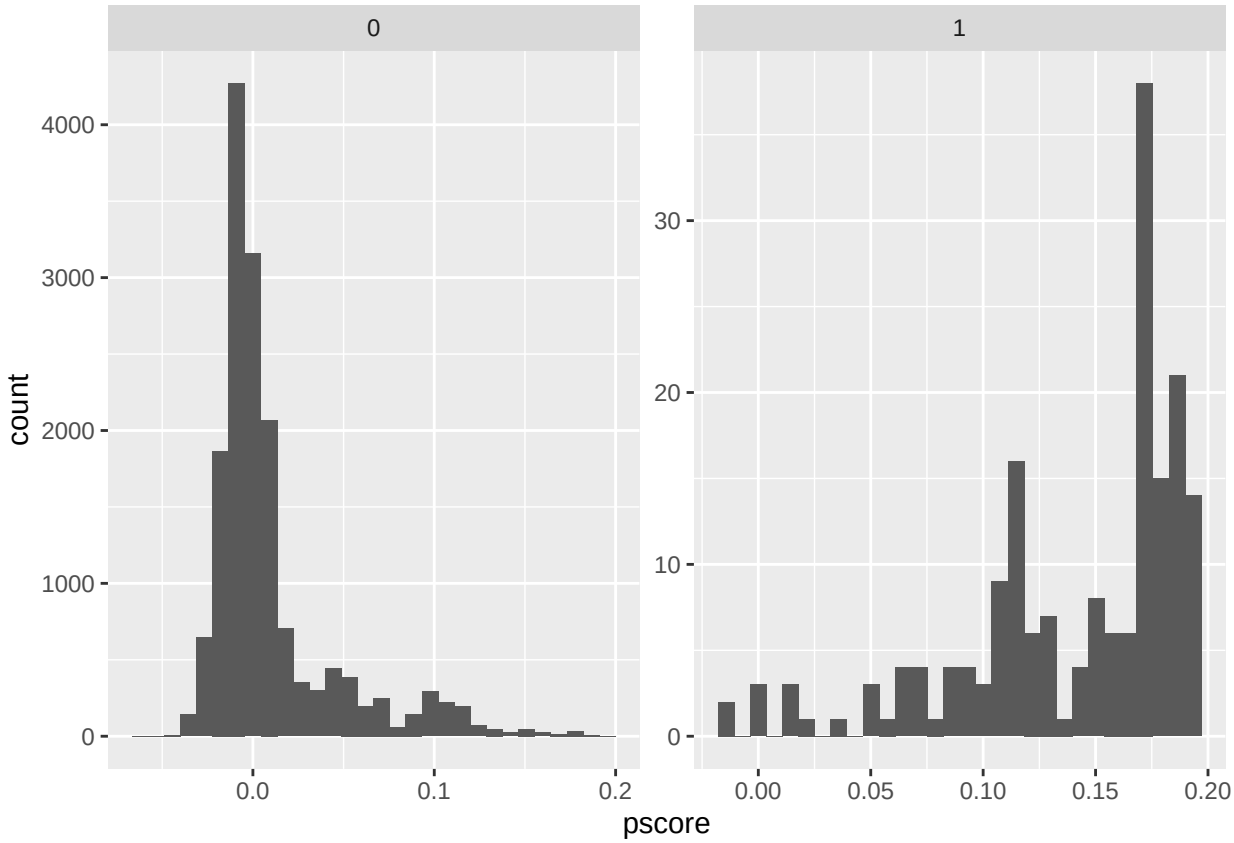
RDD2

Calculate a propensity score using OLS using up to a quadratic

```
read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw <- read_data("nsw_mixtape.dta")

mean1 <- nsw_dw %>%
  filter(treat == 1) %>%
  pull(re78) %>%
  mean()

nsw_dw$y1 <- mean1


mean0 <- nsw_dw %>%
  filter(treat == 0) %>%
  pull(re78) %>%
  mean()

nsw_dw$y0 <- mean0

ate <- unique(nsw_dw$y1 - nsw_dw$y0)

nsw_dw <- nsw_dw %>%
  filter(treat == 1) %>%
  select(-y1, -y0)


read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw_cpscontrol <- read_data("cps_mixtape.dta") %>%
  bind_rows(nsw_dw) %>%
```

```
  mutate(agesq = age^2,
         agecube = age^3,
         educsq = educ*educ,
         educcub=educsq*educ,
         u74 = case_when(re74 == 0 ~ 1, TRUE ~ 0),
         u75 = case_when(re75 == 0 ~ 1, TRUE ~ 0),
         interaction1 = educ*re74,
         re74sq = re74^2,
         re75sq = re75^2,
         interaction2 = u74*hisp)

# estimating
ols_nsw <- lm(treat ~ age + agesq + educ + educsq +
                   marr + nodegree + black + hisp + re74 + re75 + u74 +
                   u75,
                 data = nsw_dw_cpscontrol)

nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(pscore = ols_nsw$fitted.values)

# mean pscore
pscore_control <- nsw_dw_cpscontrol %>%
  filter(treat == 0) %>%
  pull(pscore) %>%
  mean()

pscore_treated <- nsw_dw_cpscontrol %>%
  filter(treat == 1) %>%
  pull(pscore) %>%
  mean()

ggplot(data=nsw_dw_cpscontrol, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Max and Min values of the scores for treated and control

```
nsw_dw_cpscontroltreated <- nsw_dw_cpscontrol %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreated$pscore)
```

[1] 0.1938788

