

废除死刑对凶杀案数量的影响——以美国各州为对象

摘要 废除死刑会不会影响凶杀案的数量，可能是很多人较为关注的问题。本文从 2014 年和 2010 年美国各州的经济、人口、教育数据入手，探究了样本中废死对凶杀案数量的平均因果作用。

1. 研究背景描述

死刑是否应该废除这个问题，一直是法律界乃至哲学界的热门讨论话题。无论是支持废死者还是反对废死者，都能从人权、国家权力、死刑的教育效果等等角度给出有利于自己观点的证据。在这些角度中，死刑对犯罪率的实际影响是重要的一环。因此，本文希望讨论的问题是：废除死刑是否会导致凶杀案数量的变化？如果会，变化趋势如何？死刑在美国更是一个充满争议的问题：美国联邦法院在上世纪六十年代一度废除了死刑，而后在七十年代又恢复了死刑。并且，美国各州州法院的死刑存废情况也各不相同，这为本文提供了很好的研究样本——美国各州的政治经济文化面貌各有不同，而又不至于差异过大，同时州之间数据的统计口径相同（而不同国家的数据统计口径可能不同），为研究上述问题提供了便利。

2. 数据介绍与预处理

2.1 原始数据集介绍

本文使用的数据集有两部分来源：其一，Kaggle 平台的“Homicide Reports, 1980–2014”数据集，提供了 1980 年–2014 年美国境内所有凶杀案的详细信息（包括受理单位、发生地点（城市和州）、发生时间、受害者及嫌疑人的数量、性别、人种等个人信息，以及案件有关的其他信息）；其二，美国人口普查局网站 www.census.gov 提供的“American Community Survey (ACS)”数据，该调查是为了弥补十年普查的限制，以抽样调查的方式每年提供美国社会人口、经济、教育等数据的估计。这种调查估计有 1 年、3 年、5 年三种，时间跨度越长，估计误差越小，但更难体现每年的数据差异，且并非所有数据都有 5 年或 3 年调查。由于本文对这方面的数据精度要求不高，因此采取了 1 年估计。另外，各州废死的时间难以通过官方文件获得，本文参考了中文维基百科的词条“美国死刑制度”下的介绍。其他详细信息见附录。

2.2 数据预处理

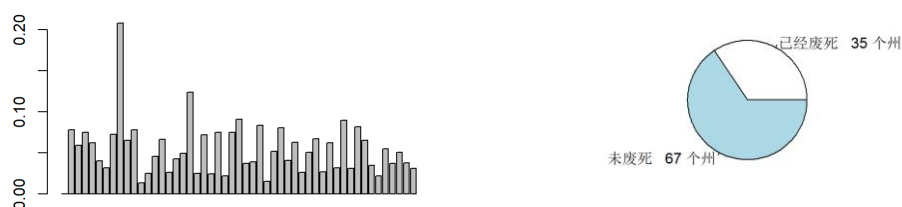
由于我们关心的数据中，美国人口普查局网站只提供 2010 年后的数据，与凶杀案数据集有交集的时间只有 2010–2014 年，并且在这段时间中，仍陆续有一些州废死，而相邻两年之间的凶杀案数量差异更可能是其他原因导致而非废死导致。又考虑到每年凶杀案的数量会受当年的经济影响，而不同地方的凶杀案数量差异更是会随着经济、教育程度等的差异而有不同，同时，美国各州的人种比例也有很大不同，这也可能成为一个影响因素。综合上述分析，本文决定选择 2010 年、2014 年各州的数据为研究样本，从原始数据集中统计、筛选出如下变量，涵盖了教育、人种、经济情况（括号内为简写）：

- 州名 (SN)：美国 50 个州加上哥伦比亚特区（波多黎各地区由于部分数据缺失被剔除），同时以年份来区分（即“Alaska14”代表 2014 年的阿拉斯加州），因此总数据量是两倍的州数量，即 102 个数据。
- 18–24 岁人口中高中毕业及以上的占比 (HS0H1824)
- 25 岁及以上人口中，高中毕业及以上的占比 (HS0H025)
- 25 岁及以上人口中，学士学位及以上的占比 (BOH025)

