

## EXAMPLE: ANALYSIS OF THE NHANES DATA

We study the publicly available (U.S.) National Health and Nutrition Examination Survey (NHANES). For this, we use the libraries `NHANES`, `tableone`, and `Matching` in R. We will focus our analysis on the question of whether current smoking affects average systolic blood pressure. The variables we will need are:

- `BPSysAve` (systolic blood pressure, response),
- `SmokeNow` (smoking status: 0–No, 1–Yes),
- `Gender`,
- `Age`,
- `Race3`,
- `Education`,
- `MaritalStatus`, and
- `Poverty`

where the first two are the outcome and exposure of interest and the remaining are potential confounders. Additionally, we will restrict our attention to 1377 adults (> 17 years old) in the second wave of the survey.

```
library(NHANES); library(tableone); library(Matching,quietly=TRUE) #Must be pre-loaded in R
NHANES$SmokeNow <- as.numeric(NHANES$SmokeNow)-1
small.nhanes <- na.omit(NHANES[NHANES$SurveyYr=="2011_12"
                             & NHANES$Age > 17,c(3,4,8:11,13,25,61)])
#dim(small.nhanes) ## 1377
vars <- c("Gender", "Age", "Race3", "Education", "MaritalStatus", "Poverty")
fit0<-lm(BPSysAve~SmokeNow,data=small.nhanes)
round(coef(summary(fit0)),5)

+           Estimate Std. Error  t value Pr(>|t|)
+ (Intercept) 125.61381    0.63365 198.23993  0.00000
+ SmokeNow     -3.67936    0.96395  -3.81696  0.00014

fmod<-formula(BPSysAve~SmokeNow+Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty)
fit1<-lm(fmod,data=small.nhanes)
round(coef(summary(fit1))[1:2,],5)

+           Estimate Std. Error  t value Pr(>|t|)
+ (Intercept) 107.25006    4.75287  22.56534  0.00000
+ SmokeNow     -1.09777    0.93042  -1.17986  0.23826
```

**Checking confounder balance:** In PS-based methods, the goal of the treatment model is to eliminate imbalance in the distribution of covariates between treated and untreated subjects; achieving balance on other covariates (particularly strong predictors of treatment) is unhelpful. The goal is *not* to build an excellent predictive model for the treatment. Common measures of balance include:

- Standardized mean difference (SMD) or proportion:

$$\frac{\bar{x}^{1,w} - \bar{x}^{0,w}}{\sqrt{(v^{1,w} + v^{0,w})/2}} \quad \text{where} \quad \bar{x}^{z,w} = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{1}_z(z_i)x_i}{f_{Z|X}^O(z_i|x_i)},$$

i.e. the weighted sample mean of variable  $X$  among those with treatment value  $z$ , and similarly  $v^{z,w}$  is the weighted variance estimate. For all methods of analysis other than IPW, the weights are taken to be 1 for all subjects. An SMD of 0.1 or less typically indicates reasonable balance.

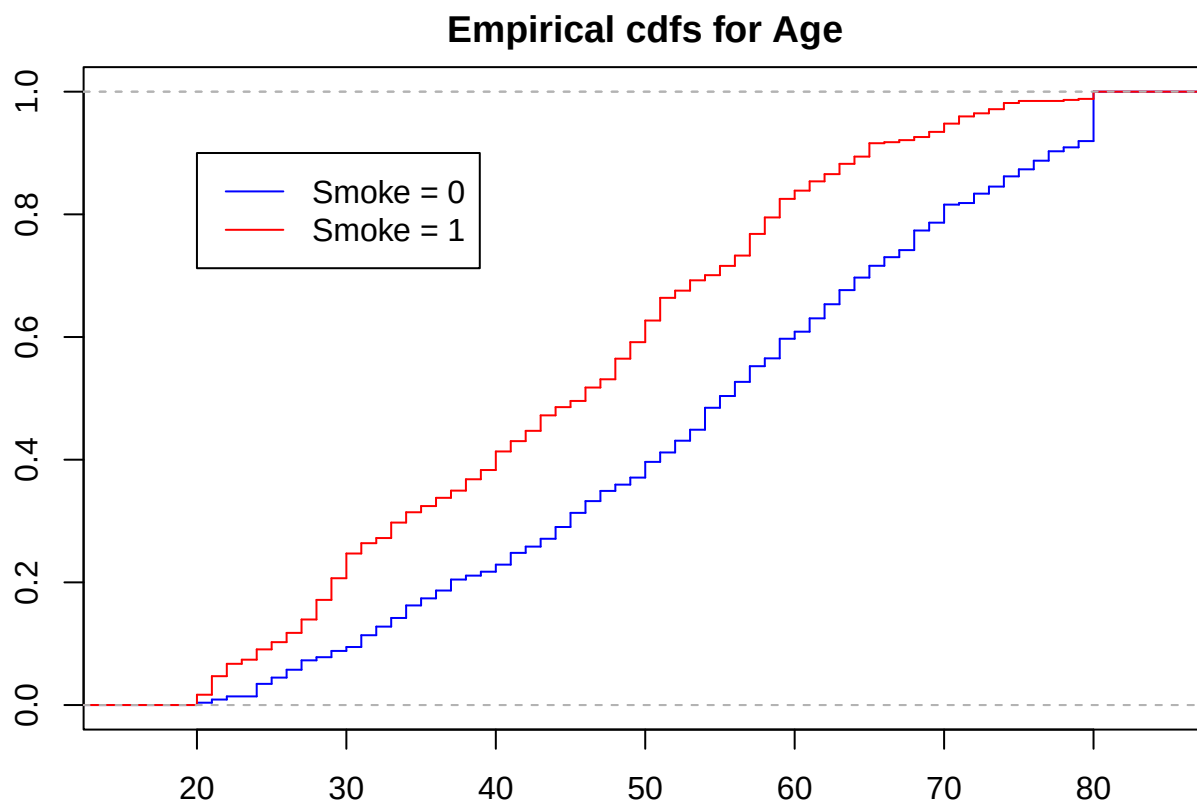
- Visual examination of weighted empirical CDFs among the treated and untreated groups.

In the NHANES example, there is clear age imbalance between the two smoking groups. This implies that if Age is also a possible predictor of the outcome, then it is a possible confounder.

```
tabUnmatched<-CreateTableOne(vars=vars, strata="SmokeNow", data=small.nhanes, test=FALSE)
print(tabUnmatched, smd = TRUE)
```

|                     | Stratified by SmokeNow |  |               |  |       |
|---------------------|------------------------|--|---------------|--|-------|
|                     | 0                      |  | 1             |  | SMD   |
| n                   | 782                    |  | 595           |  |       |
| Gender = male (%)   | 432 (55.2)             |  | 369 (62.0)    |  | 0.138 |
| Age (mean (SD))     | 54.33 (16.52)          |  | 44.96 (15.11) |  | 0.592 |
| Race3 (%)           |                        |  |               |  | 0.315 |
| Asian               | 25 ( 3.2)              |  | 15 ( 2.5)     |  |       |
| Black               | 43 ( 5.5)              |  | 64 (10.8)     |  |       |
| Hispanic            | 26 ( 3.3)              |  | 38 ( 6.4)     |  |       |
| Mexican             | 45 ( 5.8)              |  | 35 ( 5.9)     |  |       |
| White               | 630 (80.6)             |  | 416 (69.9)    |  |       |
| Other               | 13 ( 1.7)              |  | 27 ( 4.5)     |  |       |
| Education (%)       |                        |  |               |  | 0.512 |
| 8th Grade           | 59 ( 7.5)              |  | 33 ( 5.5)     |  |       |
| 9 - 11th Grade      | 71 ( 9.1)              |  | 120 (20.2)    |  |       |
| High School         | 152 (19.4)             |  | 151 (25.4)    |  |       |
| Some College        | 256 (32.7)             |  | 210 (35.3)    |  |       |
| College Grad        | 244 (31.2)             |  | 81 (13.6)     |  |       |
| MaritalStatus (%)   |                        |  |               |  | 0.488 |
| Divorced            | 85 (10.9)              |  | 77 (12.9)     |  |       |
| LivePartner         | 61 ( 7.8)              |  | 96 (16.1)     |  |       |
| Married             | 453 (57.9)             |  | 240 (40.3)    |  |       |
| NeverMarried        | 108 (13.8)             |  | 142 (23.9)    |  |       |
| Separated           | 6 ( 0.8)               |  | 14 ( 2.4)     |  |       |
| Widowed             | 69 ( 8.8)              |  | 26 ( 4.4)     |  |       |
| Poverty (mean (SD)) | 3.11 (1.65)            |  | 2.38 (1.58)   |  | 0.453 |

```
Smoke<-small.nhanes$SmokeNow; age0<-small.nhanes$Age[Smoke==0]; age1<-small.nhanes$Age[Smoke==1]
ecdf0 <- ecdf(age0); ecdf1 <- ecdf(age1)
```