```
min(nsw_dw_cpscontroltreated$pscore)
```

[1] -0.01385616

```
nsw_dw_cpscontroluntreated<- nsw_dw_cpscontrol %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreated$pscore)
```

[1] 0.1938197

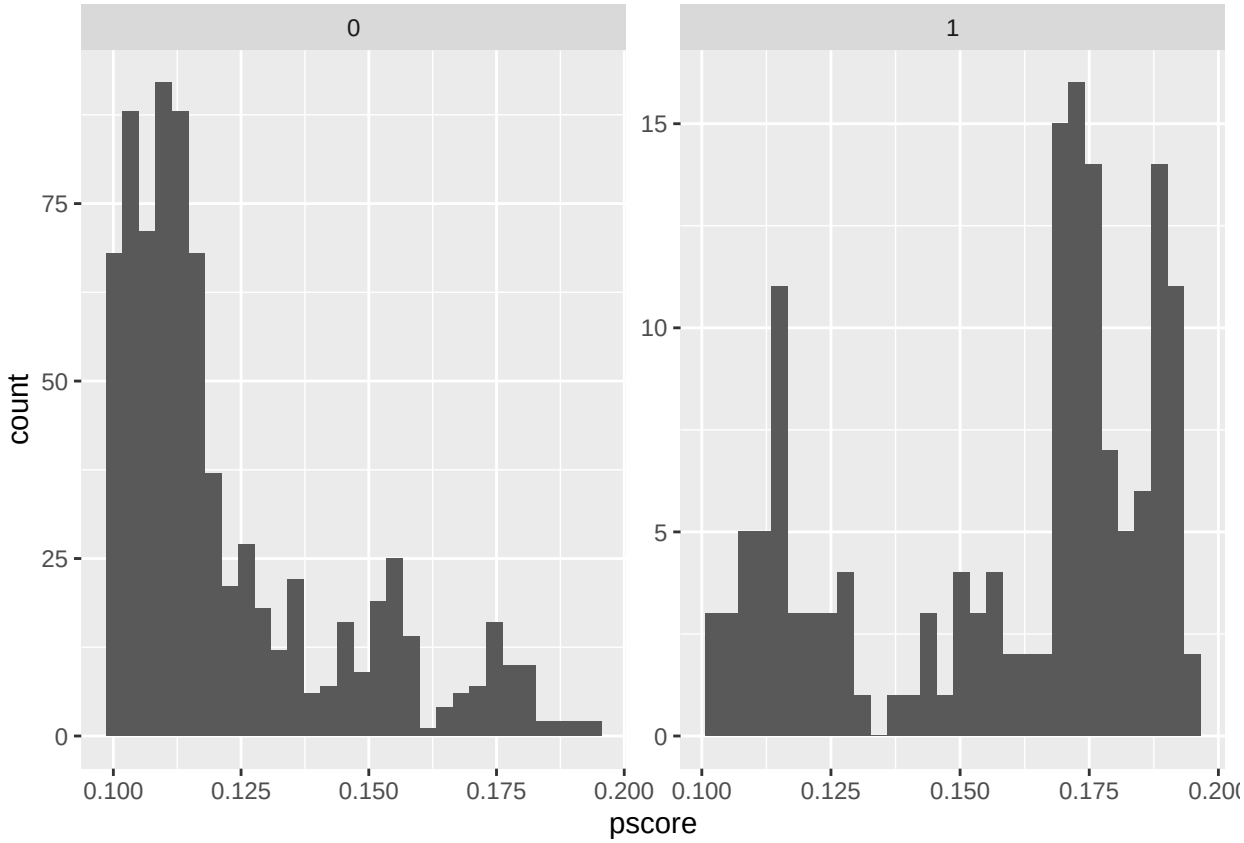
```
min(nsw_dw_cpscontroluntreated$pscore)
```

[1] -0.06371947

Droping propensity score

```
nsw_dw_cpscontrolcut <- nsw_dw_cpscontrol %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

ggplot(data=nsw_dw_cpscontrolcut, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
nsw_dw_cpscontroltreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.1938788
```

```
min(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.1011927
```

```r
nsw_dw_cpscontroluntreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.1938197
```

```r
min(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.1001003
```

```r
N <- nrow(nsw_dw_cpscontrol)
#- Manual with non-normalized weights using all data
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontrol$d1)
s0 <- sum(nsw_dw_cpscontrol$d0)


nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.15 <- nsw_dw_cpscontrol %>%
  pull(ht) %>%
  mean()

linear.16 <- nsw_dw_cpscontrol %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.15,linear.16)
stargazer(pleasework, header = FALSE)
```

Table 1:

-15,066.650	-24,136.160

```r
#-- trimming propensity score
nsw_dw_cpscontroln <- nsw_dw_cpscontrol %>%
  select(-d1, -d0, -y1, -y0, -ht, -norm) %>%
```

```
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

N <- nrow(nsw_dw_cpscontroln)

