

Quantitative Data Analysis II

SOC 781

Mediation and moderation

Today we're going to...

- Feedback on Assignment 2
- briefly discuss scale reliability and standardized coefficients
- Cover mediation and moderation
 - consider the differences between each approach

Assignment 2: Feedback

- Interpreting binary/dummy coefficients
 - NOT increase or decrease, just compare to reference group
 - A one unit increase in Sex is associated with # unit increase in DV – makes NO sense
- Start writing your assignments as if journal article
 - Find quantitative paper you're interested in (reputable journal)
 - Mimic structure and language/writing

Cronbach's alpha

- Measure of scale reliability
- Internal consistency
 - how closely related a set of items are as a group
- Higher values = greater reliability
 - 0.70 ~ psychometric threshold

Cronbach's alpha: example

```
alpha satcity sathobby satfam satfrnd sathealt if nmiss==0, item std casewise
```

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem correlation	alpha
satcity	23549	+	0.6068	0.3601	0.3389	0.6722
sathobby	23549	+	0.6873	0.4705	0.2942	0.6251
satfam	23549	+	0.6850	0.4673	0.2954	0.6265
satfrnd	23549	+	0.7294	0.5316	0.2708	0.5977
sathealt	23549	+	0.6250	0.3844	0.3288	0.6621
Test scale					0.3056	0.6876

- Standardizes variables to give average interitem correlation
 - mean = 0, variance = 1
- More important if variables measured differently

Standardized coefficients

- You may want to compare the magnitude of different X^s
 - Which has a greater impact on Y , X_1 or X_2 ?
- When X^s are measured in same units, comparisons are straightforward
 - not usually the case
- Convert X^s into z-scores: $z = \frac{X_i - \bar{X}}{s}$
 - mean = 0, SD = 1

Standardized coefficients

- Interpretation?

```
reg domsat c.age##c.age educ female nonwhite if nmiss==0, beta
```

domsat	Coef.	Std. Err.	t	P> t	Beta
age	.0695951	.0100039	6.96	0.000	.2571446
c.age#c.age	-.0006139	.0001018	-6.03	0.000	-.2242968
educ	.2296832	.0102022	22.51	0.000	.153194
female	.2383627	.0612006	3.89	0.000	.024868
nonwhite	-1.599653	.084424	-18.95	0.000	-.1222492
_cons	19.10532	.257771	74.12	0.000	.

- Only use if makes sense
 - doesn't make sense to standardize dummies (binary)

Standardized Y

- Compute z-score using egen

```
egen zdomsat=std(domsat)
reg zdomsat c.age#c.age educ female nonwhite if nmiss==0
```

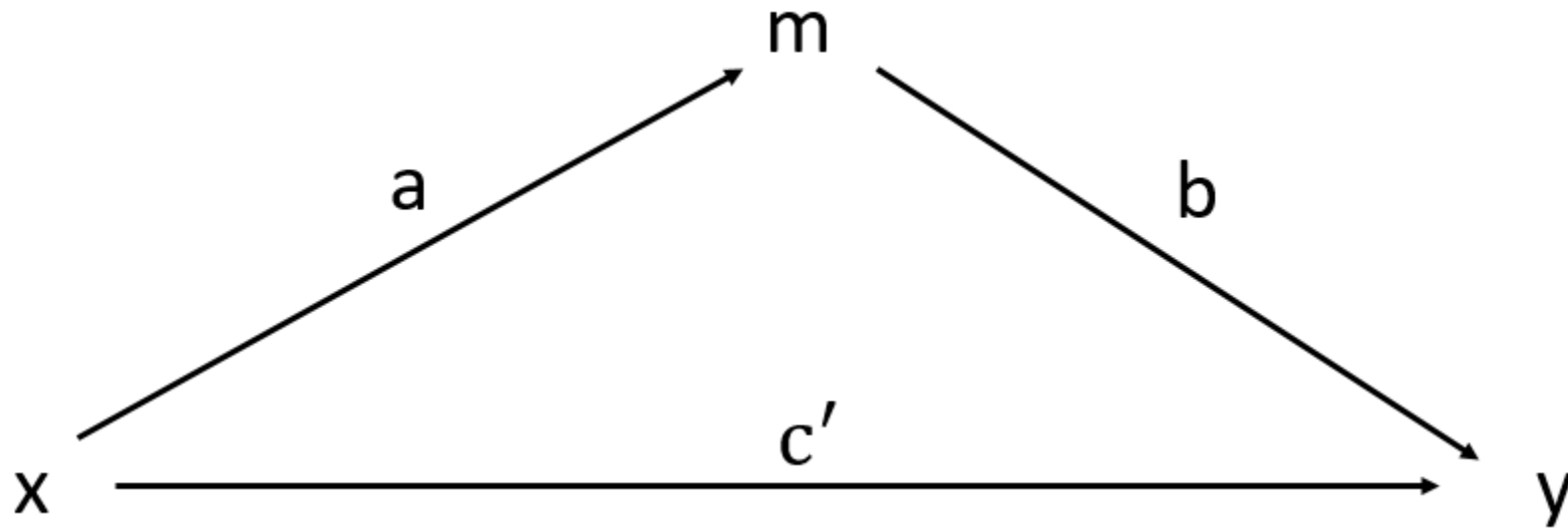
zdomsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0146228	.002102	6.96	0.000	.0105029	.0187428
c.age#c.age	-.000129	.0000214	-6.03	0.000	-.0001709	-.0000871
educ	.0482594	.0021436	22.51	0.000	.0440578	.0524611
female	.0500831	.012859	3.89	0.000	.0248785	.0752877
nonwhite	-.336108	.0177386	-18.95	0.000	-.3708768	-.3013393
_cons	-.9185514	.054161	-16.96	0.000	-1.024711	-.8123922