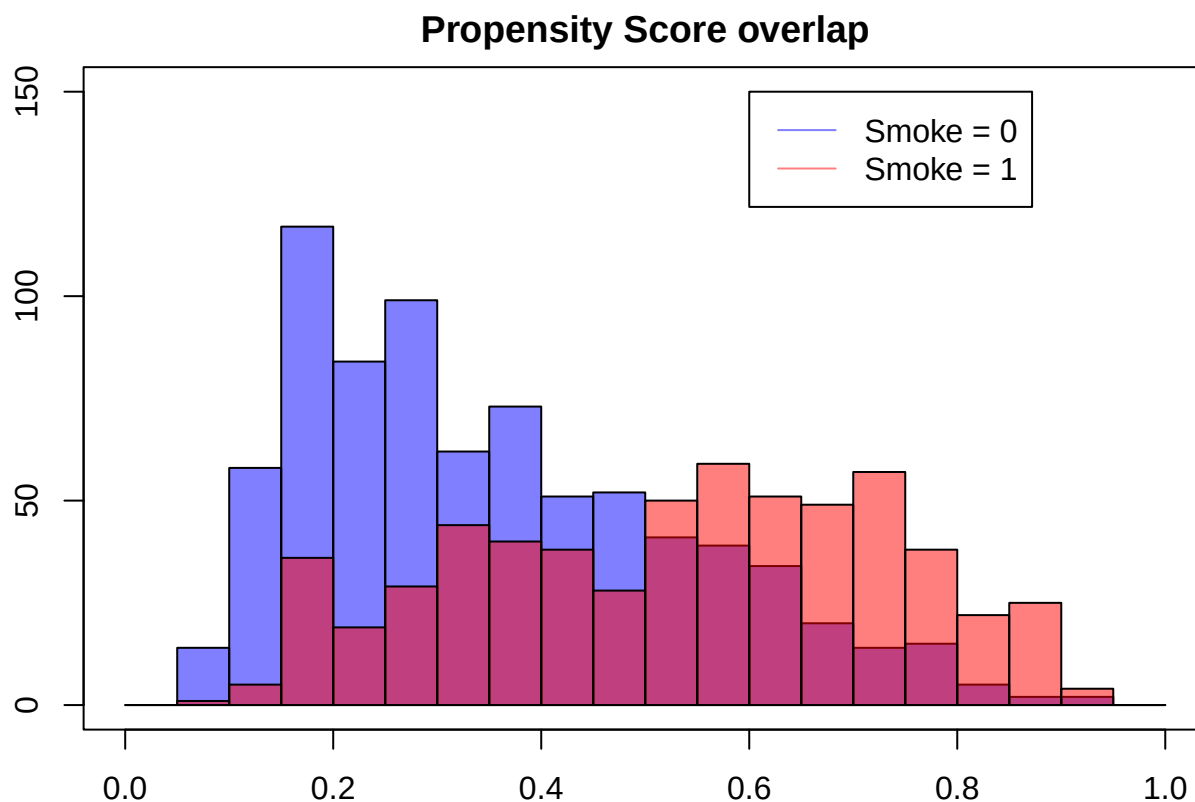
**Building the propensity score:** We attempt to build and assess a propensity score model using the available covariates:

```
fmods<-formula(SmokeNow~Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty)
ps.mod <- glm(fmods,data=small.nhanes,family="binomial")
ps.lr <- predict(ps.mod,type="response")
ps0<-ps.lr[Smoke==0];ps1<-ps.lr[Smoke==1]
quints <- c(0,quantile(ps.lr,seq(.2,1,.2)))
rbind(table(cut(ps.lr[Smoke==0],quints)),table(cut(ps.lr[Smoke==1],quints)))
```

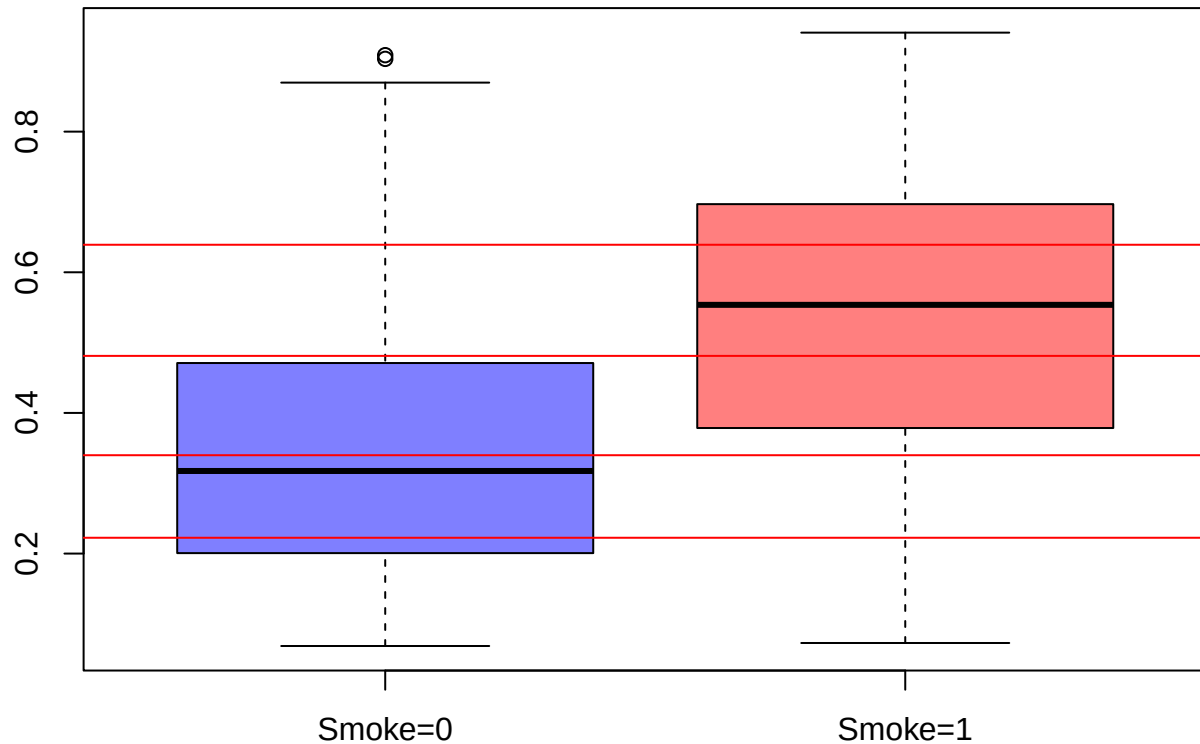
|   |           |              |              |               |               |     |
|---|-----------|--------------|--------------|---------------|---------------|-----|
| + | (0,0.222] | (0.222,0.34] | (0.34,0.481] | (0.481,0.639] | (0.639,0.941] |     |
| + | [1,]      | 231          | 194          | 167           | 121           | 69  |
| + | [2,]      | 47           | 82           | 105           | 157           | 204 |