- 白人比例 (WHITE)
- 黑人或非裔比例 (BLACK)
- 16 岁及以上劳动力人口中，失业比例 (UE016)
- 当年收入少于 1 万美金 (去除通胀后) 的家庭比例 (LT1WH)
- 平均家庭年收入 (MIH): 以 10 万美金为单位
- 没有医疗保险的人口占比 (NHIP)
- 贫困线以下家庭占比 (PF)
- 贫困线以下人口占比 (PP)
- 当年凶杀案数量 (CIPERTHOUSAND): 以每千个成年人凶杀案数量为单位，剔除了可以确定的过失杀人案。
- 是否废除死刑 (ABOLISHDP): 二元变量，1 表示当年已废除死刑。

2.3 数据简单分析

以每千成人凶杀案数量 CIPERTHOUSAND 为纵轴，可以看到大部分州的值都在 0.1 以下，下左图中条形最高的是 2010 年的哥伦比亚特区。

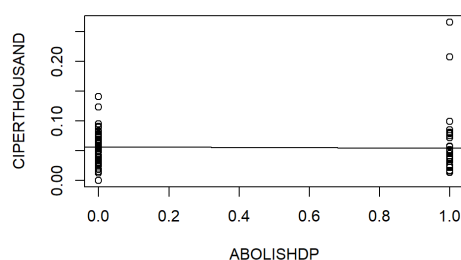


废死州和未废死州的比例和数量如右上，可以看到未废死州约为废死州数量的两倍：

3. 建模分析

3.1 建模前的简单线性回归

直接以每千成年人凶杀案数 CIPERTHOUSAND 为因变量，以是否废死 ABOLISHDP 为自变量线性回归，得到结果如下：

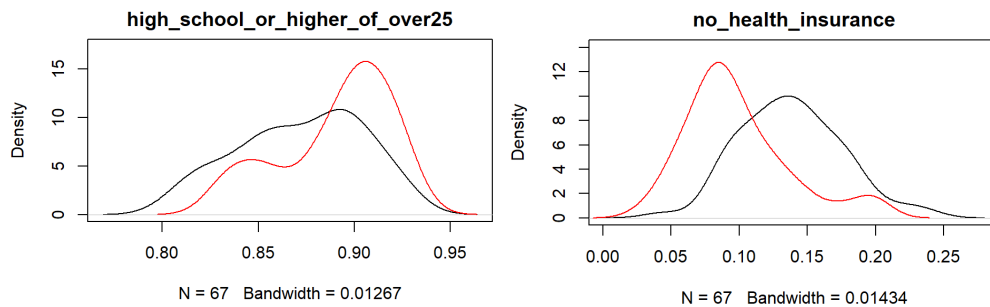


	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.056992	0.004464	12.768	<2e-16 ***
ABOLISHDP	-0.002105	0.007620	-0.276	0.783

线性回归的结果说明废死几乎没有影响，但由于线性回归与因果推断之间的不同，我们

无法在此就下定论。

计算废死州和未废死州中协变量的马氏距离，结果为 2.286972，说明两组是不平衡的。进一步画出两组各协变量的密度，红线代表废死州，选取两个代表性结果如下：



可以看到废死州的教育、经济情况要好于未废死州。

3.2 建模

3.2.1 假设分析

本文数据为观察性数据，选用 Rubin 框架进行研究。定义 Unit 为 2010 年或 2014 年的各州，Action 为废除死刑（因此处理组为废死州，控制组为未废死州），潜在结果为当年凶杀案数量（单位为每千成年人），协变量为数据集中除了潜在结果、Action 和 Unit 的其他变量，包括教育、经济、人种三个方面。平均因果作用定义为两组间凶杀案数量的平均差异。

对于 SUTVA 假设，该问题实际上不严格符合 SUTVA 假设，但可以提出一些假设使其符合 SUTVA 假设：

1. No Interference。不同地区之间的凶杀案数量的干涉，认为是存在但影响微小的。
2. No Hidden Variations of Treatments。Actions 中，“废除死刑”可能不太明确，因为例如美国有些州存在死刑，但却是长期“备而不用”的。因此这里定义为法律上明文废除死刑的情况。

