Week N: Interaction Models

POLI502

Howard Liu

December 5, 2024

University of South Carolina

Outline for Today

Interaction models

1. Numerical $\times$ binary variable
   - Estimation in R
   - Plot
2. Numerical $\times$ numerical
   - Estimation in R
   - Plot

Review: multiple regression

Multiple regression (additive model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z$$

where Z is a binary variable (0 or 1)

Review: multiple regression

Multiple regression (additive model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z$$

where Z is a binary variable (0 or 1)

$$\hat{Y} = \alpha + \beta_1 * X \text{ when Z=0}$$
$$\hat{Y} = (\alpha + \beta_2) + \beta_1 * X \text{ when Z=1}$$

Review: multiple regression

Multiple regression (additive model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z$$

where Z is a binary variable (0 or 1)

$$\hat{Y} = \alpha + \beta_1 * X \text{ when Z=0}$$
$$\hat{Y} = (\alpha + \beta_2) + \beta_1 * X \text{ when Z=1}$$

The model allows us to have different intercepts depending on Z, but the slope for X (β_1) is **assumed to be the same**.

We may want to relax this assumption.

Conditional hypotheses

We might have a third variable (Z) that not only influences the effect of X but also **conditions** it.

Conditional hypotheses

We might have a third variable (Z) that not only influences the effect of X but also **conditions** it.

Recall what we had is this (additive model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z$$

$$\hat{Y} = \quad \alpha \quad + \beta_1 * X \text{ when Z=0}$$

$$\hat{Y} = (\alpha + \beta_2) + \beta_1 * X \text{ when Z=1}$$

Conditional hypotheses

We might have a third variable (Z) that not only influences the effect of X but also **conditions** it.

Recall what we had is this (additive model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z$$
$$\hat{Y} = \alpha + \beta_1 * X \text{ when Z=0}$$
$$\hat{Y} = (\alpha + \beta_2) + \beta_1 * X \text{ when Z=1}$$

Instead, we want to have this (interaction model):

$$\hat{Y} = \alpha + \beta_1 * X + \beta_2 * Z + \beta_3 * \mathbf{XZ}$$
$$\hat{Y} = \alpha + \beta_1 * X \text{ when Z=0}$$
$$\hat{Y} = (\alpha + \beta_2) + (\beta_1 + \beta_3) * X \text{ when Z=1}$$

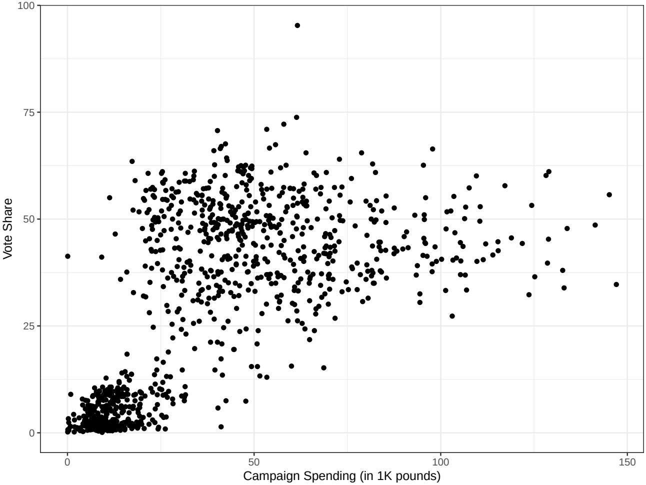
Now, both the intercept and the slope are different

Example 1: Effect of campaign spending

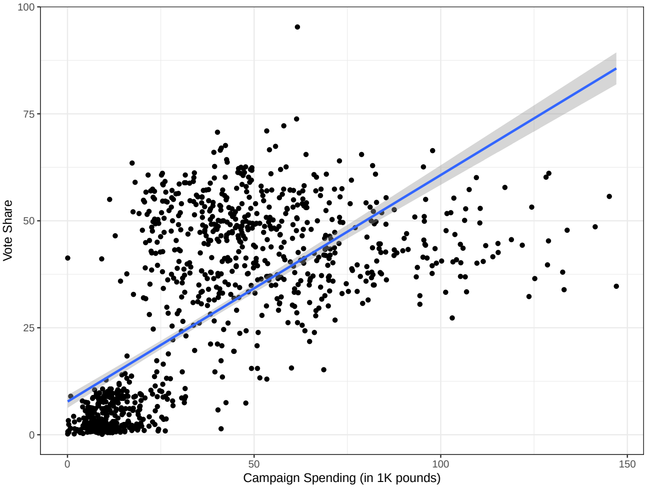
In competitive elections, the more money a candidate spends on campaigning, the more votes s/he is expected to get.