- Interpretation?

- Only use if makes sense

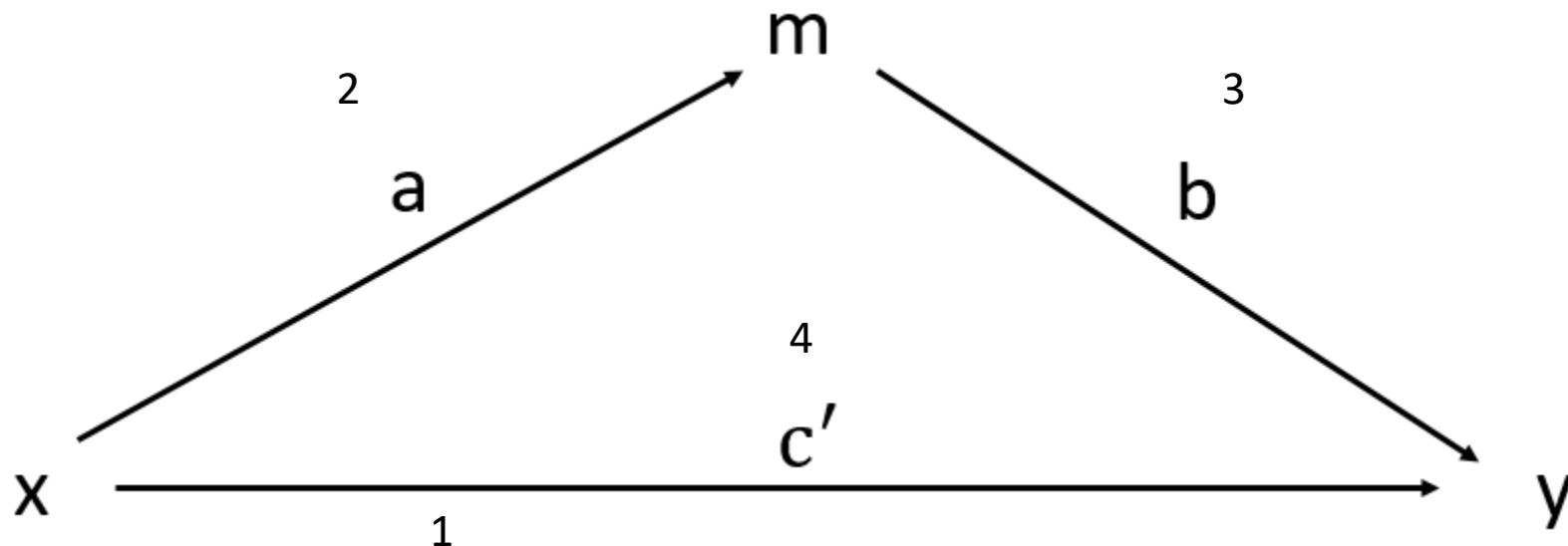
Mediation

- To what extent is the impact of X on Y direct vs. indirect?
 - What does that mean – in your own words?