#- Manual with non-normalized weights using trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontroln$d1)
s0 <- sum(nsw_dw_cpscontroln$d0)

nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights with trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.13 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.14 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.13,linear.14)
stargazer(pleasework, header = FALSE)
```

Table 2:

-3,455.285	-4,577.043

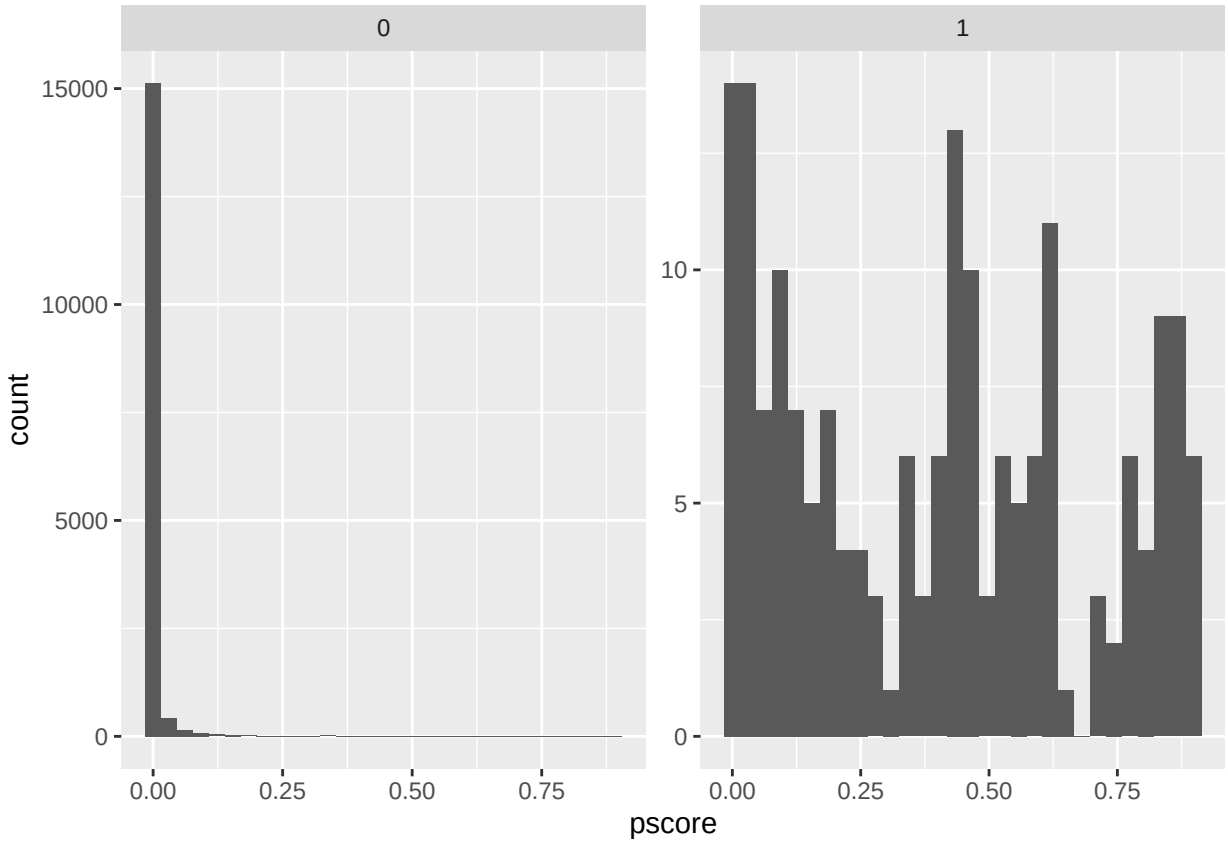
Calculate a propensity score using logit using up to a quadratic

```
read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}
```

```r
nsw_dw <- read_data("nsw_mixtape.dta")

mean1 <- nsw_dw %>%
  filter(treat == 1) %>%
  pull(re78) %>%
  mean()

nsw_dw$y1 <- mean1


mean0 <- nsw_dw %>%
  filter(treat == 0) %>%
  pull(re78) %>%
  mean()

nsw_dw$y0 <- mean0

ate <- unique(nsw_dw$y1 - nsw_dw$y0)

nsw_dw <- nsw_dw %>%
  filter(treat == 1) %>%
  select(-y1, -y0)


read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw_cpscontrol <- read_data("cps_mixtape.dta") %>%
  bind_rows(nsw_dw) %>%
  mutate(agesq = age^2,
         agecube = age^3,
         educsq = educ*educ,
         educcub=educsq*educ,
         u74 = case_when(re74 == 0 ~ 1, TRUE ~ 0),
         u75 = case_when(re75 == 0 ~ 1, TRUE ~ 0),
         interaction1 = educ*re74,
         re74sq = re74^2,
         re75sq = re75^2,
         interaction2 = u74*hisp)

# estimating
logit_nsw <- glm(treat ~ age + agesq+ educ + educsq +
                   marr + nodegree + black + hisp + re74 + re75 + u74 +
                   u75, family = binomial(link = "logit"),
                 data = nsw_dw_cpscontrol)

nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(pscore = logit_nsw$fitted.values)
```

```
# mean pscore
pscore_control <- nsw_dw_cpscontrol %>%
  filter(treat == 0) %>%
  pull(pscore) %>%
  mean()

pscore_treated <- nsw_dw_cpscontrol %>%
  filter(treat == 1) %>%
  pull(pscore) %>%
  mean()

ggplot(data=nsw_dw_cpscontrol, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Max and Min values of the scores for treated and control

```
nsw_dw_cpscontroltreated <- nsw_dw_cpscontrol %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreated$pscore)
```

[1] 0.8991878

```
min(nsw_dw_cpscontroltreated$pscore)
```

[1] 0.0007742979

```
nsw_dw_cpscontroluntreated<- nsw_dw_cpscontrol %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreated$pscore)
```

