

How the 2019 Canadian Election Result Would Have Been Different If ‘Everyone’ Had Voted

Shiqi Liu

2020/12/19

Authors: Shiqi Liu

2020-12-19

Code and data supporting this analysis is available at: <https://github.com/suki-lsq/STA304-final-report/tree/main/content/post/2015-07-23-r-rmarkdown>

Abstract

As the voter turnout rate in the Canadian Election has been decreasing over the last few decades, especially among the youth age group (Youth Voting Trends - Elections Canada, 2019), whether the turnout rate will drastically change the election result has become a crucial topic for the society as a whole. In this paper, I explore the data collected by the 2019 Canadian Election Study (Stephenson et al., 2020) and use multilevel regression and poststratification to predict the 2019 election result if “everyone” had voted. My findings have shown that Justin Trudeau would still win if “everyone” had voted, however, the support rates for different parties would have changed. My study and analysis could provide insight to political researchers, scientific researcher on election predictions, as well as the party leaders, since my analysis could help them better understand the influence of turnout rate on the election results, and make wiser decision accordingly.

Keywords

Turnout Rate, Multilevel Regression and Poststratification (MRP), Election Result Prediction

Introduction

The turnout rate has always been a significant factor in determining the election result, since not all registered voters will cast a ballot in the actual election. Some political researcher has found out that Jeremy Corbyn would hvae probably won in 2017 UK General Election if “everyone” had voted (Bailey, 2020). Besides, based on an article from The Economist, Hilary Clinton would have probably won the 2016 US Election if “everyone” had voted (The Economist, 2019).

As the Candian Election results rolled in on October 21, 2019, it was clear that Justin Trudeau Liberal government had won the election but losing the popular vote to the Andrew Scheer Conservatives. Looking backward, through the course of the election campaign, neither majority party enjoyed a significant advantage

over the other, and the trendline was characterized by indecision. Therefore, with an only 67% turnout (Elections Canada, 2019), it would be fairly meaningful to identify if the election result would have been different if “everyone” had voted.

Since it is unrealistic to have exactly every registered voter vote in the election, MRP technique could be helpful to estimate the election result if “everyone” had voted. MRP is to partition the population into cells based on demographic attributes, build multilevel regression models to estimate the response variable in each cell, and finally aggregate the cell-level estimates up to a population-level estimate by weighting each cell by its relative proportion in the population (Wang, 2014).

In this paper, I propose employing MRP to identify whether the 2019 Canadian Election result would have been different if “everyone” had voted. Specifically, I will use the multilevel logistic regression model based on provinces to estimate the voting preference of each cell, and then aggregate the estimate to all the registered voters. Through the 2019 CES data and the model built, I have found that though Trudeau would still have won if all the registered voters had voted, the support rates for different parties would have changed. Two datasets will be used in the following analysis, one of which is used for modelling, and the other is used for the poststratification and prediction. In the Methodology section (Section 2), I will describe the study which has been carried out to collect these data, the features of the data, as well as the variable selection. Furthermore, I will thoroughly discuss the model technique used in this section. Results of the prediction and the exhibition of other relative results will be provided in the Results section (Section 3). Finally, the summary, conclusion, weakness and next steps will be explained in the Discussion section (Section 4).

Methodology

Data

The two datasets of Campaign Period Survey (CPS) and Post-Election Survey (PES) used in this analysis were obtained from the 2019 Canadian Election Study (Stephenson et al., 2020). The focus of this study are vote intentions, issue positions, partisanship, and political engagement, as well as the intergroup differences. Both datasets were collected through online surveys on the Qualtrics online platform, the CPS was carried out from September 13th to October 21st, 2019, and the CES was conducted from October 24th to November 11st, 2019.

Population

The population of the two datasets are Canadian citizens and permanent residents who are aged 18 or older. The sampling frame is the Canadian general population available through Qualtrics. The sample of CPS consists of 37,822 members from the sampling frame, and the sample was obtained through stratified sampling method based by region and balanced on gender and age within each region. The sample of with 10,340 respondents for PES was randomly selected from the CPS by re-contacting the participants after the election.

The sampled respondents were composed of a two-wave panel with a modified rolling-cross section during the campaign period and a post-election recontact wave (Stephenson et al., 2020). Regarding the non-responses, incomplete responses and duplicate responses of previous respondents, they have all been excluded from the data file. Besides, a small portion of the respondents who took more than 60 minutes to complete survey have been flagged as inattentive. In the following analysis, I did not remove those inattentive responses, since I consider that some special cases have incurred during the survey period of those respondents, which is unable to suggest that these responses are invaluable.

Furthermore, the responses of PES were matched to their CPS responses by respondent’s panel IDs. Therefore, the demographics of respondents in PES are consistent with the recordings in CPS.

Key Feature, Strength & Weakness

The key feature of the dataset from this non-probability online survey is that it consists of a very large sample of Canadians across the entire country. The dataset is fairly representative of the population thanks to its large size. Moreover, the stratified sampling method makes the dataset more representative in terms of demographic distributions. The potential weakness of this study is that there are many missing values in the dataset, since one of the requirements for the 2019 CES questionnaires was that respondents had to be able to refuse to answer any given question. The questionnaires extensively use “Don’t know/ Prefer not to answer” options, or just skip that question, which is flexible for respondents, but negatively affects the quality of the dataset for analysis.

Variables Selection

Here in this paper, I will use both CPS and PES datasets to identify if the election outcome would have been different if “everyone” had voted in the 2019 Canadian Election. Since the actual votes of the respondents are available in PES dataset, instead of using their vote intentions, I will use their actual votes as the response variable. As for predictors, I will use age, gender, religion, education, last vote in 2015 (`last_vote`), and vote intention during the campaign period (`cps19_votechoice`) at the individual-level, and province at the group-level.

