Bush 631-607: Quantitative Methods

Lecture 8 (10.19.2021): Prediction vol. III

Rotem Dvir

The Bush school of Government and Public Policy

Texas A&M University

Fall 2021

What is today's plan?

- Predictions: Linear model and causal inference.
- Binary predictors and randomized experiments.
- Multiple predictors, heterogeneous treatment effects
- R work: lm(), levels(), coef().

Least squared

THE LINEAR MODEL

$$Y = \alpha + \beta * X_i + \epsilon$$

Elements of model:

- *Intercept* (α): the average value of Y when X is zero.
- *Slope* (β): the average increase in Y when X increases by 1 unit.
- *Error/disturbance term* (ϵ): the deviation of an observation from a perfect linear relationship.

Minimize the prediction error

Confused by data?

Regression to the mean - its everywhere

How sure are we?

- What does our model tell us?
- Do the results mean anything?
- **Causal inference**:
  - Predicting the counter-factual.
  - Assumptions $\rightarrow$ use regression models for prediction.

Causal inference

Randomized experiments: women politicians and policy outcomes

Causal inference

QSS example: West Bengal (1990's)

```
dim(women)
```

```
## [1] 322   6
```

```
head(women)
```

```
##   GP village reserved female irrigation water
## 1  1       2        1      1          0    10
## 2  1       1        1      1          5     0
## 3  2       2        1      1          2     2
## 4  2       1        1      1          4    31
## 5  3       2        0      0          0     0
## 6  3       1        0      0          0     0
```

Causal inference

Promoting women's issues

```r
## drinking-water facilities
mean(women$water[women$reserved == 1]) -
    mean(women$water[women$reserved == 0])
```

```
## [1] 9.252423
```

```r
## Irrigation facilities
mean(women$irrigation[women$reserved == 1]) -
    mean(women$irrigation[women$reserved == 0])
```

```
## [1] -0.3693319
```

Causal inference

Promoting women's issues: regression analysis

```r
# Drinking water model
lm(water ~ reserved, data = women)
```

```
##
## Call:
## lm(formula = water ~ reserved, data = women)
##
## Coefficients:
## (Intercept)     reserved
##      14.738        9.252
```

```r
# Irrigation facilities model
lm(irrigation ~ reserved, data = women)
```

```
##
## Call:
## lm(formula = irrigation ~ reserved, data = women)
##
## Coefficients:
## (Intercept)     reserved
##      3.3879      -0.3693
```

Causal inference

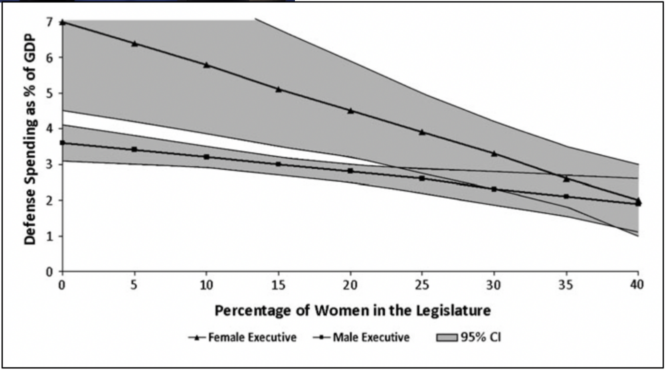
Binary dependent variable:

- slope coefficient (β) = diff-in-means estimator
- $\hat{\beta}$: estimated average treatment effect
- Effect with/without women leaders.

