Applied Controlled Comparisons

POSC 3410 – Quantitative Methods in Political Science

Steven V. Miller

Department of Political Science

Goals for Today

- *Give applied examples of controlled comparisons and controlled relationships.*
- *Show how to do some of these with R code.*
- *Discuss controlled effects and partial effects.*

Three Types of Controlled Relationships

1. Spurious relationship
2. Additive relationship
3. Interactive relationship

Notice our discussion here is *not* bivariate.

- We need to consider "control" for potential confounders in a bivariate relationship.

Applied Controlled Comparisons

We can summarize these controlled relationships with:

1. Cross-tabulation
2. Mean comparison analysis

We will also use actual data.

Example Data

We'll be going over two toy examples, largely patterned off Pollock:

1. Attitudes about gun control as a function of partisanship and gender.
2. Vote choice in 2016 as a function of abortion attitudes and issue importance.

Data come from General Social Survey (2018) and ANES (2020, ETS), respectively.

R Packages

```r
library(tidyverse) # for most things
library(stevemisc) # for recoding stuff
library(qs) # for loading my version of the GSS data
library(knitr) # for tables
library(kableExtra) # for prettier tables
```

```r
# load data, assuming this qs object
# see: http://svmiller.com/blog/2020/02/comparing-qs-fst-rds-for-bigger-datasets/
gunlaw <- qread("/home/steve/Dropbox/data/gss/GSS_spss-2018/gss7218.qs") %>%
  select(year, gunlaw, partyid, sex) %>%
  filter(year == 2018) %>%
  mutate(gender = case_when(
    sex == 2 ~ "Female",
    sex == 1 ~ "Male"
  ),
  gunlaw = carr(gunlaw, "1=1;2=0"),
  gunlawcat = ifelse(gunlaw == 1, "Favor", "Oppose"),
         pidcat = case_when(
           partyid %in% c(0,1, 2) ~ "Democrat/Lean Democrat",
           #partyid == 3 ~ "Independent", # omit indies
           partyid %in% c(4,5,6) ~ "Republican/Lean Republican"
         ))
```

```r
# load data, from here...
# https://electionstudies.org/data-center/2020-exploratory-testing-survey/
haven::read_dta("~/Dropbox/data/anes/2020-ets/anes_pilot_2020ets_dta.dta") %>%
  # select just what we want
  select(abort1, abort_imp, pid7, vote16) %>%
  # recode that we get the "never permit abortion" people as 1s
  # and whether the issue is extremely or very important
  # also lump leaners with the partisans and create vote choice variables
  mutate(abort_never = carr(abort1, "9=NA;1=1;2:4=0"),
         abort_impdum = carr(abort_imp, "9=NA;1:2=1;3:5=0"),
         vote16cat = case_when(
           vote16 == 1 ~ "Donald Trump",
           vote16 == 2 ~ "Hillary Clinton",
           vote16 == 3 ~ "Someone Else",
           vote16 == 4 ~ "Did Not Vote"
         )) %>% haven::zap_labels() %>%
  mutate(votetrump = ifelse(vote16cat == "Donald Trump", 1, 0)) -> abortimport
```

Partisanship and Gun Control (GSS, 2018)

Table 1: The Relationship Between Partisanship and Support for Gun Control

Opinion on Gun Permits	**D/Lean D**	**R/Lean R**	**Total**
Favor	82.09%	59.22%	72.43%
Oppose	17.91%	40.78%	27.57%
Total	*100%*	*100%*	*100%*

Note: Data: General Social Survey (2018)

Zero-order Relationship

This table is a zero-order relationship.

- The effect of x on y **not** controlling for z.

In 2018, the zero-order effect of partisanship on support for gun control was about -22.87 percentage points.

Partisanship and Gun Control (GSS, 2018)

Table 2: Partisanship and Support for Gun Control, Controlling for Gender

	Women			*Men*		
Opinion on Gun Permits	**D/Lean D**	**R/Lean R**	**Total**	**D/Lean D**	**R/Lean R**	**Total**
Favor	85.89%	64.91%	77.58%	76.87%	53.06%	66.05%
Oppose	14.11%	35.09%	22.42%	23.13%	46.94%	33.95%
Total	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*

Note: Data: General Social Survey (2018)

Controlled Comparison Table

That was a **controlled comparison table**.

- It shows the relationship between x and y for each (truncated) value of z.

These tables reveal two types of effects.

1. Controlled effect
2. Partial effect

Controlled Effect

A **controlled effect** is a relationship between *x* and *y* within one value of *z*.

