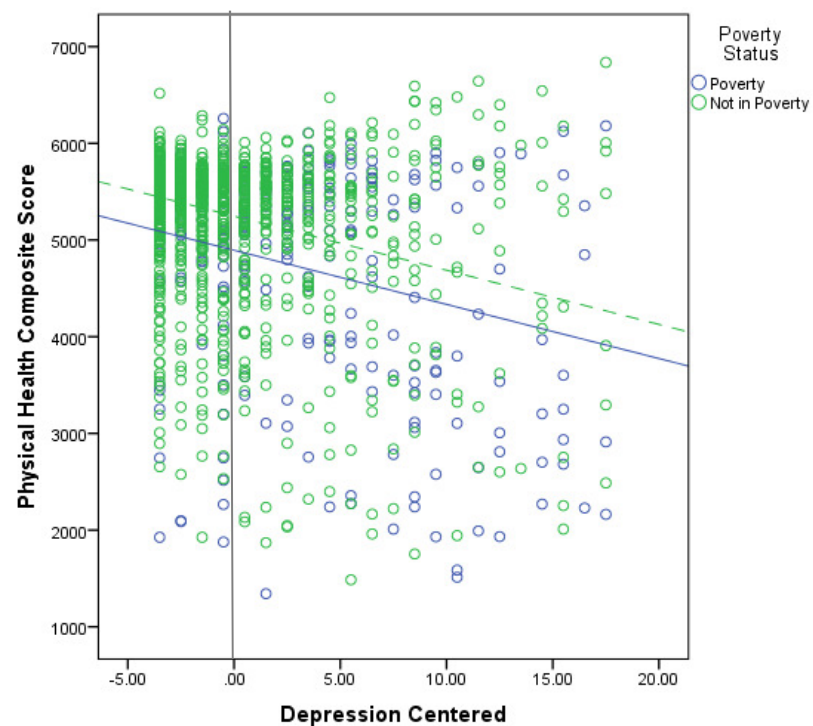
Dummy Coding Binary Variables

Why two parallel regression lines from this table?

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Dummy Coding Binary Variables

$$E(PHC|X) = 5247 - 353*Poverty - 56*Depression$$

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Poverty has two Values

1 = In Poverty
0 = Not In Poverty

Dummy Coding Binary Variables

$E(PHC|X) = 5247 - 353*Poverty - 56*Depression$

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Not in Poverty Group (Reference):

Poverty = 0

Dummy Coding Binary Variables

E(PHC|X) = 5247 − 353*Poverty − 56*Depression

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5247 − 353*0 − 56*Depression

        = 5247 − 56*Depression

Dummy Coding Binary Variables

E(PHC|X) = 5247 − 353*Poverty − 56*Depression

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Not in Poverty Group (Reference):   In Poverty Group (Comparison):

Poverty = 0   Poverty = 1

E(PHC|X) = 5247 − 56*Depression

Dummy Coding Binary Variables

$E(PHC|X) = 5247 - 353*Poverty - 56*Depression$

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Not in Poverty Group (Reference):

Poverty = 0

$E(PHC|X) = 5247 - 56*Depression$

In Poverty Group (Comparison):

Poverty = 1

$E(PHC|X) = 5247 - 353*1 - 56*Depression$

Dummy Coding Binary Variables

E(PHC|X) = 5247 − 353*Poverty − 56*Depression

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5247 − 56*Depression

In Poverty Group (Comparison):