对于分配机制的假设：

1. 个体化假设可能不符合，因为有可能其他地区的废死效果影响了该地区是否废死。所以这里假设各地区中废死与否是独立的。
2. 概率性分配是符合的，因为废死和不废死的州都有很多。
3. 无混杂假设可能不符合，因为废死的效果可能已经影响到了该地区是否废死。这里假设在所给的协变量条件下，废死的效果和废死之间是独立的。

在上述假设下，可以用 Rubin 框架进行观察性数据的研究。

3.2.2 平衡倾向得分

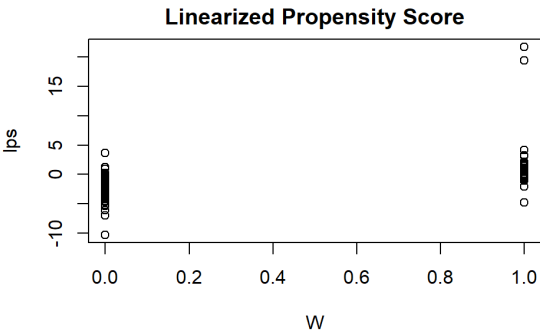
首先平衡是否废死 ABOLISHDP 的分配，选择 NHIP、HSOH025、UE016、PP、BLACK 为初始的重要变量，其中贫困个人占比 PP、黑人占比 BLACK 在潜在结果与所有协变量的线性回归中 p 值最小（分别为 0.00732 和 1.96e-06），因此可以视为影响潜在结果的重要变量，而 25 岁以上高中毕业以上占比 HSOH025 反映的教育情况很可能会影响废死的决策，另外 NHIP、UE016 代表的医疗资源、失业率也很可能影响潜在结果。

使用逻辑回归进行前向选择，直至 LRT 检验的结果小于 1 结束。对于一次项变量的选择过程和结果如下：

	Step1	Step2	Step3	Step4	Step5
HSOH1824	5.2	4.9	4.9	—	—
HSOH025	—	—	—	—	—

BOH025	0.8	0.1	1.4	6.0	—
WHITE	5.3	—	—	—	—
BLACK	—	—	—	—	—
UE016	—	—	—	—	—
LT1WH	1.2	7.0	—	—	—
MIH	4.3	0.3	0.1	0.3	0.5
NHIP	—	—	—	—	—
PF	1.2	0.3	0.7	0.9	0.6
PP	—	—	—	—	—

结果为新加入 4 个变量。本文还使用这 9 个变量的二次项进行逻辑回归前向选择，但这导致了拟合结果中 trimming 时剔除了过多样本，因此平衡倾向得分就以一次项逻辑回归的结果作为最终结果。倾向得分的平衡结果如下：



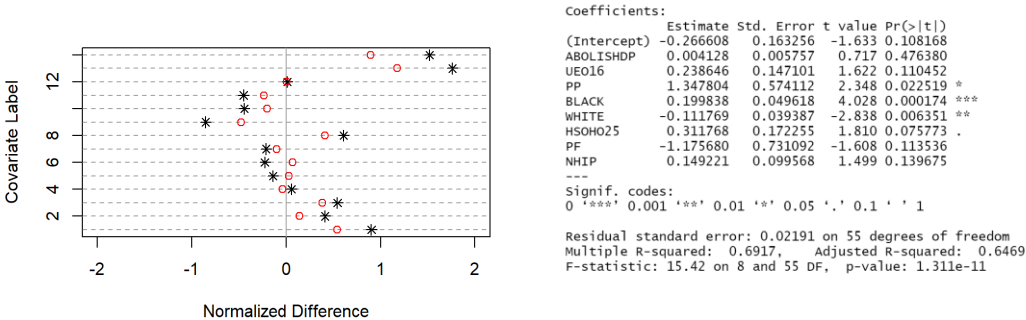
3.2.3 修整和匹配

根据上述得到的倾向评分计算模型，先进行修整（trimming），取两列公共区间 $[-4.732112, 3.685741]$ 中的行，得到未废死州 61 个，废死州 32 个，占总体的绝大部分。

3.2.3.1 无放回匹配

以倾向得分为标准，对修整得到的 32 个废死州在控制组进行一对一无放回匹配，结果得到 32 个与之匹配的未废死州。从左下图可以看出，对于 11 个协变量，匹配后（红色圆圈）相比于匹配前（黑色星号）两组的差异减小。