The table indicates that there is good representation of the exposed and unexposed groups in the five quintile-based strata. We now verify by looking for overlap of the propensity score histograms.

```
par(mar=c(2,2,2,0))
hist(ps0, col=rgb(0,0,1,0.5), breaks=seq(0,1,by=0.05), ylim=c(0,150),
     main="Propensity Score overlap", xlab="PS")
hist(ps1, col=rgb(1,0,0,0.5), breaks=seq(0,1,by=0.05), add=T);box()
legend(0.6,150,c('Smoke = 0', 'Smoke = 1'),col=c(rgb(0,0,1,0.5),rgb(1,0,0,0.5)),lty=1)
```



```
boxplot(ps0,ps1,ylab="PS",xlab="Treatment Group",names=c('Smoke=0','Smoke=1'),
       col=c(rgb(0,0,1,0.5),rgb(1,0,0,0.5)));abline(h=quints[2:5],col="red")
```



We can therefore proceed to check for balance knowing we have sufficient numbers of smokers and non-smokers in each quintile to ensure the stratum-specific estimates are not too unstable.

```
Pcat<-as.numeric(cut(ps.lnr,quints,include.lowest=T))
W<-(Smoke==0)/(1-ps.lnr)+(Smoke==1)/ps.lnr
smd.mat<-ExtractSmd(tabUnmatched)
for(k in 1:5){
  nhanesQ<-small.nhanes[Pcat == k,]
  tabQs <- CreateTableOne(vars = vars, strata = "SmokeNow", data = nhanesQ, test = FALSE)
  smd.mat<-cbind(smd.mat,ExtractSmd(tabQs))
}
colnames(smd.mat)<-c('Original','Q1','Q2','Q3','Q4','Q5')
round(smd.mat,4)
```

|                 | Original | Q1     | Q2     | Q3     | Q4     | Q5     |
|-----------------|----------|--------|--------|--------|--------|--------|
| + Gender        | 0.1379   | 0.1017 | 0.1043 | 0.0286 | 0.2003 | 0.0308 |
| + Age           | 0.5918   | 0.2574 | 0.1715 | 0.0993 | 0.3106 | 0.1642 |
| + Race3         | 0.3148   | 0.3169 | 0.1117 | 0.3444 | 0.4147 | 0.2869 |
| + Education     | 0.5119   | 0.5378 | 0.4167 | 0.2800 | 0.2378 | 0.3018 |
| + MaritalStatus | 0.4877   | 0.4320 | 0.2386 | 0.2725 | 0.2332 | 0.2608 |
| + Poverty       | 0.4530   | 0.0865 | 0.1256 | 0.1136 | 0.0041 | 0.1455 |

Balance does not appear to have been achieved: we have SMDs > 0.1 for at least three quintiles for all variables, and the empirical CDFs of age do not overlap in several quintiles.

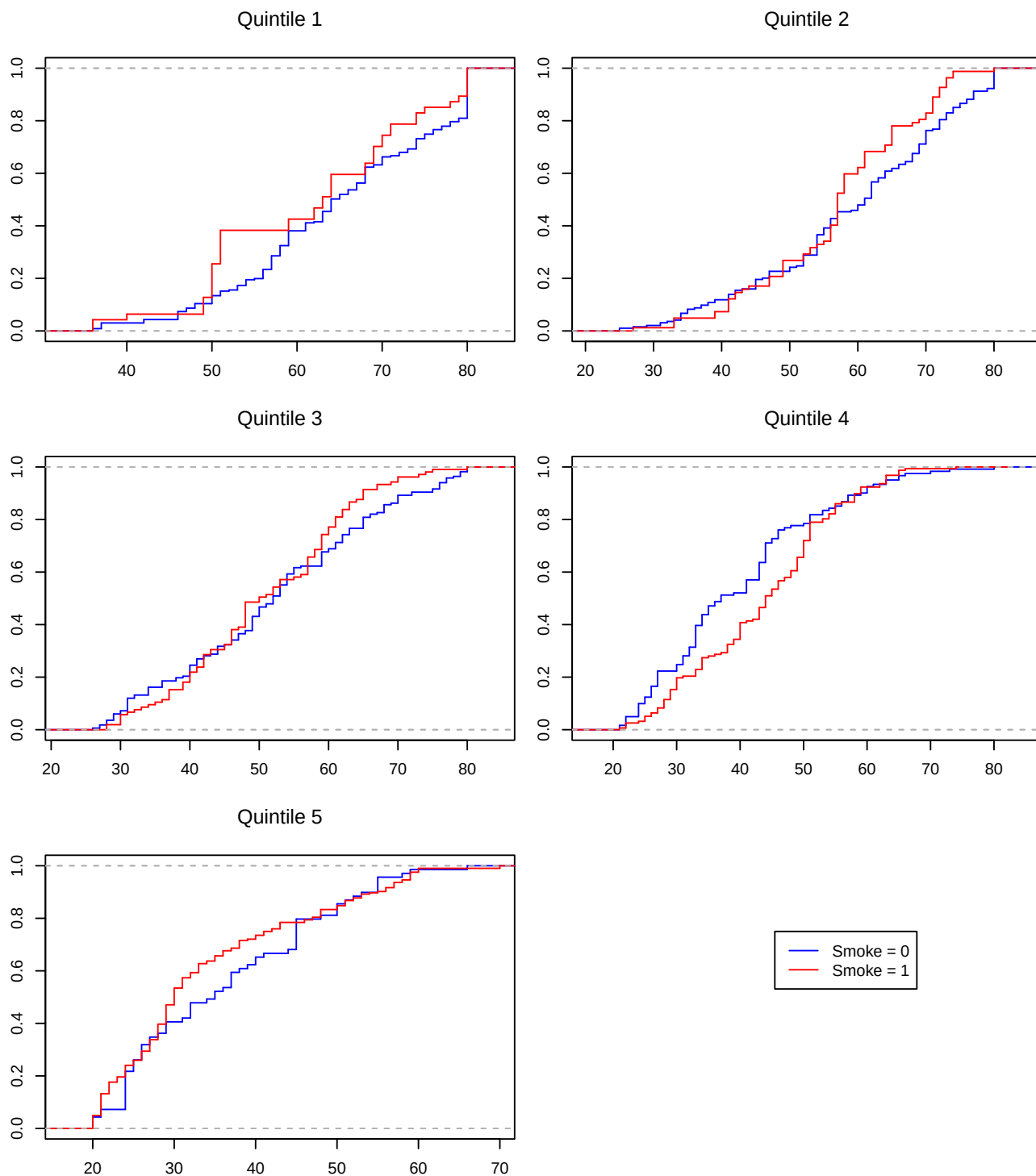
```
par(mar=c(2,3,4,0),mfrow=c(3,2),oma=c(0,0,2,0))
for(k in 1:5){
  age0 <- small.nhanes$Age[Smoke==0 & Pcat==k]
  age1 <- small.nhanes$Age[Smoke==1 & Pcat==k]
  ecdf0 <- ecdf(age0)
  ecdf1 <- ecdf(age1)
  plot(ecdf0, verticals=TRUE, do.points=FALSE,
       main=substitute(paste('Quintile ',k),list(k=k)),col='blue')
```

```

    plot(ecdf1, verticals=TRUE, do.points=FALSE, add=TRUE, col='red')
  }
  plot(age0,type='n',ylim=range(0,1),axes=FALSE)
  title("ECDFs for Age by PS quintile",outer = TRUE)
  legend(30,0.75,c('Smoke = 0', 'Smoke = 1'),col=c('blue','red'),lty=1)

```

ECDFs for Age by PS quintile



We may try to find evidence in smaller PS strata, such as those formed by deciles of the propensity score. However, here this does not assist with providing evidence of balance.

**Matching:** We now attempt matching using the propensity score. The function `MatchBalance` from the `Matching` library provides many more details than `CreateTableOne`, including:

- mean, median, and maximum difference in empirical CDF plots,
- mean, median, and maximum difference in empirical QQ plots,
- Kolmogorov-Smirnov statistics,
- ratio of variances,
- p-value for t-test.