- Why works?
  - Randomization $\rightarrow$ causal interpretation

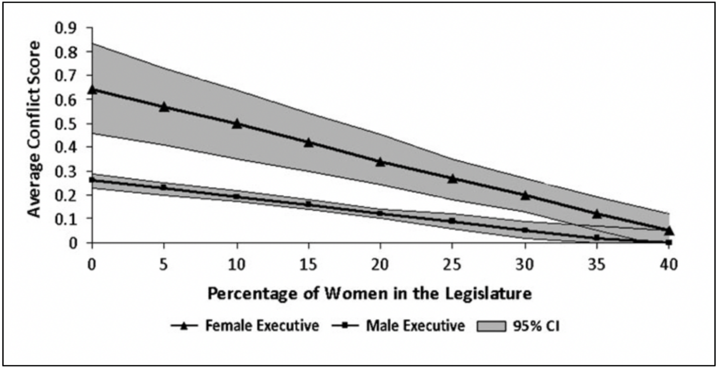
Women leaders of Government

WOMEN LEADERS & FOREIGN POLICY



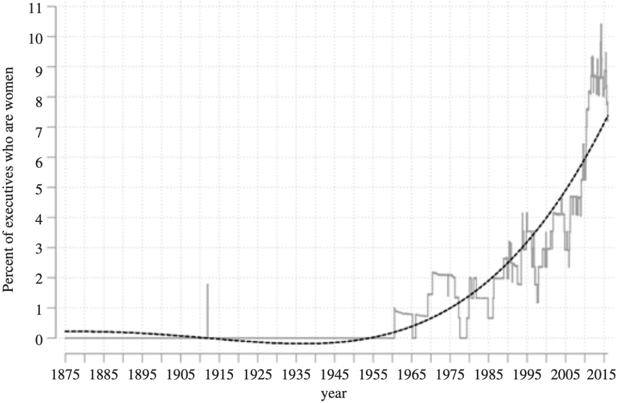
Women leaders in crisis

WOMEN LEADERS IN CONFLICT



Women leaders of Government

Schwartz and Blair (2020)



Women in crisis

Schwartz and Blair (2020)

- Audience costs $\rightarrow$ empty threat, inconsistency.
- Belligerence costs $\rightarrow$ issue a threat.
- Gender stereotypes: weak, ill-prepared, emotional.
- Leader competence: male-female dyads.

Women in crisis

- Design: experiment
- Treatments: dyads of conflict interactions.
- Outcome measures: approval (scale and binary).

```
dim(leader)
```

```
## [1] 2342   58
```

Women leaders

Gender stereotyping: small scale evidence

```r
### General: higher disapproval for women
mean(leader$Disapproval[leader$FemaleUS == 1], na.rm = T) -
  mean(leader$Disapproval[leader$FemaleUS == 0], na.rm = T)
```

```
## [1] 0.04998737
```

```r
mean(leader$DisapprovalBinary[leader$FemaleUS == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FemaleUS == 0], na.rm = T)
```

```
## [1] 0.01466212
```

```r
mean(leader$Disapproval[leader$FemaleOpp == 1], na.rm = T) -
  mean(leader$Disapproval[leader$FemaleOpp == 0], na.rm = T)
```

```
## [1] 0.131284
```

```r
mean(leader$DisapprovalBinary[leader$FemaleOpp == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FemaleOpp == 0], na.rm = T)
```

```
## [1] 0.0202939
```

Women leaders

```r
# Linear model coefficients == diff-in-means estimators
lm(DisapprovalBinary ~ FemaleUS, data = leader)
```

```
##
## Call:
## lm(formula = DisapprovalBinary ~ FemaleUS, data = leader)
##
## Coefficients:
## (Intercept)     FemaleUS
##     0.49831      0.01466
```

```r
lm(DisapprovalBinary ~ FemaleOpp, data = leader)
```

```
##
## Call:
## lm(formula = DisapprovalBinary ~ FemaleOpp, data = leader)
##
## Coefficients:
## (Intercept)    FemaleOpp
##     0.49521      0.02029
```

Gender and conflict approval

Inconsistency in male only vs. mixed dyads

```
# Male dyad <--> Male US, Female foreign
mean(leader$DisapprovalBinary[leader$MM_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$MF_NotEngage == 1], na.rm = T)
```

```
## [1] -0.05852317
```

```
# Male dyad <--> Female US, Male foreign
mean(leader$DisapprovalBinary[leader$MM_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FM_NotEngage == 1], na.rm = T)
```

```
## [1] -0.114592
```

Gender and audience costs

```r
mean(leader$DisapprovalBinary[leader$MM_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$MM_Engage == 1], na.rm = T)
```

```
## [1] 0.3262621
```

```r
mean(leader$DisapprovalBinary[leader$FM_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FM_Engage == 1], na.rm = T)
```

```
## [1] 0.5198552
```

```r
mean(leader$DisapprovalBinary[leader$MF_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$MF_Engage == 1], na.rm = T)
```

```
## [1] 0.4359946
```

```r
mean(leader$DisapprovalBinary[leader$FF_NotEngage == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FF_Engage == 1], na.rm = T)
```

```
## [1] 0.4980188
```

Gender and belligerence costs

Make a threat or not...