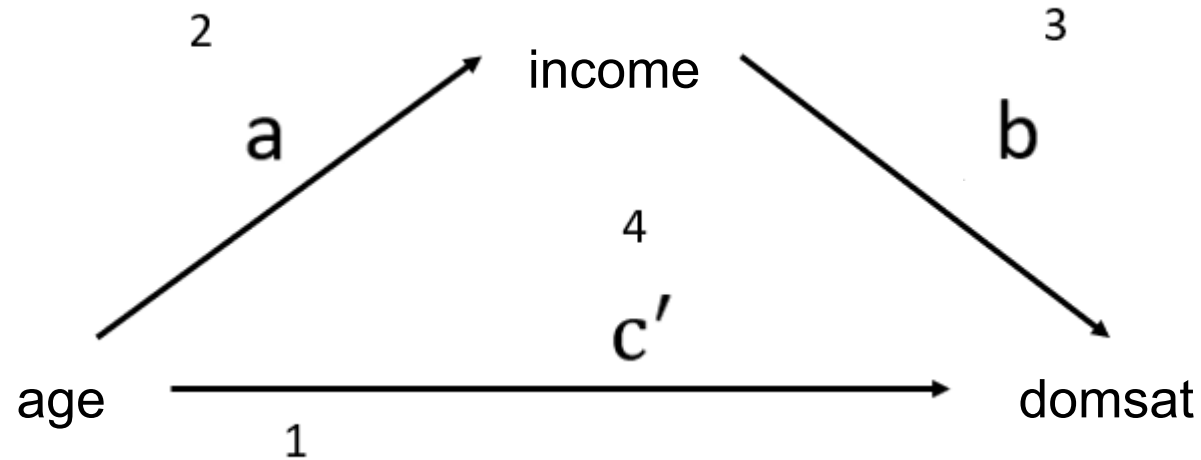
Mediation: steps

1. Show that X (causal mechanism/IV) is associated with Y (outcome/DV)
2. Show that X is correlated with m
 - essentially treat m as if it were y
3. Show that m is associated with Y net of X
 - to make sure m isn't simply associated with y because it is associated with x
4. To establish m completely mediates X - Y c' should be 0
 - rare, but also not necessary (partial mediation important)



Mediation: example

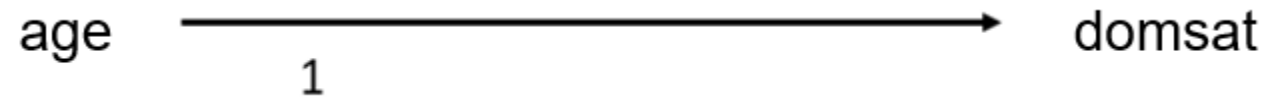
- Does income mediate the domsat-age relationship?



- Age causes domsat
- Age causes income
- Income causes domsat
- How much of the domsat-age relationship is explained by income?

Mediation example: step 1

1. Show that *age* is associated with *domsat*

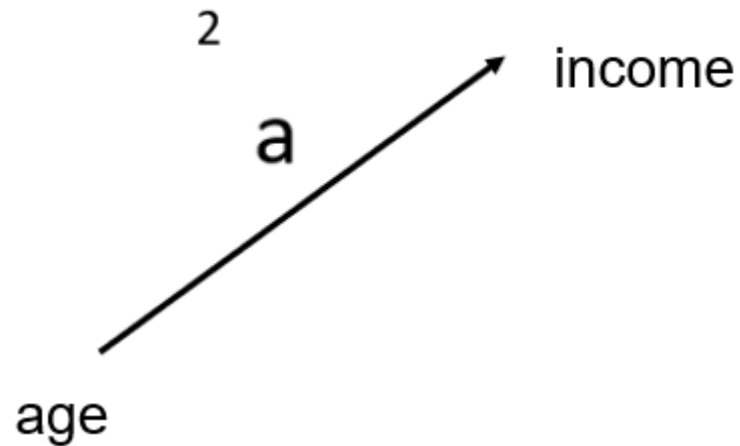


```
reg domsat age female nonwhite if age<65 & income<13 & nmiss==0
```

domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0080658	.0026886	3.00	0.003	.002796	.0133357
female	.2289447	.0683069	3.35	0.001	.0950566	.3628327
nonwhite	-1.84527	.0928829	-19.87	0.000	-2.027329	-1.66321
_cons	23.40347	.1166818	200.58	0.000	23.17476	23.63218

