Regression for Causal Inference

Satoshi Ido

```
library("ggplot2")
library("tidyverse")
library("MASS")
library("broom")
```

Data

```
email_data <- read.csv("/Users/satoshiido/Documents/programming/statistical-analysis/causal_inference/da
head(email_data)
```

```
##   recency history_segment history mens womens  zip_code newbie channel
## 1      10  2) $100 - $200  142.44    1      0 Surburban      0   Phone
## 2       6  3) $200 - $350  329.08    1      1     Rural      1     Web
## 3       7  2) $100 - $200  180.65    0      1 Surburban      1     Web
## 4       9  5) $500 - $750  675.83    1      0     Rural      1     Web
## 5       2    1) $0 - $100   45.34    1      0     Urban      0     Web
## 6       6  2) $100 - $200  134.83    0      1 Surburban      0   Phone
##         segment visit conversion spend
## 1 Womens E-Mail     0          0     0
## 2     No E-Mail     0          0     0
## 3 Womens E-Mail     0          0     0
## 4   Mens E-Mail     0          0     0
## 5 Womens E-Mail     0          0     0
## 6 Womens E-Mail     1          0     0
```

```
# create the data w/o the womens E-Mail campaign
male_df <- email_data %>%
    filter(segment != "Womens E-Mail") %>%
    mutate(treatment = if_else(segment == "Mens E-Mail", 1, 0))

# create the selection biased data set the seed
set.seed(1)

## make half depending on the condition
obs_rate_c <- 0.5
obs_rate_t <- 0.5

## create the biased data
biased_data <- male_df %>%
    mutate(obs_rate_c = if_else((history > 300) | (recency < 6) | (channel == "Multichannel"),
        obs_rate_c, 1), obs_rate_t = if_else((history > 300) | (recency < 6) | (channel ==
        "Multichannel"), 1, obs_rate_t), random_number = runif(n = NROW(male_df))) %>%
    filter((treatment == 0 & random_number < obs_rate_c) | (treatment == 1 & random_number <
```

```
        obs_rate_t))

head(biased_data, 20)
```

```
##    recency  history_segment history mens womens  zip_code newbie      channel
## 1        6   3) $200 - $350  329.08    1      1     Rural      1          Web
## 2        9   5) $500 - $750  675.83    1      0     Rural      1          Web
## 3        9   5) $500 - $750  675.07    1      1     Rural      1        Phone
## 4        2   2) $100 - $200  101.64    0      1     Urban      0          Web
## 5        4   3) $200 - $350  241.42    0      1     Rural      1 Multichannel
## 6        5     1) $0 - $100   29.99    1      0 Surburban      0        Phone
## 7        5 6) $750 - $1,000  828.42    1      0 Surburban      1 Multichannel
## 8        9     1) $0 - $100   29.99    0      1 Surburban      1        Phone
## 9       11   2) $100 - $200  182.32    1      0 Surburban      0        Phone
## 10       2   2) $100 - $200  118.40    1      0 Surburban      0          Web
## 11       2     1) $0 - $100   29.99    0      1     Urban      1        Phone
## 12       6   2) $100 - $200  139.87    0      1     Rural      1          Web
## 13       7   4) $350 - $500  435.73    0      1     Urban      1          Web
## 14       9   3) $200 - $350  334.24    1      0     Urban      0          Web
## 15       6   2) $100 - $200  128.01    0      1     Urban      0          Web
## 16       1   5) $500 - $750  514.52    0      1 Surburban      1          Web
## 17       4 6) $750 - $1,000  766.47    1      1     Urban      1 Multichannel
## 18       7   5) $500 - $750  520.43    0      1 Surburban      1          Web
## 19      11   3) $200 - $350  236.97    1      1     Urban      1        Phone
## 20       3     1) $0 - $100   99.23    1      0     Rural      0          Web
##        segment visit conversion spend treatment obs_rate_c obs_rate_t
## 1    No E-Mail     0          0     0         0        0.5        1.0
## 2  Mens E-Mail     0          0     0         1        0.5        1.0
## 3  Mens E-Mail     0          0     0         1        0.5        1.0
## 4  Mens E-Mail     1          0     0         1        0.5        1.0
## 5    No E-Mail     0          0     0         0        0.5        1.0
## 6  Mens E-Mail     0          0     0         1        0.5        1.0
## 7  Mens E-Mail     0          0     0         1        0.5        1.0
## 8    No E-Mail     0          0     0         0        1.0        0.5
## 9  Mens E-Mail     0          0     0         1        1.0        0.5
## 10 Mens E-Mail     1          0     0         1        0.5        1.0
## 11   No E-Mail     0          0     0         0        0.5        1.0
## 12 Mens E-Mail     0          0     0         1        1.0        0.5
## 13   No E-Mail     0          0     0         0        0.5        1.0
## 14 Mens E-Mail     0          0     0         1        0.5        1.0
## 15 Mens E-Mail     0          0     0         1        1.0        0.5
## 16 Mens E-Mail     0          0     0         1        0.5        1.0
## 17 Mens E-Mail     0          0     0         1        0.5        1.0
## 18 Mens E-Mail     0          0     0         1        0.5        1.0
## 19 Mens E-Mail     0          0     0         1        1.0        0.5
## 20 Mens E-Mail     1          0     0         1        0.5        1.0
##    random_number
## 1     0.26550866
## 2     0.37212390
## 3     0.57285336
## 4     0.90820779
## 5     0.20168193
## 6     0.94467527
## 7     0.06178627
```

```
## 8     0.20597457
## 9     0.17655675
## 10    0.68702285
## 11    0.38410372
## 12    0.49769924
## 13    0.38003518
## 14    0.93470523
## 15    0.21214252
## 16    0.12555510
## 17    0.26722067
## 18    0.38611409
## 19    0.01339033
## 20    0.38238796
```

