PSC 202
SYRACUSE UNIVERSITY

INTRODUCTION TO POLITICAL ANALYSIS

MULTIPLE REGRESSION, PART 1

REMINDERS

- No sections this Friday
  - Take-home assignment instead
  - Due Dec 3 (Friday after we're back)
- New Problem Set will be posted on Friday
  - Also due Dec 3

LAST TIME

- Logic of control
  - What is the relationship between X and Y when we control for *one* confounder Z?
  - Idea: Look at the effect of X among the different values of Z

LAST TIME

Afghanistan war was beneficial

	Democrats	Not Democrats	Total
Agree	19.2% (10)	41.4% (12)	27.2% (22)
Disagree	80.8% (42)	58.6% (17)	72.8% (59)
Total	100% (52)	100% (29)	100% (81)

22.2%

CONTROLLED COMPARISON TABLE

Afghanistan war was beneficial

	Female			Male		
	Dem	Non-Dem	Total	Dem	Non-Dem	Total
	22.3%			21.7%		
Agree	18.9% (7)	41.2% (7)	25.9% (14)	20.0% (3)	41.7% (5)	29.6% (8)
Disagree	81.1% (30)	58.8% (10)	74.1% (40)	80.0% (12)	58.3% (7)	70.4% (19)
Total	100% (37)	100% (17)	100% (54)	100% (15)	100% (12)	100% (27)

- Partial effect of gender, "controlling for" gender

LAST TIME

1. Are all controlled effects zero or very close to zero?
   - Yes? ⇒ relationship between x and y is spurious
   - No? ⇒ either additive or interactive
2. Are all controlled effects approximately the same size?
   - Yes? ⇒ additive relationship
   - No? ⇒ interactive relationship

REMEMBER VARIABLE LEVELS

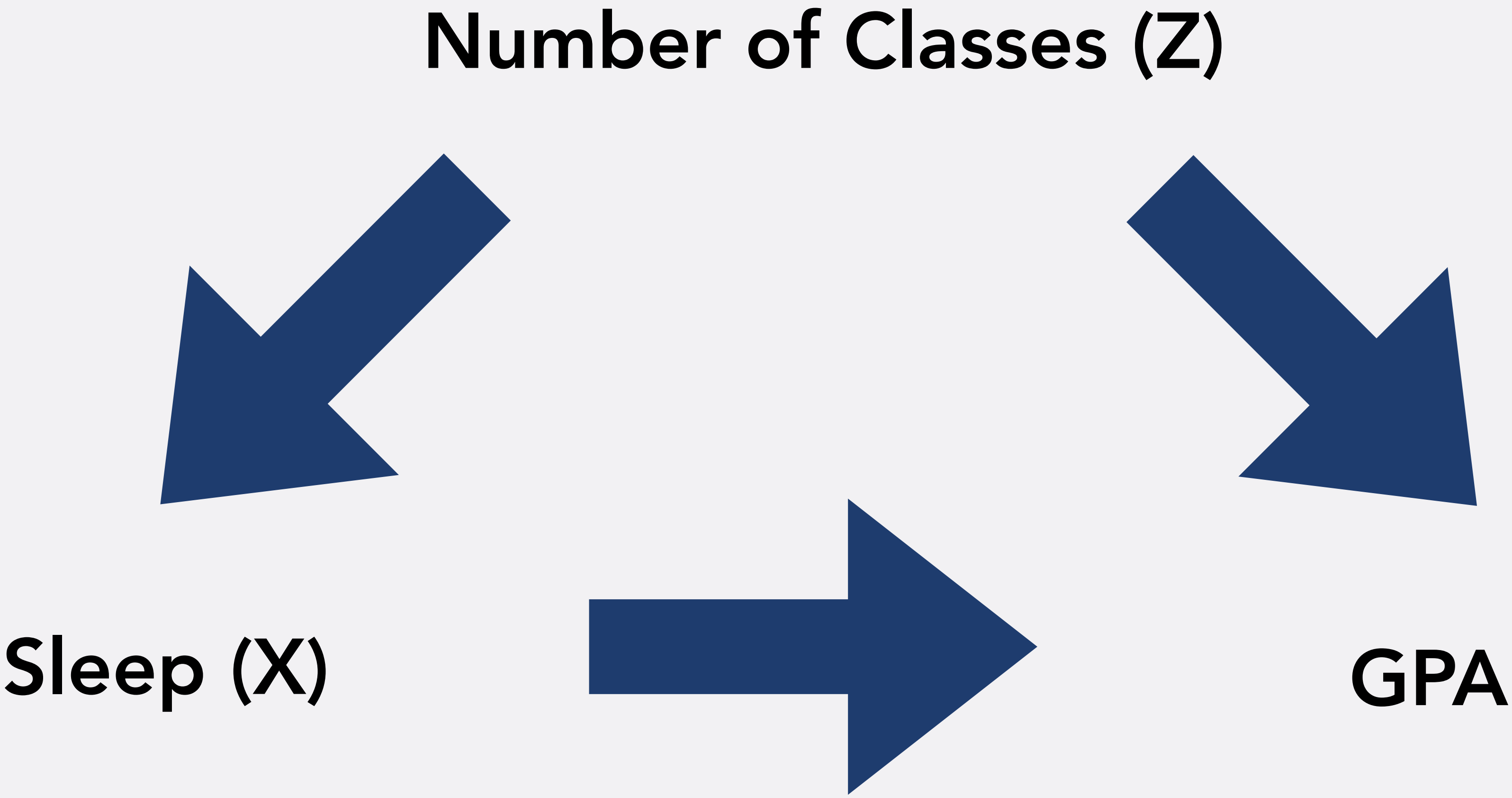
- So far: Dependent variable was nominal-level
- Now: DV is interval level
    - e.g. GPA
    - We use mean comparison
    - Determination if spurious, additive, interactive works just the same

