

The attitude of young voters on COVID-19 might be the key of election in Canada

group 80 Lan Cheng and Liyuan Cao

2020-10-07

Executive summary

This study focus on investigate the behavior of young voters in election of Canada. This study finds that young voters are would be affected seriously by event of COVID-19 in election. The proportion of voter for the Liberal Party in province ontario might be changed a lot.

Introduction

Survey methodology

The objective of the survey is to investigate the proportions of favoring Liberal in ON for 2020 election in two groups for young voters who are aged between 18 and 25 years old and once voted for Liberal before. One group is the group of the young voters who are satisfied about the work dealing with COVID-19 of the Liberal party, the other group is the group of the young voters who are not satisfied about the work dealing with COVID-19 of the Liberal party.

The target population is all of young Canadian students who are living in ON, the frame is a list of students that registered in the schools in ON, the sample is 1000 Canadian students.

The data needed to be collected in the survey is whether they will vote for Liberal or not in the 2020 election. The required accuracy of estimates is within 0.2 percent.

The sampling method selected in the survey is SRSWOR which is a method of selection of n samples out of the N population one by one, and in this sampling method, at any stage of the selection, any one of the remaining units has equal chance being selected. The details of the selection in the survey is as below:

- 1). Identify the N students in the population frame with the numbers 1 to N.
- 2). Choose any random number using R code.
- 3). Choose the sampling student whose student ID corresponds to the random number drawn from the step 2).
- 4). Removing that student from the frame list.
- 5). Repeat the previous procedures until 1000 students are obtained.

In the sampling method SRSWOR, if we denote the population characteristic of favoring Liberal in ON for 2020 election as a binary outcome Y_i , and if favoring Liberal, $Y_i = 1$, otherwise, $Y_i = 0$. Then we have the total number with the response 1 which means favoring Liberal in ON for 2020 election in the population is:

$$T = \sum_{i=1}^N Y_i$$

Where N is the population size. The proportion with the response 1 in the population is:

$$P = \frac{T}{N}$$

Now, we denote the sample characteristic of favoring Liberal in ON for 2020 election as a binary outcome y_i , and if favoring Liberal, $y_i = 1$, otherwise, $y_i = 0$, and the sample size is $n = 1000$. The proportion of sample that with the response 1 favoring Liberal in ON for 2020 election is:

$$p = \frac{\sum_{i=1}^n y_i}{n}$$

The sample variance of the proportion using the sampling method SRSWOR is:

$$var(p) = \frac{p(1-p)}{n-1} \left(1 - \frac{n}{N}\right)$$

And as N is much larger than n for this survey, the sample variance of the proportion could be approximated by:

$$var(p) = \frac{p(1-p)}{n-1}$$

The important reason of choosing the sampling method SRSWOR is that it has the most important statistical property for estimators that the sample proportion p is an unbiased estimator of the population proportion P .

This survey is going to reach the desired respondents by Mail based on the information of the students registered in the schools in ON. The cost of this survey would be about $(10 + 2) * 1000 + 500 = 12500$ dollars. For each of the sample, 10 dollars would be spent on a gift card and 2 dollars would be spent on a paid envelop. And the left 500 dollars is the fix cost of the survey design. This survey provides a gift card for each sample along with the survey which is designed to deal with the non-response problem, because the survey is short, a 10 dollar gift card should be enough for samples to feel it is worth to complete the survey. This is important because there are always issues of non-response in survey, the biggest effect of the non-response issue on the survey that it would cause non-response bias which would make the estimates not reliable. However, sometimes the non-response issue is hard to avoid, samples might not to vote or prefer not to complete the survey even after they accepted the gift card.

At last, it is very important to protect the respondent privacy in the survey, one of the important reasons is that the plotical vote is a very sensitive topic. And this survey would sign privacy contract and make a copy along with the survey before mail to the samples.