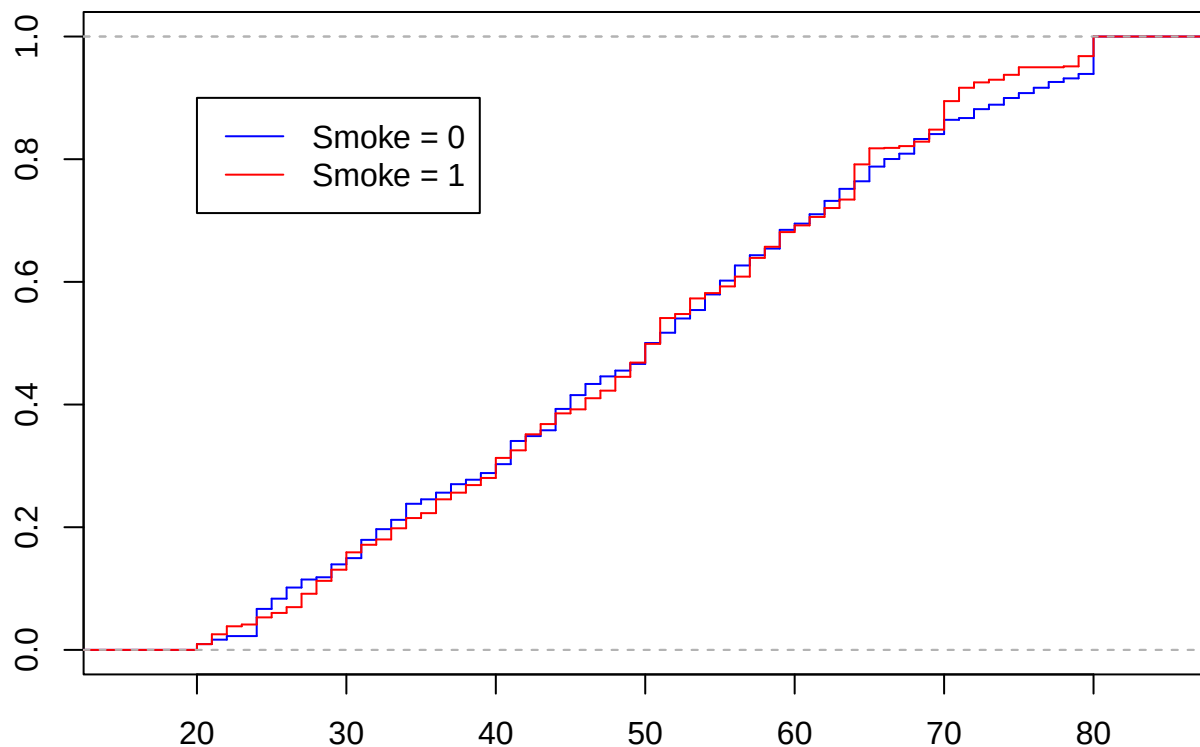
Results shown in Appendix. Note that SMDs reported are multiplied by 100.

```
small.nhanes$ps.lr<-ps.lr
ps.lr.match <- Match(Tr=small.nhanes$SmokeNow,X=small.nhanes$ps.lr,estimand="ATE",ties=FALSE)
matched.samp <- small.nhanes[c(ps.lr.match$index.control,ps.lr.match$index.treated),]
table(table(c(ps.lr.match$index.control, ps.lr.match$index.treated)))

+
+  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  23  29  38
+ 850 243 131  54  41  16   8   4   3   3   6   3   3   4   4   1   1   1   1

tabMatched<-CreateTableOne(vars = vars, strata = "SmokeNow",data = matched.samp, test = FALSE)
```

**Empirical cdfs for Age in matched sample**



Matching seems to have matched the age profiles for the smoking and non-smoking groups.

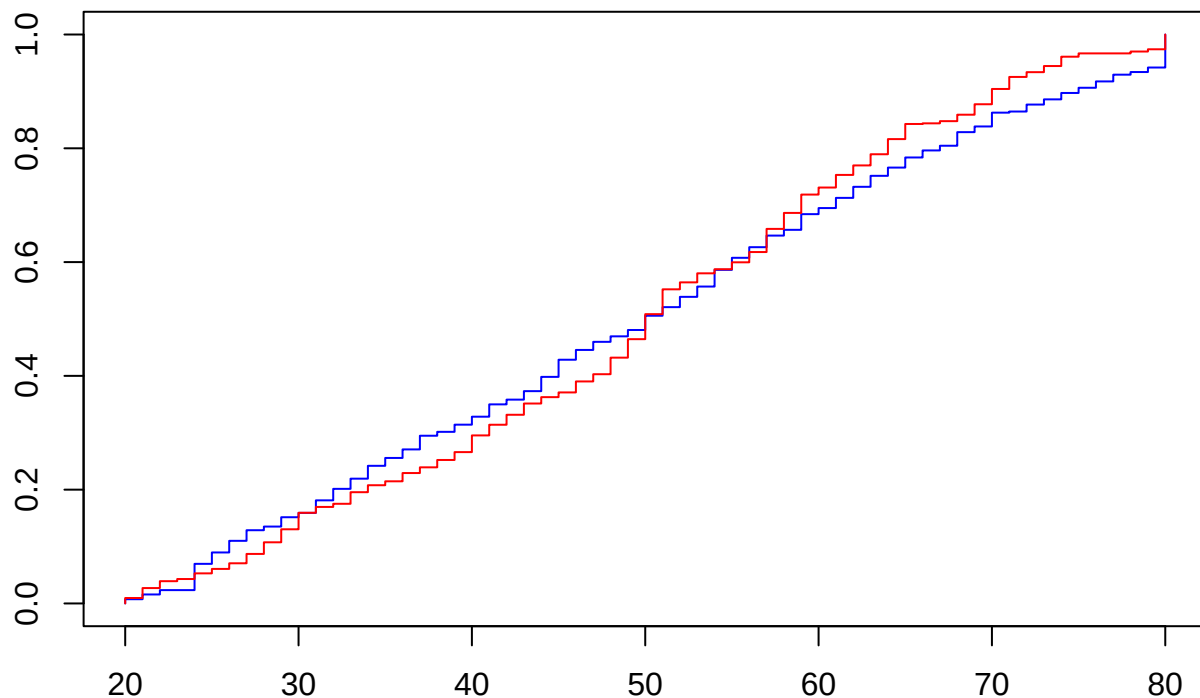
```
+
+   Original  Match
+ Gender      0.1379 0.0234
+ Age         0.5918 0.0101
+ Race3       0.3148 0.1411
+ Education   0.5119 0.1240
+ MaritalStatus 0.4877 0.1112
+ Poverty     0.4530 0.0531
```

**Inverse Weighting:** We construct an inverse weighted sample using the survey and Hmisc packages.

```
library(survey);library(Hmisc);library(htmlwidgets) #Must be pre-loaded in R
ps.lr.weight <- small.nhanes$SmokeNow/ps.lr + (1-small.nhanes$SmokeNow)/(1-ps.lr)
nhanes.IPW.lr <- svydesign(ids=~0, data=small.nhanes, weights=ps.lr.weight)
tabIPW<-svyCreateTableOne(vars = vars, strata = "SmokeNow",data = nhanes.IPW.lr, test = FALSE)
```

The summary of the reweighted samples are given in the Appendix; it is clear that in the weighted (pseudo) data balance holds well.

### Empirical cdfs for Age in weighted sample



|                 | Original | IPW    |
|-----------------|----------|--------|
| + Gender        | 0.1379   | 0.0235 |
| + Age           | 0.5918   | 0.0139 |
| + Race3         | 0.3148   | 0.0516 |
| + Education     | 0.5119   | 0.0292 |
| + MaritalStatus | 0.4877   | 0.0235 |
| + Poverty       | 0.4530   | 0.0001 |

Here the matched and inverse weighted samples exhibit good balance.

