Denis Ostroushko - HW4

Introduction

Imputation and Variable Processing

For the purpose of this assignment we retain the same imputation schemes we used in the previous two assignments. We will use imputation with the median of observed values and replace missing values with modes for categorical predictors.

Since we will need to create two models that have all possible confounders in the data set, we need to be careful with variable inclusion. We drop variable `Hisp` because it is highly correlated with other variables that contain race and ethnic information. We also `Drug.Add` due to the issues with its imputation. When imputed with the most common level "No", which indicated no drug use, this variable has one unique level. Such zero variance predictors can cause problems with fitting models, so we will avoid using it in our analyses.

Variables `BMI`, `BL.Cig.Day`, `BL.Drks.Day`, `N.living.kids` are imputed with medians like `N.prev.preg`, `Birthweight` in the previous assignments.

`Use.Alc` is imputed with a mode like `Race_ethnicity`, `Use.Tob`

Problem 1

1 - A

```
pregnancy_model <-
  glm(
    `Preg.ended...37.wk` ~
      Group + Race_ethnicity + Public.Asstce + Use.Tob + N.prev.preg +
      Live.PTB + BL.GE + BL..BOP + BL..PD.4 + BL..CAL.3,

    data = data,
```

```
    family = "binomial"
  )

all_no_treat <- data %>%
  filter(Group == "T") %>%
  select(-Group) %>%
  mutate(Group = "C")

all_no_treat$porential_no_trt <- predict(pregnancy_model, all_no_treat , type = "response"

all_treat <- data %>%
  filter(Group == "T") %>%
  select(-Group) %>%
  mutate(Group = "T")

all_treat$porential_trt <- predict(pregnancy_model, all_treat , type = "response")

preg_att <- mean(all_treat$porential_trt, na.rm = T) - mean(all_no_treat$porential_no_trt,

print("Regression ATT Pregnancy")
```

```
[1] "Regression ATT Pregnancy"
```

```
preg_att
```

```
[1] -0.01817514
```

```
[1] "Regression ATT Birthweight"
```

```
[1] 60.07997
```

1 - B

```
[1] "PSR ATT Pregnancy"
```

```
[1] -0.01344089
```

```
[1] "PSR ATT Borthweight"
```

```
[1] 50.03487
```

1 - C

```
[1] "PSS ATT Pregnancy"
```

```
[1] -0.0204806
```

```
[1] "PSS ATT Borthweight"
```

```
[1] 70.41007
```

1 - D

```
[1] "IPW ATT Pregnancy"
```

```
[1] -0.01528543
```

```
[1] "IPW ATT Borthweight"
```

```
[1] 52.0474
```

1 - E

```
[1] "1:1 ATT Pregnancy"
```

```
[1] -0.02173913
```

```
[1] "1:1 ATT Borthweight"
```

```
[1] 84.68478
```

```
[1] "2:1 ATT Pregnancy"
```

```
[1] -0.01630435
```

```
[1] "2:1 ATT Pregnancy"
```

```
[1] 65.4837
```