Mediation example: step 2

2. Show that *age* is correlated with *income*

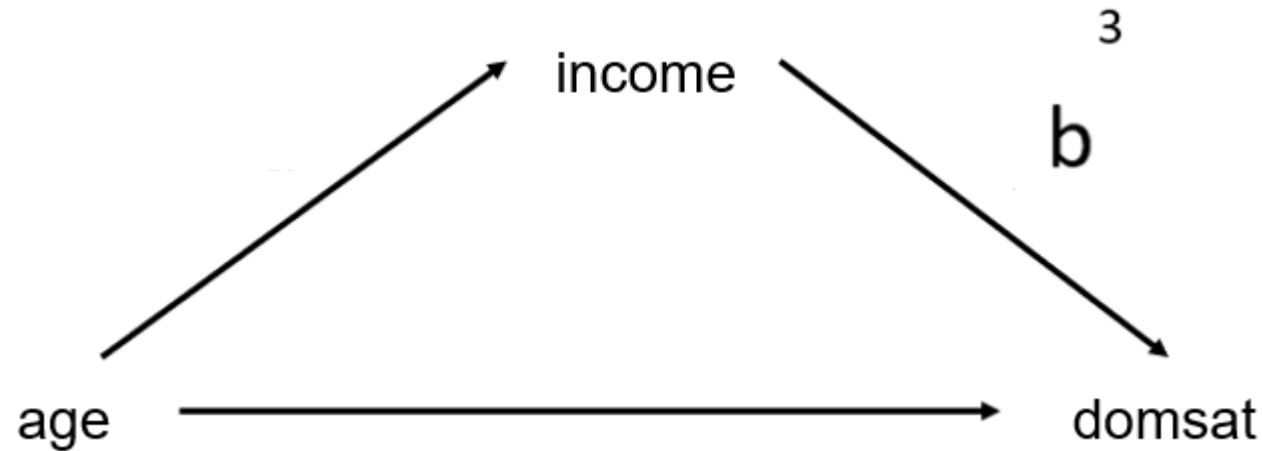


```
reg income age female nonwhite if age<65 & income<13 & nmiss==0 & domsat!=.
```

income	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0159025	.0016557	9.60	0.000	.0126572	.0191479
female	-.6278459	.0420657	-14.93	0.000	-.7102986	-.5453931
nonwhite	-1.302128	.0572004	-22.76	0.000	-1.414246	-1.19001
_cons	9.363843	.0718566	130.31	0.000	9.222997	9.504689

Mediation example: step 3

3. Show that *income* is associated with *domsat* net of *age*



```
reg domsat income age female nonwhite if age<65 & income<13 & nmiss==0
```

domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
income	.3357149	.01178	28.50	0.000	.3126249	.3588049
age	.0027271	.0026372	1.03	0.301	-.0024421	.0078963
female	.4397219	.067241	6.54	0.000	.3079232	.5715206
nonwhite	-1.408126	.0921641	-15.28	0.000	-1.588776	-1.227475
_cons	20.25989	.158748	127.62	0.000	19.94873	20.57105