For the predictor age, I divide the respondents into 4 different groups, which include 18-24, 25-39, 40-59, as well as 60+. This division could efficiently contribute to the later analysis, since it can better demonstrate the turnout rates of different age groups. As proposed in the abstract, many researchers have found out that the turnout rate of young people is the lowest among all age groups (Youth Voting Trends - Elections Canada, 2019).

According to a research paper (Fournier, 2020), the Liberals would have a dominated win over the Conservatives if only women voted. Therefore, the gender will be a significant factor in predicting the election result.

Many research have suggested that the highly educated are more likely to vote (Snyder, 2011), which means the turnout rate of the highly educated people is relatively higher than those with lower education levels. Because of this, the education level will contribute to the model prediction.

Moreover, the last vote in 2015 and vote intention during the campaign period should both be strong predictor for the actual vote outcome, since they are highly correlated.

Last but not least, according to the history data of the election (Statistics Canada, 2020), the voter turnout rates tend to vary across different provinces in Canada. To be more specific, the voter turnout in Prince Edward Island has always been the highest, while Newfoundland and Labrador has the lowest voter turnout over the last 3 elections.

Raw Data Overview

To build the multilevel logistic regression model, I clean the PES data by filtering out the cases with missing values for questions to collect their vote choice during the campaign period, and their actual vote after the election. After the removal of these respondents, there are 7,244 respondents in total. The overview of the data with the variables of interest are in the Appendix 1. Table of the Raw Data. These 7,244 cases mainly consist of middle-aged, educated respondents, with approximately equal percentage in man and woman. One factor to notice is that the youth (aged 18 to 24) only count for 1.9%, which could be a potential weakness of the following analysis.

Model

Model Specifics

I have employed the multilevel logistic regression model here to predict the probability of voters with certain demographics and vote intention voting for each of the 6 parties, as well as other parties besides these 6 parties. Hence, there are 7 models in total, which respectively predict the voting proportion for the Liberal, Conservative, NDP, Bloc Québécois, Green, People's parties and other minority parties. As specified in Section 2, I will use age, gender, religion, education, last vote in 2015 (last_vote), and vote intention during the campaign period (cps19_votechoice) at the individual-level, and province at the state-level as the explanatory variables. The multilevel logistic regression with random intercept I am using is a 2 level model, and the level-1 model is given by:

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_1 x_{ij}^{age} + \beta_2 x_{ij}^{gender} + \beta_3 x_{ij}^{last\ vote} + \beta_4 x_{ij}^{religion} + \beta_5 x_{ij}^{vote\ intention} + \beta_6 x_{ij}^{education}$$

and the level-2 model is given by:

$$\beta_{0j} = r_{00} + r_{01}W_j + u_{0j}$$

In the level-1 model, $\log(\frac{p_{ij}}{1-p_{ij}})$ represents the log-odds of the voting probability of each cell for each of the 6 parties and other specified parties, and p_{ij} will be derived from the log-odds, which is the voting probability within each cell for each of the 6 parties and other specified parties. β_{0j} represents the random intercept for the variable province, and $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ represent the coefficient for the individual level explanatory variables age, gender, last vote, religion, vote intention and education respectively.

In the level-2 model, r_{00} is the intercept value for each province, r_{01} represents the coefficient of each province, and W_j is the province level. The last term u_{0j} is the random error component for the deviation of the intercept of a each province from the overall intercept.

There are 7 multilevel logistic regression models with random intercept have been carried out in this analysis, which are to predict the voting probability for the Liberal, Conservative, NDP, Bloc Québécois, Green, People's parties and other minority parties.

Model Building and Variables

There are 7 explanatory variables in total in this model, with 1 group-level predictor province, and age, gender, last vote, religion, vote intention and education as individual predictors. All of the predictors are categorical. I have used the Rstudio to run these 7 models.

The predictor province has a level effect, since voters in different provinces are likely to have different voting probability for each party, holding other explanatory variables constant. For example, two voters of the same age, gender, with the same religion, education, the voter in Ontario will have higher probability of voting for the Liberal than the voter in Alberta.

As for the individual-level predictors, first, I use age groups rather than individual ages, because age groups can better reflect the change in voting preferences among different age groups. I divide the respondents into 4 different groups, which include 18-24, 25-39, 40-59, as well as 60+. I did not use individual ages, since there is no reason to believe that there is a linear trend between the voting probability for certain party and the age of voters. A comparison of voting probabilities for certain parties between different age groups would be more practical. Gender, religion and education are the other 3 demographic predictors, since the voting probability tends to split among different religion groups and different genders. I have also included respondents' last vote and vote intentions during the campaign period as predictor variables, because these two could both be strong predictors for their actual vote.

To further demonstrate how the predictors are chosen and why they are significant, firstly, as proposed in the abstract, many researches have found out that the turnout rate of young people is the lowest among all

age groups (Youth Voting Trends - Elections Canada, 2019). Besides, for the predictor gender, according to a research paper (Fournier, 2020), the Liberals would have a dominated win over the Conservatives if only women voted. Moreover, many research have suggested that the highly educated are more likely to vote (Snyder, 2011), which means the turnout rate of the highly educated people is relatively higher than those with lower education levels. Individuals with different education levels tend to have different ideology, thus different voting preferences.

An alternative model could be the multiple linear regression model, which will be simpler and easier to run. However, since the response variable is binary here (i.e., whether or not vote for this party), but not numerical, using a linear regression model seems inappropriate.

Last but not least, regarding the CES, the survey itself, it contains mostly the categorical variables rather than numerical variables, and responses of many variables can be coded as binary variables to make analysis more convenient. The multilevel logistic regression could also well fit the feature of this survey.

Results

The results will be presented in 2 sections, which include the overall poststratification results based on the model, and the voting probability for each party and the turnout of each variable of interest.