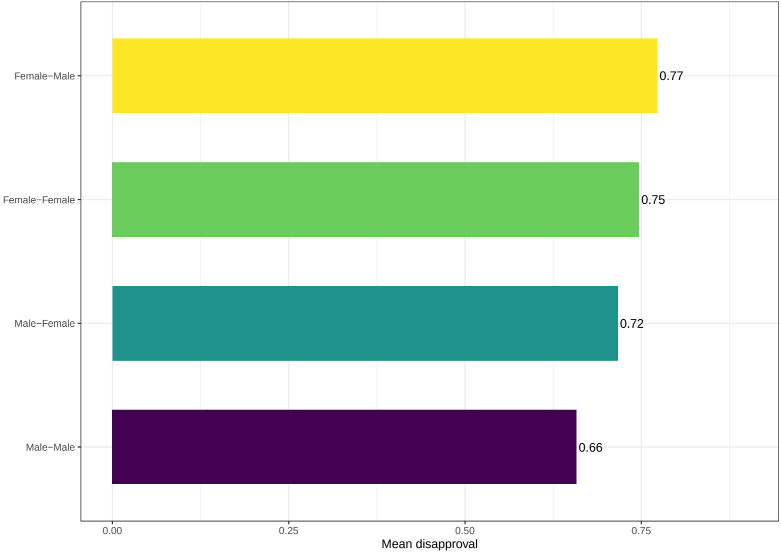
```r
# Belligerence costs by gender
mean(leader$DisapprovalBinary[leader$MM_StayOut == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$MM_Engage == 1], na.rm = T)
```

```
## [1] 0.135034
```

```r
mean(leader$DisapprovalBinary[leader$FM_StayOut == 1], na.rm = T) -
  mean(leader$DisapprovalBinary[leader$FM_Engage == 1], na.rm = T)
```

```
## [1] 0.2473684
```

Inconsistency in gender dyads



Binary predictors

Linear model elements:

- *Slope* (β): the average increase in Y when X increases by 1 unit.

When X is binary:

- Treatment: yes or no (female leader follow-through or not).
- X change by 1 unit $\rightarrow$ no to yes.
- Y (disapproval) changes as well (measured in percentages).

Regression model

Why sanctions fail?

	Likelihood of Success Versus Failure						
	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*	*Model 7*
Hypothesized Variables							
All Busters	−0.24 (0.12)**	−0.46 (0.15)***					
Black Knight Allies			0.05 (0.25)	−0.08 (0.27)			
Black Knight Great Powers					−0.27 (0.24)	−0.44 (0.40)	
HSE Black Knight							0.03 (0.67)
Control Variables							
US Cooperation		−0.99 (0.57)*		−0.93 (0.57)		−0.84 (0.57)	
IO Support		−2.76 (1.41)*		−2.56 (1.46)*		−2.17 (1.49)	
IO × Coop		1.59 (0.60)***		1.54 (0.61)**		1.37 (0.62)**	
US Defensive Alliance		−0.70 (0.76)		−0.59 (0.72)		−0.73 (0.77)	
Target Defense Alliances		0.00 (0.02)		0.00 (0.02)		0.00 (0.02)	
Modest Goal		1.82 (0.68)***		1.77 (0.68)***		1.73 (0.66)***	
Prior Relations		1.38 (0.46)***		1.37 (0.45)***		1.34 (0.46)***	
Democracy		−0.58 (0.71)		−0.46 (0.68)		−0.31 (0.71)	
Post-Cold War		−0.79 (0.64)		−0.79 (0.61)		−0.74 (0.64)	
Time	−0.08 (−0.18)	0.04 (0.77)	−0.01 (0.69)	−0.11 (0.76)	−0.09 (0.18)	−0.08 (0.76)	−0.11 (0.18)
$Time^2$	0.00 (−0.01)	0.03 (0.14)	0.04 (0.13)	0.05 (0.14)	0.00 (0.01)	0.05 (0.15)	0.00 (0.01)
$Time^3$	−0.00 (0.00)	−0.00 (0.01)	−0.00 (0.01)	−0.00 (0.01)	−.00 (.00)	−0.00 (0.01)	0.00 (0.00)
Constant	0.40 (−0.63)	−1.88 (1.59)	−0.77 (1.02)	−3.08 (1.61)*	−.12 (.51)	−2.79 (1.60)*	−0.25 (0.51)
Prob > X^2	0.02	0.00	0.07	0.00	0.02	0.00	0.01
Observations	840	753	789	753	840	753	840