[1] 0.8900935

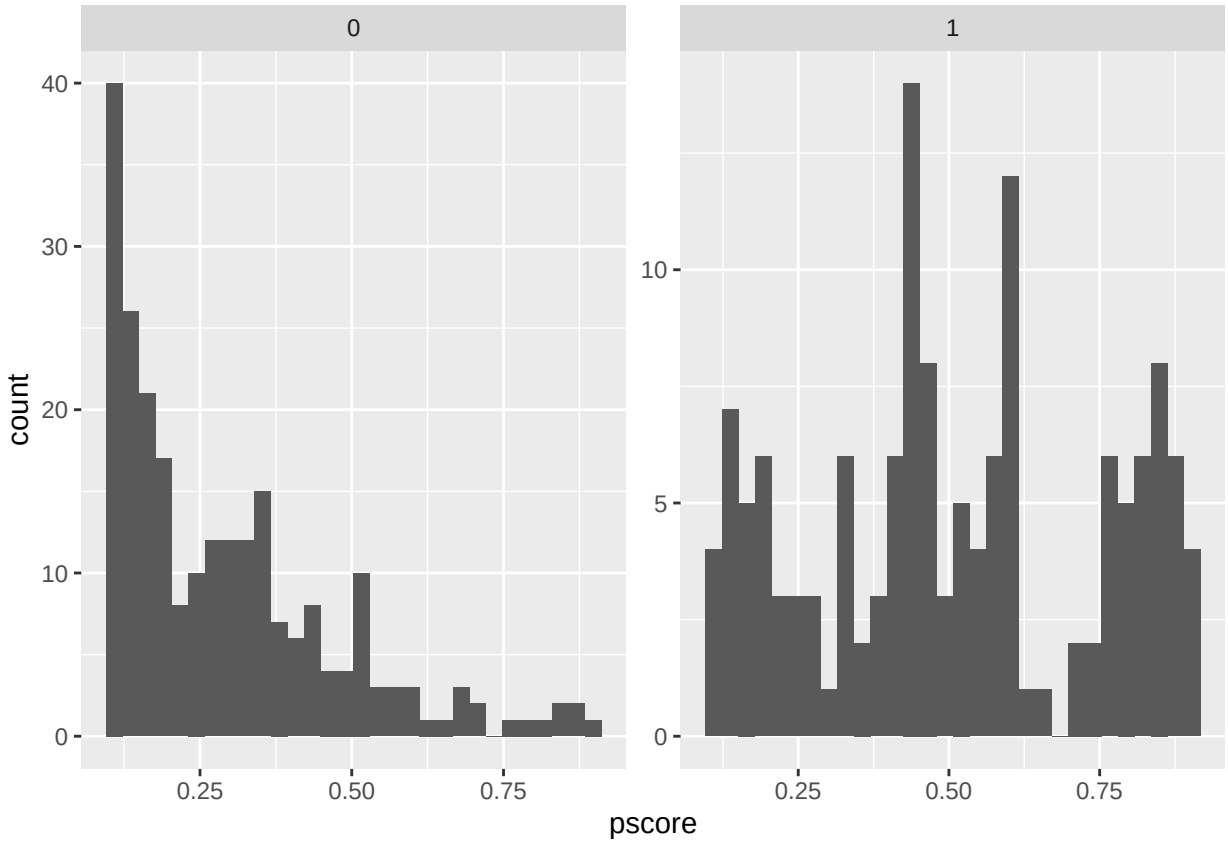
```
min(nsw_dw_cpscontroluntreated$pscore)
```

[1] 3.301435e-10

Droping propensity score

```
nsw_dw_cpscontrolcut <- nsw_dw_cpscontrol %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

ggplot(data=nsw_dw_cpscontrolcut, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
nsw_dw_cpscontroltreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.8991878
```

```
min(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.1049194
```

```
nsw_dw_cpscontroluntreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.8900935
```

```
min(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.1013622
```

```r
N <- nrow(nsw_dw_cpscontrol)
#- Manual with non-normalized weights using all data
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontrol$d1)
s0 <- sum(nsw_dw_cpscontrol$d0)


nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.11 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.12 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.11,linear.12)
stargazer(pleasework, header = FALSE)
```

Table 3:

-3,455.285	-4,577.043

```r
#-- trimming propensity score
nsw_dw_cpscontroln <- nsw_dw_cpscontrol %>%
  select(-d1, -d0, -y1, -y0, -ht, -norm) %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

N <- nrow(nsw_dw_cpscontroln)

#- Manual with non-normalized weights using trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontroln$d1)
s0 <- sum(nsw_dw_cpscontroln$d0)
```

```
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights with trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.9 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.10 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.9,linear.10)
stargazer(pleasework, header = FALSE)
```

Table 4:

1,340.551	1,280.588

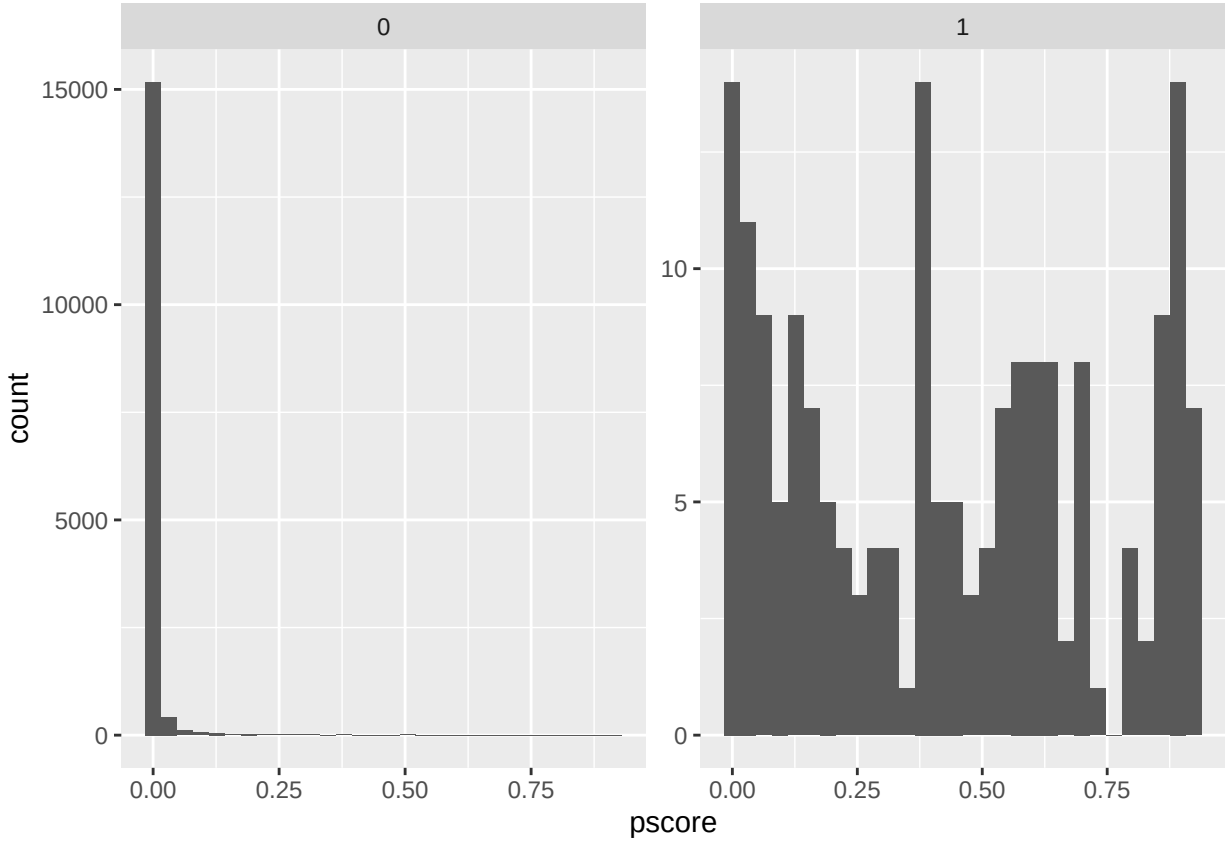
Logit Up To Cube

```
read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw <- read_data("nsw_mixtape.dta")


mean1 <- nsw_dw %>%
  filter(treat == 1) %>%
  pull(re78) %>%
  mean()

nsw_dw$y1 <- mean1

mean0 <- nsw_dw %>%
  filter(treat == 0) %>%
  pull(re78) %>%
```

```
  mean()

nsw_dw$y0 <- mean0

ate <- unique(nsw_dw$y1 - nsw_dw$y0)

nsw_dw <- nsw_dw %>%
  filter(treat == 1) %>%
  select(-y1, -y0)


read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw_cpscontrol <- read_data("cps_mixtape.dta") %>%
  bind_rows(nsw_dw) %>%
  mutate(agesq = age^2,
         agecube = age^3,
         educsq = educ*educ,
         educcub=educsq*educ,
         u74 = case_when(re74 == 0 ~ 1, TRUE ~ 0),
         u75 = case_when(re75 == 0 ~ 1, TRUE ~ 0),
         interaction1 = educ*re74,
         re74sq = re74^2,
         re75sq = re75^2,
         re74cu = re74^3,
         re75cu = re75^3,
         interaction2 = u74*hisp)

# estimating
logit_nsw <- glm(treat ~ age + agesq + agecube + educ + educsq +
                   marr + nodegree + black + hisp + re74 +re75 +re74sq +re74cu +re75cu + re75sq + u74 +
                   u75, family = binomial(link = "logit"),
                 data = nsw_dw_cpscontrol)

nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(pscore = logit_nsw$fitted.values)

# mean pscore
pscore_control <- nsw_dw_cpscontrol %>%
  filter(treat == 0) %>%
  pull(pscore) %>%
  mean()

pscore_treated <- nsw_dw_cpscontrol %>%
  filter(treat == 1) %>%
  pull(pscore) %>%
  mean()
```

```
ggplot(data=nsw_dw_cpscontrol, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
### Max and Min values of the scores for treated and control
nsw_dw_cpscontroltreated <- nsw_dw_cpscontrol %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreated$pscore)
```

```
## [1] 0.9239837
```

```
min(nsw_dw_cpscontroltreated$pscore)
```

```
## [1] 0.001153426
```

```
nsw_dw_cpscontroluntreated<- nsw_dw_cpscontrol %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreated$pscore)
```

```
## [1] 0.9150462
```

