Analysis on How the 2019 Canadian Federal Election Would Have Been Different If 'Everyone' Had Vote

Yining Chen

12/22/2020

Poststrafication technique is xxxx . . . . [@holt1979post]

Abstract

The 2019 Canadian Election Study was conducted to gather the attitudes and opinions of Canadians about some social, economic, and political issues during and after the 2019 Federal Election.The data is publicly available to researchers and professionals in lots of fields,such as political science, sociology, economics, communications, and journalism. The Canadian Election Study provides a thorough account for the reasons why people voted for one particular party and states the differences and similarities of elections between Canada and other countries. We will analysis how the 2019 Canadian Federal Election would have been different under the assumption that every Canadian citizen had voted.

Keywords: Canada, Election, Public Opinion, Politics

Introduction

The 2019 Canadian federal election was held on October 21, to elect members of the House of Commons to the 43rd Canadian Parliament. The Liberal Party, led by incumbent Prime Minister Justin Trudeau, successfully won 157 seats (around 33.12%) to form a minority government. By using the data in Canadian General Social Survey as our census data , we are interested in how the result of 2019 Canadian federal election would have been different if 'everyone' had vote.

We build a MRP model based on Canada Election Study and a post-stratification data set to essentially estimate the probability of individual voting for the Liberal Party and use demographics to 'extrapolate' how the entire population will vote.

In the Methodology section, we describe the data set and the model that was used to estimate the proportion voting for the Liberal Party. Analysis results of the model are provided in the Results section,and conclusions along with weaknesses and next steps are presented in Discussion section .

Methodology

Data:

The online survey data set contains 37822 observations of Canadian citizens and permanent residents who are 18 or older and 634 variables collected by Advanis Inc..The online sample for the 2019 Canadian Election Study was composed of a modified rolling-cross section during the campaign period and a post-election section.The campaign period survey started from 2019-09-13 until 2019-10-21 and the post-election survey started from 2019-10-24 to 2019-11-11.The data set also contains 4 weight variables which are all campaign period respondents,only high-quality campaign period respondents,all post-election survey respondents and only high-quality post-election survey respondents.The aim of weight variables is to ensure the data can represent the population well.In order to estimate the proportion of voters who will vote for Liberal Party better, we perform a post-stratification calculation. We divide the data into thousands of demographic cells and estimate the probability for each cell.Then, we aggregate all the cell level estimates to a population level estimate by weighting each cell by its relative proportion in the population. This helps for a good correction of non-probability based sampling.

Model:

Our model is used to predict the proportion of voters who will vote for the Liberal Party in 2019. We choose a multilevel regression model where the explanatory variables are sex, age and the birth country of the voter (whether the voter was born in Canada or not). We choose age because people at different ages have different opinions and preferences towards different parties. Also, males and females will probably have different attitudes about politics. The reason for choosing birth country, is that we are interested in whether the decisions made by voters who were born in and outside Canada will differ . The regression model we generated is:

$$y = \beta_0 + \beta_1 x_{age} + \beta_2 x_{sex} + \beta_3 x_{born-inCA} + \epsilon$$

y represents the probability of voters who will vote for the Liberal Party. β_0 is the intercept of the model and represents the probability of a women who were not born in Canada voting for the Liberal Party at age 0 It has no practical meaning since only citizens or permanent residents who are aged 18 or older are eligible to vote for the election.β_1 shows the relationship between age and probability of voting for the Liberal Party,so for every one unit increase in age, we expect a β_1 increase in the average probability of voting for the Liberal Party. β_2 represents the difference in the probability of voting for the Liberal Party between male and female and β_3 represents the difference in the probability of voting for the Liberal Party between people who were born in Canada and outside Canada.

```
## TO CITE THIS SURVEY FILE:
##
## - Stephenson, Laura B; Harell, Allison; Rubenson, Daniel; Loewen, Peter John, 2020, '2019 Canadian E
##
## - Stephenson, Laura, Allison Harrel, Daniel Rubenson and Peter Loewen. Forthcoming. 'Measuring Prefe
##
## LINK: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DUS88V


##
## Call:
## glm(formula = vote_party ~ age + sex + bornin_canada, family = "binomial",
##     data = ces2019_web1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9860  -0.8141  -0.7901   1.4754   1.7720
##
## Coefficients:
```

```
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.1510165  0.2027260  -5.678 1.37e-08 ***
## age              0.0031826  0.0007701   4.133 3.59e-05 ***
## sexMale          0.0248050  0.0257917   0.962   0.3362
## sexOther        -0.2625195  0.1599715  -1.641   0.1008
## bornin_canadaNo  0.3427147  0.2029198   1.689   0.0912 .
## bornin_canadaYes 0.0098928  0.2007795   0.049   0.9607
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 37640  on 31563  degrees of freedom
## Residual deviance: 37524  on 31558  degrees of freedom
##   (6258 observations deleted due to missingness)
## AIC: 37536
##
## Number of Fisher Scoring iterations: 4
```

Results

We estimate that the proportion of voters in favour of voting for the Liberal Party to be 0.2858. This is based off our post-stratification analysis of the proportion of voters, which accounted for age, sex and whether the voters were born in Canada.

However,the p-values of the estimators sex and bornin_canada are greater than the significance level of 0.05. So we do not reject the null hypothesis that $\beta2$ and $\beta3$ equal to zero which means the two estimators are not statistically significant.

Discussion

In this model, we can only get the conclusion that the probability of voting for the Liberal Party is positively associated with age.The analysis showed that for every 1 year increase in age, the average probability of voting for the Liberal Party will increase by 0.31826% .

One of the weakness about our data is that only a small amount of people were born outside Canada.The number is not large enough to support our assumption. The model is only composed of 3 estimators with 2 of them are not statistically significant.In the next steps, we can add more variables into the model and find out their relationships of the probability of voting for the Liberal Party.Also, there may be a non-response problem which means the required information is not obtained from the persons selected in the sample and the sample size will decrease because of this.The response error may also exist due to imperfect questionnaires or misinterpretation of the interviewers and interviewees. The response error reflects the lack of accuracy in responses to questions.What can be done to reduce the response error is to improve the questionnaires and clearly specify the questions.

References

TO CITE THIS SURVEY FILE: Stephenson, Laura B; Harell, Allison; Rubenson, Daniel; Loewen, Peter John, 2020, '2019 Canadian Election Study - Online Survey', https://doi.org/10.7910/DVN/DUS88V, Harvard Dataverse, V1 LINK: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DUS88V

TO CITE THIS SURVEY FILE:

- Stephenson, Laura B; Harell, Allison; Rubenson, Daniel; Loewen, Peter John, 2020, '2019 Canadian Election Study - Online Survey', https://doi.org/10.7910/DVN/DUS88V, Harvard Dataverse, V1

- Stephenson, Laura, Allison Harrel, Daniel Rubenson and Peter Loewen. Forthcoming. 'Measuring Preferences and Behaviour in the 2019 Canadian Election Study,' Canadian Journal of Political Science.

LINK: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DUS88V

LINK: https://en.wikipedia.org/wiki/2019_Canadian_federal_election