- DV: vote share (%) for candidates in the 2009 general election in Japan
- IDV: campaign spending (in 1,000 pounds)

Example 1: Effect of campaign spending



Example 1: Effect of campaign spending



Example 1: Effect of campaign spending

Table:

	Dependent variable:
	Vote Share
Campaign Spending	0.530***
	(0.017)
Constant	7.735***
	(0.757)
Observations	1,124
R^2	0.478
Adjusted R^2	0.478
Residual Std. Error	16.042 (df = 1122)
Note:	*p<0.1; **p<0.05; ***p<0.01

Example 1: Effect of campaign spending

Table:

	Dependent variable:
	Vote Share
Campaign Spending	0.530^{***}
	(0.017)
Constant	7.735^{***}
	(0.757)
Observations	1,124
R^2	0.478
Adjusted R^2	0.478
Residual Std. Error	16.042 (df = 1122)
Note:	*p<0.1; **p<0.05; ***p<0.01

$$\hat{VS} = 7.735 + 0.53 * CS$$

Example 1: Effect of campaign spending

Four hurdles to clear:

1. Causal mechanism linking $X \Rightarrow Y$
2. No reverse causality $Y \Rightarrow X$
3. X and Y covary
4. No confounding Z

Example 1: Effect of campaign spending

Four hurdles to clear:

1. Causal mechanism linking $X \Rightarrow Y$
2. No reverse causality $Y \Rightarrow X$
3. X and Y covary
4. No confounding Z

What do you think the causal mechanism is here?

Example 1: Effect of campaign spending

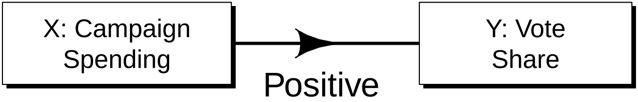
Four hurdles to clear:

1. Causal mechanism linking $X \Rightarrow Y$
2. No reverse causality $Y \Rightarrow X$
3. X and Y covary
4. No confounding Z

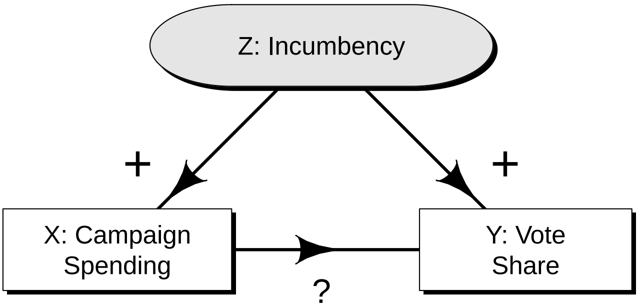
What do you think the causal mechanism is here?

What do you think potential confounders (conditioning factors) are?