首先计算未加上协变量时的平均因果作用 ACE，使用 32 个废死州潜在结果的平均值减去 32 个匹配的未废死州潜在结果的平均值，结果为 0.0008939121，方差为 8.497776×10^{-5} 。以 AIC 为度量，对潜在结果 CIPERTHOUSAND 和其余变量逐步线性回归，初始变量除了必要的分配 ABOLISHDP 外，加入了 UE016 和 PP 两个变量。



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.26608	0.163256	-1.633	0.108168
ABOLISHDP	0.004128	0.005757	0.717	0.476380
UE016	0.238646	0.147101	1.622	0.110452
PP	1.347804	0.574112	2.348	0.022519 *
BLACK	0.199838	0.049618	4.028	0.000174 ***
WHITE	-0.111769	0.039387	-2.838	0.006351 **
HSOH025	0.311768	0.172255	1.810	0.075773 .
PF	-1.175680	0.731092	-1.608	0.113536
NHIP	0.149221	0.099568	1.499	0.139675

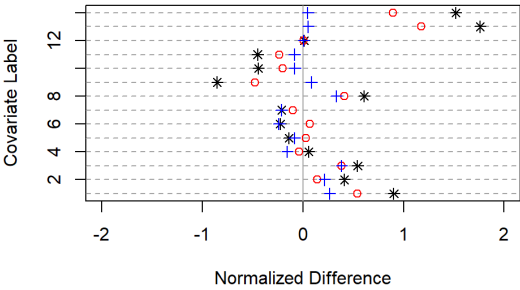
Signif. codes:
 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02191 on 55 degrees of freedom
 Multiple R-squared: 0.6917, Adjusted R-squared: 0.6469
 F-statistic: 15.42 on 8 and 55 DF, p-value: 1.311e-11

逐步回归的结果见右上，对上述 ACE 进行协变量的修整，即减去协变量差异乘以对应回归系数，得到修正后的结果为 0.004128，方差为 0.3643659。

3.2.3.2 有放回匹配

以倾向得分为标准，对修整得到的 32 个废死州在控制组进行一对一有放回匹配，结果得到 14 个未废死州，其中 2010 年的宾夕法尼亚州匹配次数最多（11 次）。下图是协变量的平衡结果，蓝色加号是有放回匹配的结果，可以看到相比于无放回匹配对协变量的平衡更好。



仍旧计算有放回匹配的 ACE，结果为-0.006522716，方差为 0.0002269029。以 AIC 为度量仍旧进行逐步回归，共选出 ABOLISHDP、UE016、PP、BLACK、NHIP、HSOH025、BOH025、HSOH1824 等协变量，用该回归模型修正上述 ACE，结果为-0.008561，方差为 0.03665166。

4. 总结

4.1 研究问题的解释

本文通过 Rubin 框架对“废除死刑对凶杀案数量的因果作用大小”进行度量，结果为：

	无放回匹配（无协变量修正）	无放回匹配（协变量修正）	有放回匹配（无协变量修正）	有放回匹配（协变量修正）
平均因果作用	0.0008939121	0.004128	-0.006522716	-0.008561
方差	8.497776e-05	0.3643659	0.0002269029	0.03665166

结论如下：由于有放回匹配和无放回匹配的度量结果中，平均因果作用都非常接近于 0，对于一个百万人口的州，凶杀案数量差异在个位数之内，且得到的方差相比于均值的绝对值都较大，本文认为，废除死刑对凶杀案数量的因果作用可以忽略不计，也就是说，在所选取的样本内，废除死刑不会导致凶杀案数量的变化。