Regression model

Multiple predictors

$$Y = \alpha + \beta_1 * X_1 + \beta_2 * X_2 + ... + \beta_p * X_p + \epsilon$$

How to interpret β_j?

- Change in Y with 1-unit increase in X_j...
- As all other predictors are **held constant**.
- Independent effect of each β.

Least squared: Multiple predictors

Sum of Squared Residuals (SSR)

$$SSR = \sum_{i=1}^{n} \hat{\epsilon}^2 = \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta}_1 * X_1 - \hat{\beta}_1 * X_1 - ... - \hat{\beta}_p * X_p)^2$$

- Estimate parameters: $\hat{\alpha}, \hat{\beta}_p$.
- Minimize SSR.

Women in crisis data

- Multiple predictors for leader's approval
- Using factor variables: binary outcome

```
### Generate a Factor variable
leader$inconsis_cond <- NA
leader$inconsis_cond[leader$MM_NotEngage == 1] <-"MM"
leader$inconsis_cond[leader$MF_NotEngage == 1] <-"MF"
leader$inconsis_cond[leader$FM_NotEngage == 1] <-"FM"
leader$inconsis_cond[leader$FF_NotEngage == 1] <-"FF"

# levels of factor
levels(factor(leader$inconsis_cond))
```

```
## [1] "FF" "FM" "MF" "MM"
```

Multiple binary predictors

$$Y(Disapproval) = \alpha + \beta_1 * MM + \beta_2 * MF + \beta_3 * FM + \beta_4 * FF + \epsilon$$

```r
fit <- lm(DisapprovalBinary ~ factor(inconsis_cond), data = leader)
fit
```

```
##
## Call:
## lm(formula = DisapprovalBinary ~ factor(inconsis_cond), data = leader)
##
## Coefficients:
##           (Intercept)  factor(inconsis_cond)FM  factor(inconsis_cond)MF
##               0.74661                  0.02588                 -0.03019
## factor(inconsis_cond)MM
##              -0.08871
```

Multiple binary predictors

Coefficients = diff-in-means??

```r
# Regression w/o the intercepts
fit3 <- lm(DisapprovalBinary ~ -1 + inconsis_cond, data = leader)
fit3
```

```
##
## Call:
## lm(formula = DisapprovalBinary ~ -1 + inconsis_cond, data = leader)
##
## Coefficients:
## inconsis_condFF  inconsis_condFM  inconsis_condMF  inconsis_condMM
##          0.7466           0.7725           0.7164           0.6579
```

Multiple binary predictors

Same with tapply()

```
tapply(leader$DisapprovalBinary, leader$inconsis_cond, mean)
```

```
##        FF        FM        MF        MM
## 0.7466063 0.7724868 0.7164179 0.6578947
```

Average treatment effect versus control (MM dyad)

```
# Using coef() function
coef(fit3)["inconsis_condFM"] - coef(fit3)["inconsis_condMM"]
```

```
## inconsis_condFM
##        0.114592
```

```
coef(fit3)["inconsis_condFF"] - coef(fit3)["inconsis_condMM"]
```

```
## inconsis_condFF
##       0.0887116
```