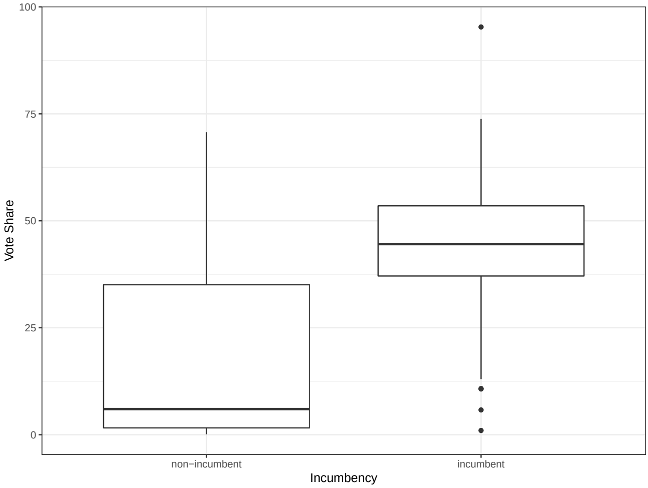
Example 1: Effect of campaign spending



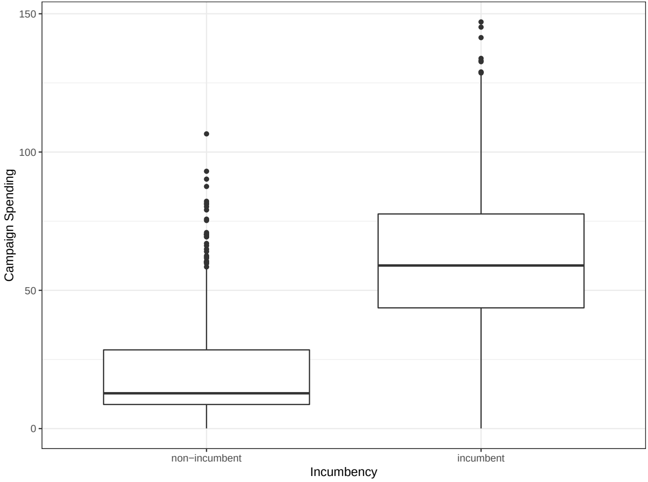
Example 1: Effect of campaign spending



Example 1: Effect of campaign spending



Example 1: Effect of campaign spending



Example 1: Effect of campaign spending

Table:

	(1)	(2)
Campaign Spending	0.530***	0.391***
	(0.017)	(0.022)
Incumbent		12.238***
		(1.335)
Constant	7.735***	8.384***
	(0.757)	(0.734)
Observations	1,124	1,124
R^2	0.478	0.515
Adjusted R^2	0.478	0.514
Residual Std. Error	16.042 (df = 1122)	15.480 (df = 1121)
Note:		*p<0.1; **p<0.05; ***p<0.01

Example 1: Effect of campaign spending

Table:

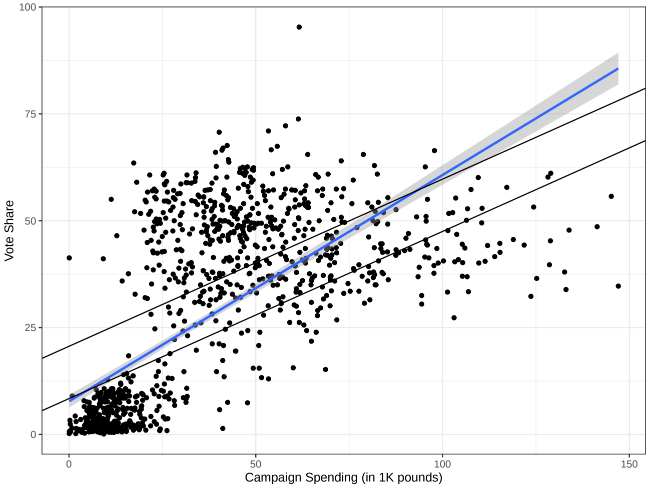
	(1)	(2)
Campaign Spending	0.530***	0.391***
	(0.017)	(0.022)
Incumbent		12.238***
		(1.335)
Constant	7.735***	8.384***
	(0.757)	(0.734)
Observations	1,124	1,124
R^2	0.478	0.515
Adjusted R^2	0.478	0.514
Residual Std. Error	16.042 (df = 1122)	15.480 (df = 1121)
Note:		*p<0.1; **p<0.05; ***p<0.01

$$\hat{VS} = 8.384 + 0.391 * CS + 12.238 * Inc$$

$$\hat{VS} = 8.384 + 0.391 * CS \text{ when Z=0 } (1)$$

$$\hat{VS} = 20.622 + 0.391 * CS \text{ when Z=1 } (2)$$

Example 1: Effect of campaign spending



Additive model: different intercept (when z= 0 or 1)

Example 1: Effect of campaign spending

The first model:

$$\hat{VS} = 7.735 + 0.530 * CS$$

The second model:

$$\hat{VS} = 8.384 + 0.391 * CS \text{ for non-incumbents}$$
$$\hat{VS} = 20.622 + 0.391 * CS \text{ for incumbents}$$

Example 1: Effect of campaign spending

The first model:

$$\hat{VS} = 7.735 + 0.530 * CS$$

The second model:

$$\hat{VS} = 8.384 + 0.391 * CS \text{ for non-incumbents}$$
$$\hat{VS} = 20.622 + 0.391 * CS \text{ for incumbents}$$

The second model is more flexible.