Poverty = 1

E(PHC|X) = 5247 − 353*1 − 56*Depression

= 4894 − 56*Depression

Dummy Coding Binary Variables

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Difference in intercepts

Not in Poverty Group:
E(PHC|X) = 5247 – 56*Depression

In Poverty Group:
E(PHC|X) = 4894 – 56*Depression
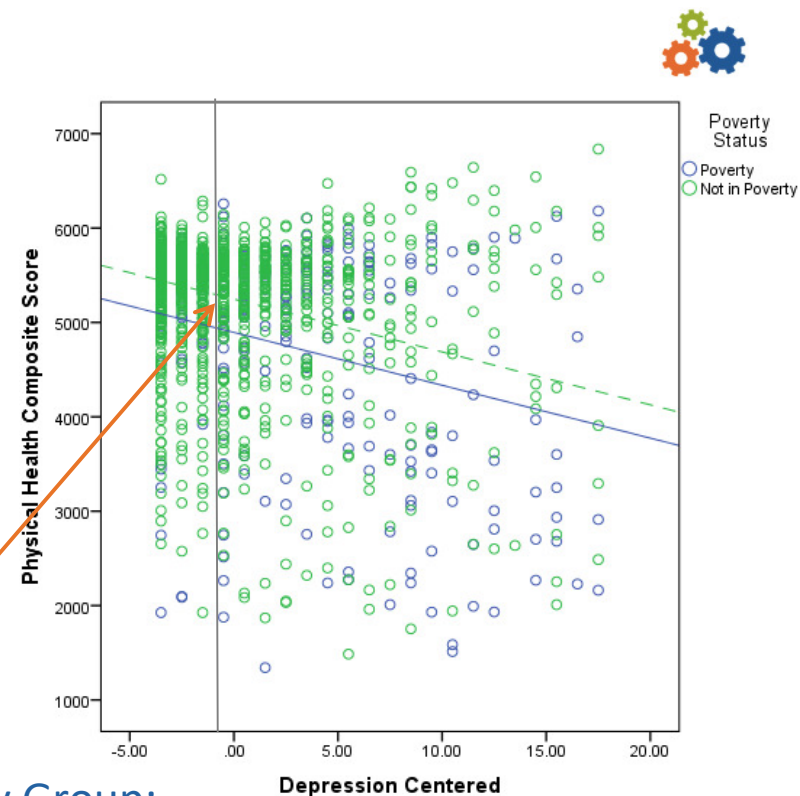
Dummy Coding Binary Variables

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5247.448	20.605	254.665	.000
Poverty	-353.131	53.819	-6.561	.000
Not In Poverty	0	.	.	.
Depression Centered	-56.011	4.434	-12.633	.000

Slope of both lines
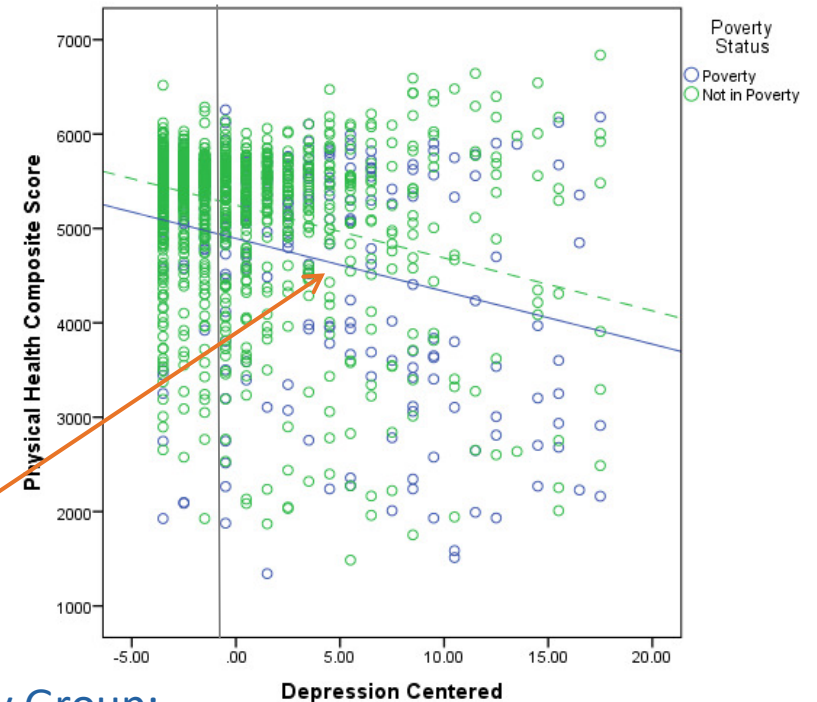
Not in Poverty Group:
E(PHC|X) = 5247 – 56*Depression

In Poverty Group:
E(PHC|X) = 4894 – 56*Depression

Interactions Between a Dummy-Coded Binary and a Numerical Predictor

Interactions Between a Binary and a Numerical Predictor

One binary predictor: Poverty Status
One numerical predictor: Depression
Interaction

ANOVA

Dependent Variable: PCS

Source	Type III Sum of Squares	df	Mean Square	F	p
Model	168549239.008[a]	3	56183079.669	94.539	.000
Poverty Status	15822101.072	1	15822101.072	26.624	.000
Depression Centered	96956714.818	1	96956714.818	163.148	.000
Poverty Status * Depression Centered	6288295.630	1	6288295.630	10.581	.001
Error	1007911146.547	1696	594287.233		
Total	1176460385.555	1699			

Interactions Between a Binary and a Numerical Predictor

One binary predictor: Poverty Status
One numerical predictor: Depression
Interaction

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Not In Poverty	0	.	.	.
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001
Not In Poverty * Depression Centered	0	.	.	.

Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Why two non-parallel regression lines from this table?
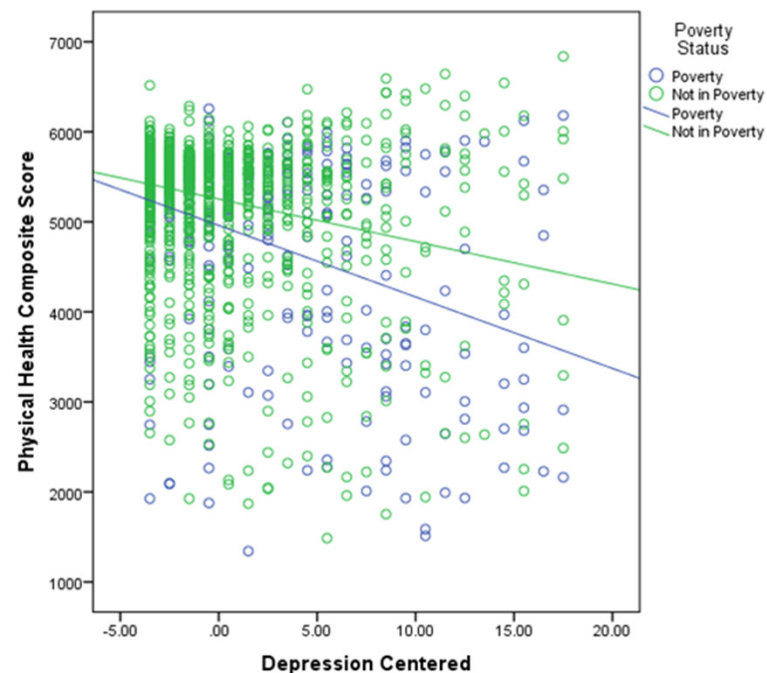
Interpreting Coefficients With an Interaction Between a Dummy-Coded Binary and a Numerical Predictor

Numerical Predictor is Key Variable
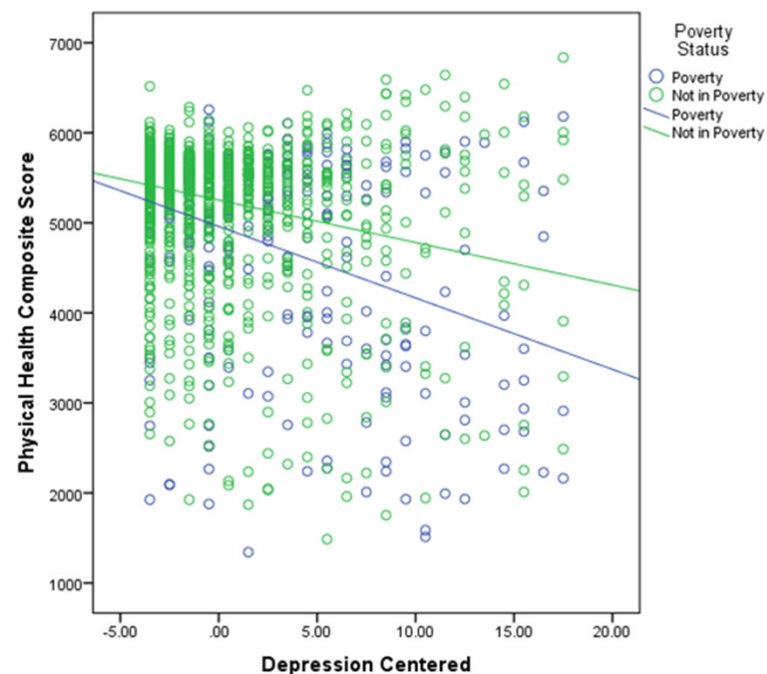
Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Focus on Slopes

Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):
Poverty = 0

Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5253 - 293*0 – 47*Depression – 32*0*Depression

Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5253 - 293*0 – 47*Depression – 32*0*Depression