Poststratification Results

To estimate the 2019 election result if “everyone” had voted, I employ the poststratification technique. I use the CPS data as the census data to predict the election result if “everyone” had voted. I have partitioned the census data into 13,940 cells based on age, gender, last vote, education level, vote intention and province, and I calculate the proportion of each cell relative to the whole census. Then I have used the multilevel logistic regression model fitted in the Model Section (Section 2) to predict the voting probability for the 6 parties and other specified parties of each cell. Finally, I aggregate the voting probability of each cell to the aggregate level. The purpose of this poststratification is to estimate the election results if every individual in the CPS data had voted in the election. The estimated election outcome by poststratification is shown in the following Figure 1.

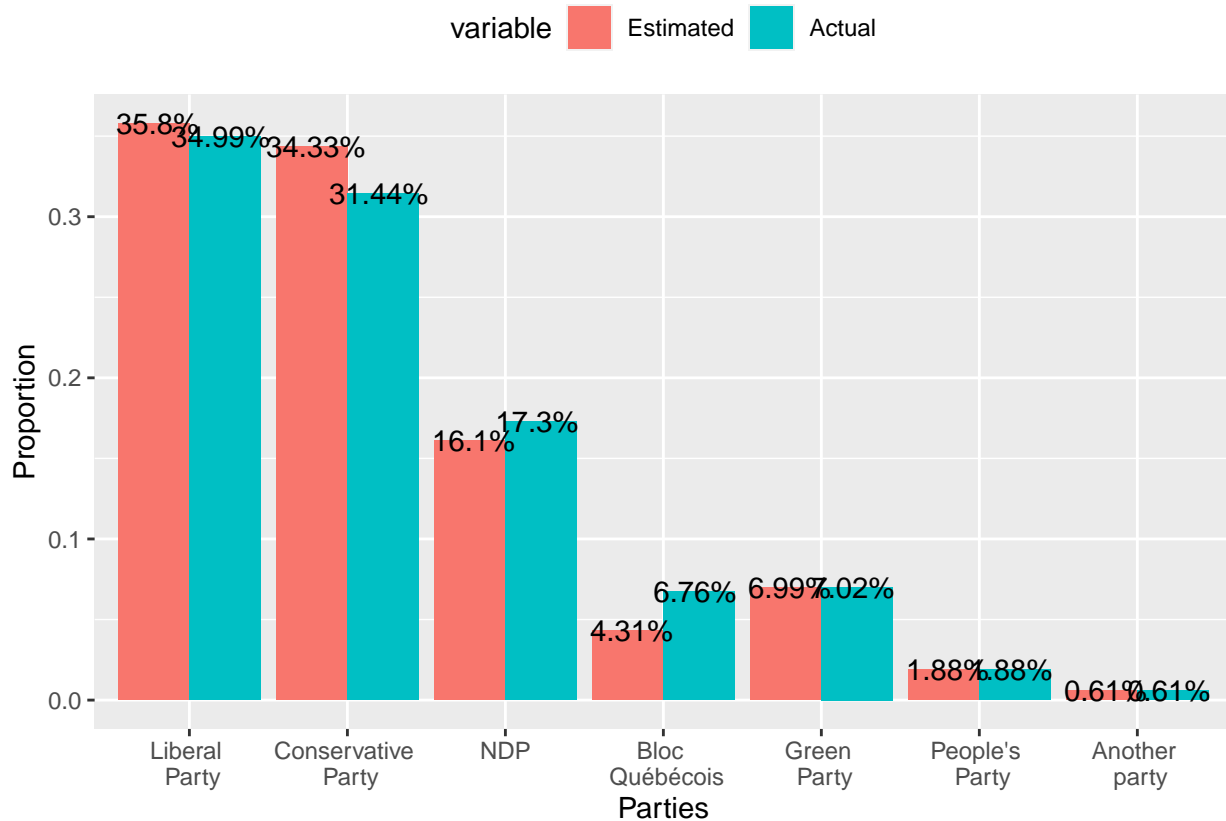


Figure 1: Estimated Poststratification Results vs. Actual Results

As we can see from the above graph, the election result if “everyone” had voted is not significantly different from the actual election result, with the Liberal still being the winner and the Conservative ranking second. Justin Trudeau However, if “everyone” had voted, then the voting proportion for both the Liberal and Conservative increase, while the voting proportion for NDP decreases.

Voting Probability For Each Party & Turnout of Each Variable of Interest

In this part, I will summarize the voting probability for each party of under each predictor variable to compare the voting preferences of individuals with different demographics. For example, will the voting probability for the Liberal vary across different genders? Furthermore, I will supplement the estimated voting probability with the turnout by each predictor variable to further illustrate the poststratification results in the last part.