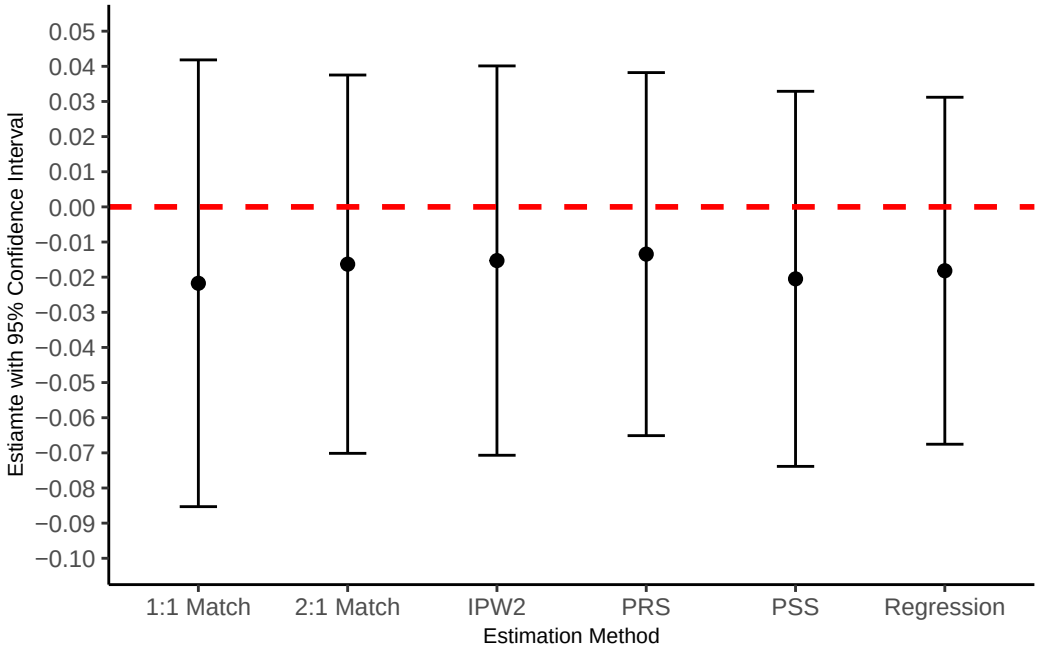

Figure 1: Comparison of Estiamtion Method Varinances for ATT of Pregnancy Outcome

Problem 2

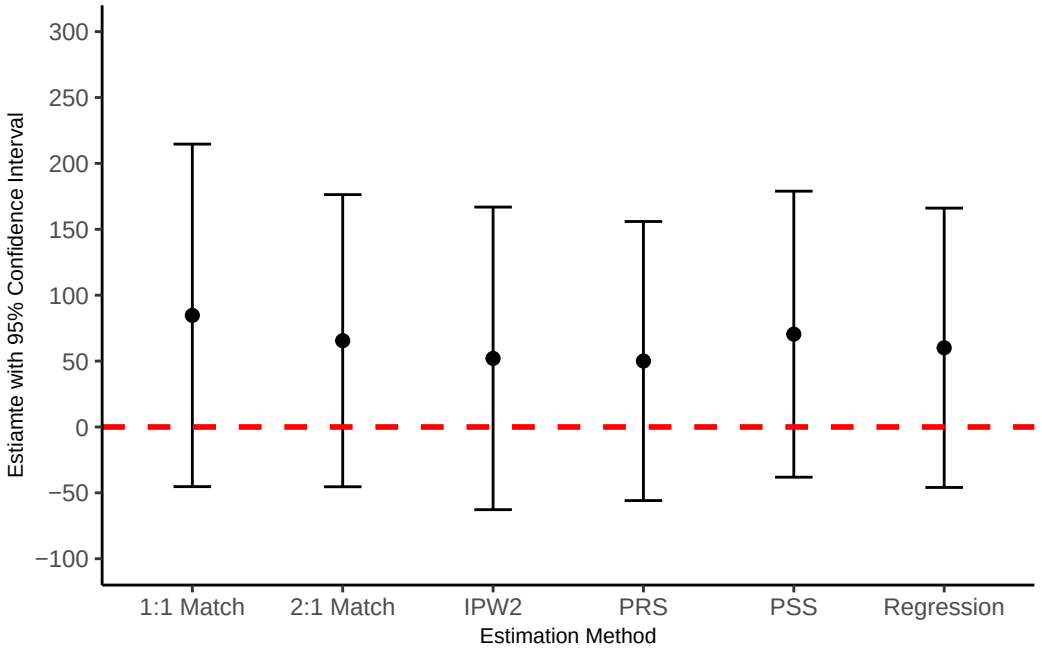

Figure 2: Comparison of Estiamtion Method Varinances for ATT of Birthweight Outcome

