Monthly Polling Updates: Green Party in Thunder Bay—Superior North

Joanne Sun, Yi Su, Leqi Sun, Tzu-Ang Su

October 7, 2020

Executive Summary

In this paper.

Introduction

Since its foundation in 1983, the Green Party of Canada has been focusing on applying principles of environmental sustainability and social justice into the reformation of democracy, and has been gaining increasing public support in the past several years as environmental issues become one of the main concerns for both the global and Canadian societies. As junior statisticians at Petit Poll, a Canadian polling company, we work with the Green Party to provide monthly estimates of public satisfaction and support rates of the Party at local communities to provide insights for future campaign promotion as well as policy formulation, and we do this through administering and analyzing surveys. In this report, we will show and interpret results from the most recent polling survey of voters in the Thunder Bay-Superior North riding in Ontario, as well as share some important implications from the results. A link to all the codes used for this report can be found at https://github.com/tomsu0826/sta304ps2polling

Survey Methodology

The Sampling Method

To address the client's needs under the pressure of budget restraints, we design the monthly survey as concise and straightforward as possible. The survey population is all qualified voters living in Thunder Bay-Superior North when we conduct the survey. The results would be more representative if we could sample randomly from all voters. Two constraints prevent us from doing so. First of all, we do not have access to a full list of all qualified voters living in the riding. The best we have is Election Canada's 2019 federal election voters' list. We think that the database is sufficient since people who do not actively participate in politics will be unlikely to change their attitudes sharply. The list will leave out voters who just moved into the riding or became eligible after the election. However, given that the election was only a year ago, this fraction of people should be minimum. So the sample frame of the survey is the 2019 voters' list. It contains all Canadian citizens living in Thunder Bay-Superior North, who voted in the 2019 federal election. The second constraint is the tight budget. It is not feasible for us to conduct a census. The 2019 voters' list registers 43,177 entries. It is possible to sample using a simple random design. However, some neighbourhoods with fewer residents can have zero samples from there. It could potentially harm the quality of the survey. We decided to stratify the riding by neighbourhoods, and then take simple random samples from each stratum. The neighbourhood's definition is vague, especially in Thunder Bay-Superior North, since the communities

are dispersed. In this survey, a neighbourhood is defined as a region that has 2,150 registered voters in it. We plot all the voters on a gridded map and then divide the riding into 20 such neighbourhoods. We randomly sample ten voters within each neighbourhood. We use Google forms to deliver the survey. The digital format makes data analysis more efficient. Five groups of surveyors will work simultaneously in the field with at least two people in each group. Each group will be responsible for four neighbourhoods. At least one person in the group will need to drive. Surveyors are supposed to finish the fieldwork within one day (six working hours). The cost of each surveyor per day is 6 x $15 = $90. We decided to involve volunteers in the data collection process. Ideally, volunteers will consist of 90% of the total workforce. Surveyors are supposed to bring their own electronic devices to access the form and record the survey data. The estimated total costs will be less than $200.

The Survey Questions

The survey contains six questions starting from a mandatory one asking for the eligibility of voting. We also want to figure out the relationship between voters' choices and their top concern issues. It may provide some insights for deciding the focus of the party's future policies. Above all, we want to know if the Green party can count on voters' support in a future election. So we include the question in the survey. At any point, if the respondent feels uncomfortable answering survey questions, they are free to quit.

Statistical Properties

Since all of the survey questions are multiple-choice, we will not be able to calculate the means. We assume that the sample provides an appropriate representation of the population. The regional imbalance within the riding will be taken care of by the stratification. The probability of a unit being sampled is 200/43,177 before and after the stratification.

Non-response

The questions are carefully designed, and the survey has been made as accessible as possible to minimize non-response. If the non-response rate is less than 5% of the sample size, we simply drop the non-responses. If it is larger than 5%, we will consider methods for imputation of missing data. Several R packages offer imputation functions. We choose "Hmisc" to deal with these issues if the scenario happens.

Results

Discussion

Survey

Results

Weakness and areas for future

Appendix

Survey

Add a link to the survey and screen shots of the questions.

References

- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
- Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686
- Frank E Harrell Jr, with contributions from Charles Dupont and many others. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-1. https://CRAN.R-project.org/package=Hmisc