- the first one is based on an assumption that incumbents and non-incumbents have the same intercept;
- the second one relaxes that assumption.

Example 1: Effect of campaign spending

But the second one is still based on an untested assumption.

Example 1: Effect of campaign spending

But the second one is still based on an untested assumption.

- The effect ("slope") of Campaign Spending on Vote Share is the same for incumbents and non-incumbents.

Example 1: Effect of campaign spending

But the second one is still based on an untested assumption.

- The effect ("slope") of Campaign Spending on Vote Share is the same for incumbents and non-incumbents.

- If the effect of Campaign Spending on Vote Share is *through* increased publicity, the effect could be bigger for non-incumbents
    - Campaign spending may increase vote share in general;

Example 1: Effect of campaign spending

But the second one is still based on an untested assumption.

- The effect ("slope") of Campaign Spending on Vote Share is the same for incumbents and non-incumbents.

- If the effect of Campaign Spending on Vote Share is *through* increased publicity, the effect could be bigger for non-incumbents
  - Campaign spending may increase vote share in general;
  - Yet, an additional spending will increase vote share more for non-incumbents;
    An additional spending may have little effect on vote share for incumbents, as they are relatively well known already;

Example 1: Effect of campaign spending

But the second one is still based on an untested assumption.

- The effect ("slope") of Campaign Spending on Vote Share is the same for incumbents and non-incumbents.

- If the effect of Campaign Spending on Vote Share is *through* increased publicity, the effect could be bigger for non-incumbents
    - Campaign spending may increase vote share in general;
    - Yet, an additional spending will increase vote share more for non-incumbents;
      An additional spending may have little effect on vote share for incumbents, as they are relatively well known already;
    - → depending on the incumbency status, not only the intercept but also the slope for spending may differ.

Example 1: Effect of campaign spending

To relax this assumption, we include a **product** of Campaign Spending and Incumbency status:

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \underline{\beta_3 * CS * Inc}$$

Example 1: Effect of campaign spending

To relax this assumption, we include a **product** of Campaign Spending and Incumbency status:

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \underline{\beta_3 * CS * Inc}$$

The term $CS * Inc$ is called a (multiplicative) **interaction term**

Example 1: Effect of campaign spending

To relax this assumption, we include a **product** of Campaign Spending and Incumbency status:

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \underline{\beta_3 * CS * Inc}$$

The term $CS * Inc$ is called a (multiplicative) **interaction term**

- When $Inc = 0$, the model simplifies to

$$\begin{aligned} \hat{VS} &= \alpha + \beta_1 * CS + \beta_2 * (0) + \beta_3 * CS * (0) \\ &= \alpha + \beta_1 * CS \end{aligned}$$

Example 1: Effect of campaign spending

To relax this assumption, we include a **product** of Campaign Spending and Incumbency status:

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \underline{\beta_3 * CS * Inc}$$

The term $CS * Inc$ is called a (multiplicative) **interaction term**

- When $Inc = 0$, the model simplifies to

$$\begin{aligned} \hat{VS} &= \alpha + \beta_1 * CS + \beta_2 * (0) + \beta_3 * CS * (0) \\ &= \alpha + \beta_1 * CS \end{aligned}$$

- When $Inc = 1$, the model reduces to

$$\begin{aligned} \hat{VS} &= \alpha + \beta_1 * CS + \beta_2 * (1) + \beta_3 * CS * (1) \\ &= \alpha + \beta_1 * CS + \beta_2 + \beta_3 * CS \\ &= (\alpha + \beta_2) + (\beta_1 + \beta_3) * CS \end{aligned}$$

Example 1: Effect of campaign spending

Table:

	(1)	(2)	(3)
Campaign Spending	0.530***	0.391***	0.864***
	(0.017)	(0.022)	(0.026)
Incumbency		12.238***	46.758***
		(1.335)	(1.791)
Campaign Spending × Incumbency			−0.866***
			(0.036)
Constant	7.735***	8.384***	−1.502**
	(0.757)	(0.734)	(0.721)
Observations	1,124	1,124	1,124
R^2	0.478	0.515	0.681
Adjusted R^2	0.478	0.514	0.680
Residual Std. Error	16.042 (df = 1122)	15.480 (df = 1121)	12.548 (df = 1120)
Note:			*p<0.1; **p<0.05; ***p<0.01

Example 1: Effect of campaign spending

Table:

	(1)	(2)	(3)
Campaign Spending	0.530***	0.391***	0.864***
	(0.017)	(0.022)	(0.026)
Incumbency		12.238***	46.758***
		(1.335)	(1.791)
Campaign Spending × Incumbency			−0.866***
			(0.036)
Constant	7.735***	8.384***	−1.502**
	(0.757)	(0.734)	(0.721)
Observations	1,124	1,124	1,124
R^2	0.478	0.515	0.681
Adjusted R^2	0.478	0.514	0.680
Residual Std. Error	16.042 (df = 1122)	15.480 (df = 1121)	12.548 (df = 1120)
Note:			*p<0.1; **p<0.05; ***p<0.01

$$\hat{VS} = -1.502 + 0.864 * CS + 46.758 * Inc - 0.866 * CS * Inc$$

Example 1: Effect of campaign spending

$$\hat{VS} = -1.502 + 0.864 * CS + 46.758 * Inc - 0.866 * CS * Inc$$

Example 1: Effect of campaign spending

$$\hat{VS} = -1.502 + 0.864 * CS + 46.758 * Inc - 0.866 * CS * Inc$$

- When $Inc = 0$, the model simplifies to

$$\hat{VS} = -1.502 + 0.864 * CS$$

- When $Inc = 1$, the model simplifies to

$$\begin{aligned}\hat{VS} &= -1.502 + 0.864 * CS + 46.758 - 0.866 * CS \\ &= (-1.502 + 46.758) + (0.864 - 0.866) * CS \\ &= 45.256 - 0.002 * CS\end{aligned}$$

Estimation in R

When you want to include an interaction term between `x` and `z`, you write

```
lm(y ~ x + z + x*z)
```

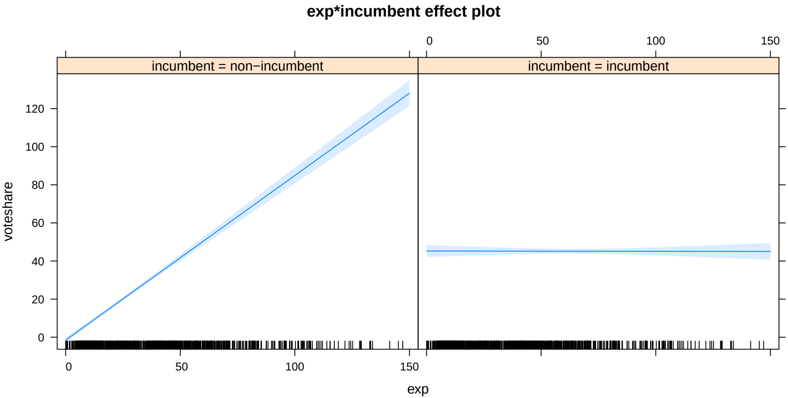
(The third term that combines `x` and `z` with a colon or asterisk ($X \star Z$) is the interaction term.)

Whenever you estimate an interactive model, make sure you interpret the results graphically using the effect function.

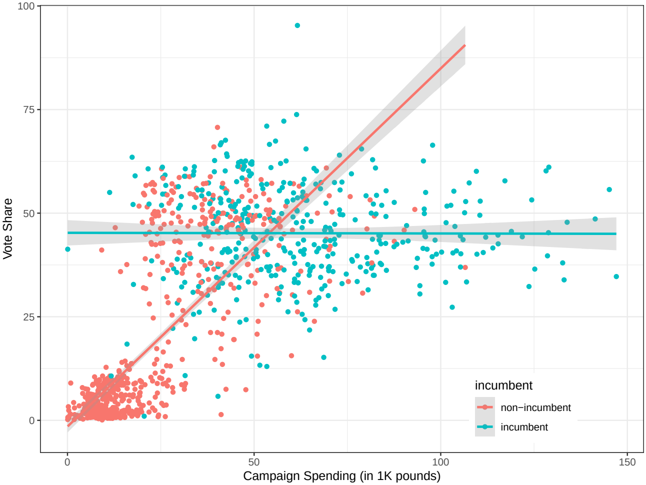
```
plot(effect(term = "x:z", mod = fit))
```

Or even better, use ggplot to plot the effect plots.