Results

Table 1 and Figure 1 show the estimated proportions of favoring Liberal in ON for 2020 election grouped by gender and whether voted for Liberal before. From the table 1 and figure 1, it can be found that for female young voters who are not ever voted for Liberal shows an estimated propotion of favoring Liberal about 0.350 with a $SD = 0.0443$, it is close to the male young voters who are not ever voted for Liberal with estimated propotion of favoring Liberal about 0.376 with a $SD = 0.0484$. And for female young voters who voted for Liberal before shows an estimated propotion of favoring Liberal about 0.358 with a $SD = 0.0665$, it is higher than the male young voters who voted for Liberal with estimated propotion of favoring Liberal about 0.309 with a $SD = 0.0629$. The conclusion is that male young voters who voted for Liberal appear to show a lower proportion of favoring Liberal in ON for 2020 election compared with other groups.

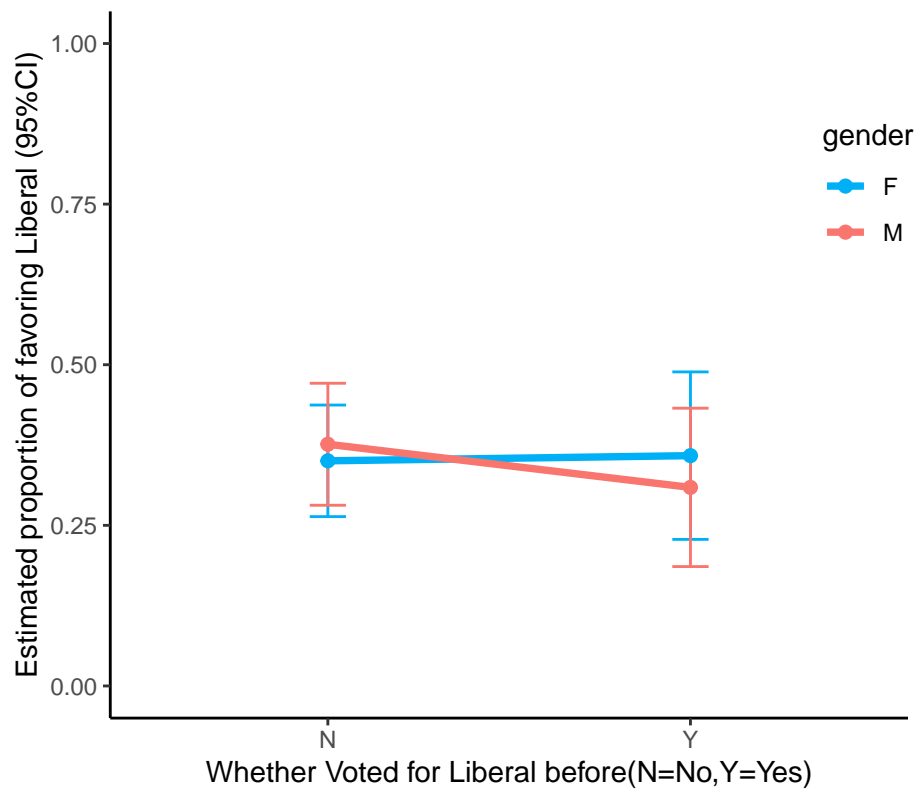


Figure 1: An investigation of estimated proportions of favoring Liberal (95%CI) in ON for 2020 election grouped by gender(F=female,M=male) and whether voted for Liberal before

Table 1: Estimated proportion of favoring Liberal (95%CI) grouped by gender and whether voted for Liberal before

| gender | voted | mean | SD |
|--------|-------|-----------|-----------|
| F | N | 0.3504274 | 0.0442980 |
| F | Y | 0.3584906 | 0.0665026 |
| M | N | 0.3762376 | 0.0484441 |
| M | Y | 0.3090909 | 0.0628864 |

Table 2 and Figure 2 show the estimated proportions of favoring Liberal in ON for 2020 election grouped by whether satisfied of the work for COVID-19 and whether voted for Liberal before. The original scale of satisfaction is a linear scale from 1 to 5, in this part, it is converted into a binary outcome group that the satisfaction group is the group which shows a score no less than 4 while the non-satisfaction group is the group which shows a score no higher than 3.