Gender

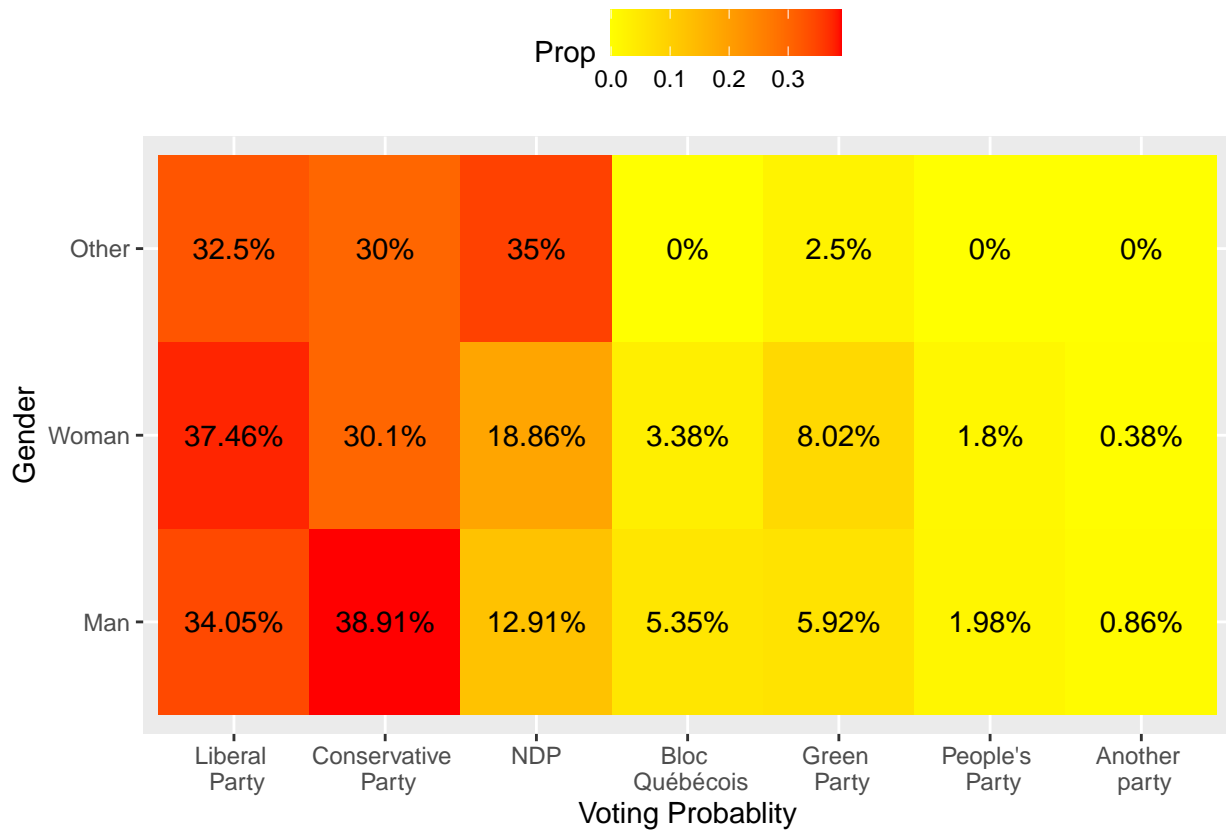


Figure 2: Voting Probability by Gender

As shown in Figure 2, women has the highest voting rate for the Liberal, while men has the highest voting probability for the Conservative. This suggests that women are more likely to vote for the Liberal, while men are more likely to vote for the Conservative.

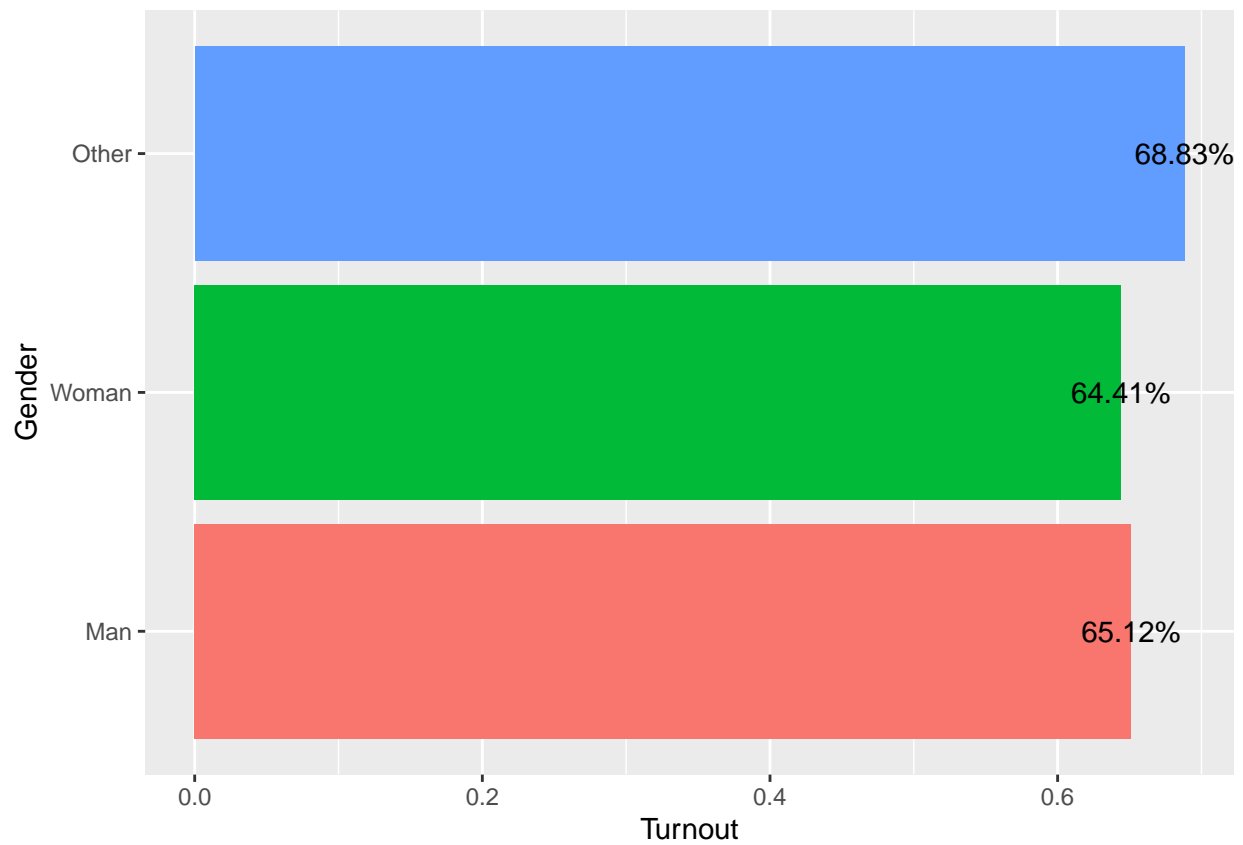


Figure 3: Turnout by Gender

As shown in Figure 3, the turnouts for both male and female are approximately the same, with male having slightly higher turnout.