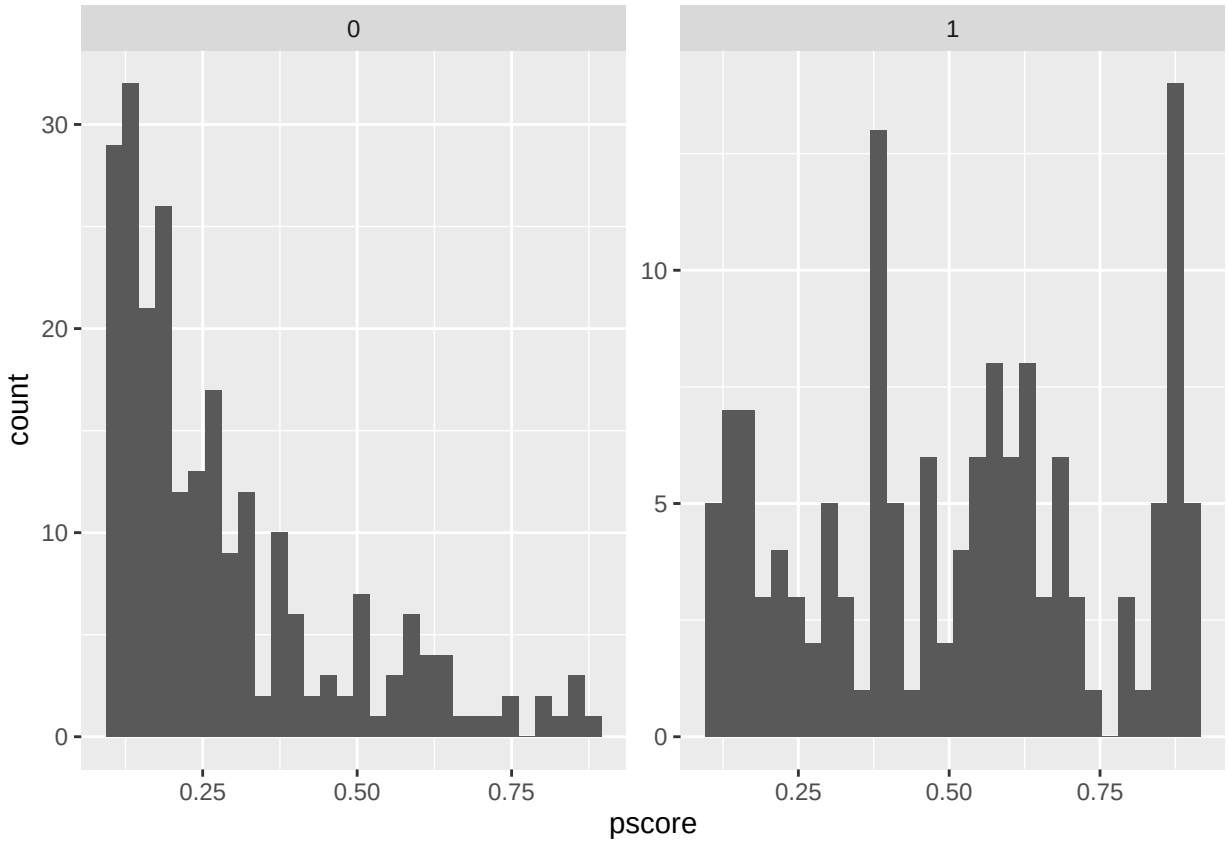
```
min(nsw_dw_cpscontroluntreated$pscore)
```

```
## [1] 7.87669e-09
```

Droping propensity score

```
nsw_dw_cpscontrolcut <- nsw_dw_cpscontrol %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

ggplot(data=nsw_dw_cpscontrolcut, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
nsw_dw_cpscontroltreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.8967915
```

```
min(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.1031797
```

```
nsw_dw_cpscontroluntreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.8768247
```

```
min(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.1018253
```

```
N <- nrow(nsw_dw_cpscontrol)
#- Manual with non-normalized weights using all data
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontrol$d1)
s0 <- sum(nsw_dw_cpscontrol$d0)


nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.7 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.8 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.7,linear.8)
stargazer(pleasework, header = FALSE)
```

Table 5:

1,340.551	1,280.588

```
#-- trimming propensity score
nsw_dw_cpscontroln <- nsw_dw_cpscontrol %>%
  select(-d1, -d0, -y1, -y0, -ht, -norm) %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

N <- nrow(nsw_dw_cpscontroln)

#- Manual with non-normalized weights using trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontroln$d1)
s0 <- sum(nsw_dw_cpscontroln$d0)

nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights with trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.5 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.6 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.5,linear.6)
stargazer(pleasework, header = FALSE)
```

Table 6:

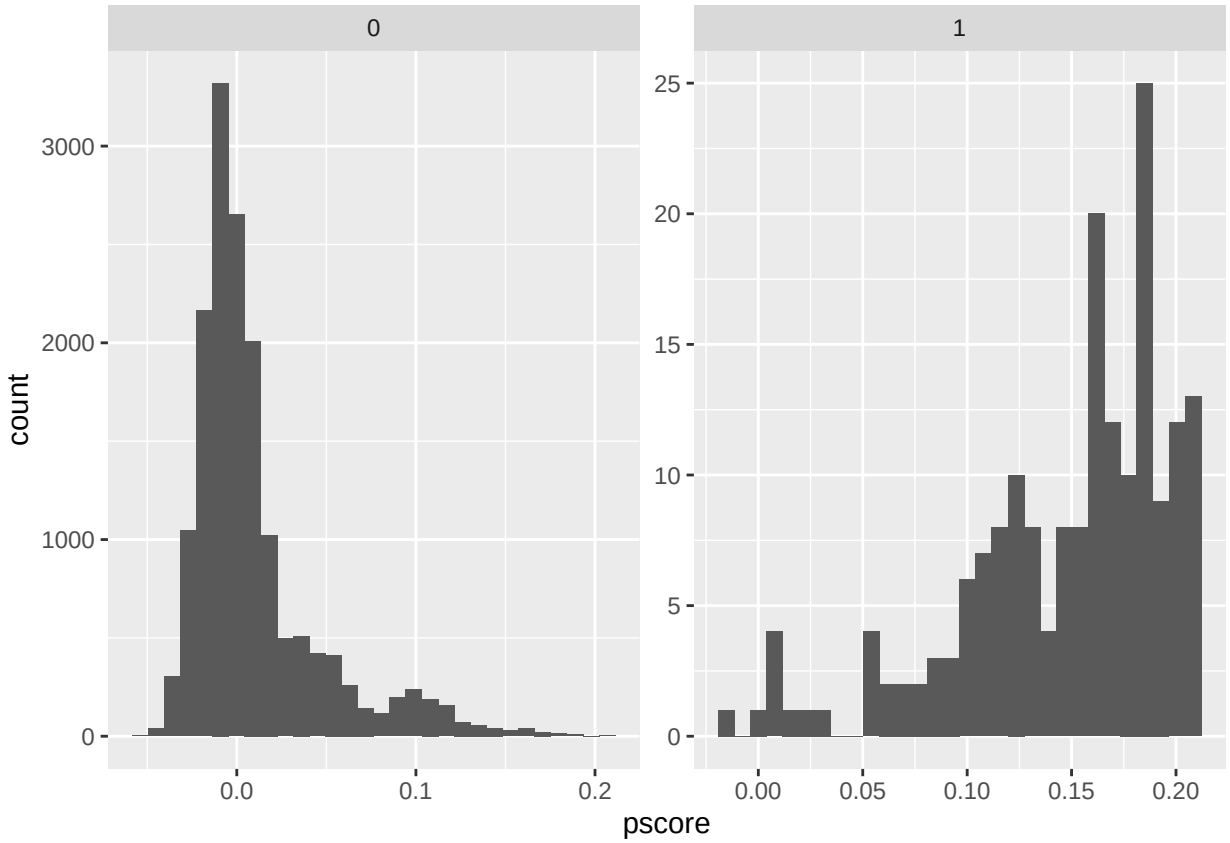
1,887.267	1,601.571

OLS up to a cubic

```
read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw <- read_data("nsw_mixtape.dta")

mean1 <- nsw_dw %>%
  filter(treat == 1) %>%
  pull(re78) %>%
  mean()

nsw_dw$y1 <- mean1


mean0 <- nsw_dw %>%
  filter(treat == 0) %>%
  pull(re78) %>%
  mean()

nsw_dw$y0 <- mean0

ate <- unique(nsw_dw$y1 - nsw_dw$y0)

nsw_dw <- nsw_dw %>%
  filter(treat == 1) %>%
  select(-y1, -y0)


read_data <- function(df)
{
  full_path <- paste("https://raw.github.com/scunning1975/mixtape/master/",
                     df, sep = "")
  df <- read_dta(full_path)
  return(df)
}

nsw_dw_cpscontrol <- read_data("cps_mixtape.dta") %>%
  bind_rows(nsw_dw) %>%
  mutate(agesq = age^2,
         agecube = age^3,
         educsq = educ*educ,
         educcub=educsq*educ,
         u74 = case_when(re74 == 0 ~ 1, TRUE ~ 0),
         u75 = case_when(re75 == 0 ~ 1, TRUE ~ 0),
         interaction1 = educ*re74,
         re74sq = re74^2,
         re75sq = re75^2,
```

```
         re74cu = re74^3,
         re75cu = re75^3,
         interaction2 = u74*hisp)

# estimating
ols_nsw <- lm(treat ~ age + agesq+agecube +educcub+ educ + educsq +
                   marr + nodegree + black + hisp + re74 +re75 +re74sq +re74cu +re75cu + re75sq + u74 +
                   u75,
                 data = nsw_dw_cpscontrol)


nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(pscore = ols_nsw$fitted.values)
# mean pscore
pscore_control <- nsw_dw_cpscontrol %>%
  filter(treat == 0) %>%
  pull(pscore) %>%
  mean()

pscore_treated <- nsw_dw_cpscontrol %>%
  filter(treat == 1) %>%
  pull(pscore) %>%
  mean()

ggplot(data=nsw_dw_cpscontrol, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Max and Min values of the scores for treated and control

```
nsw_dw_cpscontroltreated <- nsw_dw_cpscontrol %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreated$pscore)
```

[1] 0.20838

```
min(nsw_dw_cpscontroltreated$pscore)
```

[1] -0.01526928

```
nsw_dw_cpscontroluntreated<- nsw_dw_cpscontrol %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreated$pscore)
```

[1] 0.2072448

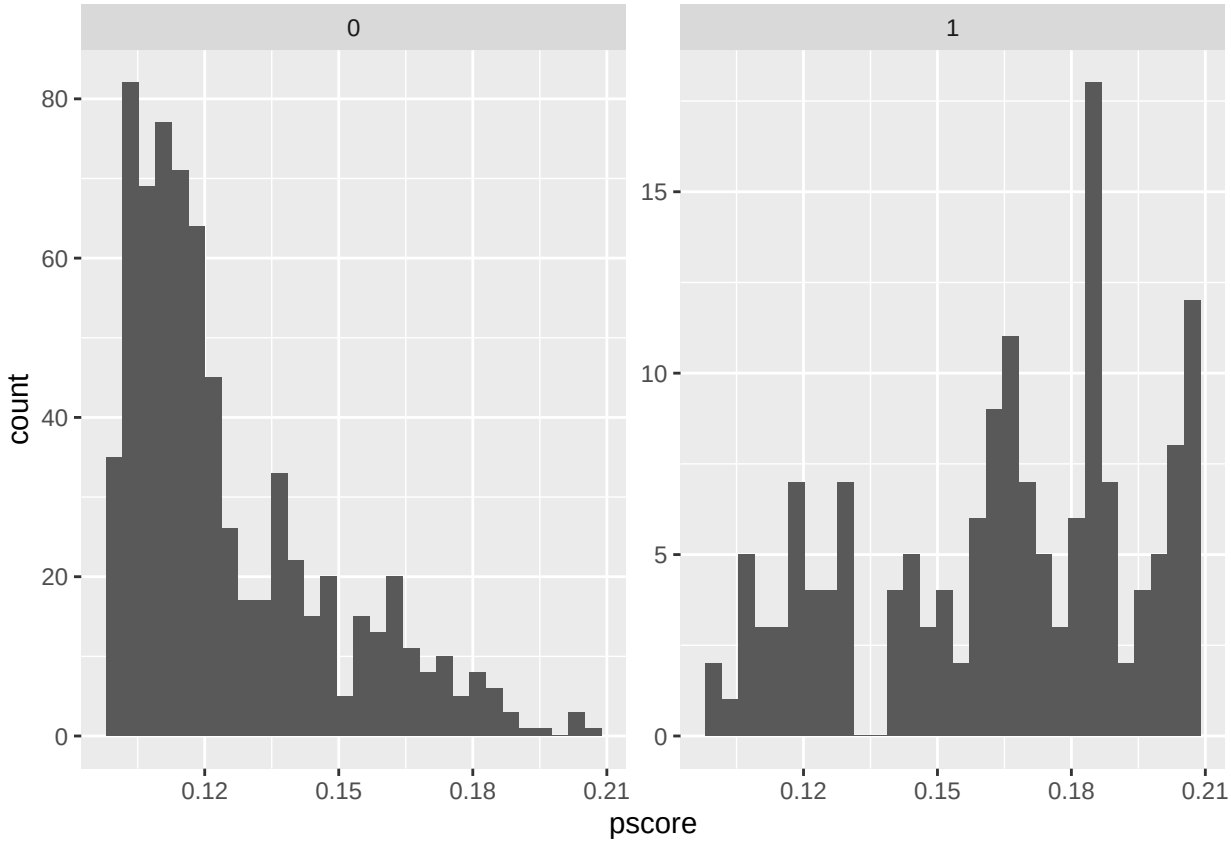
```
min(nsw_dw_cpscontroluntreated$pscore)
```

[1] -0.05352829

Droping propensity score