ZERO-ORDER RELATIONSHIP

	More Than 7 Hours/Night		7 Or Fewer Hours/Night
Average Gpa	3.54 (40)	0.09	3.45 (42)

- Frequency in parentheses

STUDYING



- Spurious? Additive? Interactive?

CONTROLLED COMPARISON

	5 Or Fewer Classes		6 Or More Classes	
Sleep	More Than 7 Hours/Night	7 Or Fewer Hours/Night	More Than 7 Hours/Night	7 Or Fewer Hours/Night
Average Gpa	3.53 (30)	3.42 (29)	3.55 (10)	3.52 (13)

- Frequency in parentheses

HOW CAN WE TELL WHICH ONE?

1. Are all controlled effects zero or very close to zero?
   - Yes? ⇒ relationship between x and y is spurious
   - No? ⇒ either additive or interactive
2. Are all controlled effects approximately the same size?
   - Yes? ⇒ additive relationship
   - No? ⇒ interactive relationship

CONTROLLED COMPARISON

	5 Or Fewer Classes		6 Or More Classes	
Sleep	More Than 7 Hours/Night	7 Or Fewer Hours/Night	More Than 7 Hours/Night	7 Or Fewer Hours/Night
Average Gpa	$\boxed{3.53}$	$\boxed{3.42}$	$\boxed{3.55}$	$\boxed{3.52}$
	(30)	(29)	(10)	(13)
	0.11		0.03	

- Frequency in parentheses

HOW CAN WE TELL WHICH ONE?

1. Are all controlled effects zero or very close to zero?
    - Yes? ⇒ relationship between x and y is spurious
    - No? ⇒ either additive or interactive
2. Are all controlled effects approximately the same size?
    - Yes? ⇒ additive relationship
    - No? ⇒ interactive relationship