E(PHC|X) = 5253 – 47*Depression

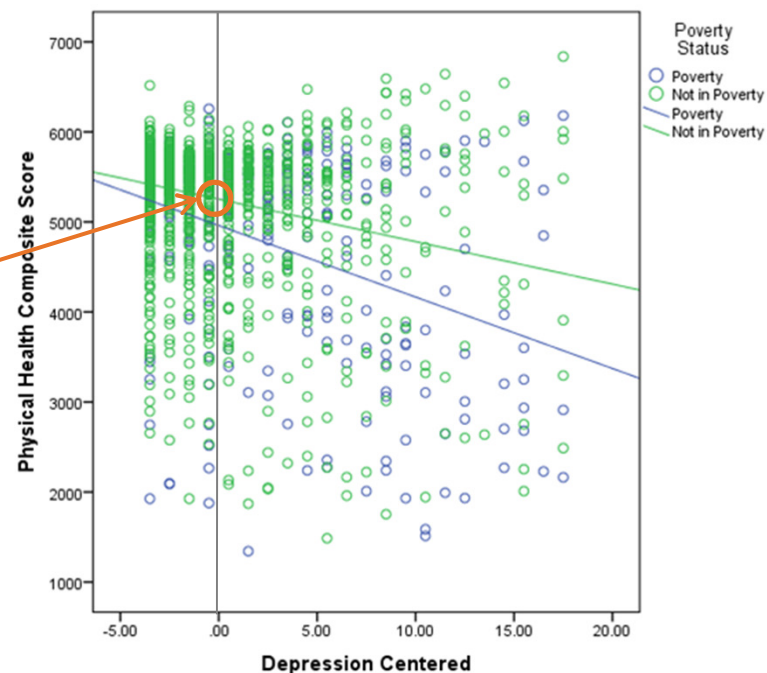# Interactions Between a Binary and a Numerical Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Reference Group Intercept

Not in Poverty Group (Reference):

E(PHC|X) = 5253 – 47*Depression

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients
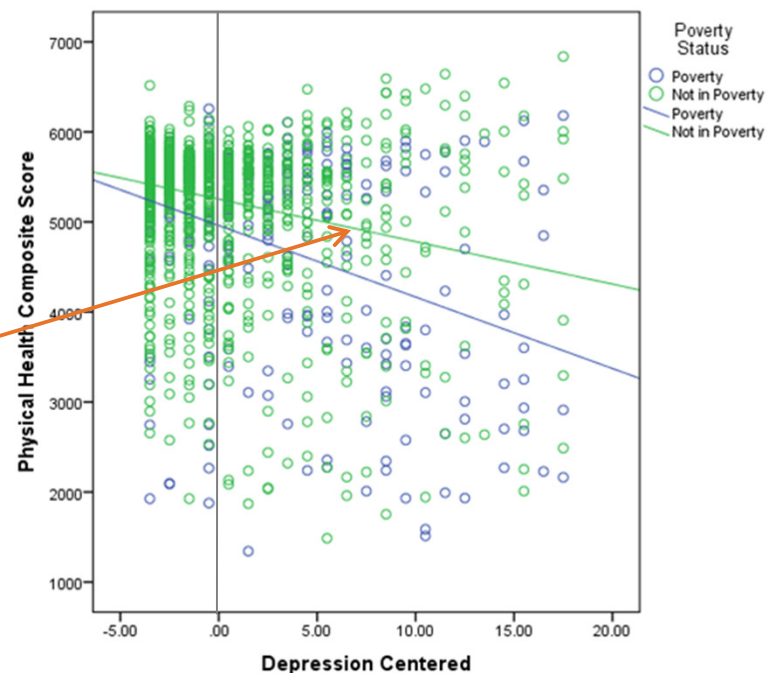
Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Reference Group Slope of Depression

Not in Poverty Group (Reference):

E(PHC|X) = 5253 – 47*Depression

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):
Poverty = 0

In Poverty Group (Comparison):
Poverty = 1

E(PHC|X) = 5253 - 47(Depression)

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5253 - 47(Depression)

In Poverty Group (Comparison):

Poverty = 1

E(PHC|X) = [5253 – 293*1] – [47*Depression + 32*1*Depression]

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

E(PHC|X) = 5253 - 293*Poverty - 47*Depression - 32*Poverty*Depression

Not in Poverty Group (Reference):

Poverty = 0

E(PHC|X) = 5253 - 47(Depression)

In Poverty Group (Comparison):

Poverty = 1

E(PHC|X) = [5253 – 293*1] – [47*Depression + 32*1*Depression]

= 4960 – 79*Depression

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Difference in Intercepts

Not in Poverty Group (Reference):
E(PHC|X) = 5253 - 47(Depression)