Education Level

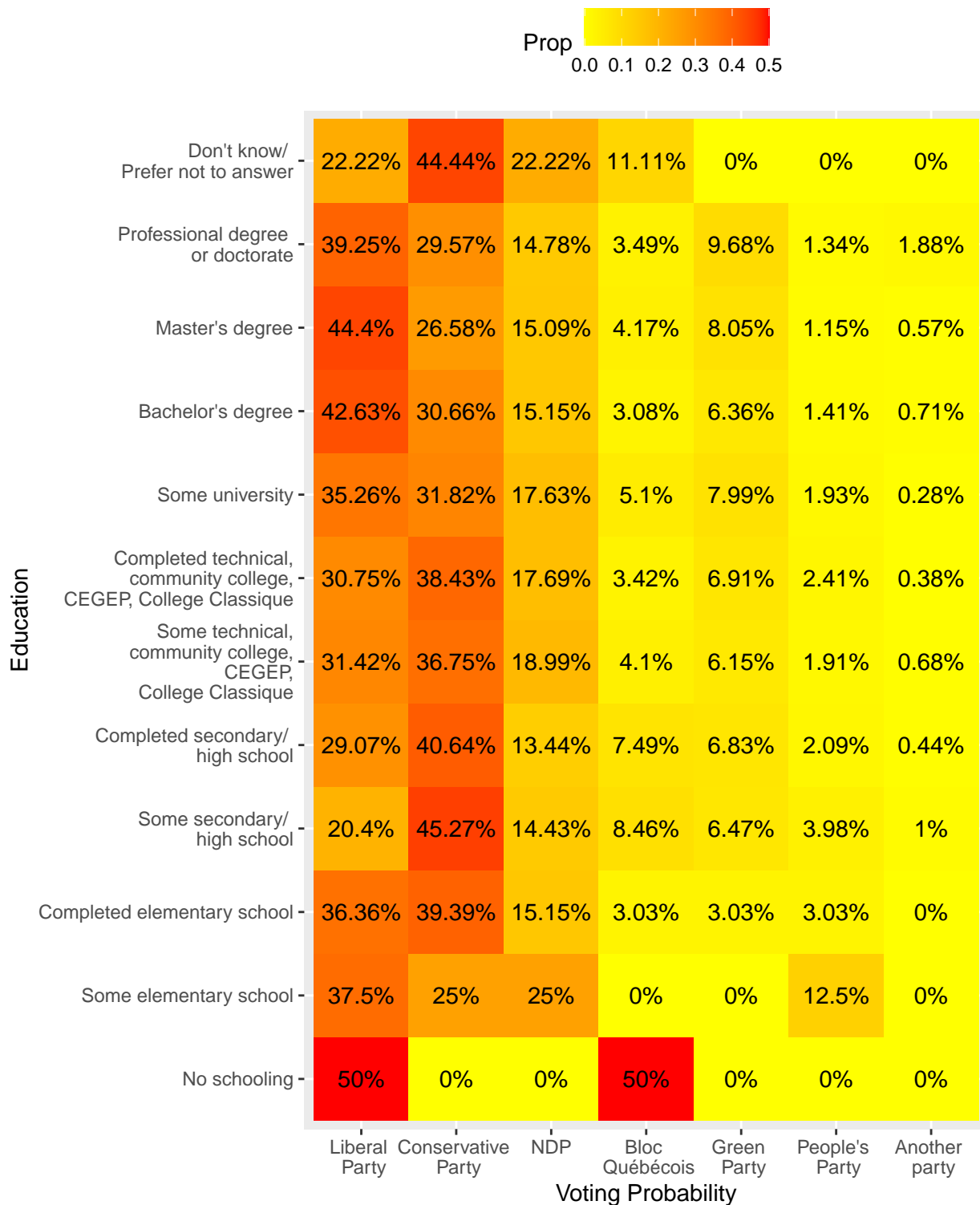


Figure 4: Voting Probability by Education Level

As shown in the Figure 4 above, we can see that the individual with higher education levels (i.e., with college degree and above) tend to vote for the Liberal, besides, the individuals with limited education level are also more likely to vote for the Liberal. Meanwhile, individuals with degree from secondary high school

to technical college are more supportive for the Conservative.