CONTROLLED COMPARISON

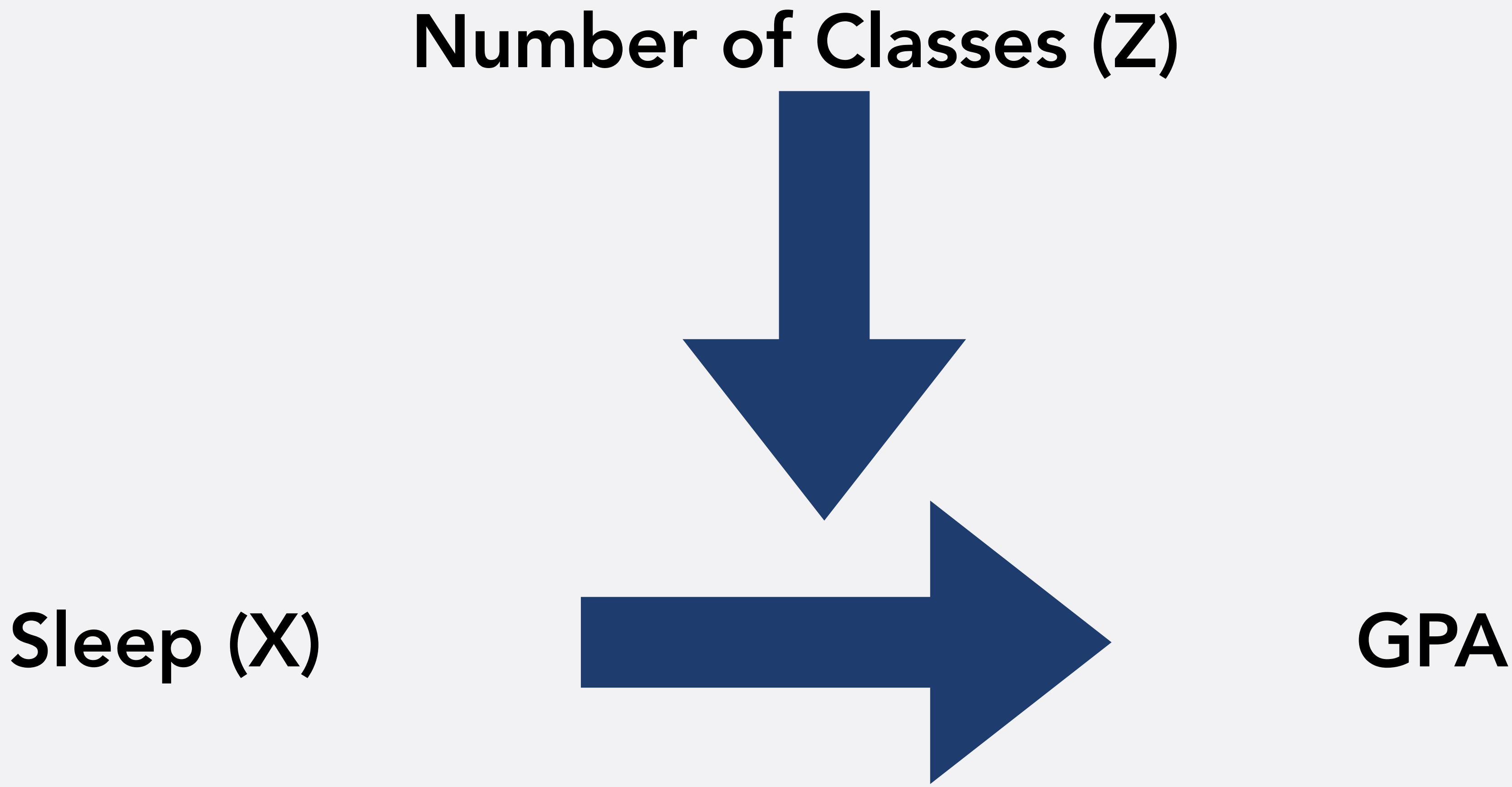
	5 Or Fewer Classes		6 Or More Classes	
Sleep	More Than 7 Hours/Night	7 Or Fewer Hours/Night	More Than 7 Hours/Night	7 Or Fewer Hours/Night
Average Gpa	$\boxed{3.53}$ (30)	$\boxed{3.42}$ (29)	$\boxed{3.55}$ (10)	$\boxed{3.52}$ (13)
	0.11		0.03	

- Frequency in parentheses

HOW CAN WE TELL WHICH ONE?

1. Are all controlled effects zero or very close to zero?
   - Yes? ⇒ relationship between x and y is spurious
   - No? ⇒ either additive or interactive
2. Are all controlled effects approximately the same size?
   - Yes? ⇒ additive relationship
   - No? ⇒ interactive relationship

INTERACTIVE RELATIONSHIP



- Number of classes determines how much sleep affects GPA
  - Sleep matters quite a bit among students who take 5 or fewer classes
  - Sleep doesn't matter for students who take more classes

ANOTHER EXAMPLE

	Democrats		Non-Democrats
Rating Of J. Biden	63.3		35.5
	(52)	27.8	(29)

- Frequency in parentheses

ANOTHER EXAMPLE

	Male		Female	
Partisanship	Democrats	Non-Democrats	Democrats	Non-Democrats
Rating Of J. Biden	68.6 (15)	32.0 (12)	61.6 (37)	37.5 (17)

- Frequency in parentheses

ANOTHER EXAMPLE

	Male		Female	
Partisanship	Democrats	Non-Democrats	Democrats	Non-Democrats
Rating Of J. Biden	68.6 (15)	32.0 (12)	61.6 (37)	37.5 (17)
	36.6		24.1	

- **Frequency in parentheses**

HURDLES TO CAUSALITY

- Is there a credible causal mechanism that connects X to Y?
- Can we rule out the possibility that Y could cause X?
- Is there covariation between X and Y?
- Have we controlled for all confounding variables (Z) that might make the association between X and Y spurious?