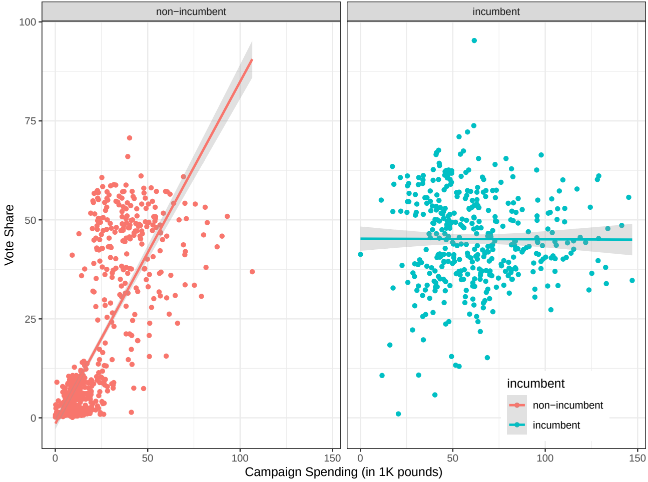
Example 1: Effect of campaign spending



Example 1: Effect of campaign spending



Example 1: Effect of campaign spending



Estimation in R: factor vs numeric

Recall that, when including a binary variable in a regression model, we could do so in one of two ways.

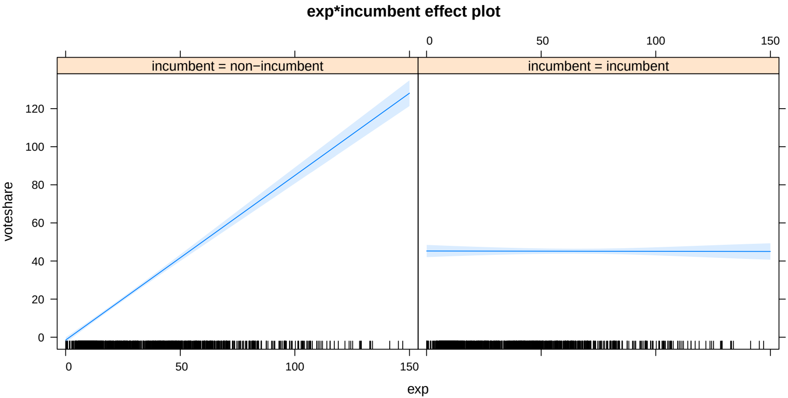
- Include the original factor variable as is
- Include a numerical binary variable

We saw this when dealing with the NorthSouth binary in the Putnam data set:

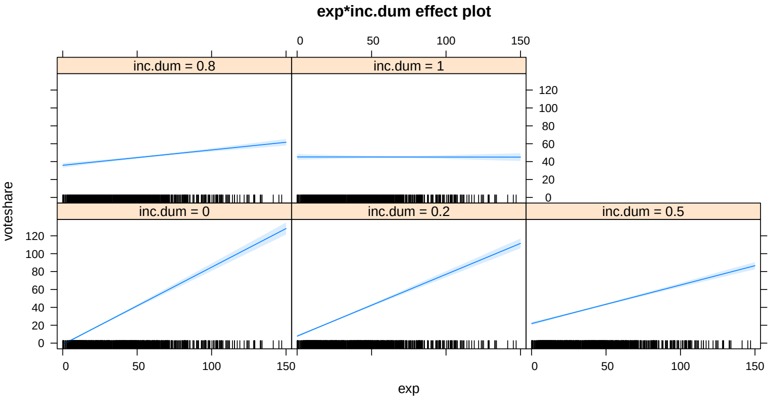
- Include the original factor variable is necessary to create an effect "plot" (o.w. R will assume it is a numeric variable)
- Include a numerical binary variable is preferable in order to produce an intuitive regression "table"

The same applies here.