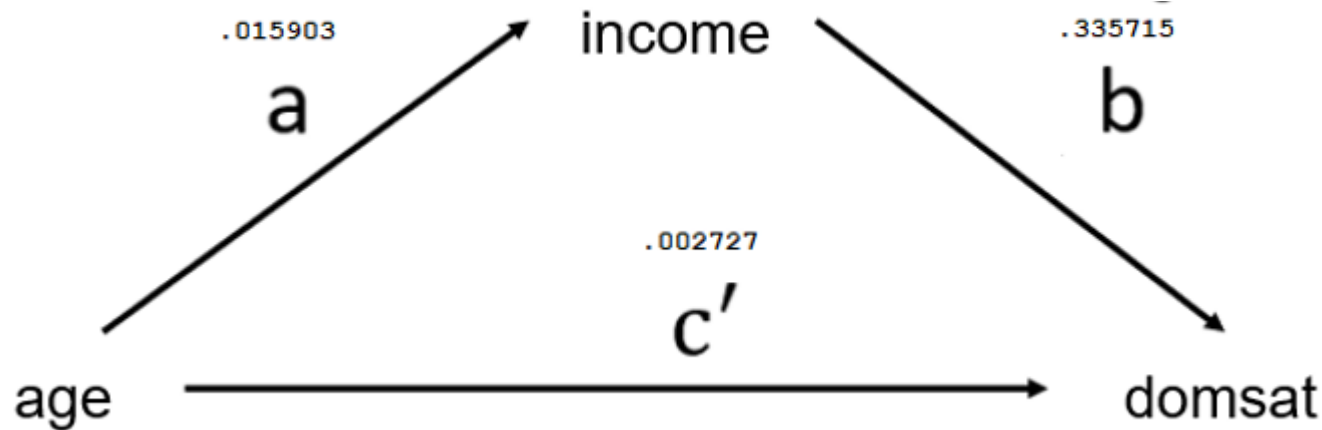
Mediation example: step 4

4. Establish how much *income* mediates *domsat-age* relationship

- proportion mediated = indirect effect / total effect
- This gets tricky: can't base on change in coef. and p-value of *age* (IV)
 - we need statistical tests
- Luckily, Stata can do this

Mediation: example

- total effect = direct effect + indirect effect
 - $c = c' + ab$



```
sgmediation domsat if age<65 & nmiss==0, mv(income) iv(age) cv(female nonwhite)
```

Sobel-Goodman Mediation Tests

	Coef	Std Err	Z	P> Z
Sobel	.00533871	.00058657	9.102	0
Goodman-1 (Aroian)	.00533871	.00058689	9.097	0
Goodman-2	.00533871	.00058624	9.107	0

	Coef	Std Err	Z	P> Z
a coefficient =	.015903	.001656	9.60463	0
b coefficient =	.335715	.01178	28.4986	0
Indirect effect =	.005339	.000587	9.10164	0
Direct effect =	.002727	.002637	1.03409	.301096
Total effect =	.008066	.002689	3.00004	.002699

Proportion of total effect that is mediated: .66189297
Ratio of indirect to direct effect: 1.9576433
Ratio of total to direct effect: 2.9576433

Direct effect = .002727

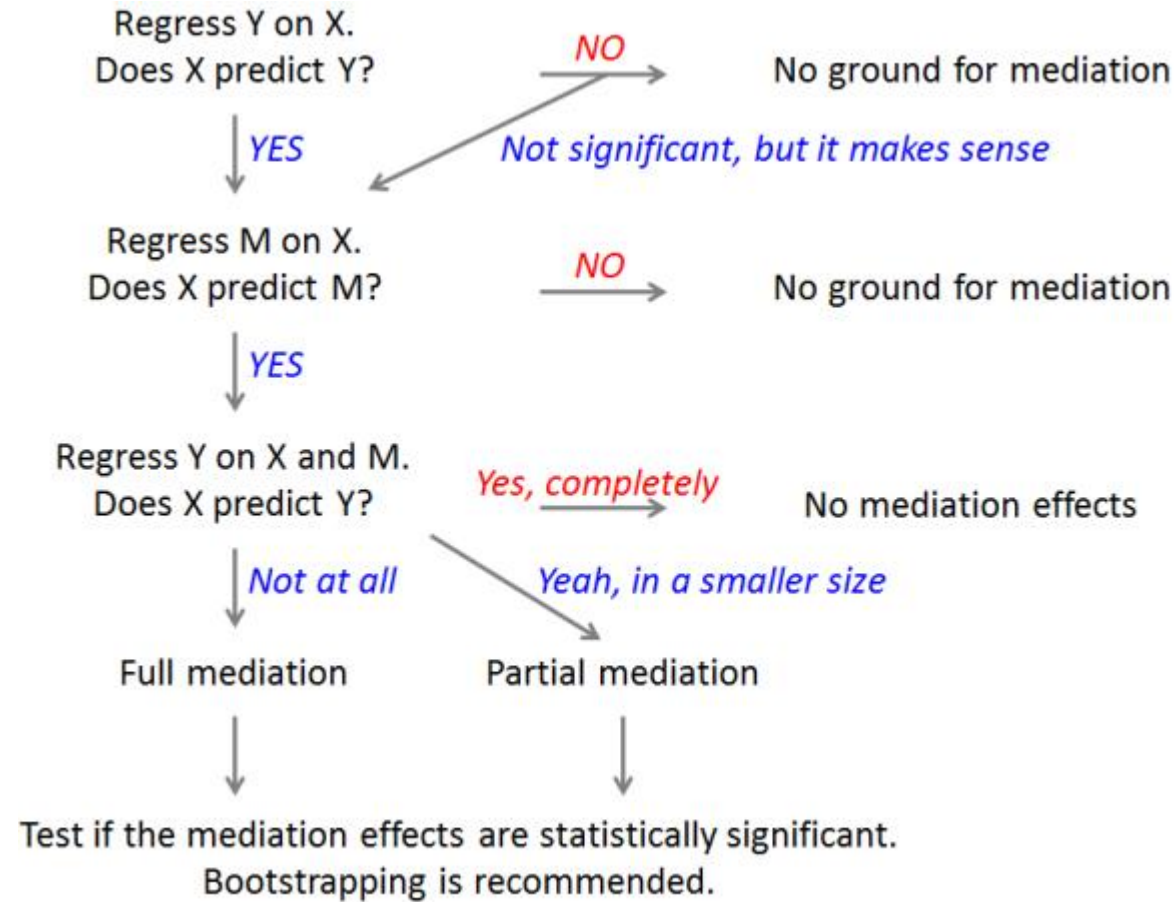
Indirect effect = .015903 * .335715 = .005339