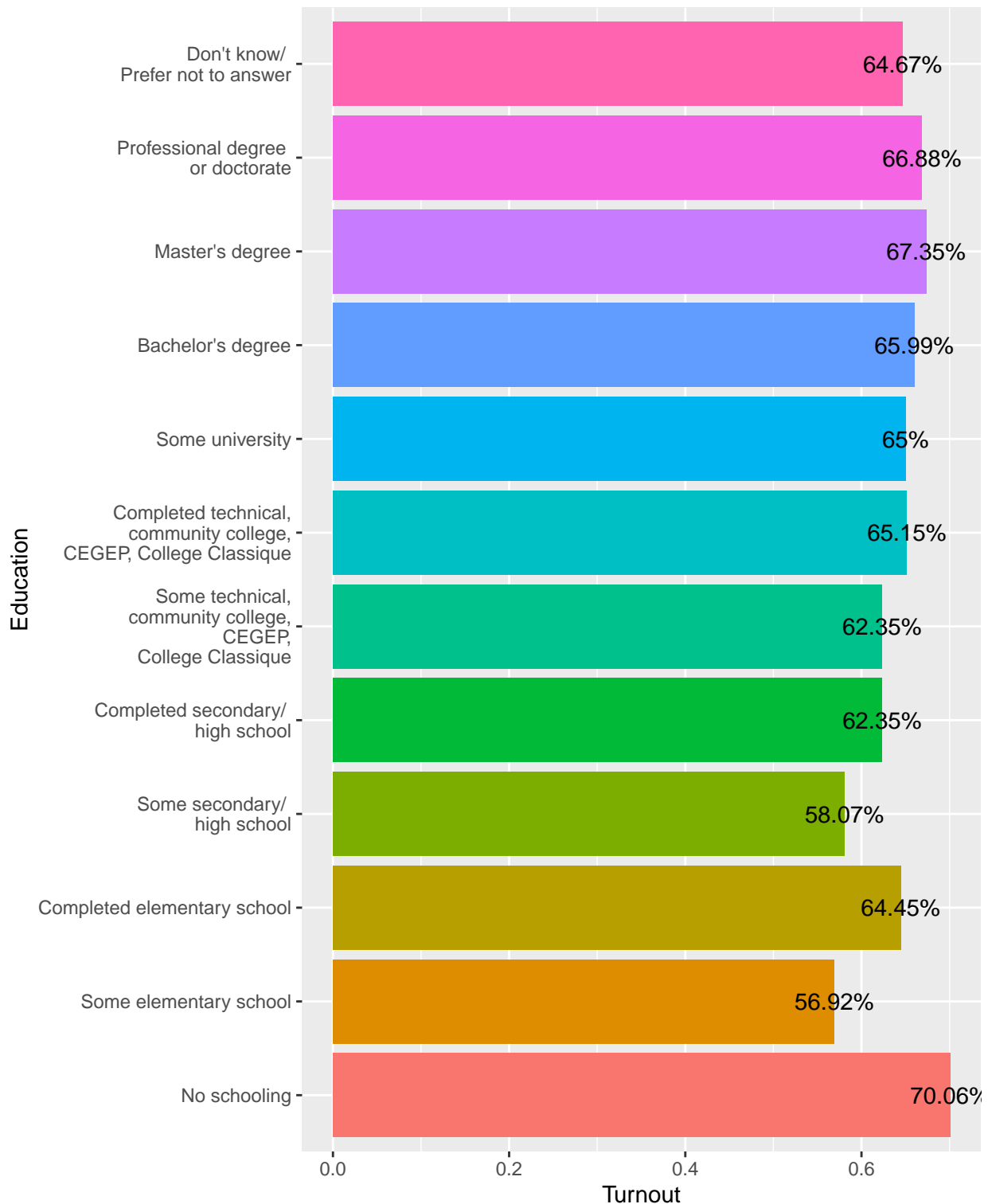


Figure 5: Turnout by Education Level

As shown in Figure 5, the higher educated are more likely to participate in the election vote, which means the turnout of the individuals with higher education levels are higher than those with lower education levels.

Age

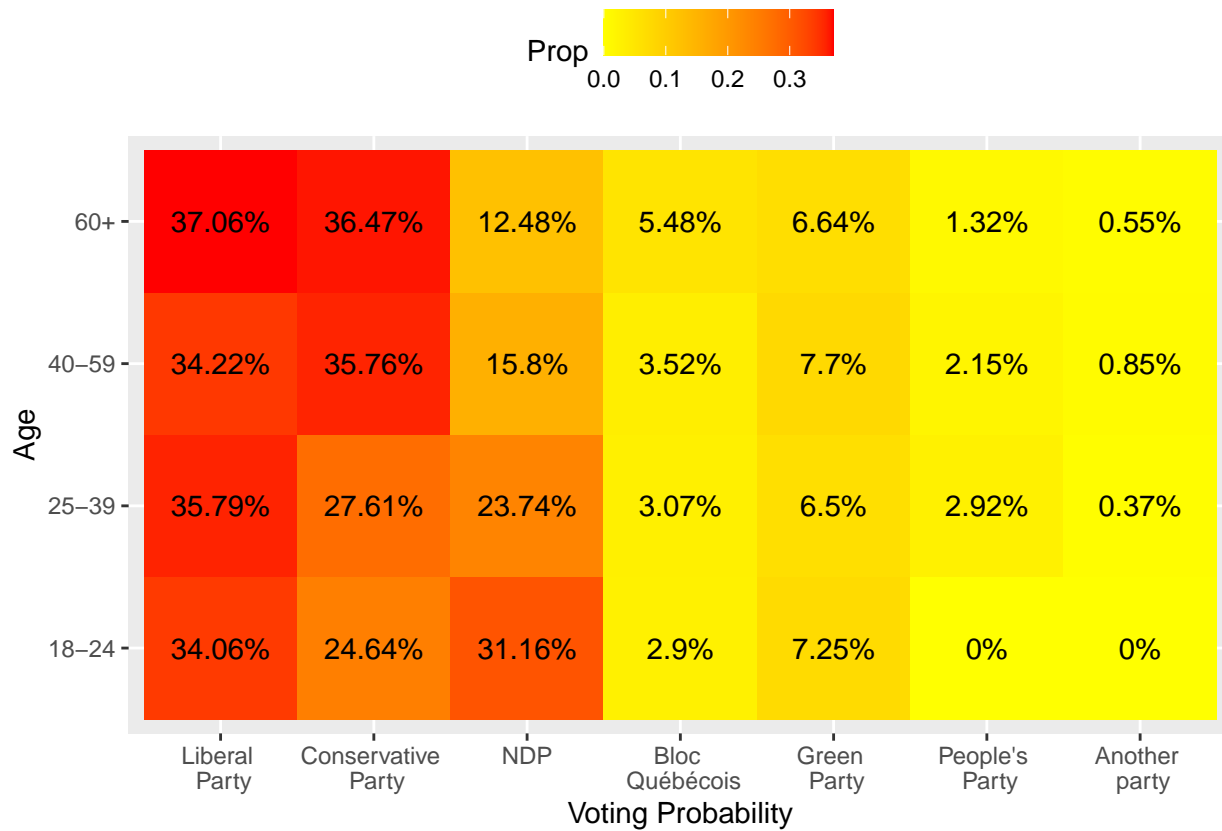


Figure 6: Voting Probability by Age

As shown in the Figure 6, we can see that youth are more likely to vote for the Liberal and NDP, rather than the Conservative, while the middle-aged and the elderly tend to be more supportive of the Conservative. However, there is no linear trend between the voting probability for the Liberal or the Conservative and the age, which means the results do not show as age increases, the voting probability for a certain party increases. The results only justifies that the voting probability for a certain party is the highest within each age group.