Model fit: multiple predictors

R^2 with multiple predictors $\rightarrow$ Adjusted R^2

Degrees of freedom (DOF):

- How many observations vary ‘freely’?
- DOF: $(n - p - 1) = n - (p + 1)$
- Multiple predictors $\rightarrow$ larger R^2
- Large sample (data) $\rightarrow$ not much difference b-w R^2 and adjusted R^2

Model fit: multiple predictors

R^2 and adjusted R^2 in regression model

```r
# summary() model
summary(lm(DisapprovalBinary ~ MF_NotEngage + FM_NotEngage +
             FF_NotEngage, data = leader))
```

```
##
## Call:
## lm(formula = DisapprovalBinary ~ MF_NotEngage + FM_NotEngage +
##     FF_NotEngage, data = leader)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7725 -0.4211  0.2275  0.5789  0.5789
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.42114    0.01153  36.532  < 2e-16 ***
## MF_NotEngage  0.29527    0.03574   8.262 2.38e-16 ***
## FM_NotEngage  0.35134    0.03674   9.562  < 2e-16 ***
## FF_NotEngage  0.32546    0.03426   9.499  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4796 on 2338 degrees of freedom
## Multiple R-squared:  0.08127,    Adjusted R-squared:  0.08009
## F-statistic: 68.94 on 3 and 2338 DF,  p-value: < 2.2e-16
```

Heterogenous treatment effects

- Variation in effect of main predictor
- When?
- ATE vary among individuals: positive/negative
- Experiments: differences guide treatment assignment

Women leaders:

- Respondents' gender and views of leader
- Do women judge female leaders more harshly?

Heterogenous treatment effects

Leader criticism by respondents gender

```r
# Subset of female respondents
lead.gen <- subset(leader, Gender == 1)

# Diff-in-means: support for female versus male leader
mean(lead.gen$Disapproval[lead.gen$FemaleUS == 1], na.rm = T) -
  mean(lead.gen$Disapproval[lead.gen$FemaleUS == 0], na.rm = T)
```

```
## [1] -0.06103819
```

```r
# Subset of male respondents
lead.gen2 <- subset(leader, Gender == 0)

# Diff-in-means: support for female versus male leader
mean(lead.gen2$Disapproval[lead.gen2$FemaleUS == 1], na.rm = T) -
  mean(lead.gen2$Disapproval[lead.gen2$FemaleUS == 0], na.rm = T)
```

```
## [1] 0.1652623
```

Estimated ATE

```r
# Estimated treatment effect for gender
(mean(lead.gen$Disapproval[lead.gen$FemaleUS == 1], na.rm = T) -
  mean(lead.gen$Disapproval[lead.gen$FemaleUS == 0], na.rm = T)) -
(mean(lead.gen2$Disapproval[lead.gen2$FemaleUS == 1], na.rm = T) -
  mean(lead.gen2$Disapproval[lead.gen2$FemaleUS == 0], na.rm = T))
```

```
## [1] -0.2263005
```

- Women respondents are less critical on female leaders

Regression model: conditional effects

- Add predictor to the model

$$Y(Disapproval) = \alpha + \beta_1 * LeaderDyad + \beta_2 * RespondentGender + \epsilon$$

- However, *conditional effect* $\rightarrow$ Interaction model

$$Y(Disapproval) = \alpha + \beta_1 * LeaderDyad + \beta_2 * RespondentGender + \\ \beta_3 * LeaderDyad * RespondentGender + \epsilon$$

Interaction models

$$Y = \alpha + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_1 * X_2 + \epsilon$$

- Coefficient β_3: How X_1 depends on X_2.
- Average effect of women respondent (and leader): $\beta_2 + \beta_3$.
- Average effect of men respondent: β_2.