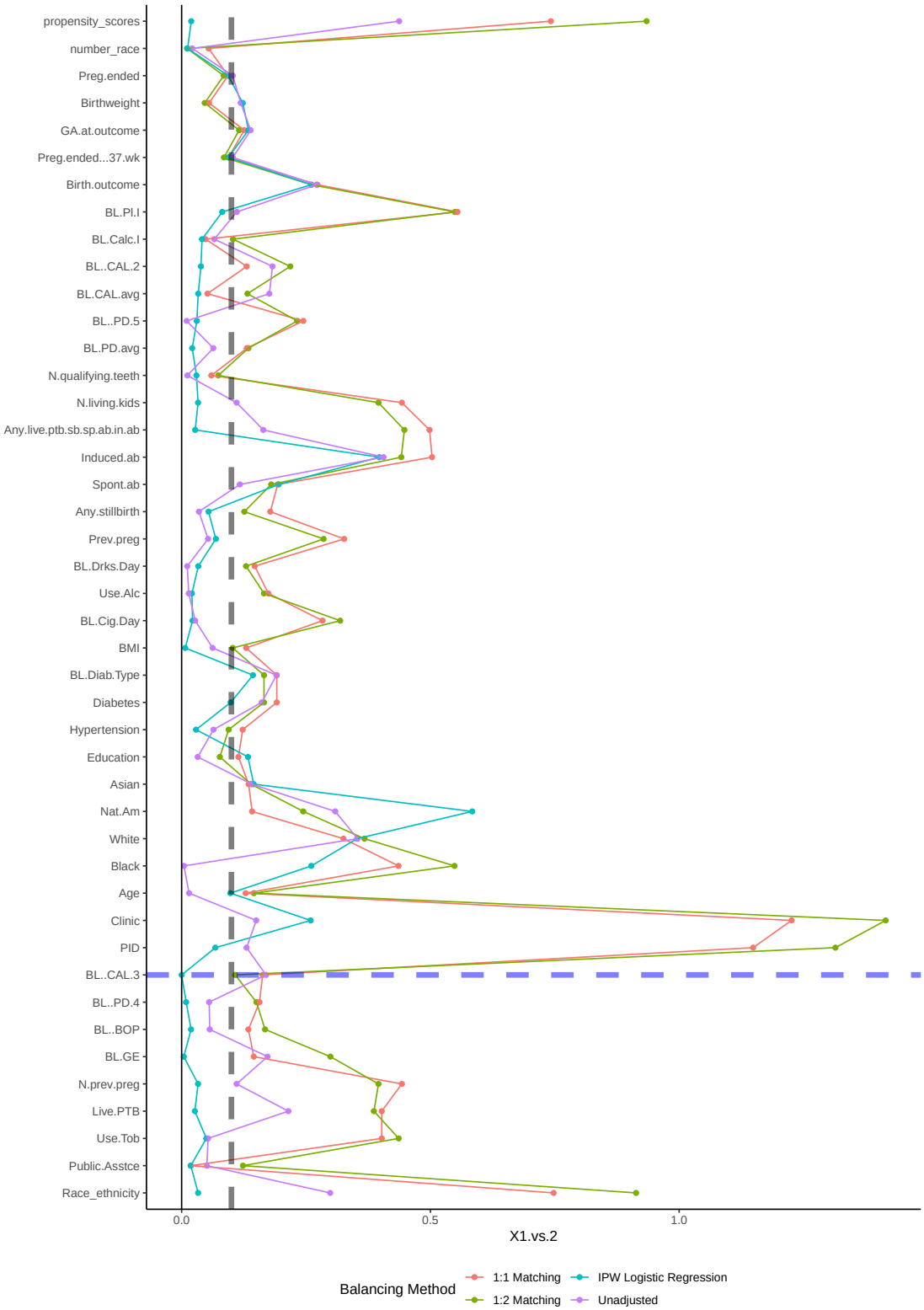

Figure 3: Comparion of confounding adjustment methods and their impact on SMDs

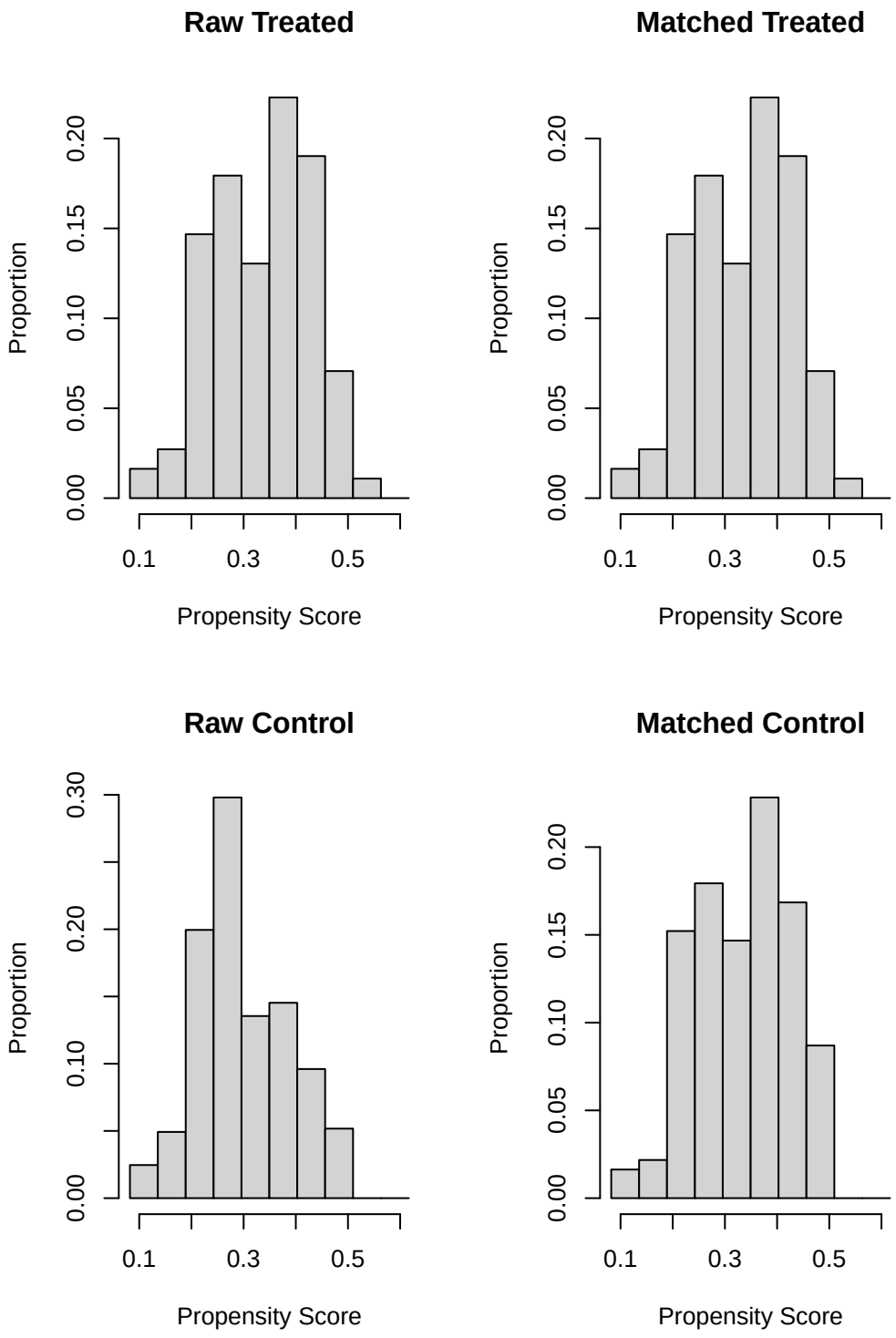