In Poverty Group (Comparison):
E(PHC|X) = 4960 – 79*Depression
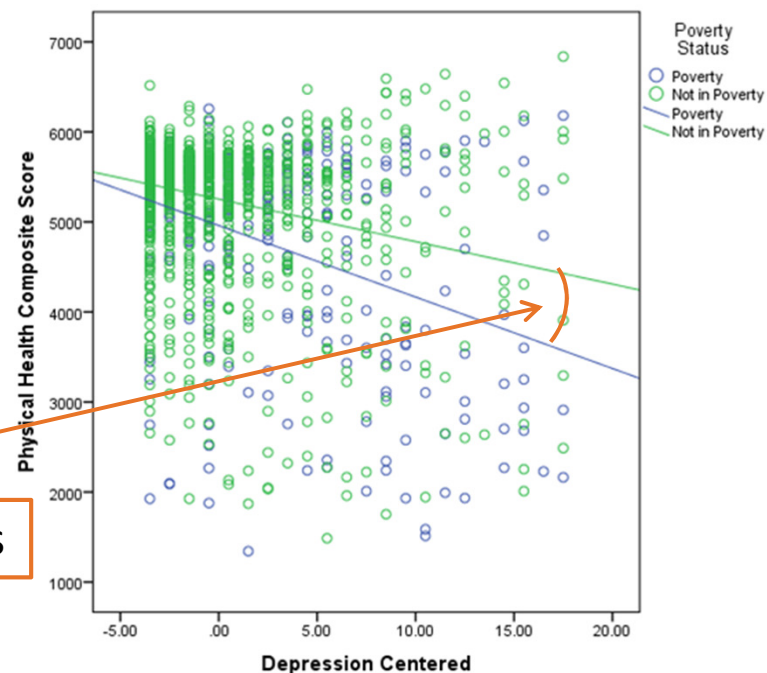
Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Difference in Slopes

Not in Poverty Group (Reference):
E(PHC|X) = 5253 - 47(Depression)

In Poverty Group (Comparison):
E(PHC|X) = 4960 – 79*Depression

Interpreting Coefficients With an Interaction Between a Dummy-Coded Binary and a Numerical Predictor

Binary Predictor is Key Variable

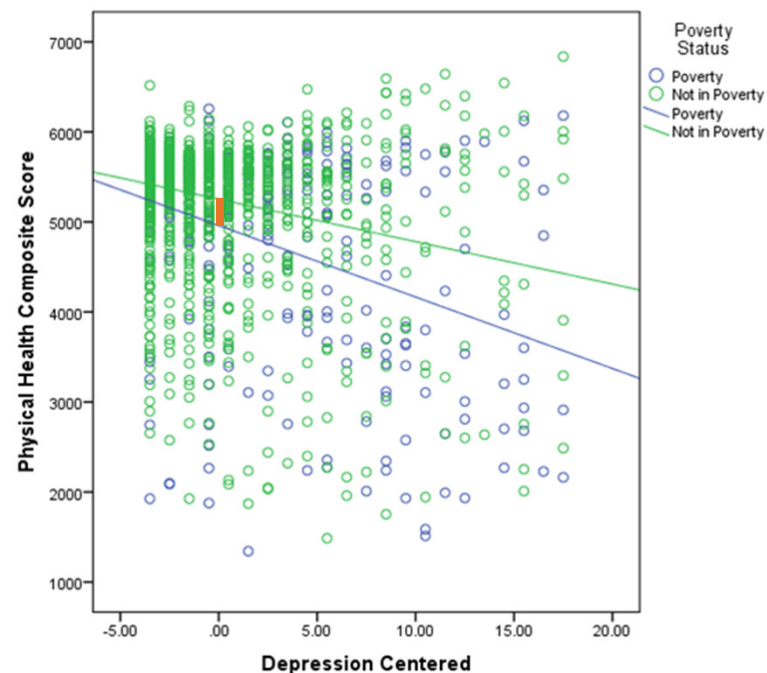# Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Focus on mean differences at specific values of Depression

