Causal Inference

Minerva University

CS130: Statistical Modeling - Prediction and Causal Inference

Prof. A. Diamond

April 14, 2022

Causal Inference

Question 1:

The first step is to conduct a regression analysis to understand the relationship between different predictors and the response variable (*nowtot* here). For regression specification, we only include 6 predictors, namely the party (*Dems, Repubs*), the religion (*Christian*), age, service length (*srvlng*), Democratic votes (*Demvote*) and the treatment here which is whether the participant has girl children (*hasgirls*). The regression gives us results:

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  38.6991    18.6306   2.077 0.038390 *
Dems         -8.1022    17.5861  -0.461 0.645238
Repubs      -55.1069    17.6340  -3.125 0.001901 **
Christian   -13.3961     3.7218  -3.599 0.000357 ***
age           0.1260     0.1117   1.128 0.259938
srvlng       -0.2251     0.1355  -1.662 0.097349 .
demvote      87.5501     8.4847  10.319  < 2e-16 ***
hasgirls     -0.4523     1.9036  -0.238 0.812322
```

Table 1. The Table prints out results from our linear regression analysis. We can learn more about coefficients of each predictor variable and their statistical significance

From here, we can derive that our treatment effect is - 0.4523. To extract the 95% confidence interval, we use the function confint() and get the result [- 4.194, 3.289]. To evaluate, the coefficient of *hasgirls* is not statistically significant because the p-value is quite large (0.812). It is consistent with the confidence interval when the 95% confidence interval includes 0. This means there is insufficient evidence to claim the difference between the treatment and control groups. However, we have not conducted matching techniques yet.

Therefore, this result is quite invalid because it is influenced by multiple potential confounders. We check the balance between two groups and derive that following result:

Predictor Variables	Before Matching p-value
Dems	0.35571
Repubs	0.3873
Christian	0.68107
age	0.0020402
srvlng	0.85956
demvote	0.61103

Table 2. The Table reports covariate balance through the metric of Before Matching p-value

From Table 2, the Minimum p-value is currently too low (0.002). This weakens the validity of our treatment effect derived from the regression model.

To improve our results, we will conduct Genetic Matching to improve covariate balance between the treatment and control groups. The balance will give a more accurate treatment effect. We decide to match on 6 covariates including (*Dems, Repubs, Christian, age, srvlng, demvote*). By improving covariate balance between the treatment and control groups, we expect to see higher p-value of predictors after matching, especially an improvement in the Minimum p-value after matching. We start with GenMatch() which searches for the most optimal set of weights assigned to different predictors. These weights serve as scalars to adjust the most appropriate standardization for predictors' ranges of values. As a result, the set of weights derived from this process enables us to achieve the optimal balance. We use this set of weights to input in the Match() function which helps us compute the average treatment effect and standard

error. This function also provides information about which observation gets matched with which or the number of observations dropped. Finally, to understand the level of balance achieved after matching, we use the MatchBalance() function. Notice that during our procedure, we will hide the outcome values, which means that we do not provide Y values in the function. This will help the model less subject to bias and provide a more objective evaluation. Our result for Covariate Balance after matching is reported below:

Predictor Variables	Before Matching p-value	After Matching p-value
Dems	0.35571	0.31731
Repubs	0.3873	1
Christian	0.68107	1
age	0.0020402	0.43069
srvlng	0.85956	0.42669
demvote	0.61103	0.63267

Table 3. The Table reports results Before and After Matching for covariate balance

From Table 3, we can observe that some predictors even achieve complete balance after matching. Furthermore, there is a great improvement in *age* balance, from the smallest p-value 0.002 to more than 0.4. Most importantly, the Minimum p-value After Matching is much higher than that Before Matching (0.317 vs 0.002).

We will experiment with different values of M in our GenMatch() and Match() functions to explore different results. M here refers to the number of matches that should be found for each treated unit.We only use Minimum p-value After Matching as the metric here for comparison:

	M = 1	M = 2	M = 3
After Matching Minimum p-value	0.31731	0.186	0.07

Table 4. The Table compares After Matching p-value when we experiment with different values of M. In other words, we try to understand the covariate balance achieved from different M values

We can see that as M increases, the covariate balance decreases. Therefore, we will still stick to the original value of M. We are satisfied with the covariate balance achieved from M = 1, so we will plug Y values in the matching procedure to understand the treatment effect and standard error.

After plugging Y outcomes in the Match() function, we learn that the treatment effect is 1.0737 and the standard error is 2.213. To compute the 95% confidence interval, we will use the formula $Lower\ Bound = Estimate - 1.96 * Standard\ Error$ and $Upper\ Bound = Estimate + 1.96 * Standard\ Error$. We get the 95% confidence interval after matching for treatment effect: [- 3.26369, 5.411126]. To compare with the confidence interval before matching, this one after matching has a slight move to the right, yet still involves 0. Therefore, we cannot make a sharp conclusion here yet.

We now redefine our treatment and control groups. We preprocess data by eliminating all participants who have both girls and boys. We now only view the treatment groups as having only more than 2 girls and the control group as having only more than 2 boys. As a result, in our remaining data, participants with *hasgirls* = 1 surely means that they have more than 2 girls and no boys. This new definition creates a sharper difference between two groups, signifying a more extreme treatment.

We now conduct the regression analysis and receive the following results:

```
Coefficients: (1 not defined because of singularities)
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -18.4332    16.3697  -1.126  0.26335
Dems         49.9284     4.4002  11.347  < 2e-16 ***
Repubs           NA         NA      NA       NA
Christian    -3.4303     7.7211  -0.444  0.65798
age          -0.2558     0.2180  -1.173  0.24395
srvlng        0.3908     0.2654   1.473  0.14461
demvote      86.6631    17.9813   4.820 6.32e-06 ***
hasgirls     12.2925     3.5008   3.511  0.00072 ***
```

Table 5. The Table prints out results from our regression analysis

We can clearly see the difference between two regression models. When we redefine treatment, the coefficient for *hasgirls* is much higher (12.2925) and the estimate is also much more statistically significant. The 95% confidence interval is [5.33, 19.25]. This interval is way more than 0, which shows that when the treatment is defined to be more extreme, its effect is also larger. We now conduct Genetic Matching as before to achieve covariate balance, thus getting a more credible treatment effect estimate.

Predictor Variables	Before Matching p-value	After Matching p-value
Dems	0.04806	1
Repubs	0.04806	1
Christian	0.1887	1
age	0.71354	0.60035
srvlng	0.13925	0.66963
demvote	0.23199	0.44246

Table 6. The Table reports results Before and After Matching for covariate balance between two newly defined

groups

Genetic Matching also helped this time when half of covariates achieved the maximum balance and the Minimum After Matching p-value is much higher than Before Matching (0.44246 vs 0.13925). We can also notice that the Minimum After Matching p-value for the new dataset is also higher than that for the old dataset. Therefore, the result of the treatment effect can also be more reliable. We achieve a treatment effect estimate of 11.383 and the standard error of 4.0247. We compute the 95% confidence interval: [3.49, 19.27]. This interval is much different from the one we got in the old treatment. This interval no longer contains 0. Therefore, we can understand there is some kind of causal path here between having more than 2 girls and no boys and greater agreement with the National Organization for Women.

We also try accounting for different predictors, namely *white* and *female*. We use personal knowledge that race and gender also play a role in being liberal and open-minded. Therefore, we want to balance on these two covariates and observe the treatment effect.

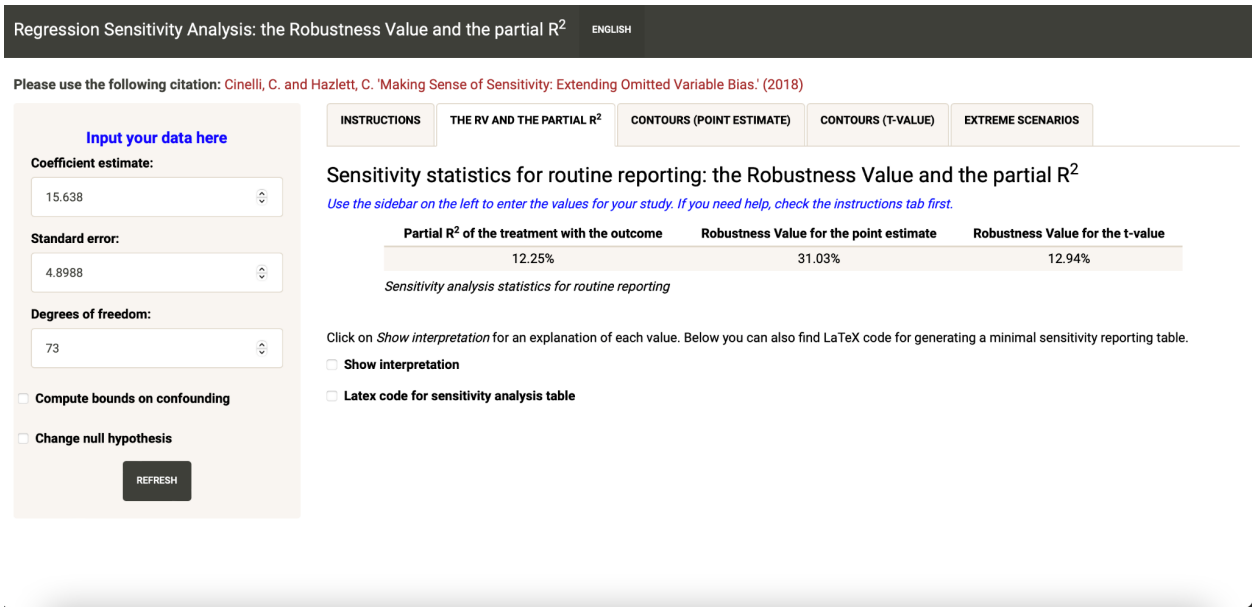
Predictor Variables	Before Matching p-value	After Matching p-value
Dems	0.04806	0.31739
Repubs	0.04806	0.31739
Christian	0.1887	0.31739
age	0.71354	0.5885
srvlng	0.13925	0.48832
demvote	0.23199	0.43356
white	0.80701	0.31739
female	0.095836	0.15515

Table 7. The Table reports results Before and After Matching for covariate balance between two newly defined groups. We also include 2 new variables which are *female* and *white*.