#### Summary:

- Creating or restoring confounder balance is essential to estimating a causal effect.
- It can be hard to assess overlap or achieve balance in high dimensions.
- The propensity score, a scalar summary of confounding variables, simplifies this task.
- However:
  - fitting a model for treatment does not guarantee balance,
  - fitting a model that predicts treatment with a high degree of precision can be unhelpful.

## ESTIMATING THE ATE

We proceed now to estimating the average treatment effect (ATE), using:

- outcome regression:
- PS stratification,
- PS matching,
- PS regression,
- IPW & AIPW

We will use the PS estimated via logistic regression, as this provided the best balance. Note that the IPW and AIPW estimators can suffer from the fact that the inverse weights based on the propensity score can get large if the propensity score is near zero or one. For these estimators, weight truncation can be used, where propensity score values more extreme than a given threshold (0.001 or 0.999 say) are set equal to these thresholds.

- **Outcome regression, unadjusted:**

```
coef(summary(lm(BPSysAve~SmokeNow,data=small.nhanes)))[2,]  
+      Estimate      Std. Error      t value      Pr(>|t|)  
+ -3.6793569602  0.9639505436 -3.8169561548  0.0001411118
```

- **Outcome regression no interaction:** We use the model

BPSysAve ~ SmokeNow+Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty

```
fmod<-formula(BPSysAve~SmokeNow+Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty)  
coef(lm(fmod,data=small.nhanes))[2]  
+ SmokeNow  
+ -1.097768
```

```
nhanes.allsmoke <- small.nhanes  
nhanes.allsmoke$SmokeNow <- 1  
nhanes.nosmoke <- small.nhanes  
nhanes.nosmoke$SmokeNow <- 0  
mod1.lm <- lm(fmod,data=small.nhanes)  
APO.lm.1 <- mean(predict(mod1.lm,nhanes.allsmoke))  
APO.lm.0 <- mean(predict(mod1.lm,nhanes.nosmoke))  
APO.lm.1 - APO.lm.0 #Note: same as for previous answer regression  
+ [1] -1.097768
```

The naive conditional effect estimate is more than 3 times greater than its confounder-adjusted counterpart.

- **Outcome regression with interactions:** We now add interaction terms

BPSysAve ~SmokeNow+Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty  
+SmokeNow:HHIncome+SmokeNow:Gender+SmokeNow:Age

```
fmodi<-formula(BPSysAve~SmokeNow+Gender+Age+Race3+Education+  
MaritalStatus+HHIncome+Poverty+SmokeNow:HHIncome+SmokeNow:Gender+SmokeNow:Age)  
mod1.lmX <- lm(fmodi,data=small.nhanes)  
coef(summary(mod1.lmX))[c(2,10,11,12),]  
+      Estimate Std. Error      t value      Pr(>|t|)  
+ SmokeNow      -14.280706   6.041067 -2.3639377  0.01822449  
+ Education9 - 11th Grade -2.953706   2.070564 -1.4265223  0.15395163  
+ EducationHigh School   -3.063842   1.958024 -1.5647624  0.11787581  
+ EducationSome College  -1.228311   1.931240 -0.6360218  0.52487131  
APO.lmX.1 <- mean(predict(mod1.lmX,nhanes.allsmoke))  
APO.lmX.0 <- mean(predict(mod1.lmX,nhanes.nosmoke))  
APO.lmX.1 - APO.lmX.0  
+ [1] -1.402538
```

Here the coefficients alone **do not** provide the ATE estimate, and we need to use the predict method to compute the ATE estimate as

$$\frac{1}{n} \sum_{i=1}^n \mu(x_i, 1; \hat{\beta}) - \frac{1}{n} \sum_{i=1}^n \mu(x_i, 0; \hat{\beta})$$

- **PS stratification:** We first fit use the propensity score model based on the logistic regression with all possible confounders included

BPSysAve ~ Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty

and then used the stratification based on PS quintiles.

```
fmods<-formula(SmokeNow~Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty)
ps.mod <- glm(fmods,data=small.nhanes,family="binomial")
ps.lr <- predict(ps.mod,type="response")
Y<-small.nhanes$BPSysAve
ps.lr.quints <- cut(ps.lr,quints,labels=1:5)
p.strat <- table(ps.lr.quints)/length(ps.lr.quints)
ATE.strat <- rep(NA,5)
for(j in 1:5) {
  ATE.strat[j]<-mean(Y[Smoke==1 & ps.lr.quints==j])-mean(Y[Smoke==0 & ps.lr.quints==j])
}
sum(ATE.strat*p.strat)
+ [1] -1.816879
```

- **PS matching:** We now use the same PS model for matching.

```
ps.lr.match <- Match(Tr=small.nhanes$SmokeNow,
  X=small.nhanes$ps.lr,estimand="ATE",ties=FALSE)
matched.samp <- small.nhanes[c(ps.lr.match$index.control,ps.lr.match$index.treated),]
mean(matched.samp$BPSysAve[matched.samp$SmokeNow == 1]) -
  mean(matched.samp$BPSysAve[matched.samp$SmokeNow == 0])
+ [1] -0.8366013
```

- **PS regression:** We now attempt three propensity score regression analyses based on increasingly complex specifications:

$$\text{PSR 1 } Y = \beta_0 + \phi_0 e(X; \hat{\alpha}) + \psi_0 Z + \epsilon$$

$$\text{PSR 2 } Y = \beta_0 + \phi_0 e(X; \hat{\alpha}) + \phi_1 \{e(X; \hat{\alpha})\}^2 + Z(\psi_0 + \psi_1 e(X; \hat{\alpha})) + \epsilon$$

$$\text{PSR 3 } Y = s(e(X; \hat{\alpha})) + Z(\psi_0 + \psi_1 s(e(X; \hat{\alpha}))) + \epsilon$$

where  $s(\cdot)$  is a flexible spline function. The inclusion of extra terms including the propensity score attempts to explain the treatment free component, and the addition of interactions with  $Z$  aims to allow for non-additive treatment effects.

```
library(splines)
mod1.PS1m1 <- lm(BPSysAve~SmokeNow+ps.lr,data=small.nhanes)
APO.PS1m1.1 <- mean(predict(mod1.PS1m1,nhanes.allsmoke))
APO.PS1m1.0 <- mean(predict(mod1.PS1m1,nhanes.nosmoke))
APO.PS1m1.1 - APO.PS1m1.0
+ [1] -1.10791
mod1.PS1m2 <- lm(BPSysAve~SmokeNow*(1+ps.lr)+(ps.lr+I(ps.lr^2)),data=small.nhanes)
APO.PS1m2.1 <- mean(predict(mod1.PS1m2,nhanes.allsmoke))
APO.PS1m2.0 <- mean(predict(mod1.PS1m2,nhanes.nosmoke))
APO.PS1m2.1 - APO.PS1m2.0
+ [1] -1.30571
mod1.PS1m3 <- lm(BPSysAve~SmokeNow*(1+bs(ps.lr,df=4,Boundary.knots = range(0,1)))+
  bs(ps.lr,df=4,Boundary.knots = range(0,1)),
  data=small.nhanes)
APO.PS1m3.1 <- mean(predict(mod1.PS1m3,nhanes.allsmoke))
APO.PS1m3.0 <- mean(predict(mod1.PS1m3,nhanes.nosmoke))
APO.PS1m3.1 - APO.PS1m3.0
+ [1] -1.971314
```

- **IPW:** The IPW estimator can be computed using a direct calculation

$$\frac{1}{n} \sum_{i=1}^n \frac{Z_i Y_i}{e(X_i; \hat{\alpha})} - \frac{1}{n} \sum_{i=1}^n \frac{(1 - Z_i) Y_i}{1 - e(X_i; \hat{\alpha})}$$

or in its standardized weight form via weighted least squares.

```
ps.lr.weight <- Smoke/ps.lr + (1-Smoke)/(1-ps.lr)
mean(Smoke*Y*ps.lr.weight) - mean((1-Smoke)*Y*ps.lr.weight) #IPW0
+ [1] -1.928655
coef(lm(Y ~ Smoke, weights = ps.lr.weight))[2] #IPW
+ Smoke
+ -1.991233
```