Total effect = .002727 + .005339 = .008066

Proportion of total effect that is mediated = indirect effect / total effect (.005339 / .008066 = 0.6619)

So, the mediating effect of income on happiness accounts for ~66% of the total effect of age on domsat

Mediation: additional resources

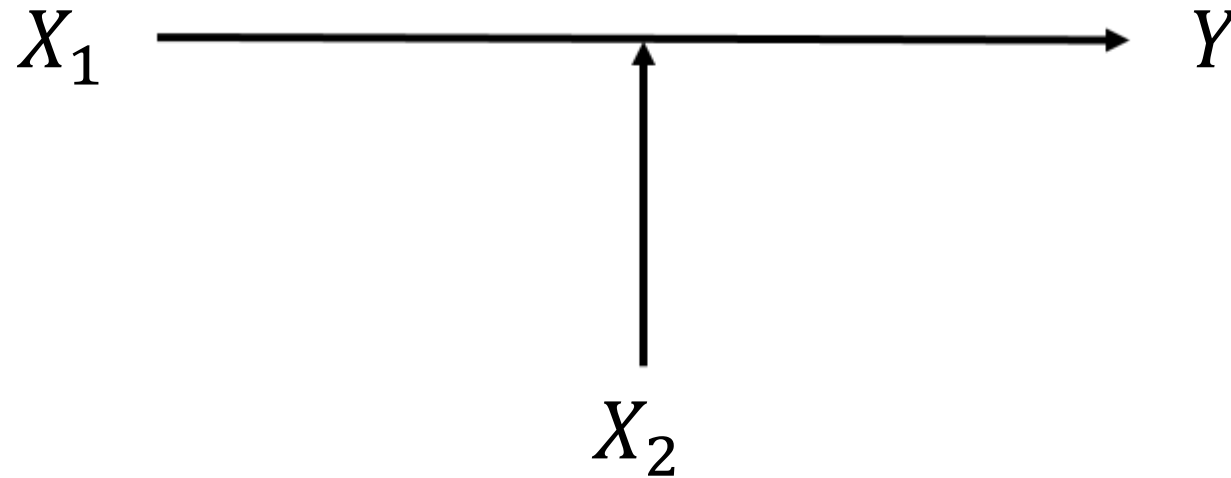


- [Source link](#)

- DV and M have to be continuous for Sobel test
 - [See for additional resources](#)

Moderation

- The effect of one predictor X_1 depends on the value of another X_2
 - in terms of strength and/or direction



- Can test with interaction terms
 - $Y = B_0 + B_1 X_1 + B_2 X_2 + B_3 X_{1*2} + \varepsilon$