Regression

```
# regression
biased_reg <- lm(data = biased_data, formula = spend ~ treatment + recency + history)
summary(biased_reg)
```

```
##
## Call:
## lm(formula = spend ~ treatment + recency + history, data = biased_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.55  -1.49  -1.17  -0.49 497.99
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.6166555  0.2377393   2.594  0.00950 **
## treatment    0.8446137  0.1782792   4.738 2.17e-06 ***
## recency     -0.0401840  0.0259462  -1.549  0.12145
## history      0.0009723  0.0003455   2.815  0.00489 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.36 on 31859 degrees of freedom
## Multiple R-squared:  0.001414,   Adjusted R-squared:  0.00132
## F-statistic: 15.04 on 3 and 31859 DF,  p-value: 8.92e-10
```

```
biased_reg <- lm(data = biased_data, formula = spend ~ treatment + history)
summary(biased_reg)
```

```
##
## Call:
## lm(formula = spend ~ treatment + history, data = biased_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.74  -1.46  -1.26  -0.48 497.74
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) 0.3241996  0.1444390   2.245  0.02480 *
## treatment   0.9026109  0.1743057   5.178 2.25e-07 ***
## history     0.0010927  0.0003366   3.246  0.00117 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.36 on 31860 degrees of freedom
## Multiple R-squared:  0.001339,   Adjusted R-squared:  0.001276
## F-statistic: 21.35 on 2 and 31860 DF,  p-value: 5.406e-10
```

```
# check only for treatment coefficient
biased_reg_coef <- tidy(biased_reg)
biased_reg_coef
```

```
## # A tibble: 3 x 5
##   term        estimate std.error statistic     p.value
##   <chr>          <dbl>     <dbl>     <dbl>       <dbl>
## 1 (Intercept)  0.324    0.144         2.24 0.0248
## 2 treatment    0.903    0.174         5.18 0.000000225
## 3 history      0.00109  0.000337      3.25 0.00117
```

Biases in regression

```
# simple regression with RCT data
rct_reg <- lm(data = male_df, formula = spend ~ treatment)
rct_reg_coef <- summary(rct_reg) %>%
    tidy()

# simple regression with biased data
nonrct_reg <- lm(data = biased_data, formula = spend ~ treatment)
nonrct_reg_coef <- summary(nonrct_reg) %>%
    tidy()

rct_reg_coef
```

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)    0.653     0.103      6.36 2.09e-10
## 2 treatment      0.770     0.145      5.30 1.16e- 7
```

```
nonrct_reg_coef
```

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic      p.value
##   <chr>          <dbl>     <dbl>     <dbl>        <dbl>
## 1 (Intercept)    0.548     0.127      4.32 0.0000156
## 2 treatment      0.979     0.173      5.67 0.0000000143
```

Regression with some covariates

```
nonrct_mreg <- lm(data = biased_data, formula = spend ~ treatment + recency + channel +
    history)
nonrct_mreg_coef <- summary(nonrct_mreg)
# suppress selection bias a little by controlling for covariates
nonrct_mreg_coef
```

```
##
## Call:
## lm(formula = spend ~ treatment + recency + channel + history,
##     data = biased_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.62  -1.51  -1.17  -0.51 497.88
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.5024129  0.3793847   1.324  0.18542
## treatment     0.8465757  0.1784760   4.743 2.11e-06 ***
## recency      -0.0402666  0.0259470  -1.552  0.12070
## channelPhone -0.0017789  0.3040193  -0.006  0.99533
## channelWeb    0.2261596  0.3034664   0.745  0.45612
## history       0.0010299  0.0003754   2.744  0.00608 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.36 on 31857 degrees of freedom
## Multiple R-squared:  0.001467,   Adjusted R-squared:  0.00131
## F-statistic:  9.36 on 5 and 31857 DF,  p-value: 6.335e-09
```