Interaction model in R

Syntax: use the (*) or (:) between factors

```r
# Female leader and respondents gender: Interaction model
summary(lm(Disapproval ~ FemaleUS * Gender, data = leader))
```

```
##
## Call:
## lm(formula = Disapproval ~ FemaleUS * Gender, data = leader)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5809 -1.4157  0.4191  1.4398  2.5843
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.41567    0.06740  65.511   <2e-16 ***
## FemaleUS         0.16526    0.09664   1.710   0.0874 .
## Gender           0.14453    0.09488   1.523   0.1278
## FemaleUS:Gender -0.22630    0.13515  -1.674   0.0942 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.633 on 2334 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.001536,   Adjusted R-squared:  0.0002522
## F-statistic: 1.197 on 3 and 2334 DF,  p-value: 0.3096
```

Interaction model: continous predictors

- How the average treatment effect varies along age scale?
- Linearity assumption: one-unit increase in predictor $\rightarrow$ similar increase in outcome.

- Data: ICB (observational).
- Variables:
  - International crises: 1918-2015.
  - Y: Crisis management technique (how to respond).
  - X_1: Trigger event severity/type
  - X_2: Leaders' age.
  - Model: how response varies based on tirgger event (and leader's age).

Interaction model: ICB data

$$CrisisAction = \alpha + \beta_1 * Trigger + \beta_2 * Age + \beta_3 * Trigger * Age + \epsilon$$

	cracid	actor	systrgyr	systrgda	crisname	leader	cris_date	triggr	crismg	lead_age
1	2	USA	1937	12	PANAY INCIDENT	Roosevelt, F.	12/12/37	9	1	55
2	2	USA	1946	7	TURKISH STRAITS	Truman	8/7/46	2	4	62
3	2	USA	1947	21	TRUMAN DOCTRINE	Truman	2/21/47	2	4	63
4	2	USA	1948	24	BERLIN BLOCKADE	Truman	6/24/48	3	4	64
5	2	USA	1948	23	CHINA CIVIL WAR	Truman	9/23/48	8	1	64
6	2	USA	1950	25	KOREAN WAR I	Truman	6/25/50	8	8	66
7	2	USA	1950	30	KOREAN WAR II	Truman	9/30/50	9	8	66
8	2	USA	1953	16	KOREAN WAR III	Eisenhower	4/16/53	9	7	63
9	2	USA	1953	12	GUATEMALA	Eisenhower	12/12/53	7	4	63
10	2	USA	1954	13	DIEN BIEN PHU	Eisenhower	3/13/54	2	1	64
11	2	USA	1954	3	TAIWAN STRAIT I	Eisenhower	9/3/54	8	4	64
12	2	USA	1956	29	SUEZ NATN.-WAR	Eisenhower	10/29/56	5	6	66
13	2	USA	1957	18	SYRIA/TURKEY CONFRNT.	Eisenhower	8/18/57	2	4	67
14	2	USA	1958	8	IRAQ/LEB. UPHEAVAL	Eisenhower	5/8/58	2	6	68
15	2	USA	1958	17	TAIWAN STRAIT II	Eisenhower	7/17/58	8	1	68
16	2	USA	1958	27	BERLIN DEADLINE	Eisenhower	11/27/58	2	1	68
17	2	USA	1961	9	PATHET LAO OFFENSIVE	Kennedy	3/9/61	8	1	44
18	2	USA	1961	15	BAY OF PIGS	Kennedy	4/15/61	2	5	44

Interaction model: ICB data

Outcome - crisis management method:

- Negotiation, mediation
- Non-military pressure (economic)
- Non-violent military
- Violence

Predictor - triggering event: Verbal/political act, violent act.

```
summary(mydata$lead_age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   18.00   48.00   56.00   55.84   64.00   91.00       2
```