Genetic Matching improves the Minimum p-value from 0.04806 to 0.15515. It makes sense that when we involve more variables, it is harder to achieve better covariate balance. With 2 more covariates, the treatment effect estimate is now 15.638 and the standard error is 4.8988. We compute the 95% confidence interval: [6.036, 25.239]. When we involve 2 more covariates, the confidence interval is even farther away from 0. This really supports our belief that there is a causal relationship between the legislator's having girl children and his being more liberal. For future replication, we should involve more covariates to ensure balance between two groups.

Question 3:

Because our method still only accounts for observable covariates, it is still subject to bias and the influence of unobserved covariates can mislead results. Therefore, we need to conduct a sensitivity analysis to understand how reliable our results are currently. We use an online tool to compute different statistics with inputs as our results above.



We use results from redefined treatment and the degree of freedom here is 73 because we have only 91 observations satisfying the new requirements and we have 8 confounders. From the online tool, we derive results for our sensitivity analysis test:

Partial R^2	Robustness Value for point estimate	Robustness Value for t-value
12.25%	31.03%	12.94%

Table 8. The Table reports results derived from the Sensitivity Analysis test

This test will take into account all observed variables to evaluate the sensitivity to unobserved confounders. Partial R^2 gives us information about the proportion of variation in the outcome explained by all explanatory variables, both observed and unobserved, that cannot be explained by observed explanatory variables only. In other words, to fully account for the estimated treatment effect, all explanatory variables in the full model need to explain at least 12.25% of the variation of the treatment. When it comes to the second statistic, we will interpret that unobserved confounders need to explain more than 31.03% of the residual variance of both the outcome and the treatment to bring the treatment effect to 0. In other words, the current treatment effect estimate is much larger than 0. And to refuse that there is a treatment effect, unobserved covariates have to explain at least 31.03% of the residual variance of the outcome and the treatment. For the last statistic, if unobserved confounders can explain more than 12.94% of the residual variance of both the outcome and the treatment, they will lead the result to fall into a no longer statistically significant range, which means the treatment effect back then is no longer valid.

Appendix: The code for this entire assignment can be found here.

References

Cinelli, C. and Hazlett, C. 'Making Sense of Sensitivity: Extending Omitted Variable Bias.' (2018) https://carloscinelli.shinyapps.io/robustness_value/