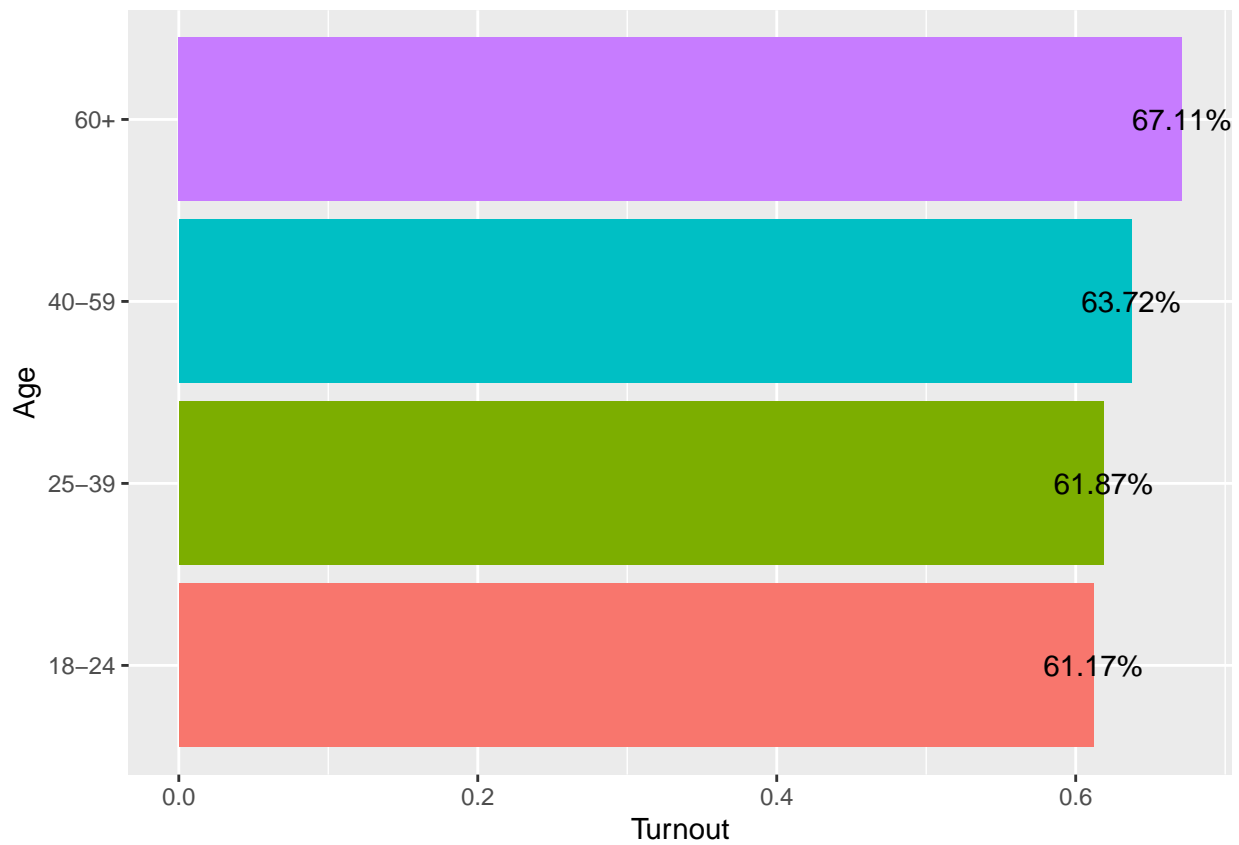


Figure 7: Turnout by Age

As shown in the Figure 7, the turnout of the young-aged is the lowest, while the turnout of people of age 60+ is the highest.