Example 1: factor vs numeric



Example 1: factor vs numeric



Splitting the sample

There is another way to obtain a similar regression results, **but you need to be careful about sample size**:

- Split the data into two subsets (incumbents and non-incumbents);
- Regress Vote Share on Campaign Spending on each subset.

Splitting the sample

There is another way to obtain a similar regression results, **but you need to be careful about sample size**:

- Split the data into two subsets (incumbents and non-incumbents);
- Regress Vote Share on Campaign Spending on each subset.

	(1)	(2)
Campaign Spending	-0.002	0.864^{***}
	(0.023)	(0.027)
Constant	45.256^{***}	-1.502^{**}
	(1.558)	(0.740)
Observations	392	732

- Results for $Inc = 0$

$$\hat{VS} = -1.502 + 0.864 * CS$$

- Results for $Inc = 1$

$$\hat{VS} = 45.256 - 0.002 * CS$$

Interaction terms Do's and Don'ts

(1) Whenever you have an interaction term, you must include all the constitutive terms as well.

Interaction terms Do's and Don'ts

(1) Whenever you have an interaction term, you must include all the constitutive terms as well.

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \beta_3 * CS * Inc$$

Interaction terms Do's and Don'ts

(1) Whenever you have an interaction term, you must include all the constitutive terms as well.

$$\hat{VS} = \alpha + \beta_1 * CS + \beta_2 * Inc + \beta_3 * CS * Inc$$

- What happens if we (mistakenly) drop $\beta_2 * Inc$...?

Interaction terms Do's and Don'ts

(2) Caveat: **The numerical results (coefficients) in interaction models can often be misleading**; always interpret the results graphically (draw implied regression lines)

- Just because you get a statistically significant coefficient on an interaction term, it does NOT automatically mean that you find a meaningful conditional relationship! $\rightarrow$ look at the slope/effect size
- Likewise, even if your interaction term is statistically significant, it does NOT mean that your conditional relationship will NOT be substantial!

(3) Z (modifying variable) does not have to be a binary variable. It can be a continuous variable as well.

Example 2: Female representation

Let's say we want to test the following hypotheses:

- As the level of ethnic fractionalization increases, female representation goes down.

  This relationship may not hold universally;

  - The above relationship will not exist in poorer countries;
  - The relationship will be stronger in wealthier countries.

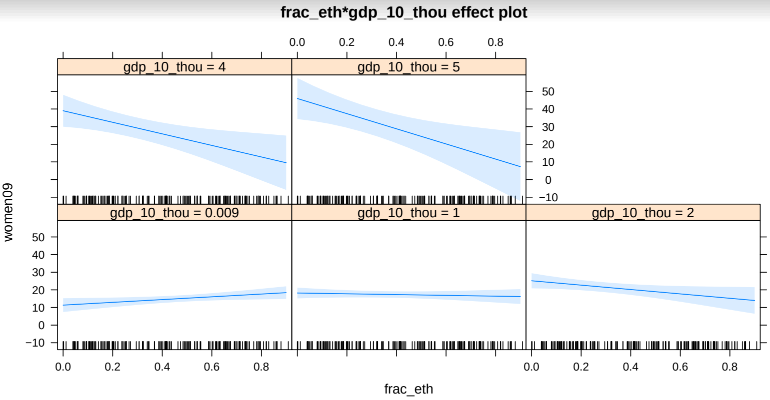
To test these hypotheses, we could estimate the following model:

```
lm(women09 ~ frac_eth + gdp_10_thou +
frac_eth*gdp_10_thou)
```