- **AIPW:** The augmented IPW estimator adds a proposed mean model for the APOs, that is

$$\hat{\mu}(1) = \frac{1}{n} \sum_{i=1}^n \frac{Z_i (Y_i - \mu(X_i, 1; \hat{\beta}))}{e(X_i; \hat{\alpha})} + \frac{1}{n} \sum_{i=1}^n \mu(X_i, 1; \hat{\beta})$$

$$\hat{\mu}(0) = \frac{1}{n} \sum_{i=1}^n \frac{(1 - Z_i) (Y_i - \mu(X_i, 0; \hat{\beta}))}{1 - e(X_i; \hat{\alpha})} + \frac{1}{n} \sum_{i=1}^n \mu(X_i, 0; \hat{\beta})$$

or computed via an augmented outcome regression with the two additional predictors

$$\frac{Z_i}{e(X_i; \hat{\alpha})} \quad \frac{(1 - Z_i)}{1 - e(X_i; \hat{\alpha})}.$$

```
mod1 <- lm(fmodi, data=small.nhanes)
m1<-predict(mod1, nhanes.allsmoke)
m0<-predict(mod1, nhanes.nosmoke)
ps.trunc<-pmin(pmax(0.001, ps.lr), 0.999)
APO.1<-mean(Smoke*(Y-m1)/ps.trunc)+mean(m1)
APO.0<-mean((1-Smoke)*(Y-m0)/(1-ps.trunc))+mean(m0)
APO.1-APO.0
+ [1] -1.854776
small.nhanes$ps <- ps.lr
fmodia<-formula(BPSysAve~SmokeNow+Gender+Age+Race3+Education+
  MaritalStatus+HHIncome+Poverty+
  SmokeNow:HHIncome+SmokeNow:Gender+SmokeNow:Age+
  I(SmokeNow/ps)+I((1-SmokeNow)/(1-ps)))
mod1.lmX <- lm(fmodia, data=small.nhanes)
nhanes.allsmoke <- nhanes.nosmoke <- small.nhanes;
nhanes.allsmoke$SmokeNow <- 1
nhanes.nosmoke$SmokeNow <- 0
mean(predict(mod1.lmX, nhanes.allsmoke))-mean(predict(mod1.lmX, nhanes.nosmoke))
+ [1] -1.818062
```

#### Additional considerations:

- All of the PS approaches considered rely on substitution estimators.
  - In PS regression, we plug in an estimated PS as a covariate.
  - In IPW, we plug in estimated weights.
- We need to account for this when estimating standard errors and/or confidence intervals.
- Analytically derived asymptotic variances can be used, but are not provided in many standard software packages.
- The easiest approach is to bootstrap; however, that the bootstrap is *not* in general valid for matching.

## BOOTSTRAP ANALYSIS

2000 bootstrap samples are used to approximate the sampling distribution of the various estimators.

```
nboot<-2000;set.seed(45);ests.mat<-matrix(0,nrow=nboot,ncol=10)
n<-nrow(small.nhanes)
library(splines)
for(ib in 1:nboot){
  ind<-sample(1:n,size=n,rep=T)
  boot.nhanes<-small.nhanes[ind,]
  boot.allsmoke <- boot.nosmoke <- boot.nhanes
  boot.allsmoke$SmokeNow <- 1
  boot.nosmoke$SmokeNow <- 0

  ests.mat[ib,1]<-coef(lm(BPSysAve~SmokeNow,data=boot.nhanes))[2]
  ests.mat[ib,2]<-coef(lm(fmod,data=boot.nhanes))[2]

  mod1.lmX <- lm(fmodi,data=boot.nhanes)
  APO.lmX.1 <- mean(predict(mod1.lmX,boot.allsmoke))
  APO.lmX.0 <- mean(predict(mod1.lmX,boot.nosmoke))
  ests.mat[ib,3]<-APO.lmX.1 - APO.lmX.0

  ps.mod <- glm(fmods,data=boot.nhanes,family="binomial")
  ps.lr <- predict(ps.mod,type="response")
  ps0<-ps.lr[boot.nhanes$SmokeNow==0];ps1<-ps.lr[boot.nhanes$SmokeNow==1]
  quints <- c(0,quantile(ps.lr,seq(.2,1,.2)))
  Y<-boot.nhanes$BPSysAve
  ps.lr.quints <- cut(ps.lr,quints,labels=1:5)
  p.strat <- table(ps.lr.quints)/length(ps.lr.quints)
  ATE.strat <- rep(NA,5)
  for(j in 1:5) {
    ATE.strat[j] <- mean(Y[boot.nhanes$SmokeNow == 1 & ps.lr.quints==j])-
      mean(Y[boot.nhanes$SmokeNow == 0 & ps.lr.quints==j])
  }
  ests.mat[ib,4]<-sum(ATE.strat*p.strat)

  boot.nhanes$ps.lr<-ps.lr
  ps.lr.match <- Match(Tr=boot.nhanes$SmokeNow,X=boot.nhanes$ps.lr,estimand="ATE",ties=FALSE)
  matched.samp <- boot.nhanes[c(ps.lr.match$index.control,ps.lr.match$index.treated),]
  ests.mat[ib,5]<-mean(matched.samp$BPSysAve[matched.samp$SmokeNow == 1]) -
    mean(matched.samp$BPSysAve[matched.samp$SmokeNow == 0])

  mod1.PS1m1 <- lm(BPSysAve~SmokeNow+ps.lr,data=boot.nhanes)
  APO.PS1m1.1 <- mean(predict(mod1.PS1m1,boot.allsmoke))
  APO.PS1m1.0 <- mean(predict(mod1.PS1m1,boot.nosmoke))
  ests.mat[ib,6]<-APO.PS1m1.1 - APO.PS1m1.0

  mod1.PS1m2 <- lm(BPSysAve~ SmokeNow*(1+ps.lr)+(ps.lr+I(ps.lr^2)),data=boot.nhanes)
  APO.PS1m2.1 <- mean(predict(mod1.PS1m2,boot.allsmoke))
  APO.PS1m2.0 <- mean(predict(mod1.PS1m2,boot.nosmoke))
  ests.mat[ib,7]<-APO.PS1m2.1 - APO.PS1m2.0

  mod1.PS1m3 <- lm(BPSysAve~SmokeNow*(1+bs(ps.lr,df=4,Boundary.knots = range(0,1)))+
    bs(ps.lr,df=4,Boundary.knots = range(0,1)),
    data=boot.nhanes)
  APO.PS1m3.1 <- mean(predict(mod1.PS1m3,boot.allsmoke))
  APO.PS1m3.0 <- mean(predict(mod1.PS1m3,boot.nosmoke))
  ests.mat[ib,8]<-APO.PS1m3.1 - APO.PS1m3.0

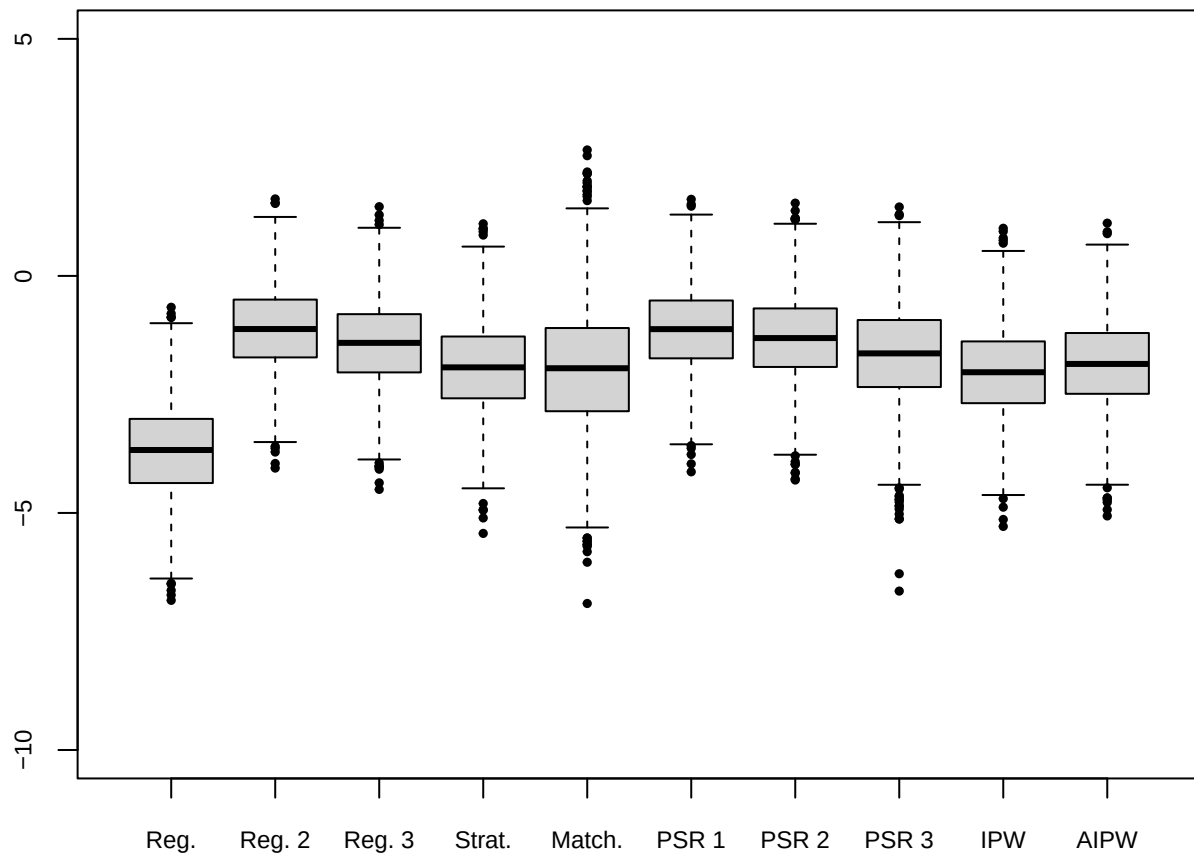
  ps.lr.weight <- boot.nhanes$SmokeNow/ps.lr + (1-boot.nhanes$SmokeNow)/(1-ps.lr)
  ests.mat[ib,9]<-coef(lm(Y ~ boot.nhanes$SmokeNow, weights = ps.lr.weight))[2]
```

```

boot.allsmoke <- boot.nosmoke <- boot.nhanes
boot.allsmoke$SmokeNow <- 1
boot.nosmoke$SmokeNow <- 0

mod1 <- lm(fmodi,data=boot.nhanes)
m1<-predict(mod1,boot.allsmoke)
m0<-predict(mod1,boot.nosmoke)
Smoke<-boot.nhanes$SmokeNow
ps.trunc<-pmin(pmax(0.001,ps.lr),0.999)
AP0.1<-mean(Smoke*(Y-m1)/ps.trunc)+mean(m1)
AP0.0<-mean((1-Smoke)*(Y-m0)/(1-ps.trunc))+mean(m0)
ests.mat[ib,10]<-AP0.1-AP0.0
}