Moderation example: dummy * dummy

- Does the effect of marital status on domsat differ by sex?
- First, examine main effects

```
reg domsat female married age age2 nonwhite educ if nmiss==0
```

domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	.3672061	.0606168	6.06	0.000	.2483932	.4860189
married	1.626516	.0648008	25.10	0.000	1.499502	1.753529
age	-.0054657	.0103158	-0.53	0.596	-.0256854	.014754
age2	.0001543	.000105	1.47	0.142	-.0000516	.0003601
nonwhite	-1.264869	.0843791	-14.99	0.000	-1.430257	-1.09948
educ	.2361874	.0100719	23.45	0.000	.2164457	.2559291
_cons	19.50611	.2548955	76.53	0.000	19.0065	20.00573

- Being married, vs. not, increases domsat by 1.627 units on average,
 - holding all else constant

Moderation example: dummy * dummy

- Does the effect of marital status on domsat differ by sex?

```
reg domsat female married marriedfemale age age2 nonwhite educ if nmiss==0
```

domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
female	.811013	.0986965	8.22	0.000	.6175616	1.004464
married	2.04618	.098094	20.86	0.000	1.853909	2.238451
marriedfemale	-.7186664	.1261773	-5.70	0.000	-.965982	-.4713508
age	-.0071566	.0103132	-0.69	0.488	-.0273712	.013058
age2	.0001536	.000105	1.46	0.143	-.0000521	.0003593
nonwhite	-1.280488	.0843674	-15.18	0.000	-1.445853	-1.115122
educ	.2350914	.0100671	23.35	0.000	.2153593	.2548235
_cons	19.32432	.2567175	75.27	0.000	18.82113	19.8275

- Being married increases domsat by 2.046 units for males, but only 1.327 units for females, holding all else constant
 - $2.046 - 0.719 = 1.327$
- Being married is beneficial for domsat
 - but more so for males compared to females

Moderation: dummy * interval

- Does the effect of education on domsat differ by sex?

```
reg domsat c.educ##i.female nonwhite married age age2 if nmiss==0
```