OVB

```
# pull `treatment` parameters from the model A, B, C
treatment_coef <- df_results %>%
    filter(term == "treatment") %>%
    pull(estimate)

# pull `history` parameters from the model B
history_coef <- df_results %>%
    filter(model_index == "reg_B", term == "history") %>%
    pull(estimate)

# check OVB (beta_4 * gamma_1)
OVB <- history_coef * treatment_coef[3]
coef_gap <- treatment_coef[1] - treatment_coef[2]
OVB
```

```
## [1] 0.02805398
```

```
coef_gap
```

```
## [1] 0.02805398
```

Post treatment bias

```
# add the non-recommended variable to the model
cor_visti_treatment <- lm(data = biased_data, formula = treatment ~ visit + recency +
    history + channel) %>%
    tidy()
cor_visti_treatment
```

```
## # A tibble: 6 x 5
##   term          estimate std.error statistic   p.value
##   <chr>            <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept)   0.726    0.0112        65.0  0
## 2 visit         0.144    0.00761       18.9  2.30e- 79
## 3 recency      -0.0292   0.000795     -36.7  3.36e-289
## 4 history       0.000109 0.0000117      9.31 1.41e- 20
## 5 channelPhone -0.0751   0.00948       -7.92 2.51e- 15
## 6 channelWeb   -0.0738   0.00947       -7.80 6.38e- 15
```

```
bad_control_reg <- lm(data = biased_data, formula = spend ~ treatment + channel +
    recency + history + visit) %>%
    tidy()
bad_control_reg
```

```
## # A tibble: 7 x 5
##   term          estimate std.error statistic   p.value
##   <chr>            <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept)  -0.438     0.376       -1.16  2.44e-  1
## 2 treatment     0.294     0.177        1.66  9.68e-  2
## 3 channelPhone  0.121     0.300        0.403 6.87e-  1
## 4 channelWeb    0.117     0.299        0.392 6.95e-  1
## 5 recency       0.00988   0.0257       0.385 7.00e-  1
## 6 history       0.000525  0.000371     1.42  1.57e-  1
## 7 visit         7.16      0.242       29.6   3.85e-190
```

Regression EDA with vouchers data

```
remotes::install_github("itamarcaspi/experimentdatar")
library("experimentdatar")
data(vouchers)
vouchers
```

```
## # A tibble: 25,330 x 89
##       ID BOG95SMP BOG97SMP JAM93SMP   SEX   AGE  AGE2 HSVISIT SCYFNSH INSCHL
##    <dbl>    <dbl>    <dbl>    <dbl> <dbl> <dbl> <dbl>   <dbl>   <dbl>  <dbl>
##  1    NA        0        0        0    NA    NA    NA      NA       5     NA
##  2     1        0        0        0     1    NA    12      NA       5     NA
##  3     2        0        0        0     0    NA    13      NA       5     NA
##  4     3        1        0        0     0    14    12       0       8      1
##  5     4        1        0        0     1    14    12       0       8      1
##  6     5        1        0        0     0    14    12       0       8      1
##  7     6        1        0        0     0    12    10       0       7      1
##  8     7        0        0        0     1    NA    13      NA       5     NA
##  9     8        0        0        0     1    NA    12      NA       5     NA
## 10     9        0        0        0     1    NA    13      NA       5     NA
## # i 25,320 more rows
## # i 79 more variables: PRSCH_C <dbl>, PRSCHA_1 <dbl>, PRSCHA_2 <dbl>,
## #   VOUCH0 <dbl>, BOG95ASD <dbl>, BOG97ASD <dbl>, JAM93ASD <dbl>,
## #   DBOGOTA <dbl>, DJAMUNDI <dbl>, D1995 <dbl>, D1997 <dbl>, RESPONSE <dbl>,
## #   TEST_TAK <dbl>, SEX_NAME <dbl>, SVY <dbl>, D1993 <dbl>, PHONE <dbl>,
## #   DAREA1 <dbl>, DAREA2 <dbl>, DAREA3 <dbl>, DAREA4 <dbl>, DAREA5 <dbl>,
## #   DAREA6 <dbl>, DAREA7 <dbl>, DAREA8 <dbl>, DAREA9 <dbl>, DAREA10 <dbl>, ...
```

```
# prepare the regression prepare the character vectors for the regression
formula_x_base <- "VOUCH0"
formula_x_covariate <- "SVY + HSVISIT + AGE + STRATA1 + STRATA2 + STRATA3 + STRATA4 + STRATA5 + STRATA6
formula_y <- c("TOTSCYRS", "INSCHL", "PRSCH_C", "USNGSCH", "PRSCHA_1", "FINISH6",
    "FINISH7", "FINISH8", "REPT6", "REPT", "NREPT", "MARRIED", "HASCHILD", "HOURSUM",
    "WORKING3")

## create the simple regression formulas for each element in formula_y without
## covariates
base_reg_formula <- paste(formula_y, "~", formula_x_base)
names(base_reg_formula) <- paste(formula_y, "base", sep = "_")

## create the multiple regression formulas for each element in formula_y with
## covariates
covariate_reg_formula <- paste(formula_y, "~", formula_x_base, "+", formula_x_covariate)
names(covariate_reg_formula) <- paste(formula_y, "covariate", sep = "_")

## create the vectors for the models
table3_formula <- c(base_reg_formula, covariate_reg_formula)

## enframe the vectors
models <- table3_formula %>%
    enframe(name = "model_index", value = "formula")

# map the regression extract the data
regression_data <- vouchers %>%
    filter(TAB3SMPL == 1, BOG95SMP == 1)

df_models <- models %>%
    mutate(model = map(.x = formula, .f = lm, data = regression_data)) %>%
    mutate(lm_result = map(.x = model, .f = tidy))

# unnest the result
df_results <- df_models %>%
    mutate(formula = as.character(formula)) %>%
    dplyr::select(formula, model_index, lm_result) %>%
    unnest(cols = c(lm_result))
```