```
nsw_dw_cpscontrolcut <- nsw_dw_cpscontrol %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

ggplot(data=nsw_dw_cpscontrolcut, aes(x = pscore)) +
  geom_histogram() +
  facet_wrap(~treat, scales = "free")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
nsw_dw_cpscontroltreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 1)

max(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.20838
```

```
min(nsw_dw_cpscontroltreatedcut$pscore)
```

```
## [1] 0.1011328
```

```
nsw_dw_cpscontroluntreatedcut <- nsw_dw_cpscontrolcut %>%
  filter(treat == 0)

max(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.2072448
```

```
min(nsw_dw_cpscontroluntreatedcut$pscore)
```

```
## [1] 0.1000906
```

```
N <- nrow(nsw_dw_cpscontrol)
#- Manual with non-normalized weights using all data
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontrol$d1)
s0 <- sum(nsw_dw_cpscontrol$d0)


nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = treat * re78,
         y0 = (1-treat) * re78,
         ht = y1 - y0)


#- Manual with normalized weights
nsw_dw_cpscontrol <- nsw_dw_cpscontrol %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.1 <- nsw_dw_cpscontroln %>%
  pull(ht) %>%
  mean()

linear.2 <- nsw_dw_cpscontroln %>%
  pull(norm) %>%
  mean()

pleasework <- c(linear.1,linear.2)
stargazer(pleasework, header = FALSE)
```

```
##
## \begin{table}[!htbp] \centering
```

```
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}} cc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
## $1,887.267$ & $1,601.571$ \\
## \hline \\[-1.8ex]
## \end{tabular}
## \end{table}
```

```
#-- trimming propensity score
nsw_dw_cpscontroln <- nsw_dw_cpscontrol %>%
  select(-d1, -d0, -y1, -y0, -ht, -norm) %>%
  filter(!(pscore >= 0.9)) %>%
  filter(!(pscore <= 0.1))

N <- nrow(nsw_dw_cpscontroln)

#- Manual with non-normalized weights using trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(d1 = treat/pscore,
         d0 = (1-treat)/(1-pscore))

s1 <- sum(nsw_dw_cpscontroln$d1)
s0 <- sum(nsw_dw_cpscontroln$d0)

nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = treat * re78/pscore,
         y0 = (1-treat) * re78/(1-pscore),
         ht = y1 - y0)

#- Manual with normalized weights with trimmed data
nsw_dw_cpscontroln <- nsw_dw_cpscontroln %>%
  mutate(y1 = (treat*re78/pscore)/(s1/N),
         y0 = ((1-treat)*re78/(1-pscore))/(s0/N),
         norm = y1 - y0)

linear.3 <- mean(nsw_dw_cpscontroln$ht)
linear.4 <- mean(nsw_dw_cpscontroln$norm)

pleasework <- c(linear.3,linear.4)
stargazer(pleasework, header = FALSE,title="Results From Cubic OLS")
```

Table 7: Results From Cubic OLS

-2,254.879	-3,918.455

Comparing results

So when I compare my results to what was found in the mixtape I can see a clear pattern. The first pattern is that I consistently get results that are nowhere near to what I should find the mixtape when I use the

OLS. This makes sense because we are doing a linear probability model which means that we can predict probabilities that are less than zero and more than 1, as we scan see from the minimum propensity scores when we do OLS. On the other hand, we get propensity scores that are in between 0 and 1 when we run a logit regression. Even though we are trimming the data, we still have the issue that our data set has been severely reduced, thus impacting our results.

When we run the logit regressions and we trim the data, we can see that our results are a lot closer to what we are supposed to get in the mixtape but they are not quite there. The reason for this is simple, we are not running the exact same regression as the on in the mixtape as we are not include the interaction variables. If we were to run the logit regression with the exact same covariates as in the mixtape, then we would get exactly the same response as he does.