4.2 问题与可能的改进方向

- 本文的结论无法对废除死刑或保留死刑提供任何（即使是极小的）论据支撑，因为废死还包括经济废死，这方面本文并没有研究，且本文仅仅研究了凶杀案数量变化这众多结果中的一个，得出的结论也只是样本中的因果作用。
- 无混杂假设可能是本研究中最弱的一环，在所给协变量条件下，其实仍无法认为凶杀案数量和废死的决策能够独立。
- 本文的样本量过小，且只集中在美国各州，无论是研究的广度和深度都受限。这是由于本文作者无法获得 2010 年以前各州的经济人口教育数据。如果有更多时间，可以考虑搜集世界各国或上世纪及本世纪初的美国各州数据。
- 废死州的数据来自中文维基百科，可能不准确，如果有更多时间应该寻找权威数据。
- 数据的预处理可能存在问题，因为废死的影响可能是长远的、厚积薄发的，单独一年的数据（且未剔除新近废死州）很可能误差过大。如果能得到时间跨度更大的数据，可以考虑改进数据预处理为多年的均值。

5 附录

几个原始数据集的链接：

<https://www.kaggle.com/datasets/murderaccountability/homicide-reports>

<https://zh.m.wikipedia.org/zh-hans/%E7%BE%8E%E5%9B%BD%E6%AD%BB%E5%88%91%E5%88%B6%E5%BA%A6>

以及 data.census.gov。原始数据集以及处理后的数据集已打包在./data 文件夹下，由于页数限制，预处理代码不在此展示，而是打包在./data 文件夹下。

在数据预处理部分使用 python 语言，其他使用 R 语言。代码如下：

#R 语言部分

```
data <- read.csv("./data/mydata.csv")
head(data)
attach(data)
barplot(data[1:51,]$CIPERTHOUSAND)
data$SN[which.max(data$CIPERTHOUSAND)]
lbs1 = c("已经废死", "未废死")
co = c(sum(data$ABOLISHDP[data$ABOLISHDP==1]), 102 -
sum(data$ABOLISHDP[data$ABOLISHDP==1]))
lbs2 = paste(lbs1, " ", co, "个州")
pie(c(sum(data$ABOLISHDP[data$ABOLISHDP==1]), 102 -
sum(data$ABOLISHDP[data$ABOLISHDP==1])), labels=lbs2)

fit1 <- lm(CIPERTHOUSAND~ABOLISHDP, data=data)
summary(fit1)
plot(ABOLISHDP, CIPERTHOUSAND)
abline(fit1)

x0 <- as.matrix(data[data$ABOLISHDP==0, 3:13])
x1 <- as.matrix(data[data$ABOLISHDP==1, 3:13])
mahalanobis(colMeans(x1), colMeans(x0), (var(x0)+var(x1))/2)

plot_difference = function(i, name) {
  ymax <- max(c(density(x0[,i])$y, density(x1[,i])$y))
  plot(density(x0[,i]), ylim = c(0, 1.1*ymax), main=name)
  lines(density(x1[,i]), col="red")
}
plot_difference(2, "high_school_or_higher_of_over25")
plot_difference(4, "white")
plot_difference(9, "no_health_insurance")
fit2 <- lm(CIPERTHOUSAND~.-X-SN-CIPERTHOUSAND, data=data)
summary(fit2)
md1 <- glm(ABOLISHDP~NHIP+HSOH025+UE016+PP+BLACK, data = data, family =
binomial)
summary(md1)
```

```

fw_LRT <-function(object, scope, C) {
  step <- 0
  cov_name <- names(coef(scope))[-1]
  tab <- matrix(NA, nrow=length(cov_name),ncol=length(cov_name)-
length(coef(object))+1)
  rownames(tab) <- cov_name
  repeat {
    if (setequal(names(coef(object)),names(coef(scope)))) {
      cat("\nWARNING: Maximun Model Reached.\n")
      break
    }
    step <- step+1
    cat("\n>>> Step =", step, "\n")
    res <-add1(object, scope = scope, test="LRT")
    print(res)
    tab[rownames(res)[-1], step] <- res$LRT[-1]
    add_id <-which.max(res$LRT)
    if(res$LRT[add_id]<C)
      break
    object <- update(object,as.formula(paste("~ .
+",rownames(res)[add_id])))
  }
  cat("\n>>> Final Model\n")
  print(object)
  return(list(md = object, tab = tab[, 1:step]))
}

mdL <- glm(ABOLISHDP~.-X-SN-CIPERTHOUSAND, data = data, family =
binomial)
res <-fw_LRT(md1, scope = mdL, C = 1)
summary(res$md)
print(round(res$tab, 1), na.print = "-")
md2 <- res$md
summary(md2)
plot(data$ABOLISHDP, predict(md2, type = "link"), xlab = "W", ylab =
"lps",main = "Linearized Propensity Score")
data$le = md2$linear.predictor
lbound = min(data$le[data$ABOLISHDP==1])
ubound = max(data$le[data$ABOLISHDP==0])
print(lbound)
print(ubound)
data.trim = data[data$le >= lbound & data$le <= ubound, ]
number = nrow(data.trim)
number.t = sum(data.trim$ABOLISHDP)
number.c = number - number.t