- It could also be between *z* and *y* within one value of *x*.

We obtain the controlled effect of partisanship for both women and men.

- 85.89% of female Dems/Dem leaners favor these permits to 64.91% of Rs/R leaners.
  - The controlled effect is -20.98 percentage points.
- 76.87% of male Dems/Dem leaners favor these permits to 53.06% of Rs/R leaners.
  - The controlled effect is -23.81 percentage points.

Partial Effect

We summarize controlled effects as **partial effect**.

- Summarizes a relationship between two variables taking into account rival variables.

It's tempting, but *don't average the two controlled effects*.

- Doing so assumes the samples are equal.

Instead: weight the controlled effect by percentage of the sample.

- Then: add those.

```r
gunlaw %>%
  select(gunlawcat, gender, pidcat) %>%
  na.omit %>%
  group_by(gender) %>% tally() %>%
  mutate(prop = n/sum(n),
         contreff = c(-20.98, -23.81),
         product = prop*contreff,
         parteff = sum(product))
```

```
## # A tibble: 2 x 6
##   gender     n  prop contreff product parteff
##   <chr>  <int> <dbl>    <dbl>   <dbl>   <dbl>
## 1 Female   669 0.554    -21.0   -11.6   -22.2
## 2 Male     539 0.446    -23.8   -10.6   -22.2
```

The partial effect of partisanship on gun control opinions is about -22.2 percentage points.

Partial Effect

What is the partial effect of *gender* on gun control attitudes?

- Sounds weird to ask. No one “increases” in gender.

Follow the **rule of direction for nominal relationships**.

- Treat the left-most column as the base category (here: women).

Partial Effect

Controlled effects:

- 85.89% of female Dems/leaners favor these permits to 76.87% of male Dems/leaners (-9.02%).
- 64.91% of female Rs/leaners favor permits to 53.06% of male Republicans (-11.85%).

Weight the controlled effects (*by partisanship*) to get a partial effect.

```r
gunlaw %>%
  select(gunlawcat, gender, pidcat) %>%
  na.omit %>%
  group_by(pidcat) %>% tally() %>%
  mutate(prop = n/sum(n),
         contreff = c(-9.02, -11.85),
         product = prop*contreff,
         parteff = sum(product))
```

```
## # A tibble: 2 x 6
##   pidcat                          n  prop contreff product parteff
##   <chr>                       <int> <dbl>    <dbl>   <dbl>   <dbl>
## 1 Democrat/Lean Democrat        698 0.578    -9.02   -5.21   -10.2
## 2 Republican/Lean Republican    510 0.422   -11.8    -5.00   -10.2
```

The partial effect of gender on gun control opinions (controlling for partisanship) is about -10.2 percentage points.

Identifying the Pattern

Ask the following three questions for relationships among x, y, and z.

1. Does a relationship exist between x and y in at least one value of z?
2. Is the tendency (i.e. positive or negative) the same at all values of z?
3. Is the magnitude effect the same or close to it in all values of z?

Identifying the Pattern

If the answer to the first question is no, you can stop there.

- It's a spurious relationship.

If the answer to the second question is no, you can stop there.

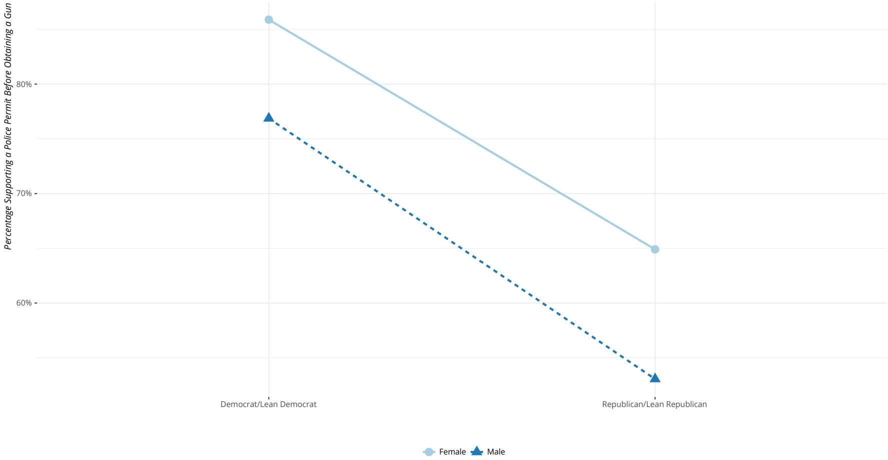
- There's an interaction effect.

If the answer to the third question is no, there's an interaction.