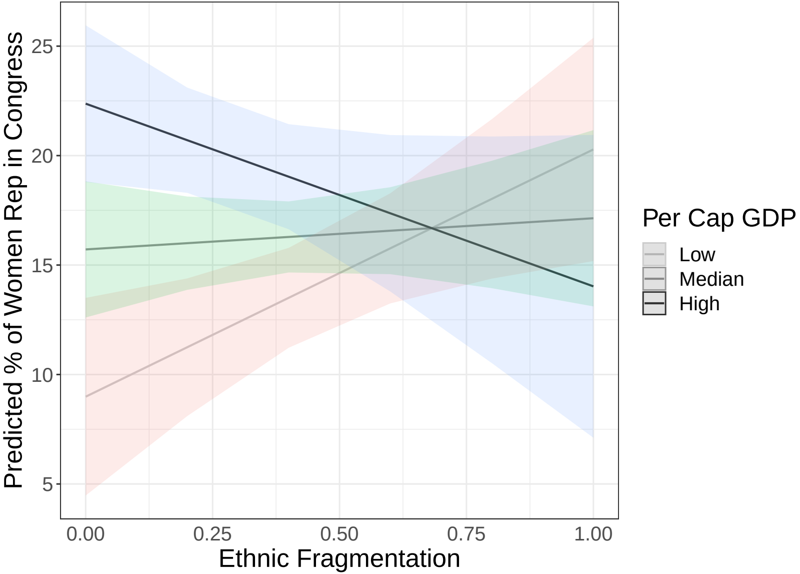
Example 2: Female representation

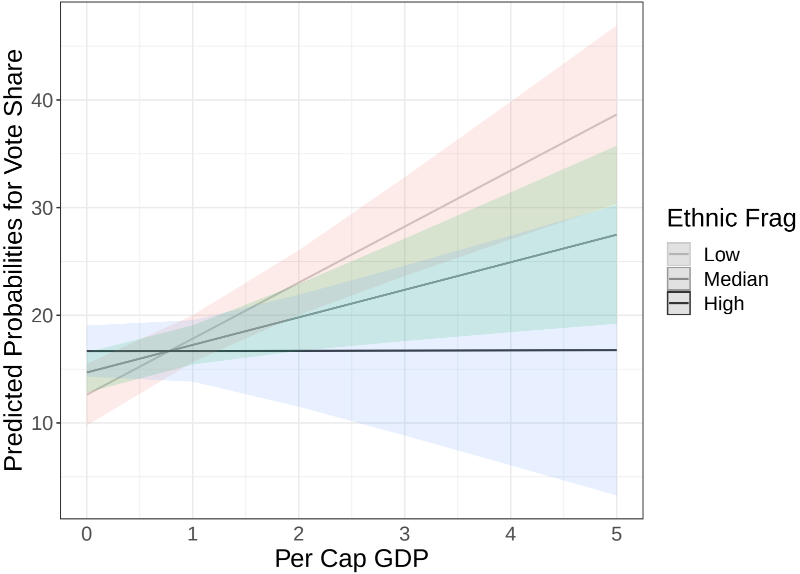
Table:

	(1)	(2)
Ethnic Fractionalization	2.406	7.937**
	(3.300)	(3.697)
Per capita GDP	3.610***	6.936***
	(0.880)	(1.390)
Ethnic Fractionalization × Per capita GDP		−10.177***
		(3.344)
Constant	13.776***	11.275***
	(1.855)	(1.987)
Observations	166	166
Adjusted R^2	0.084	0.128
Residual Std. Error	10.359 (df = 163)	10.106 (df = 162)
Note:		*p<0.1; **p<0.05; ***p<0.01



frac_eth*gdp_10_thou effect plot
gdp_10_thou = 4
gdp_10_thou = 5
gdp_10_thou = 0.009
gdp_10_thou = 1
gdp_10_thou = 2
women09
frac_eth
0.0
0.2
0.4
0.6
0.8
50
40
30
20
10
0
−10


Predicted % of Women Rep in Congress
25
20
15
10
5
0.00
0.25
0.50
0.75
1.00
Ethnic Fragmentation
Per Cap GDP
Low
Median
High


Predicted Probabilities for Vote Share
40
30
20
10
0
1
2
3
4
5
Per Cap GDP
Ethnic Frag
Low
Median
High

Final Advice: all interactions are "symmetric"

When you propose a conditional hypothesis/theory, you actually should test this following pair of hypotheses:

- $H_{X|Z}$: The marginal effect of X on Y is positive at all values of Z; this effect is strongest when Z is at its lowest and declines in magnitude as Z increases.
- $H_{Z|X}$: The marginal effect of Z on Y is positive at all values of X; this effect is strongest when X is at its lowest and declines in magnitude as X increases.