

Homework 9

Isaiah Benny

EID: ieb357

2024-04-18

Github

Problem 1

Part A

The proportion of those receiving a GOTV call who voted in 1998 is 0.648. The proportion of those not receiving a GOTV call who voted in 1998 is 0.444.

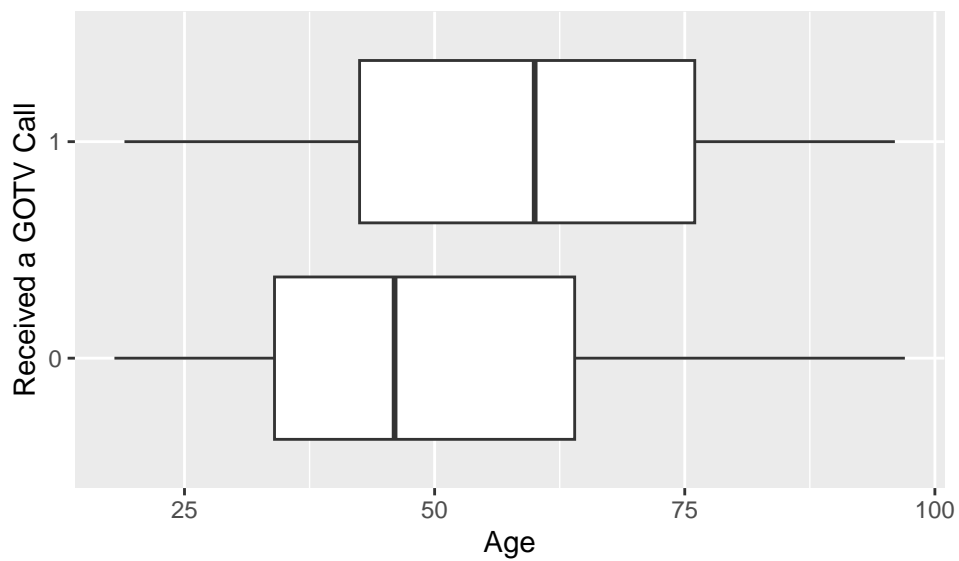
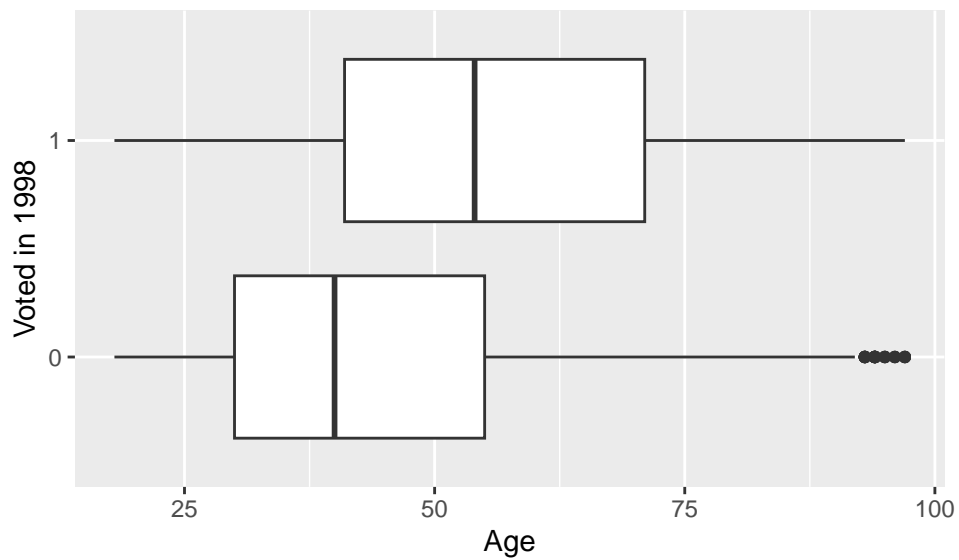
The confidence interval for the difference in these proportions of voting in 1998 (GOTV call - no GOTV call) is $[0.141, 0.266]$

Part B

Voted in 1996	Proportion Who Voted in 1998
0	0.229
1	0.640

Voted in 1996	Proportion Who Received a Call
0	0.014
1	0.030

These tables above show that an individual who voted in 1996 is more likely to vote in 1998 and is more likely to have received a GOTV call. Thus, since voting in 1996 is correlated with the treatment and the outcome, it is a confounder.



These boxplots show that those who voted in 1998 tend to be older than those who didn't, and that those who received a GOTV call tend to be older than those who didn't. Therefore, age is a confounding variable, since it is correlated with both voting in 1998 and receiving a GOTV call.

MAJORPTY	voted1998	n	prop
0	0	1787	0.650
0	1	963	0.350
1	0	4181	0.518
1	1	3898	0.482

