Drew Levitt, Damario Abdalla, and Si Thu Aung

Concise Overview

The *potential outcomes* of the treatment and control groups were arranged to fall under the *ceteris paribus* condition, to determine the causal effect of wellness programs on healthcare spending. The treatment must be given randomly, without considering whether or not the subject of the experiment would have wanted the treatment to avoid selection bias. The study randomly assigned financial incentives to 4,834 employees, resulting in 3,300 employees in the treatment group and another randomly selected 1,534 employees in the control group. This screening was a key part of the intervention aimed at improving employee health outcomes through targeted incentives. Conclusively, none of the differences between the control group means and treatment group means were very large, and none of the p-values for any of the variables were below 0.05. This level of confidence can assure us that the sample was not biased, as it does not give a statistically significant p-value. Similarly, whether demographics are taken into consideration, the effect of workplace wellness programs is similar.

Case Description

1. In order for one to describe the causal effect of wellness programs on healthcare spending between firms, one should make sure that all conditions between the two firms fall under the *ceteris paribus* condition and not differ (systematically) from one another prior to the introduction of a treatment. For instance, both firms should have identical cultures, demographics, and other pre-conditions However, such a condition is difficult to be satisfied, as one firm may be healthier, younger, and/or more salaried than another and vice versa.

2. In order to compare the actual outcomes of employees in the Illinois study, it will be important that the *potential outcomes* of the treatment and control groups also meet the *ceteris paribus* condition. However, since one cannot observe potential outcomes lest he be omnipotent, it will be enough to satisfy this requirement that we find similarities among the qualities and characters exterior to whether or not they received the treatment.

3. In the context of the Illinois study, the treatment was being a part of the wellness program while the controlled condition was an absence of that program. This is not the same difference as between participation and non-participation, however. While having a treatment group and control group is necessary to conduct such an experiment, participation is not an acceptable condition to be had under such an experiment because introducing the option to participate introduces bias and will be detrimental to the experiment's credibility. Instead, the treatment must be given randomly and without considering whether or not the subject of the experiment would have wanted the treatment. The study mentioned in the report randomly assigned financial incentives to 4,834 employees. Of these, 3,300 employees were in a randomly assigned treatment group while the other 1,534 people were assigned to the control group.

4. The results in our table indicate that the treatment group and the control group were assigned randomly because none of the differences between the control group mean and the treatment group mean were very large, and none of the p-values for all of the variables were below 0.05. The hypothesized value lies within the 95% confidence interval, assuring us that the sample is not biased.

5. Between the column that takes into consideration demographics and the column that does not take into consideration demographics, we should expect differences since grouping participants by demographics introduces bias into the experiment. This does happen

significantly. Where means and standard errors do change between columns, averages between the two columns do not differ greatly, and with some variables, like “pos_er_crtitical”, which remains at –0.02 between the two columns, the average doesn’t change at all. What these estimates indicate about causal treatment effects is that the effect of workplace wellness programs is similar whether or not demographics are taken into consideration.

6. In an experiment, participation intrinsically introduces bias. This is because participation is a matter of preference, and preference carries bias over into the experiment through the pre-existent qualities of its subjects. However, just as with the previous table, the differences between averages without demographic control and averages with demographic control do not differ substantially. This indicates that randomization was successful and that selection bias did not happen. The reason for this is that there are no substantial differences in the estimates when we introduced demographics to the prediction model.

Conclusion

When, in the Illinois experiment, researchers compared the control group and the treatment group, there showed to be no significant differences between the two. The p-values arrived at in the experiment also showed that there were no significant differences between the two groups, either, as none of the p-values were below 0.05 and outside of the confidence interval. This signals that there was not a large causal effect in people being a part of the wellness program versus not being in it and that spending on healthcare remained roughly the same, regardless of demographics.