```



The boxplot demonstrates largely similar estimates from quite different analysis methods. The estimates and standard errors are as follows.

|          | Est.      | s.e.      |
|----------|-----------|-----------|
| + Reg.   | -3.685990 | 0.9625842 |
| + Reg. 2 | -1.114372 | 0.8956875 |
| + Reg. 3 | -1.425917 | 0.9196439 |
| + Strat. | -1.931379 | 0.9697013 |
| + Match. | -1.972286 | 1.3268084 |
| + PSR 1  | -1.131600 | 0.8922795 |
| + PSR 2  | -1.317842 | 0.9287679 |
| + PSR 3  | -1.656501 | 1.0705033 |
| + IPW    | -2.020559 | 0.9881891 |
| + AIPW   | -1.853941 | 0.9734414 |

## APPENDIX: SUMMARY OF MATCHED AND WEIGHTED SAMPLES

### Matched sample summary:

```
MatchBalance(SmokeNow~Gender+Age+Education+MaritalStatus+Poverty,
              data=small.nhanes,match.out=ps.lr.match)
```

```
+
+ ***** (V1) Gendermale *****
+               Before Matching      After Matching
+ mean treatment.....      0.62017      0.59768
+ mean control.....      0.55243      0.55047
+ std mean diff.....      13.945      9.6228
+
+ mean raw eQQ diff.....      0.068908      0.047204
+ med  raw eQQ diff.....      0      0
+ max  raw eQQ diff.....      1      1
+
+ mean eCDF diff.....      0.033869      0.023602
+ med  eCDF diff.....      0.033869      0.023602
+ max  eCDF diff.....      0.067738      0.047204
+
+ var ratio (Tr/Co).....      0.9531      0.97174
+ T-test p-value.....      0.011311      0.012015
+
+ ***** (V2) Age *****
+               Before Matching      After Matching
+ mean treatment.....      44.958      50.299
+ mean control.....      54.327      50.685
+ std mean diff.....      -62.015      -2.1631
+
+ mean raw eQQ diff.....      9.3294      1.2571
+ med  raw eQQ diff.....      10      1
+ max  raw eQQ diff.....      14      4
+
+ mean eCDF diff.....      0.1536      0.020608
+ med  eCDF diff.....      0.15554      0.020334
+ max  eCDF diff.....      0.2521      0.060276
+
+ var ratio (Tr/Co).....      0.83608      1.1481
+ T-test p-value..... < 2.22e-16      0.56376
+ KS Bootstrap p-value.. < 2.22e-16      0.012
+ KS Naive p-value..... < 2.22e-16      0.013437
+ KS Statistic.....      0.2521      0.060276
+
+ ***** (V3) Education9 - 11th Grade *****
+               Before Matching      After Matching
+ mean treatment.....      0.20168      0.14161
+ mean control.....      0.090793      0.13508
+ std mean diff.....      27.612      1.874
+
+ mean raw eQQ diff.....      0.11092      0.0065359
+ med  raw eQQ diff.....      0      0
+ max  raw eQQ diff.....      1      1
+
+ mean eCDF diff.....      0.055444      0.003268
+ med  eCDF diff.....      0.055444      0.003268
+ max  eCDF diff.....      0.11089      0.0065359
+
+ var ratio (Tr/Co).....      1.9512      1.0405
```

```

+ T-test p-value..... 0.000000014534      0.62087
+
+
+ ***** (V4) EducationHigh School *****
+               Before Matching      After Matching
+ mean treatment.....      0.25378      0.23965
+ mean control.....      0.19437      0.2077
+ std mean diff.....      13.64      7.4828
+
+ mean raw eQQ diff..... 0.060504      0.031954
+ med  raw eQQ diff.....      0      0
+ max  raw eQQ diff.....      1      1
+
+ mean eCDF diff..... 0.029704      0.015977
+ med  eCDF diff..... 0.029704      0.015977
+ max  eCDF diff..... 0.059408      0.031954
+
+ var ratio (Tr/Co).....      1.2098      1.1073
+ T-test p-value..... 0.009248      0.044937
+
+
+ ***** (V5) EducationSome College *****
+               Before Matching      After Matching
+ mean treatment.....      0.35294      0.30356
+ mean control.....      0.32737      0.36456
+ std mean diff.....      5.3473      -13.262
+
+ mean raw eQQ diff..... 0.026891      0.061002
+ med  raw eQQ diff.....      0      0
+ max  raw eQQ diff.....      1      1
+
+ mean eCDF diff..... 0.012788      0.030501
+ med  eCDF diff..... 0.012788      0.030501
+ max  eCDF diff..... 0.025575      0.061002
+
+ var ratio (Tr/Co).....      1.0375      0.91261
+ T-test p-value..... 0.32201      0.00084705
+
+
+ ***** (V6) EducationCollege Grad *****
+               Before Matching      After Matching
+ mean treatment.....      0.13613      0.24982
+ mean control.....      0.31202      0.22948
+ std mean diff.....     -51.246      4.6954
+
+ mean raw eQQ diff..... 0.17479      0.020334
+ med  raw eQQ diff.....      0      0
+ max  raw eQQ diff.....      1      1
+
+ mean eCDF diff..... 0.087943      0.010167
+ med  eCDF diff..... 0.087943      0.010167
+ max  eCDF diff..... 0.17589      0.020334
+
+ var ratio (Tr/Co).....      0.54806      1.0599
+ T-test p-value..... 1.3323e-15      0.21318
+
+
+ ***** (V7) MaritalStatusLivePartner *****
+               Before Matching      After Matching
+ mean treatment.....      0.16134      0.088598
+ mean control.....      0.078005      0.11184