Figure 4: Distribution of Propensity Scores Before and After 1:1 Matching

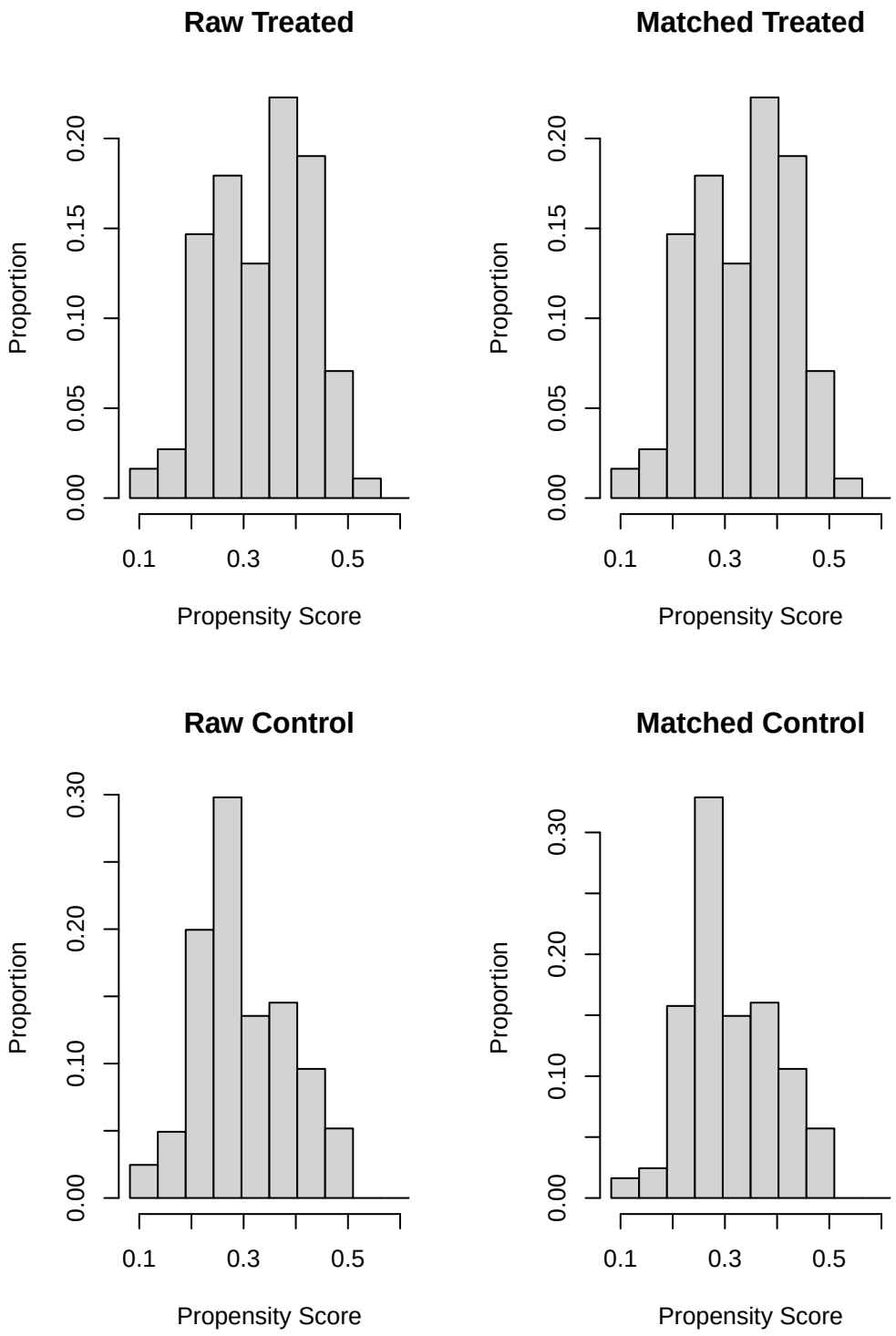

Figure 5: Distribution