Analysis of private school attendance and use of vouchers

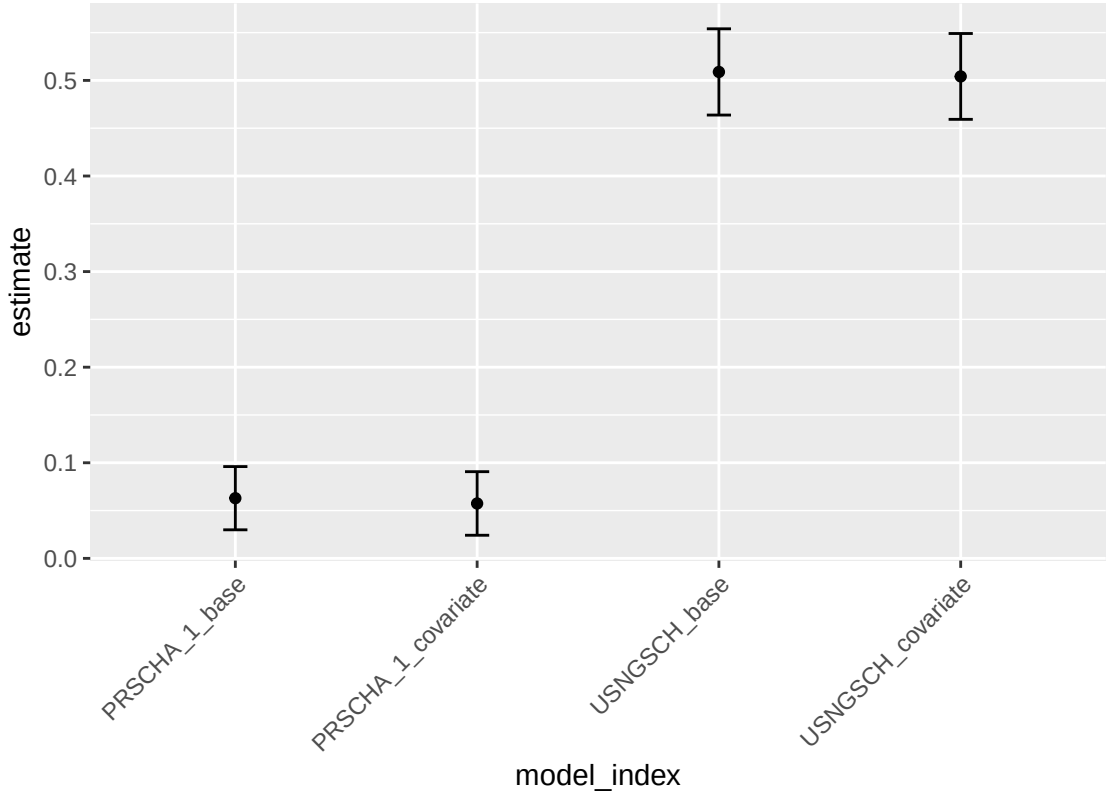
Was the voucher used and did it help private school enrollment?

```
# enrollment and voucher usage (PRSCHA_1 = if the student used the voucher,
# USNGSCH = if the student enrolled the private school)
using_voucher_results <- df_results %>%
    filter(term == "VOUCH0", str_detect(model_index, "PRSCHA_1|USNGSCH")) %>%
    dplyr::select(model_index, term, estimate, std.error, p.value) %>%
    arrange(model_index)
using_voucher_results
```

```
## # A tibble: 4 x 5
##   model_index        term   estimate std.error  p.value
##   <chr>              <chr>     <dbl>     <dbl>    <dbl>
## 1 PRSCHA_1_base      VOUCH0   0.0629    0.0169 2.00e- 4
```

```
## 2 PRSCHA_1_covariate VOUCHO   0.0574    0.0170 7.36e- 4
## 3 USNGSCH_base       VOUCHO   0.509     0.0230 1.80e-90
## 4 USNGSCH_covariate  VOUCHO   0.504     0.0229 1.49e-89
```

```
# plot the result
using_voucher_results %>%
    ggplot(aes(y = estimate, x = model_index)) + geom_point() + geom_errorbar(aes(ymax = estimate +
    std.error * 1.96, ymin = estimate - std.error * 1.96, width = 0.1)) + theme(axis.text.x = element_te
    hjust = 1), plot.title = element_text(hjust = 0.5), legend.position = "bottom",
    plot.margin = margin(0.5, 1, 0.5, 1, "cm"))
```