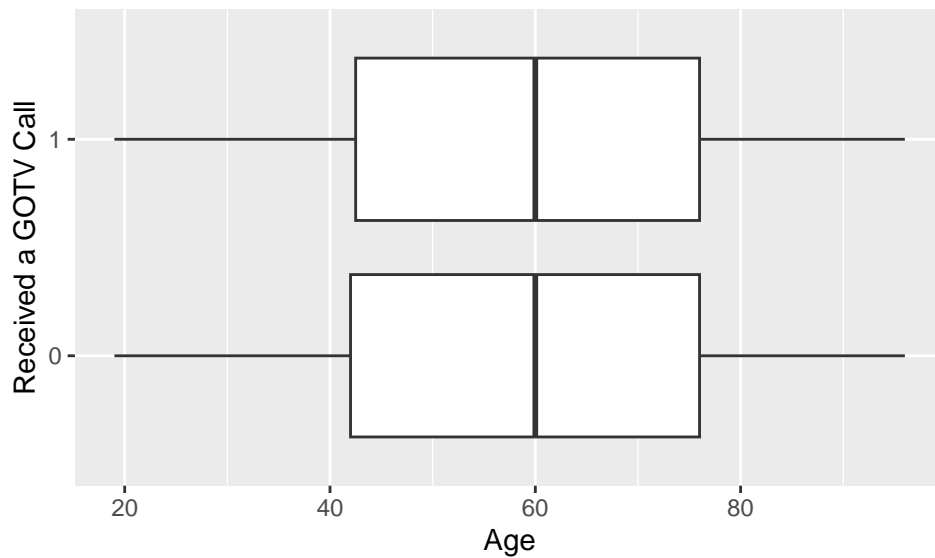
MAJORPTY	GOTV_call	n	prop
0	0	2701	0.982
0	1	49	0.018
1	0	7881	0.975
1	1	198	0.025

These tables show that voters who were a part of a major party were more likely to voted in 1998 and slightly more likely to have received a GOTV call, making MAJORPTY a confounding variable.

Part C

To show that these variables are no longer confounders, we need to show that they are not correlated with the treatment variable.

Voted in 1996	Proportion Who Received a Call
0	0.167
1	0.167



MAJORPTY	GOTV_call	n	prop
0	0	238	0.829
0	1	49	0.171
1	0	997	0.834
1	1	198	0.166

The two tables and the boxplots show that the variables that were confounders in the original dataset no longer correlate with receiving a GOTV call in the matched dataset. This is because a change in the value of each variable is associated with no change in the proportion of those who received a GOTV call.

Using the matched data, the proportion of those receiving a GOTV call who voted in 1998 is 0.648 and the proportion of those not receiving a GOTV call who voted in 1998 is 0.574.

The confidence interval for the difference in these proportions (GOTV call - no GOTV call) is $[0.006, 0.142]$

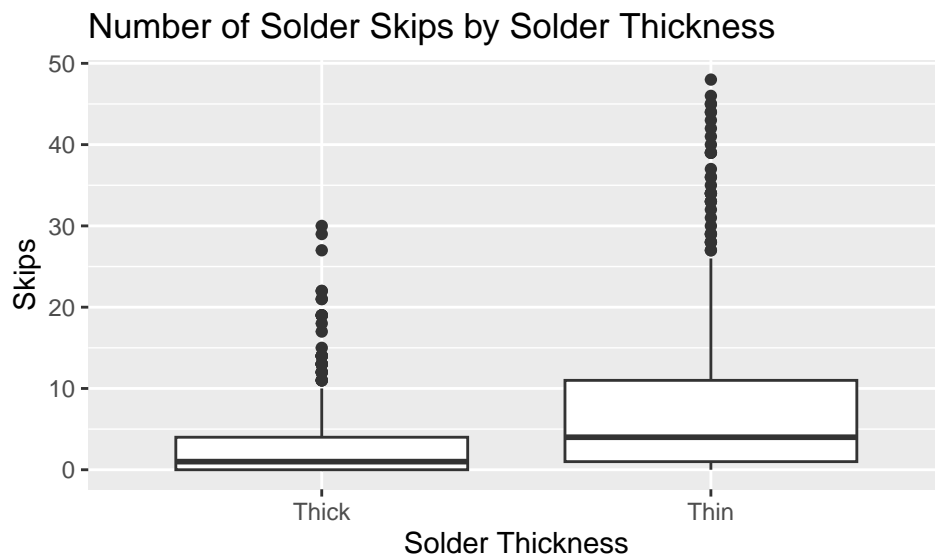
With this confidence interval, we can conclude that receiving a GOTV call likely has increased the chance of voting in 1998, since the interval only contains positive values. However, this effect is not nearly as strong as the original confidence interval implied. In fact, GOTV calls may have a negligible effect on one's likelihood of voting in 1998, since the confidence interval contains values that are close to 0.

Problem 2

Part A



This plot shows the number of solder skips for each opening size (large, medium, and small). The solder skips tend to increase as the size decreases.



This plot shows the number of solder skips for each solder thickness (thick or thin). Thinner alloys tend to have more skips.

Part B

Table 7: 95% Confidence Interval for Each Coefficient

	lower	upper	estimate
(Intercept)	-0.63	1.41	0.39
OpeningM	0.96	3.85	2.41
OpeningS	3.68	6.57	5.13
SolderThin	0.84	3.72	2.28
OpeningM:SolderThin	-2.78	1.30	-0.74
OpeningS:SolderThin	7.61	11.70	9.65

Part C

The intercept represents the baseline solder skips for solder guns with large openings and thick alloys and is 0.39 skips. The main effect of a medium sized opening is 2.41 additional skips from the baseline. The effect of a small opening in isolation is 5.13 skips. The effect of using a thin alloy for soldering in isolation is an additional 2.28 skips from the baseline. The interaction effect of a medium sized opening and thin soldering alloy is -0.74 skips; a setup that uses a medium sized opening and a thin soldering alloy has .74 skips less than the sum of the individual effects. The interaction effect between a small opening and a thin soldering alloy is 9.65 skips. That is, when both a small opening and a thin alloy are used, there is an additional 9.65 skips added to the sum of the individual effects.

Part D

The combination that I would recommend would be the baseline, which is a large opening with a thick soldering alloy. This is because the baseline leads to the minimum value in the linear model, which is 0.39. Any other combination of size openings and thickness would lead to a number of average skips that is higher than 0.39.