```



```

tstat = (mean(data.trim$le[data.trim$ABOLISHDP==1]) -
mean(data.trim$le[data.trim$ABOLISHDP==0]))/sqrt((var(data.trim$le[data
.trim$ABOLISHDP==1])*(number.t - 1)
+ var(data.trim$le[data.trim$ABOLISHDP==0])*(number.c - 1))/(number -
2)*(1/number.c + 1/number.t))
data2 = data[data$le >= lbound & data$le <= ubound, ]
library(Matching)
match_ps <- Match(Y = data2$CIPERTHOUSAND, Tr = data2$ABOLISHDP, X =
data2$le, M = 1, replace = FALSE, distance.tolerance = 0)
summary(match_ps)
MatchBalance(data2$ABOLISHDP ~ ., data = data2[,3:13], match.out =
match_ps)
get_table <- function(X, W, lps, basic = TRUE, alpha = 0.05,
left.include = FALSE) {
  K <- ncol(X)
  ps <- exp(lps)/(1 + exp(lps)) # propensity score in [0, 1]
  ind <- list("0" = which(W == 0), "1" = which(W == 1))
  X_aug <- cbind(X, NA, ps, lps)
  colnames(X_aug)[K + (1:3)] <- c("Multivar_measure", "pscore",
"linear_pscore")
  tab <- t(apply(X_aug, 2, function(x) # basic summary statistics
unlist(tapply(x, W, function(z) c(mean(z), sd(z))))))
  mu_dif <- tab[, 3] - tab[, 1] # mean(Xt) - mean(Xc)
  NorDif <- mu_dif/sqrt((tab[, 2]^2 + tab[, 4]^2)/2)
  LR_STD <- log(tab[, 4]) - log(tab[, 2])
  get_pi <- function(x, grp) { # measure M3 & M4
    if (anyNA(x))
      return(NA)
    else {
      Fun <- function(xx, boundary = TRUE) {
        if (boundary)
          return(mean(x[ind[[grp + 1]]] <= xx))
        else
          return(mean(x[ind[[grp + 1]]] < xx))
      }
      qtl <- quantile(x[ind[[2 - grp]]], c(1 - alpha/2, alpha/2), type = 1)
      return(1 - Fun(qtl[1]) + Fun(qtl[2], boundary = left.include))
    }
  }
  pi_c <- apply(X_aug, 2, get_pi, grp = 0)
  pi_t <- apply(X_aug, 2, get_pi, grp = 1)
  # Mahalanobis distance
  Sigma_ct <- (cov(X[ind$"0", ]) + cov(X[ind$"1", ]))/2
  mv_ct <- sqrt(t(mu_dif[1:K]) %*% (Sigma_ct) %*% mu_dif[1:K])