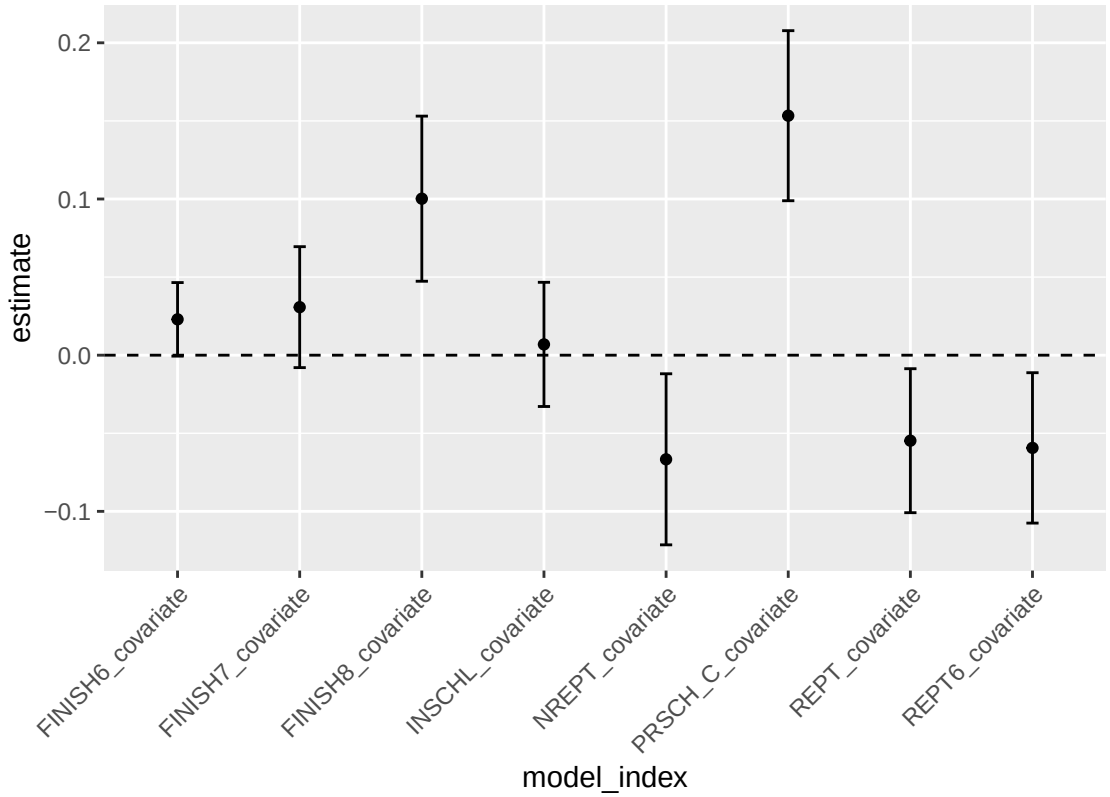


Did voucher help the students to finish high school?

ggplot

```
# extract the effect of VOUCHO on PRSCH_C, INSCHL, FINISH6-8, REPT PRSCH_C =
# show if the student still attend the private school after 3 yrs, INSCHL =
# show if the student still attend the school (public or pirvate) after 3 yrs,
# FINISH6-8 = show if the student was retained in the 6th grade.  REPT = show
# if the student was retained at least once before the survey
going_private_results <- df_results %>%
    filter(term == "VOUCHO", str_detect(model_index, "PRSCH_C|INSCHL|FINISH|REPT")) %>%
    dplyr::select(model_index, term, estimate, std.error, p.value) %>%
    arrange(model_index)

going_private_results %>%
    filter(str_detect(model_index, "covariate")) %>%
    ggplot(aes(y = estimate, x = model_index)) + geom_point() + geom_errorbar(aes(ymax = estimate +
    std.error * 1.96, ymin = estimate - std.error * 1.96, width = 0.1)) + geom_hline(yintercept = 0,
    linetype = 2) + theme(axis.text.x = element_text(angle = 45, hjust = 1), plot.title = element_text(
```

```
    legend.position = "bottom", plot.margin = margin(0.5, 1, 0.5, 1, "cm"))
```