From the table 2 and figure 2, it can be found that for satisfied young voters who are not ever voted for Liberal shows an estimated proportion of favoring Liberal about 0.356 with a SD = 0.0513, it is very close to the not satisfied young voters who are not ever voted for Liberal shows an estimated proportion of favoring Liberal about 0.366 with a SD = 0.0421. And for satisfied young voters who voted for Liberal before shows an estimated proportion of favoring Liberal about 0.364 with a SD = 0.0837, it is higher than the not satisfied young voters who voted for Liberal with estimated proportion of favoring Liberal about 0.320 with a SD = 0.0539. The conclusion is that there is no big difference between the two groups of whether satisfied of the work for COVID-19 in proportions of favoring Liberal if the young voters are not ever voted for Liberal, but there is a big difference between the two groups of whether satisfied of the work for COVID-19 in proportions of favoring Liberal if the young voters voted for Liberal before.

Table 2: Estimated proportion of favoring Liberal (95%CI) grouped by satisfaction and whether voted for Liberal before

| satisfied | voted | mean | SD |
|-----------|-------|-----------|-----------|
| N | N | 0.3664122 | 0.0420972 |
| N | Y | 0.3200000 | 0.0538640 |
| Y | N | 0.3563218 | 0.0513447 |
| Y | Y | 0.3636364 | 0.0837393 |

Discussion

This survey applied the sampling method SRSWOR,

References

1. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
2. Hadley Wickham, Romain Francois, Lionel Henry and Kirill Muller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. <https://CRAN.R-project.org/package=dplyr>
3. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
4. Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595

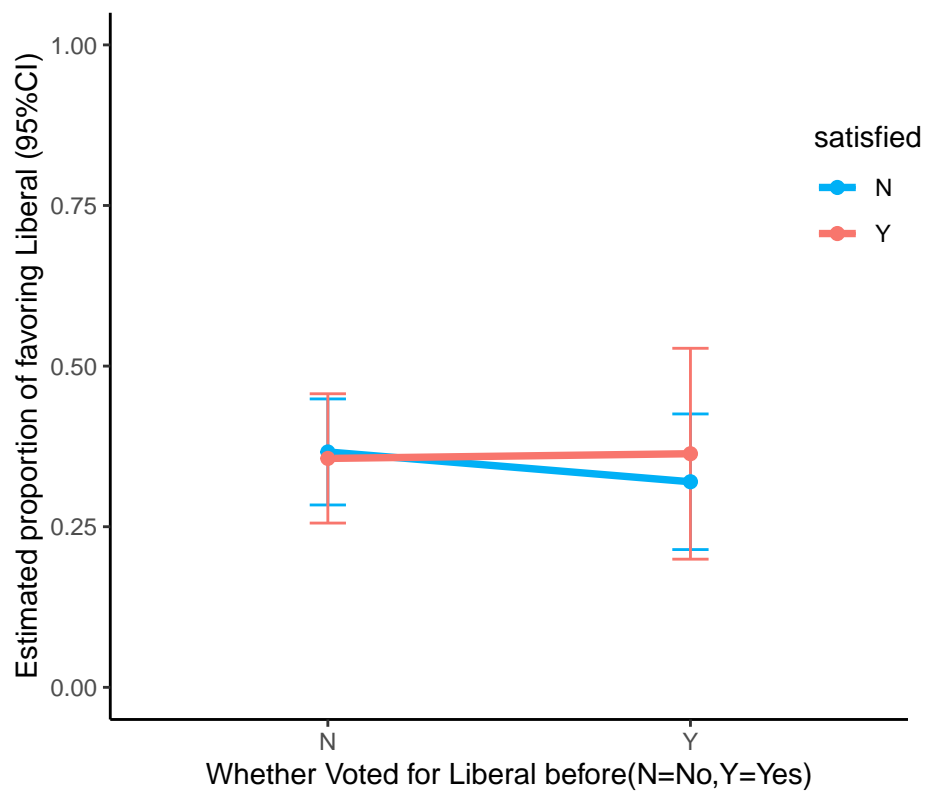


Figure 2: An investigation of estimated proportions of favoring Liberal (95%CI) in ON for 2020 election grouped by satisfaction of work for COVID-19 (N= score ≤ 3 , Y = score ≥ 4) using a linear scale 1-5 and whether voted for Liberal before

Appendix