of Propensity Scores Before and After 2:1 Matching

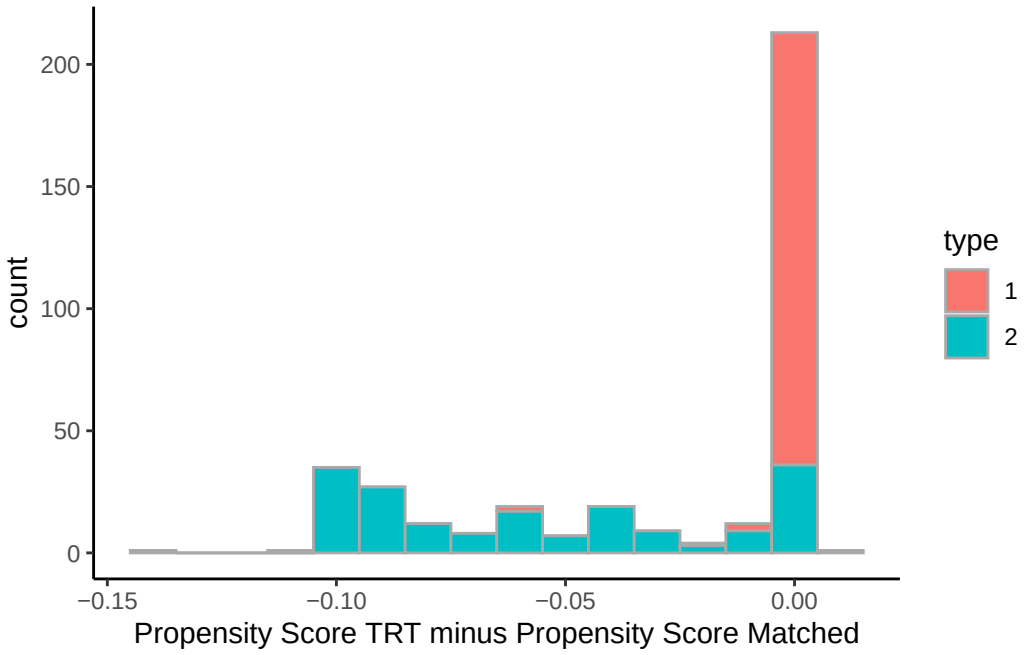

Figure 6: Difference between the Treated ith Matched Control