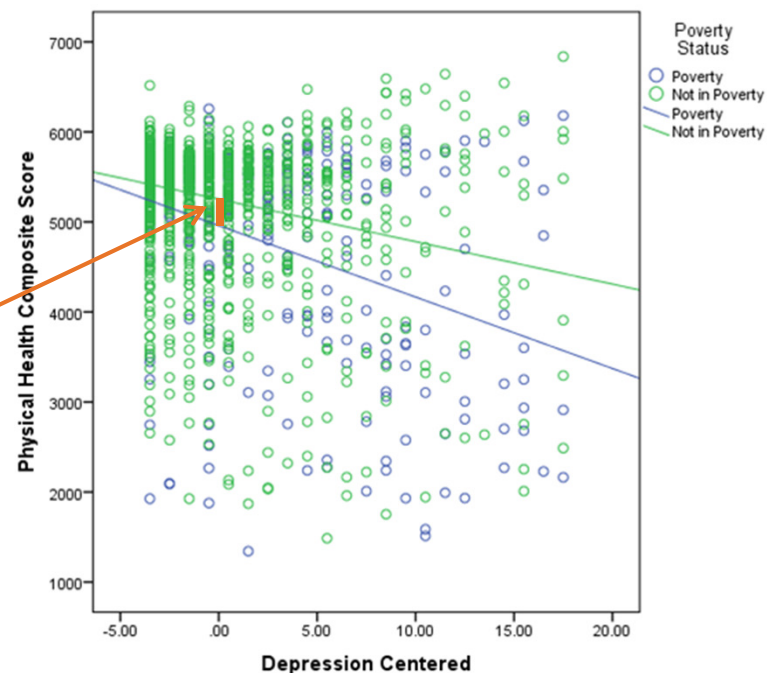# Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Difference in
E(PCS|Depression = Mean)

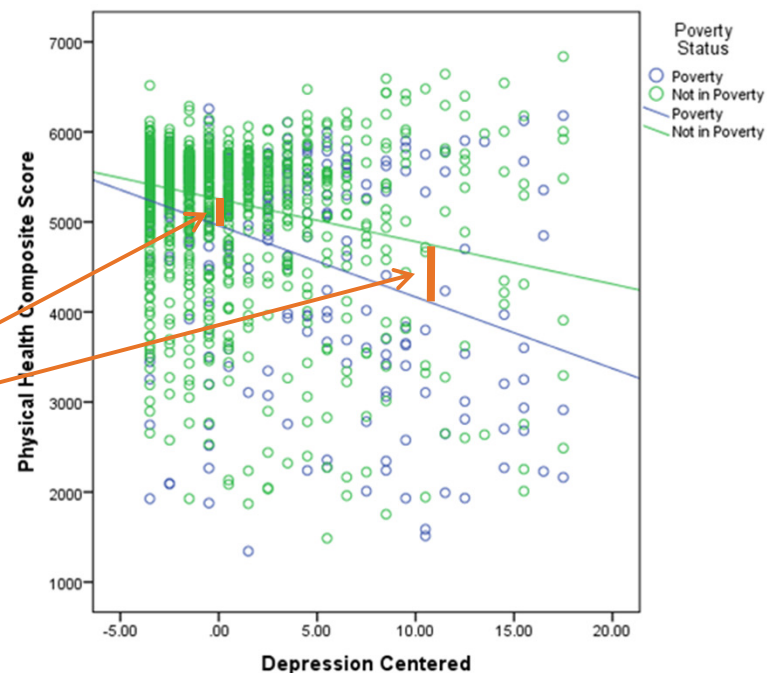# Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

How much different are the mean PCS differences between In Poverty and Not in Poverty groups when
Depression = mean
            and
Depression = mean + 10?

Interactions Between a Binary and a Continuous Predictor
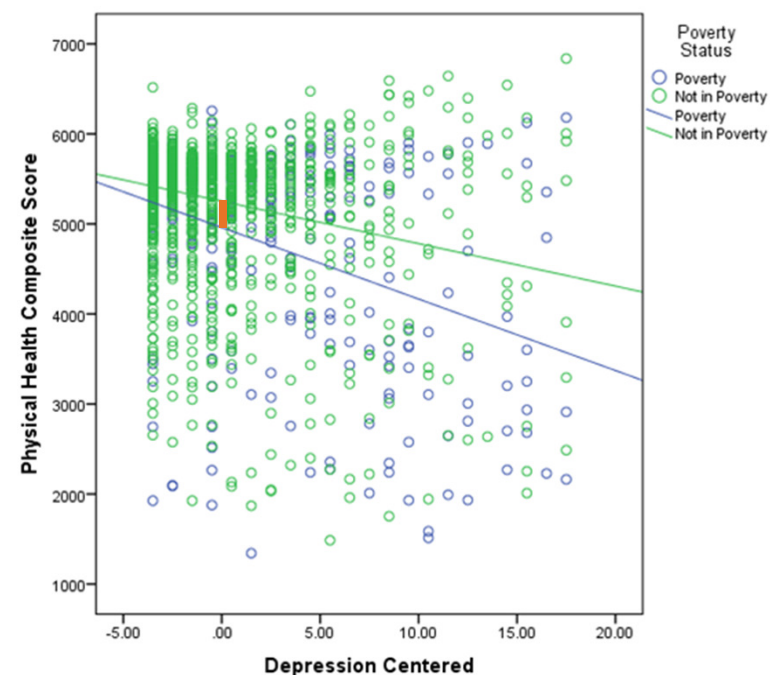
Marginal Means