EXAMPLE

	Male	Female	Total
Democrats			
Non-Democrats			
Total			

2 INDEPENDENT VARIABLES

	Religious		Not Religious		
	Male	Female	Male	Female	Total
Democrats					
Non-Democrats					
Total					

3 INDEPENDENT VARIABLES

	White				Black/African-American				Asian				
	Religious		Not Religious		Religious		Not Religious		Religious		Not Religious		
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Total
Democrats													
Non-Democrats													
Total													

- As we control for more potential confounders, table gets increasingly unwieldy (and few/no observations in some cells)

ANOTHER ISSUE

- What if a control is interval-level?
  - e.g liberal-conservative (0-100 scale)
  - A table with 100 categories?
  - Again, unwieldy table with many cells where there are no observations

TODAY

- Multiple regression
    - Extends bivariate regression to incorporate not just one, but many independent variables

CONTROLLED COMPARISON

- Logic of controlled comparison:
  - Separate men and women
  - Among men: What is the effect of partisanship on ratings of J. Biden?
  - Among women: What is the effect of partisanship on ratings of J. Biden?
- Gives us the partial effects of partisanship, holding gender constant

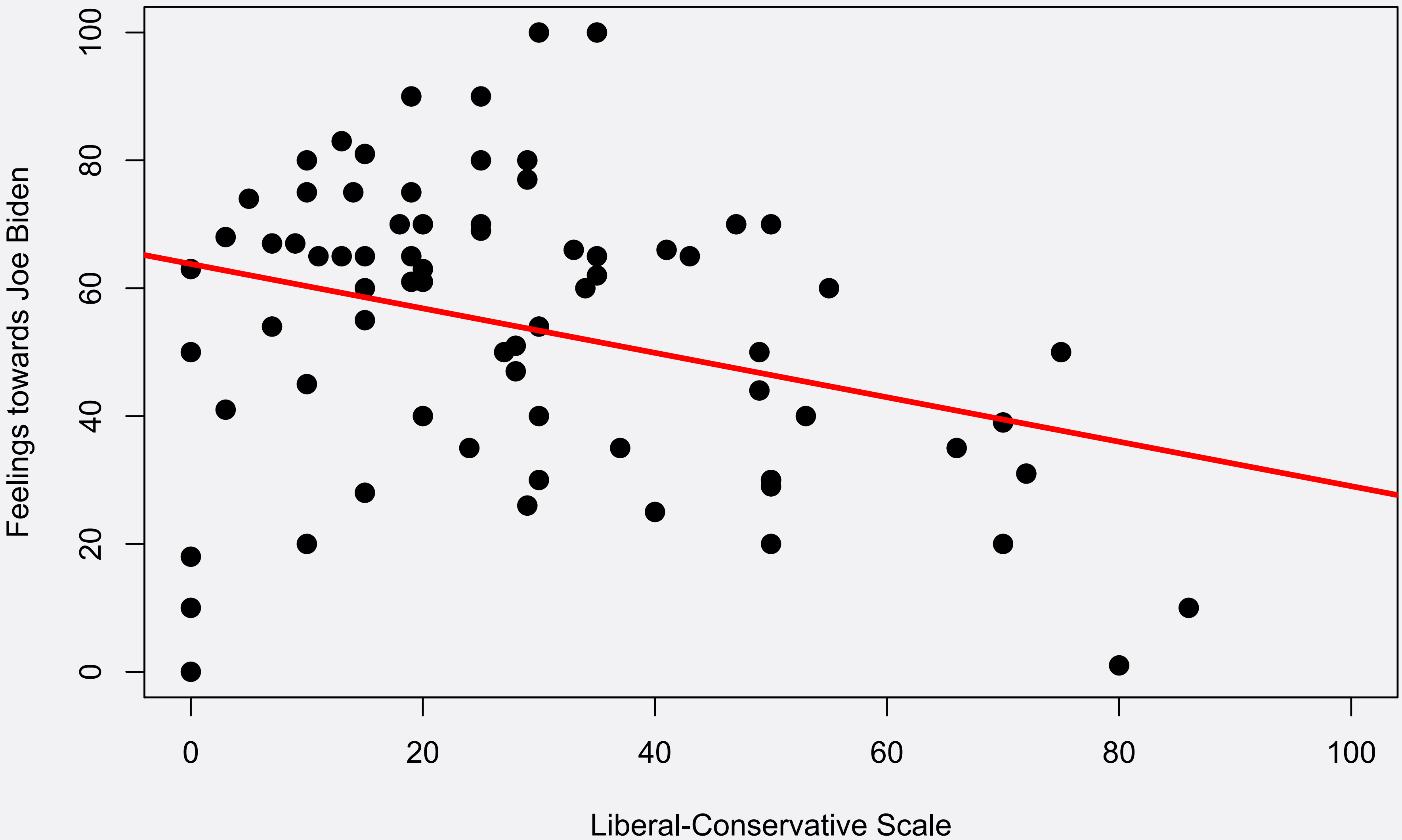
MULTIPLE REGRESSION

- Multiple regression does something similar
- Can estimate the effect of two variables on dependent variable (ratings of J. Biden)
  - Gives the partial effect of gender, holding partisanship constant
  - And: gives the partial effect of partisanship, holding gender constant

MULTIPLE REGRESSION

- Can include more than 2 independent variables
  - e.g. age, partisanship, gender
  - Gives the partial effect of age, holding constant partisanship and gender
  - Gives the partial effect of partisanship, holding constant gender and age
  - Gives the partial effect of gender, holding constant partisanship and age

BIVARIATE REGRESSION



BIVARIATE REGRESSION

	Coefficient	Standard Error	T-Value
Intercept	63.80	4.35	14.64
Liberal-Conservative	-0.35	0.12	-2.84

R^2: 0.10

LINEAR REGRESSION

- Let's add age as a second control
- $y = a + b_1 * x_1 + b_2 * x_2$
  - y: evaluation of J. Biden
  - x_1: liberal-conservative
  - x_2: age

EFFECT OF LIB/CONS

	Coefficient	Standard Error	T-Value
Intercept	95.2	61.9	1.54
Liberal-Conservative	-0.33	0.13	-2.46
Age	-1.55	3.13	-0.50

R^2: 0.10

EFFECT OF LIB/CONS

- Coefficient: -0.33 (SE 0.13, t-value -2.46)