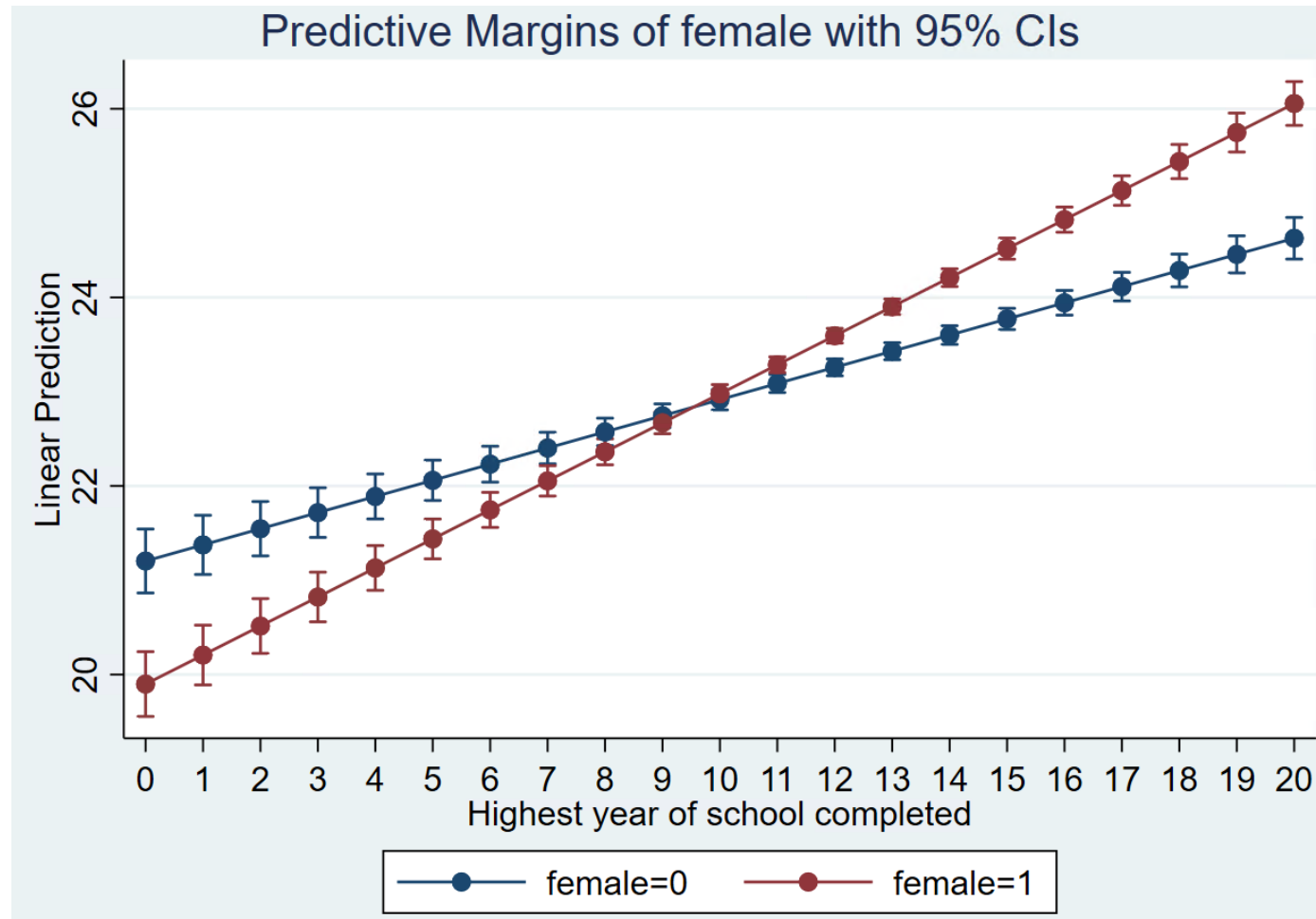
domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.1711588	.0134944	12.68	0.000	.1447088	.1976087
1.female	-1.305091	.2390649	-5.46	0.000	-1.773673	-.8365078
female#c.educ						
1	.1367026	.0189052	7.23	0.000	.0996471	.1737581
nonwhite	-1.269813	.0842901	-15.06	0.000	-1.435027	-1.104599
married	1.606453	.0647897	24.79	0.000	1.479461	1.733445
age	-.0046298	.0103053	-0.45	0.653	-.0248288	.0155692
age2	.0001514	.0001049	1.44	0.149	-.0000543	.000357
_cons	20.29507	.2770112	73.26	0.000	19.75211	20.83803

- A one-year increase in education increases domsat by 0.171 units for males, but 0.308 units for females, holding all else constant
 - $0.171 + 0.137 = 0.308$
- Educational attainment is beneficial for domsat
 - but more so for females compared to males

Moderation: dummy * interval (graph)

- Does the effect of education on domsat differ by sex?

```
reg domsat c.educ##i.female nonwhite married age age2 if nmiss==0  
margins female, at(educ=(0(1)20))  
marginsplot
```



Moderation: interval * interval

- Does the effect of age on domsat differ by education?

```
reg domsat c.age#c.educ female nonwhite married if nmiss==0
```

domsat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0375679	.006733	5.58	0.000	.0243709	.050765
educ	.3560612	.0298595	11.92	0.000	.2975347	.4145877
c.age#c.educ	-.0024062	.0005556	-4.33	0.000	-.0034952	-.0013173
female	.3772904	.0606453	6.22	0.000	.2584216	.4961591
nonwhite	-1.276308	.0841642	-15.16	0.000	-1.441275	-1.111341
married	1.624428	.0622491	26.10	0.000	1.502416	1.74644
_cons	17.74567	.3878287	45.76	0.000	16.9855	18.50584

- A one-year increase in age increases domsat by 0.038 units, but this decreases by 0.002 units for every additional year of education
- Age is beneficial for domsat
 - but more so among those with less education: doesn't make sense
 - selection effects: characteristics of low educated that survived to older ages

Predicted profiles

- Sometimes you want to know the value of an outcome given certain circumstances or for groups with specific characteristics
- Manipulate ALL x^s to reflect your desired circumstances and/or group
- How would you do this?
 - set value of all x^s to reference profile
 - E.g., age = 40 then create new var age40 = age – 40
 - Recall: reference group for constant/intercept is ALL 0s in model
- Practice with your data

Next class we will...

- discuss basics of GLM
- read Long & Freese CH3 & CH4 before class