Dependent Variable: PCS

Poverty Status	Mean	se
Poverty	4960.099[a]	52.893
Not in Poverty	5253.021[a]	20.619

a. Covariates appearing in the model are evaluated at the following values: CESDCen = .00

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-292.922*	56.770	.000

Interactions Between a Binary and a Continuous Predictor
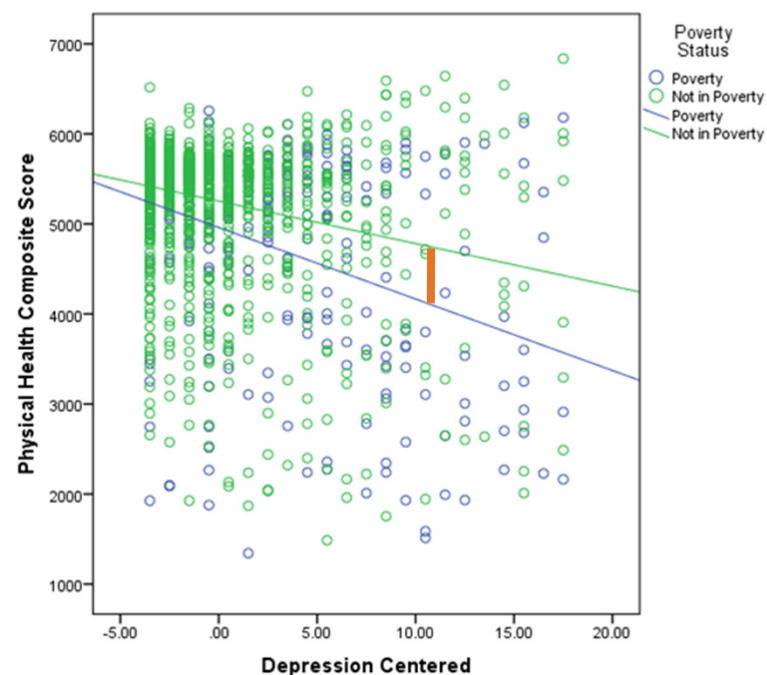
Marginal Means

Dependent Variable: PCS

Poverty Status	Mean	se
Poverty	4166.056[a]	76.917
Not in Poverty	4781.324[a]	58.806

a. At Depression Centered = 10

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-615.268*	96.821	.000

Interactions Between a Binary and a Continuous Predictor

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-292.922*	56.770	.000

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-615.268*	96.821	.000

Difference in mean differences when depression increases by 10:

-292.922 - 615.268 = -322.346
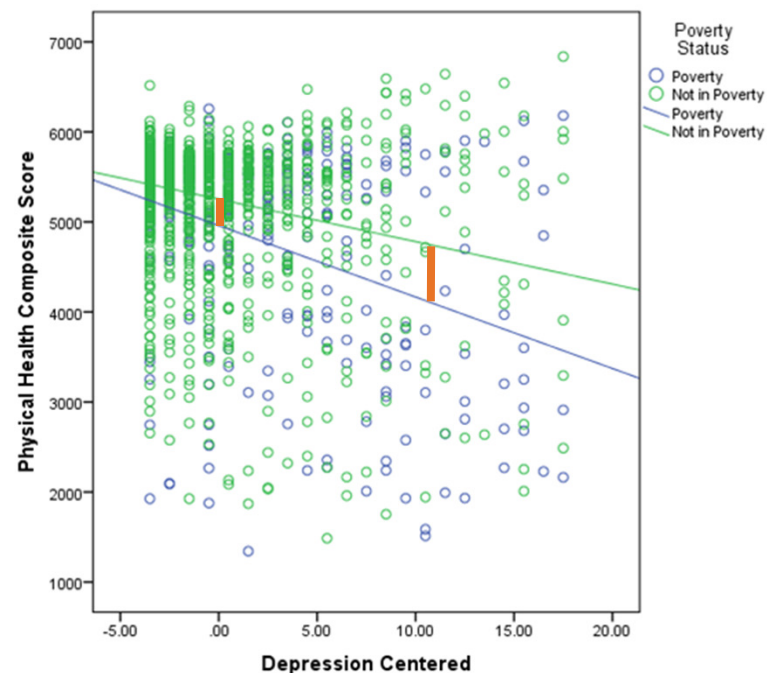
Interactions Between a Binary and a Continuous Predictor