- Interpretation: For every one point increase on the liberal-conservative scale, the evaluation of J. Biden decreases by 0.33 points, *holding all other variables constant*
- We reject H_0, so negative effect of liberal-conservative on evaluation is significant at the 5% level
  - even when controlling for age

EFFECT OF AGE

	Coefficient	Standard Error	T-Value
Intercept	95.2	61.9	1.54
Liberal-Conservative	-0.33	0.13	-2.46
Age	-1.55	3.13	-0.50

R^2: 0.10

EFFECT OF AGE

- Coefficient: -1.55 (SE 3.13, t-value -0.50)
- Interpretation: For every one year increase in age, the evaluation of J. Biden decreases by 1.55 points, *holding all other variables constant*
- We cannot reject H_0, so effect of age on evaluation is *not* significant at the 5% level
    - Age does not have an independent effect from lib/cons

INTERCEPT

	Coefficient	Standard Error	T-Value
Intercept	95.2	61.9	1.54
Liberal-Conservative	-0.33	0.13	-2.46
Age	-1.55	3.13	-0.50

R^2: 0.10

INTERCEPT

- Intercept: 95.2
- Gives expected feeling score when *both* lib/cons=0 *and* age=0

R-SQUARED

	Coefficient	Standard Error	T-Value
Intercept	95.2	61.9	1.54
Liberal-Conservative	-0.33	0.13	-2.46
Age	-1.55	3.13	-0.50

R^2: 0.10

CETERIS PARIBUS

- Linear regression allows us to estimate the effect of several independent variables on the dependent variable
- Gives us the effect of an independent variable on the dependent variable, *holding all other variables constant*
  - "ceteris paribus"
- We can assess the effect of the variables independently of each other

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- What is the predicted evaluation for a person with a lib/cons score of 50 and an age of 20?

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- What is the predicted evaluation for a person with a lib/cons score of 50 and an age of 20?
- Evaluation = 95.2 - 0.33*50 - 1.55*20 = 47.7

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- What if that person had a lib/cons score of 100 instead?

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- What if that person had a lib/cons score of 100 instead?
- Evaluation = 95.2 - 0.33*100 - 1.55*20 = 31.2

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- And what if they were 25 years old?

CETERIS PARIBUS

- Evaluation = 95.2 - 0.33*Lib/Cons - 1.55*Age
- And what if they were 25 years old?
- Evaluation = 95.2 - 0.33*100 - 1.55*25 = 23.5

CETERIS PARIBUS

- Of course, being older might also make someone more conservative
- But the linear regression estimates the effects of liberal/conservative and age *independently of each other*