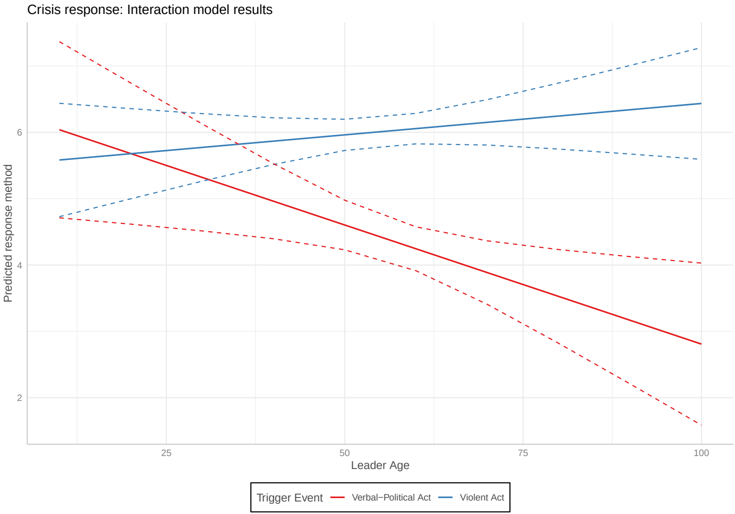
Interaction model: ICB data

```r
summary(fit.age <- lm(crismg ~ triggr * lead_age, data = mydata))
```

```
##
## Call:
## lm(formula = crismg ~ triggr * lead_age, data = mydata)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.2086 -1.6012  0.9619  1.8246  4.0730
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      6.512835   0.935138   6.965 6.24e-12 ***
## triggr          -0.113761   0.134857  -0.844  0.39913
## lead_age        -0.041579   0.016074  -2.587  0.00984 **
## triggr:lead_age  0.005672   0.002337   2.427  0.01541 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.387 on 927 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.06487,    Adjusted R-squared:  0.06184
## F-statistic: 21.44 on 3 and 927 DF,  p-value: 1.984e-13
```

Interaction model: ICB data

Heterogeneous treatment effects: trigger over age



Causality with observational data

- The problem of *free riding*

Leaders and alliance contribution

Business experience and military alliances (Fuhrmann 2020):

- Leader experience explain variations.
- Business: executive level.
- Smaller contributions (defense expenditures), Why?
- Egoistic tendencies.
- Belief in self-efficacy and power.

Our goals:

1. Evaluate casual effect with linear regression (Δ spending per year).
2. Run *placebo test*: strengthen the proposed causal links.

Alliance contribution

NATO Defense spending data (1949-2020)

```
head(matt1)
```

```
## # A tibble: 6 x 74
##   Country ccode  `1949`  `1950`  `1951`  `1952`  `1953`  `1954`  `1955`  `19
##   <chr>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <d
## 1 Canada     20      NA   3809.   7718.  12405.  14234.  13242.  13113.  133
## 2 USA         2 147593. 158620. 339387. 478080. 492223. 424699. 402015. 4072
## 3 Czechia   316      NA      NA      NA      NA      NA      NA      NA
## 4 Hungary   310      NA      NA      NA      NA      NA      NA      NA
## 5 Poland    290      NA      NA      NA      NA      NA      NA      NA
## 6 Belgium   211   2074.   2092.   3095.   4574.   4554.   4698.   3891.   37
## # ... with 64 more variables: 1957 <dbl>, 1958 <dbl>, 1959 <dbl>, 1960 <dbl>
## #   1961 <dbl>, 1962 <dbl>, 1963 <dbl>, 1964 <dbl>, 1965 <dbl>, 1966 <dbl>,
## #   1967 <dbl>, 1968 <dbl>, 1969 <dbl>, 1970 <dbl>, 1971 <dbl>, 1972 <dbl>,
## #   1973 <dbl>, 1974 <dbl>, 1975 <dbl>, 1976 <dbl>, 1977 <dbl>, 1978 <dbl>,
## #   1979 <dbl>, 1980 <dbl>, 1981 <dbl>, 1982 <dbl>, 1983 <dbl>, 1984 <dbl>,
## #   1985 <dbl>, 1986 <dbl>, 1987 <dbl>, 1988 <dbl>, 1989 <dbl>, 1990 <dbl>,
## #   1991 <dbl>, 1992 <dbl>, 1993 <dbl>, 1994 <dbl>, 1995 <dbl>, 1996 <dbl>,
```