```



```

res <- cbind(NorDif, LR_STD, pi_c, pi_t) # merge tables
if (basic) {
  colnames(tab) <- c("Mean_c", "(S.D.)_c", "Mean_t", "(S.D.)_t")
  res <- cbind(tab, res)
}
res[K + 1, ncol(res) - 3] <- mv_ct
return(res)
}
tab <- get_table(data2[, 3:13], data2$ABOLISHDP, data2$le)
dat_ps <- data2[c(match_ps$index.treated, match_ps$index.control), ]
tab_ps <- get_table(dat_ps[, 3:13], dat_ps$ABOLISHDP,
data2$le[c(match_ps$index.treated, match_ps$index.control)], basic =
FALSE)
tab_ps
no_Covariate_tau = mean(data2[match_ps$index.treated,]$CIPERTHOUSAND) -
mean(data2[match_ps$index.control,]$CIPERTHOUSAND)
no_Covariate_tau
no_Covariate_var = 2 * (var(data2[cbind(match_ps$index.treated,
match_ps$index.control),]$CIPERTHOUSAND)) /
(nrow(data2[match_ps$index.treated,]))
no_Covariate_var
library(MASS)
match_data = data2[cbind(match_ps$index.treated,
match_ps$index.control), 3:15]
lmo<-lm(CIPERTHOUSAND~ABOLISHDP+UE016+PP, data=match_data)
lmfor<-
step(lmo, scope=list(upper=~ABOLISHDP+WHITE+BLACK+PP+UE016+HSOH025+NHIP+
LT1WH+HSOH1824+BOH025+PF+MIH, lower=~ABOLISHDP+UE016+PP), direction="both
")
summary(lmfor)
stj1 = c(0.147101, 0.574112, 0.049618, 0.039387, 0.172255, 0.731092,
0.099568)
Covariate_tau = 0.004128
Covariate_tau
Covariate_var = sqrt(sum(stj1^2)/7)
Covariate_var
match_ps2 <- Match(Y = data2$CIPERTHOUSAND, Tr = data2$ABOLISHDP, X =
data2$le, M = 1, replace = TRUE, distance.tolerance = 0)
tab2 <- get_table(data2[, 3:13], data2$ABOLISHDP, data2$le)
dat_ps2 <- data2[c(match_ps2$index.treated, match_ps2$index.control), ]
tab_ps2 <- get_table(dat_ps2[, 3:13], dat_ps2$ABOLISHDP,
data2$le[c(match_ps2$index.treated, match_ps2$index.control)], basic =
FALSE)
tab_ps2

```

```

match_ps2$index.treated
match_ps2$index.control
K <- ncol(data2) - 2
par(mfrow = c(1, 1))
plot(NULL, xlim = c(-2, 2), ylim = c(1, K), xlab = "Normalized
Difference",
ylab = "Covariate Label")
abline(h = 1:K, lty = 2, col = "gray65")
abline(v = 0, col = "gray65")
points(tab2[1:K, "NorDif"], 1:K, pch = 8)
points(tab_ps[1:K, "NorDif"], 1:K, pch = 1, col = "red")
points(tab_ps2[1:K, "NorDif"], 1:K, pch = 3, col = "blue")
no_Covariate_tau2 = mean(data2[match_ps2$index.treated,]$CIPERTHOUSAND)
- mean(data2[match_ps2$index.control,]$CIPERTHOUSAND)
no_Covariate_tau2
all_var = var(data2[cbind(match_ps2$index.treated,
match_ps2$index.control),]$CIPERTHOUSAND)
indicator<-duplicated(match_ps2$index.control)
data2[match_ps2$index.control,]$SN
table(match_ps2$index.control[indicator])+1
no_duplicated = 32 + 32 - 23
no_Covariate_var2 = all_var * (no_duplicated
+2^2+2^2+4^2+2^2+2^2+11^2)/ (32^2)
no_Covariate_var2
match_data2 = data2[cbind(match_ps2$index.treated,
match_ps2$index.control),3:15]
lmo2<-lm(CIPERTHOUSAND~ABOLISHDP+UE016+PP,data=match_data2)
lmfor2<-
step(lmo2,scope=list(upper=~ABOLISHDP+WHITE+BLACK+PP+UE016+HSOH025+NHIP
+LT1WH+HSOH1824+BOHO25+PF+MIH,lower=~ABOLISHDP+UE016+PP),direction="bot
h")
summary(lmfor2)
lmfor2$coefficients[3:9]
stj2 = c(0.108004, 0.106363,
0.026918,0.097420,0.109739,0.049959,0.134507)
D_bar2 = colMeans(data2[match_ps2$index.treated,c(8,13,7,11,4,5,3)]) -
colMeans(data2[match_ps2$index.control,c(8,13,7,11,4,5,3)])
D_bar2
stj2
Covariate_tau2 = no_Covariate_tau2 - D_bar2 %*% c(0.2614760, 0.5910031,
0.2978525, 0.4287860, 0.3776103, 0.1150685, 0.1995873 )
Covariate_tau2
Covariate_var2 = sqrt(sum(stj2^2))/7
Covariate_var2

```