```

```

+ std mean diff.....      22.637      -8.1751
+
+ mean raw eQQ diff.....    0.084034      0.023239
+ med raw eQQ diff.....      0          0
+ max raw eQQ diff.....      1          1
+
+ mean eCDF diff.....       0.04167      0.011619
+ med eCDF diff.....       0.04167      0.011619
+ max eCDF diff.....       0.083339      0.023239
+
+ var ratio (Tr/Co).....     1.8822      0.81294
+ T-test p-value..... 0.0000035776      0.039567
+
+
+ ***** (V8) MaritalStatusMarried *****
+               Before Matching      After Matching
+ mean treatment.....      0.40336      0.5374
+ mean control.....      0.57928      0.49964
+ std mean diff.....     -35.831      7.5711
+
+ mean raw eQQ diff.....    0.17479      0.037763
+ med raw eQQ diff.....      0          0
+ max raw eQQ diff.....      1          1
+
+ mean eCDF diff.....     0.087961      0.018882
+ med eCDF diff.....     0.087961      0.018882
+ max eCDF diff.....     0.17592      0.037763
+
+ var ratio (Tr/Co).....     0.98787      0.99441
+ T-test p-value..... 7.3435e-11      0.043159
+
+
+ ***** (V9) MaritalStatusNeverMarried *****
+               Before Matching      After Matching
+ mean treatment.....      0.23866      0.20479
+ mean control.....      0.13811      0.18083
+ std mean diff.....      23.569      5.9364
+
+ mean raw eQQ diff.....    0.10084      0.023965
+ med raw eQQ diff.....      0          0
+ max raw eQQ diff.....      1          1
+
+ mean eCDF diff.....     0.050274      0.011983
+ med eCDF diff.....     0.050274      0.011983
+ max eCDF diff.....     0.10055      0.023965
+
+ var ratio (Tr/Co).....     1.5271      1.0994
+ T-test p-value..... 0.0000029703      0.10766
+
+
+ ***** (V10) MaritalStatusSeparated *****
+               Before Matching      After Matching
+ mean treatment.....      0.023529      0.0050835
+ mean control.....      0.0076726      0.019608
+ std mean diff.....      10.452      -20.416
+
+ mean raw eQQ diff.....    0.016807      0.014524
+ med raw eQQ diff.....      0          0
+ max raw eQQ diff.....      1          1
+
+ mean eCDF diff.....     0.0079284      0.0072622

```

```

+ med eCDF diff..... 0.0079284    0.0072622
+ max eCDF diff..... 0.015857    0.014524
+
+ var ratio (Tr/Co)..... 3.0189    0.2631
+ T-test p-value..... 0.022929    0.00058884
+
+
+ ***** (V11) MaritalStatusWidowed *****
+               Before Matching    After Matching
+ mean treatment..... 0.043697    0.072622
+ mean control..... 0.088235    0.078431
+ std mean diff..... -21.769    -2.2379
+
+ mean raw eQQ diff..... 0.043697    0.0058097
+ med raw eQQ diff..... 0    0
+ max raw eQQ diff..... 1    1
+
+ mean eCDF diff..... 0.022269    0.0029049
+ med eCDF diff..... 0.022269    0.0029049
+ max eCDF diff..... 0.044538    0.0058097
+
+ var ratio (Tr/Co)..... 0.51964    0.93176
+ T-test p-value..... 0.00073816    0.56972
+
+
+ ***** (V12) Poverty *****
+               Before Matching    After Matching
+ mean treatment..... 2.3776    2.8431
+ mean control..... 3.1091    2.7958
+ std mean diff..... -46.293    2.8585
+
+ mean raw eQQ diff..... 0.728    0.065243
+ med raw eQQ diff..... 0.7    0.03
+ max raw eQQ diff..... 1.65    0.46
+
+ mean eCDF diff..... 0.13667    0.012574
+ med eCDF diff..... 0.15502    0.010893
+ max eCDF diff..... 0.20409    0.041394
+
+ var ratio (Tr/Co)..... 0.91853    1.0081
+ T-test p-value..... 2.2204e-16    0.45178
+ KS Bootstrap p-value.. < 2.22e-16    0.174
+ KS Naive p-value..... 1.1907e-12    0.18878
+ KS Statistic..... 0.20409    0.041394
+
+
+ Before Matching Minimum p.value: < 2.22e-16
+ Variable Name(s): Age Poverty Number(s): 2 12
+
+ After Matching Minimum p.value: 0.00058884
+ Variable Name(s): MaritalStatusSeparated Number(s): 10

```

# Weighted sample summary:

```
print(tabIPW, smd = TRUE)
```

|                     | Stratified by SmokeNow |               |        |
|---------------------|------------------------|---------------|--------|
|                     | 0                      | 1             | SMD    |
| n                   | 1379.0                 | 1379.7        |        |
| Gender = male (%)   | 797.4 (57.8)           | 781.8 (56.7)  | 0.023  |
| Age (mean (SD))     | 50.06 (17.16)          | 49.83 (15.42) | 0.014  |
| Race3 (%)           |                        |               | 0.052  |
| Asian               | 40.4 ( 2.9)            | 37.9 ( 2.7)   |        |
| Black               | 119.0 ( 8.6)           | 103.3 ( 7.5)  |        |
| Hispanic            | 75.5 ( 5.5)            | 70.6 ( 5.1)   |        |
| Mexican             | 81.2 ( 5.9)            | 87.9 ( 6.4)   |        |
| White               | 1019.8 (74.0)          | 1038.8 (75.3) |        |
| Other               | 43.0 ( 3.1)            | 41.2 ( 3.0)   |        |
| Education (%)       |                        |               | 0.029  |
| 8th Grade           | 89.4 ( 6.5)            | 96.2 ( 7.0)   |        |
| 9 - 11th Grade      | 184.6 (13.4)           | 185.8 (13.5)  |        |
| High School         | 293.9 (21.3)           | 285.5 (20.7)  |        |
| Some College        | 482.7 (35.0)           | 474.2 (34.4)  |        |
| College Grad        | 328.4 (23.8)           | 338.1 (24.5)  |        |
| MaritalStatus (%)   |                        |               | 0.023  |
| Divorced            | 161.0 (11.7)           | 165.1 (12.0)  |        |
| LivePartner         | 154.6 (11.2)           | 150.8 (10.9)  |        |
| Married             | 695.0 (50.4)           | 690.1 (50.0)  |        |
| NeverMarried        | 249.9 (18.1)           | 254.0 (18.4)  |        |
| Separated           | 23.1 ( 1.7)            | 20.5 ( 1.5)   |        |
| Widowed             | 95.4 ( 6.9)            | 99.2 ( 7.2)   |        |
| Poverty (mean (SD)) | 2.80 (1.67)            | 2.80 (1.63)   | <0.001 |