Leaders and military alliances expenditures

NATO leaders and defense spending data

ccode \textsc{COW numeric country code}	year \textsc{year}	leadername \textsc{leader name}	business \textsc{business experience}	Country	def.exp	def.delta
2	2003	G.W. Bush	1	USA	612232.612	13.81651492
2	2004	G.W. Bush	1	USA	667284.639	8.99201159
2	2005	G.W. Bush	1	USA	698019.039	4.60589054
2	2006	G.W. Bush	1	USA	708077.303	1.44097276
2	2007	G.W. Bush	1	USA	726971.529	2.66838457
2	2008	G.W. Bush	1	USA	779854.123	7.27436936
2	2009	Obama	0	USA	841220.473	7.86895241
2	2010	Obama	0	USA	865268.025	2.85865034
2	2011	Obama	0	USA	855022.313	-1.18410840
2	2012	Obama	0	USA	807530.267	-5.55448034
2	2013	Obama	0	USA	745415.975	-7.69188406
2	2014	Obama	0	USA	699563.842	-6.15121420
20	1949	St. Laurent	0	Canada	*NA*	*NA*
20	1950	St. Laurent	0	Canada	3808.656	*NA*
20	1951	St. Laurent	0	Canada	7718.028	102.64439720
20	1952	St. Laurent	0	Canada	12404.681	60.72344453
20	1953	St. Laurent	0	Canada	14234.412	14.75032982

Testing a causal mechanism

Does business experience matter?

```r
# subsets by business experience
no.business <- subset(def.matt, subset = (business == 0))
business <- subset(def.matt, subset = (business == 1))

## Diff-in-means estimator
mean(business$def.delta, na.rm = T) -
  mean(no.business$def.delta, na.rm = T)
```

```
## [1] -2.134511
```

```r
# Regression model
lm(def.delta ~ business, data = def.matt)
```

```
##
## Call:
## lm(formula = def.delta ~ business, data = def.matt)
##
## Coefficients:
## (Intercept)     business
##       2.847       -2.135
```

The Placebo test

- Data: non-defense related expenses
- Business experience matters $\rightarrow$ not on other issues.

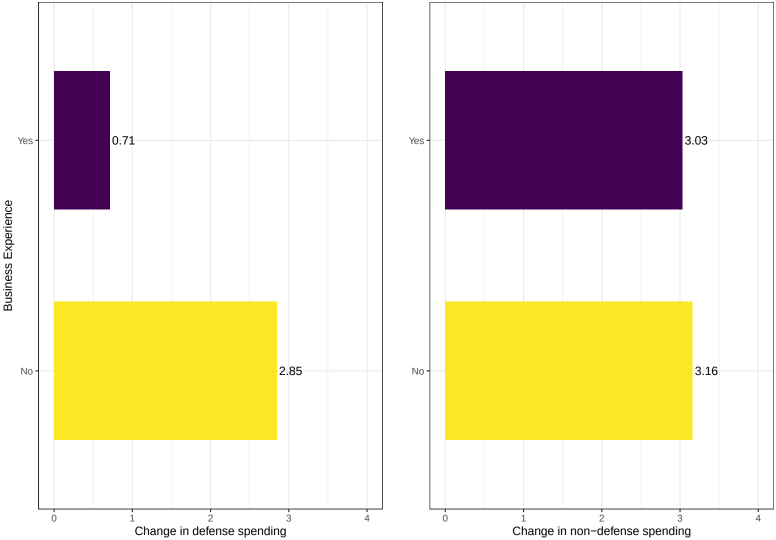
```r
## Diff-in-means estimator: non-defense spending
mean(business$nondefspend_ch, na.rm = T) -
  mean(no.business$nondefspend_ch, na.rm = T)
```

```
## [1] -0.1239881
```

```r
# Regression model
lm(nondefspend_ch ~ business, data = def.matt)
```

```
##
## Call:
## lm(formula = nondefspend_ch ~ business, data = def.matt)
##
## Coefficients:
## (Intercept)     business
##       3.164       -0.124
```

Businessmen, politicians and spending



Wrapping up week 8

Summary:

- Prediction and causal inference.
- Binary predictors and linear regression models.
- Multiple predictors.
- Heterogeneous effects: interaction models.
- Causal inference with observational data.

Task 3