Difference in Effectiveness by Gender

Replicate Angrist (2002) Table 4 & 6 bogota 1995

```
# create the data for table 4
data_tbl4_bog95 <- vouchers %>%
    filter(BOG95SMP == 1, TAB3SMPL == 1, !is.na(SCYFNSH), !is.na(FINISH6), !is.na(PRSCHA_1),
        !is.na(REPT6), !is.na(NREPT), !is.na(INSCHL), !is.na(FINISH7), !is.na(PRSCH_C),
        !is.na(FINISH8), !is.na(PRSCHA_2), !is.na(TOTSCYRS), !is.na(REPT)) %>%
    dplyr::select(VOUCH0, SVY, HSVISIT, DJAMUNDI, PHONE, AGE, STRATA1:STRATA6, STRATAMS,
        DBOGOTA, D1993, D1995, D1997, DMONTH1:DMONTH12, SEX_MISS, FINISH6, FINISH7,
        FINISH8, REPT6, REPT, NREPT, SEX2, TOTSCYRS, MARRIED, HASCHILD, HOURSUM,
        WORKING3, INSCHL, PRSCH_C, USNGSCH, PRSCHA_1)
```

Women data

```
# exctract women data
regression_data <- data_tbl4_bog95 %>%
    filter(SEX2 == 0)

# run the regression all together
df_models <- models %>%
    mutate(model = map(.x = formula, .f = lm, data = regression_data)) %>%
    mutate(lm_result = map(.x = model, .f = tidy))

# format the result
df_results_female <- df_models %>%
```

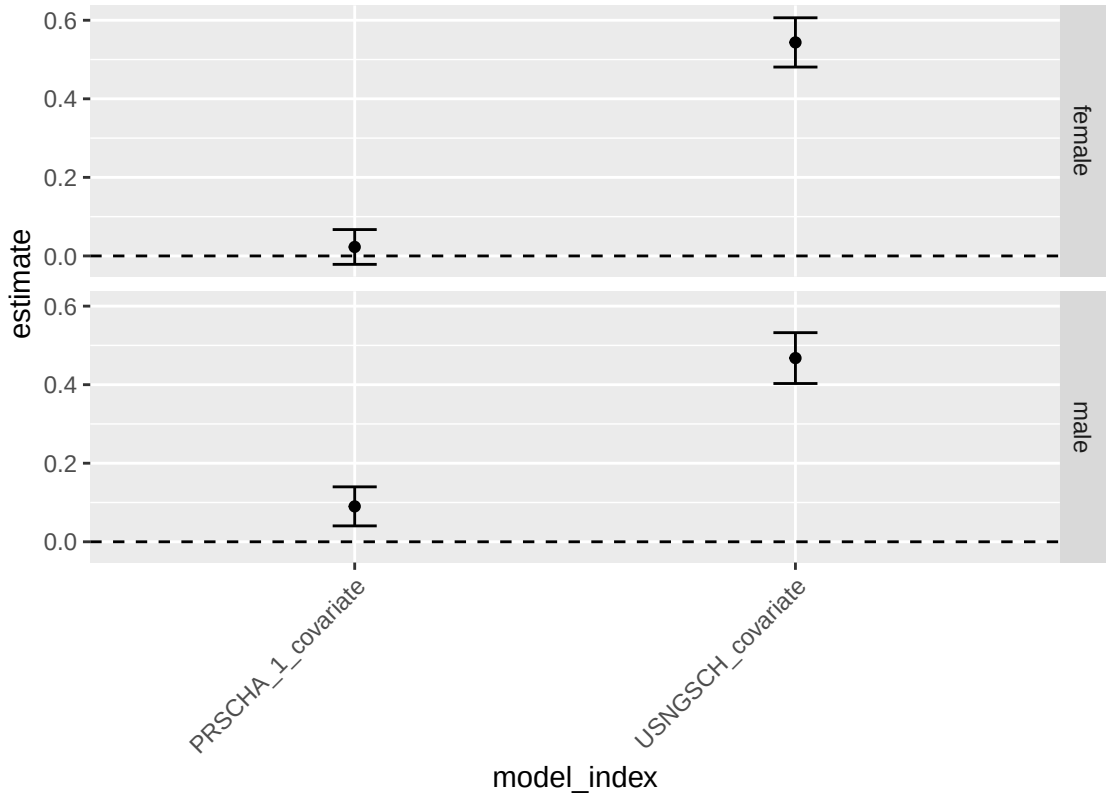
```
    mutate(formula = as.character(formula), gender = "female") %>%
    dplyr::select(formula, model_index, lm_result, gender) %>%
    unnest(cols = c(lm_result))
```