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-292.922*	56.770	.000

Pairwise Comparisons

Dependent Variable: PCS

Comparison	Mean Difference	se	p
In Poverty - Not in Poverty	-615.268*	96.821	.000

Difference in mean differences when depression increases by 10:

-292.922 - 615.268 = -32.235*10
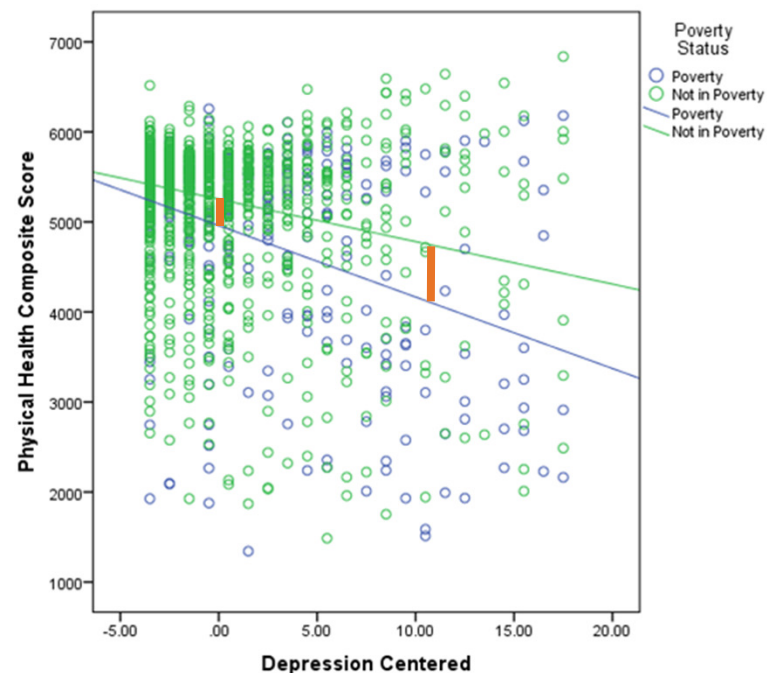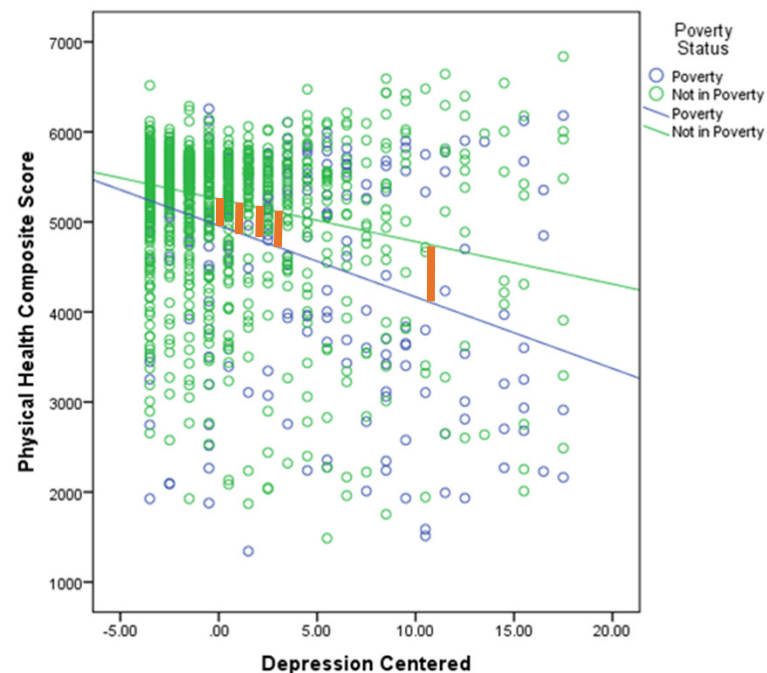
Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Difference in mean differences when depression increases by 1:

= -32.235

THE ANALYSIS
F A C T O R

Interactions Between a Binary and a Continuous Predictor

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	5253.021	20.619	254.770	.000
Poverty	-292.922	56.770	-5.160	.000
Depression Centered	-47.170	5.190	-9.089	.000
Poverty * Depression Centered	-32.235	9.910	-3.253	.001

Regression Coefficients

Dependent Variable: PCS

Variable	B	se	t	p
Intercept	4781.324	58.806	81.307	.000
Poverty	-615.268	96.821	-6.355	.000
Depression Centered at 13.5	-47.170	5.190	-9.089	.000
Poverty * Depression Centered at 13.5	-32.235	9.910	-3.253	.001