- If "yes", it's an additive relationship.

A Line Chart of Partisanship and Support for Gun Control, Controlling for Gender

The lines are effectively the same, implying an additive relationship. The distance between two points on the same part of the x-axis is the controlled effect of gender.



Data: General Social Survey

Abortion Opinions, Issue Salience, and Vote Choice (ANES, 2020)

Table 3: Abortion Permissibility and the Trump Vote, Controlling for Issue Importance

	Not Very Important			*Very Important*		
Voted for Trump?	**Permit**	**Never Permit**	**Total**	**Permit**	**Never Permit**	**Total**
Yes	33.33%	42.37%	33.79%	33.48%	62.13%	40.16%
No	66.67%	57.63%	66.21%	66.52%	37.87%	59.84%
Total	*100%*	*100%*	*100%*	*100%*	*100%*	*100%*

Note: Data: American National Election Studies (2020 Exploratory Testing Survey)

Abortion Opinions, Salience, and Vote Choice

The controlled effect of abortion opinion:

- 33.33 -> 42.37 for “not very important” = 9.04
- 33.48 -> 62.13 for “very important” = 28.65

Controlled effect of issue importance:

- 33.33 -> 33.48 for “permit” = 0.15
- 42.37 -> 62.13 for “never permit” = 19.76

Something already looks a lot different here.

```r
abortimport %>%
  select(abort_impdum, votetrump, abort_never) %>%
  na.omit %>%
  group_by(abort_impdum) %>% tally() %>%
  mutate(prop = n/sum(n),
         contreff = c(9.04, 28.65),
         product = prop*contreff,
         parteff = sum(product))
```

```
## # A tibble: 2 x 6
##   abort_impdum     n  prop contreff product parteff
##          <dbl> <int> <dbl>    <dbl>   <dbl>   <dbl>
## 1            0  1322 0.431     9.04    3.90    20.2
## 2            1  1742 0.569    28.6    16.3     20.2
```

The partial effect of issue importance is about 20.2 percentage points.

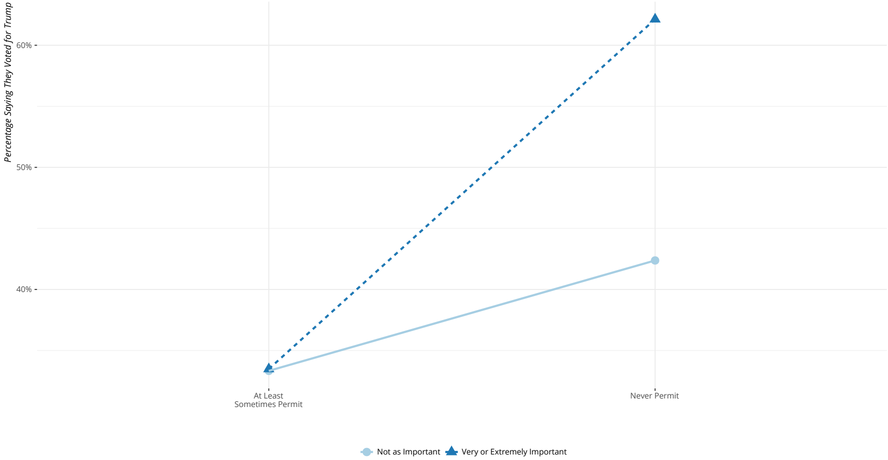
```r
abortimport %>%
  select(abort_impdum, votetrump, abort_never) %>%
  na.omit %>%
  group_by(abort_never) %>% tally() %>%
  mutate(prop = n/sum(n),
         contreff = c(.15, 19.76),
         product = prop*contreff,
         parteff = sum(product))
```

```
## # A tibble: 2 x 6
##   abort_never     n  prop contreff product parteff
##         <dbl> <int> <dbl>    <dbl>   <dbl>   <dbl>
## 1           0  2601 0.849     0.15   0.127    3.11
## 2           1   463 0.151    19.8    2.99     3.11
```

The partial effect of attitudes about abortion is about 3.11 percentage points.

A Line Chart of Abortion Attitudes and the Trump Vote, Controlling for the Importance of Abortion as an Issue

The lines show a different effect of abortion attitudes by importance of the issue, suggesting a clear interactive effect.



Data: American National Election Studies (2020 Exploratory Testing Survey)

Conclusion

No causal statement can be made as a zero-order relationship

- This will get more complicated in multiple regression.
- Fortunately, computers do the heavy lifting for us.

Get comfortable making these types of controlled comparisons within a simple three-variable context.

Table of Contents