Men data

```
# extract the men data
regression_data <- data_tbl4_bog95 %>%
    filter(SEX2 == 1)

# run the regression all together
df_models <- models %>%
    mutate(model = map(.x = formula, .f = lm, data = regression_data)) %>%
    mutate(lm_result = map(.x = model, .f = tidy))

# format the result
df_results_male <- df_models %>%
    mutate(formula = as.character(formula), gender = "male") %>%
    dplyr::select(formula, model_index, lm_result, gender) %>%
    unnest(cols = c(lm_result))
```

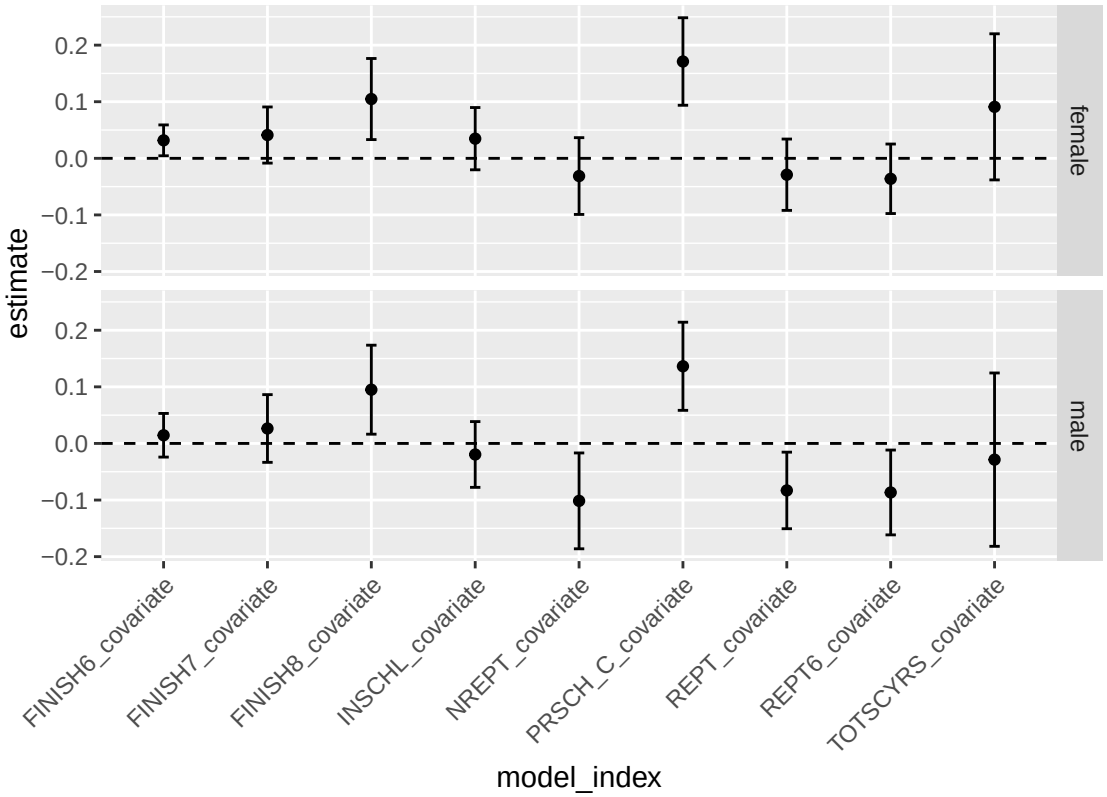
Analysis of private school attendance and use of vouchers by gender

```
# visualize analysis results to school attendance trends extract results for
# PRSCHA_1, USNGSCH
using_voucher_results_gender <- rbind(df_results_male, df_results_female) %>%
    filter(term == "VOUCHO", str_detect(model_index, "PRSCHA_1|USNGSCH")) %>%
    dplyr::select(gender, model_index, term, estimate, std.error, p.value) %>%
    # reorder the outputed dataframe
arrange(gender, model_index) %>%
    filter(str_detect(model_index, "covariate"))

## ggplot
using_voucher_results_gender %>%
    filter(str_detect(model_index, "covariate")) %>%
    ggplot(aes(y = estimate, x = model_index)) + geom_point() + geom_errorbar(aes(ymax = estimate +
    std.error * 1.96, ymin = estimate - std.error * 1.96, width = 0.1)) + geom_hline(yintercept = 0,
    linetype = 2) + theme(axis.text.x = element_text(angle = 45, hjust = 1), plot.title = element_text(
    legend.position = "bottom", plot.margin = margin(0.5, 1, 0.5, 1, "cm")) + facet_grid(gender ~
    .)
```

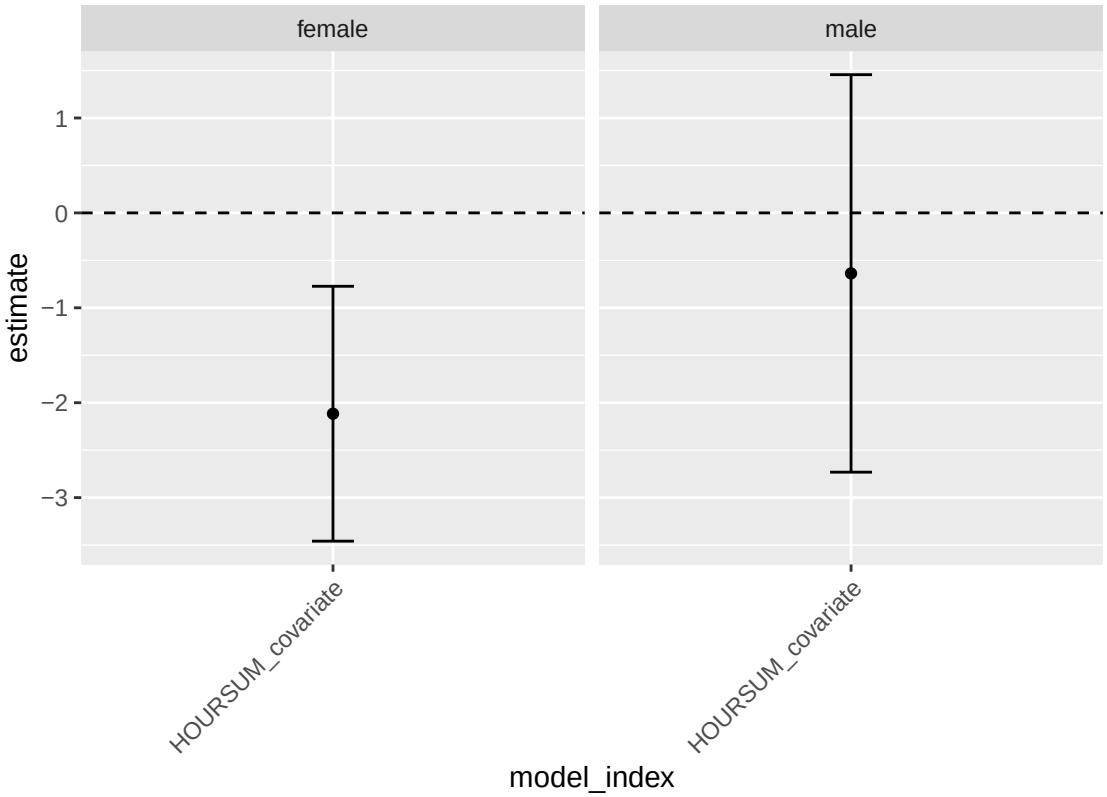


How Did voucher affect the students by gender to finish high school?

```
# visualize analysis results to retention and years of school attendance
# extract the results for PRSCH_C,INSCHL,REPT,TOTSCYRS,FINISH
going_private_results_gender <- rbind(df_results_male, df_results_female) %>%
    filter(term == "VOUCHO", str_detect(model_index, "PRSCH_C|INSCHL|REPT|TOTSCYRS|FINISH")) %>%
    dplyr::select(gender, model_index, term, estimate, std.error, p.value) %>%
    arrange(model_index)

## ggplot
going_private_results_gender %>%
    filter(str_detect(model_index, "covariate")) %>%
    ggplot(aes(y = estimate, x = model_index)) + geom_point() + geom_errorbar(aes(ymax = estimate +
    std.error * 1.96, ymin = estimate - std.error * 1.96, width = 0.1)) + geom_hline(yintercept = 0,
    linetype = 2) + theme(axis.text.x = element_text(angle = 45, hjust = 1), plot.title = element_text(
    legend.position = "bottom", plot.margin = margin(0.5, 1, 0.5, 1, "cm")) + facet_grid(gender ~
    .)
```




```
# explore other factors because the above analysis showed that girls' academic
# persistence is less correlated with academic achievement, retention, and
# winning vouchers visualize the results of the analysis against HOUR extract
# the results of the analysis against HOUR
working_hour_results_gender <- rbind(df_results_male, df_results_female) %>%
    filter(term == "VOUCHO", str_detect(model_index, "HOUR")) %>%
    dplyr::select(gender, model_index, term, estimate, std.error, p.value) %>%
    arrange(gender, model_index)

## ggplot
working_hour_results_gender %>%
    filter(str_detect(model_index, "covariate")) %>%
    ggplot(aes(y = estimate, x = model_index)) + geom_point() + geom_errorbar(aes(ymax = estimate +
    std.error * 1.96, ymin = estimate - std.error * 1.96, width = 0.1)) + geom_hline(yintercept = 0,
    linetype = 2) + theme(axis.text.x = element_text(angle = 45, hjust = 1), plot.title = element_text(
    legend.position = "bottom", plot.margin = margin(0.5, 1, 0.5, 1, "cm")) + facet_grid(. ~
    gender)
```


female
male
1
0
−1
−2
−3
estimate
HOURSUM_covariate
HOURSUM_covariate
model_index