Province

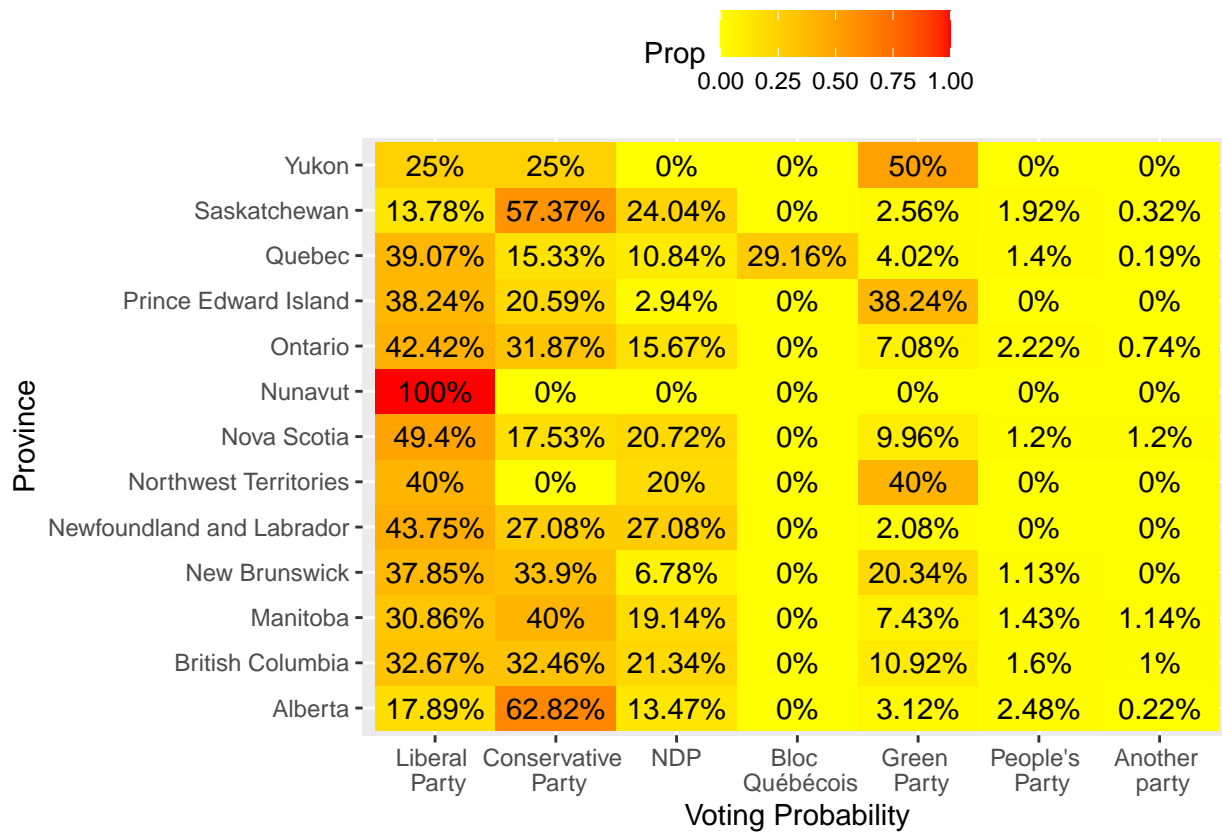


Figure 8: Voting Probability by Province

From the Figure 8, it could be inferred that the voting rates for the Liberal, Conservative and NDP are the highest among all parties. The majority provinces have higher voting proportion for the Liberal, while it is noticable that Alberta, Manitoba and Saskatchewan have obviously higher voting proportion for the Conservative.

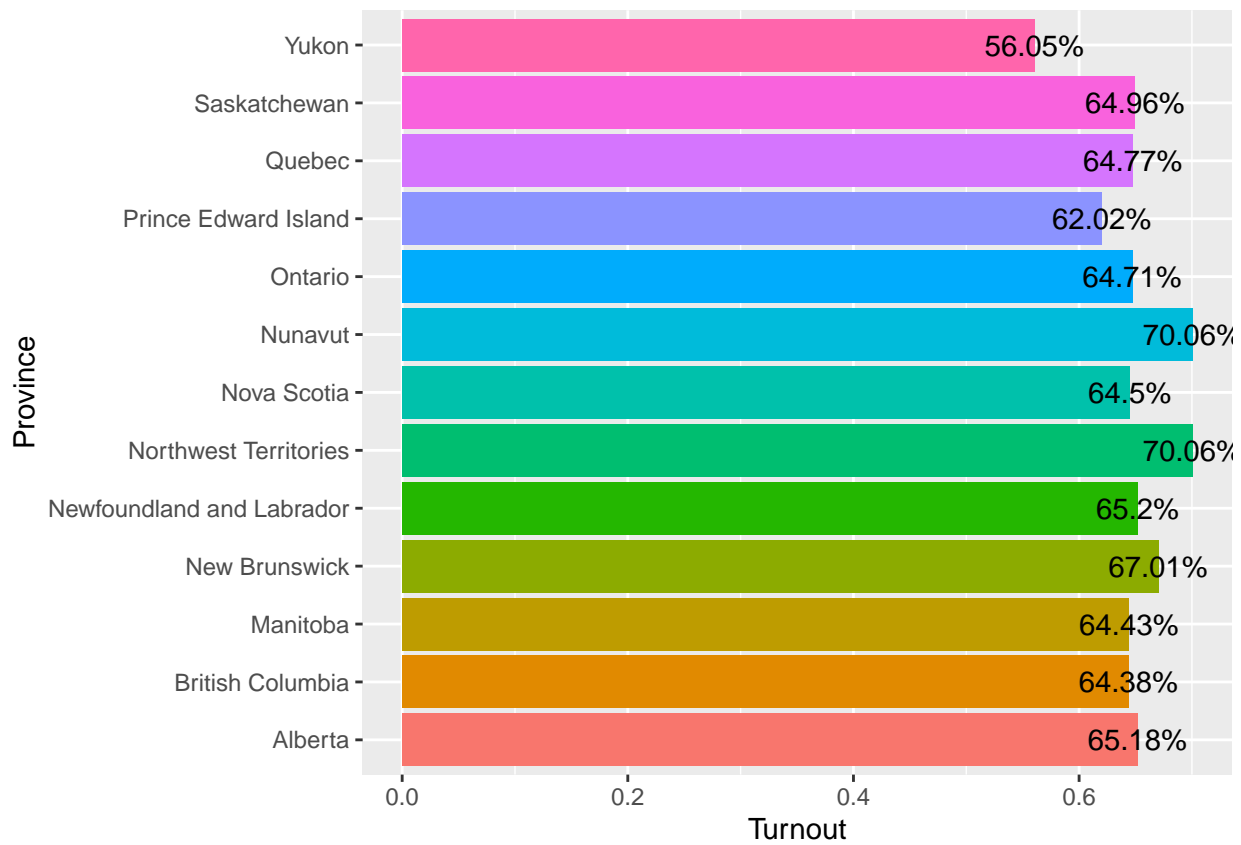


Figure 9: Turnout by Province

The Figure 9 exhibits the turnout of every province. It has shown that the turnout of the majority of provinces are similar with each other, while the turnouts of PEI and Yukon are relatively lower.

Religion

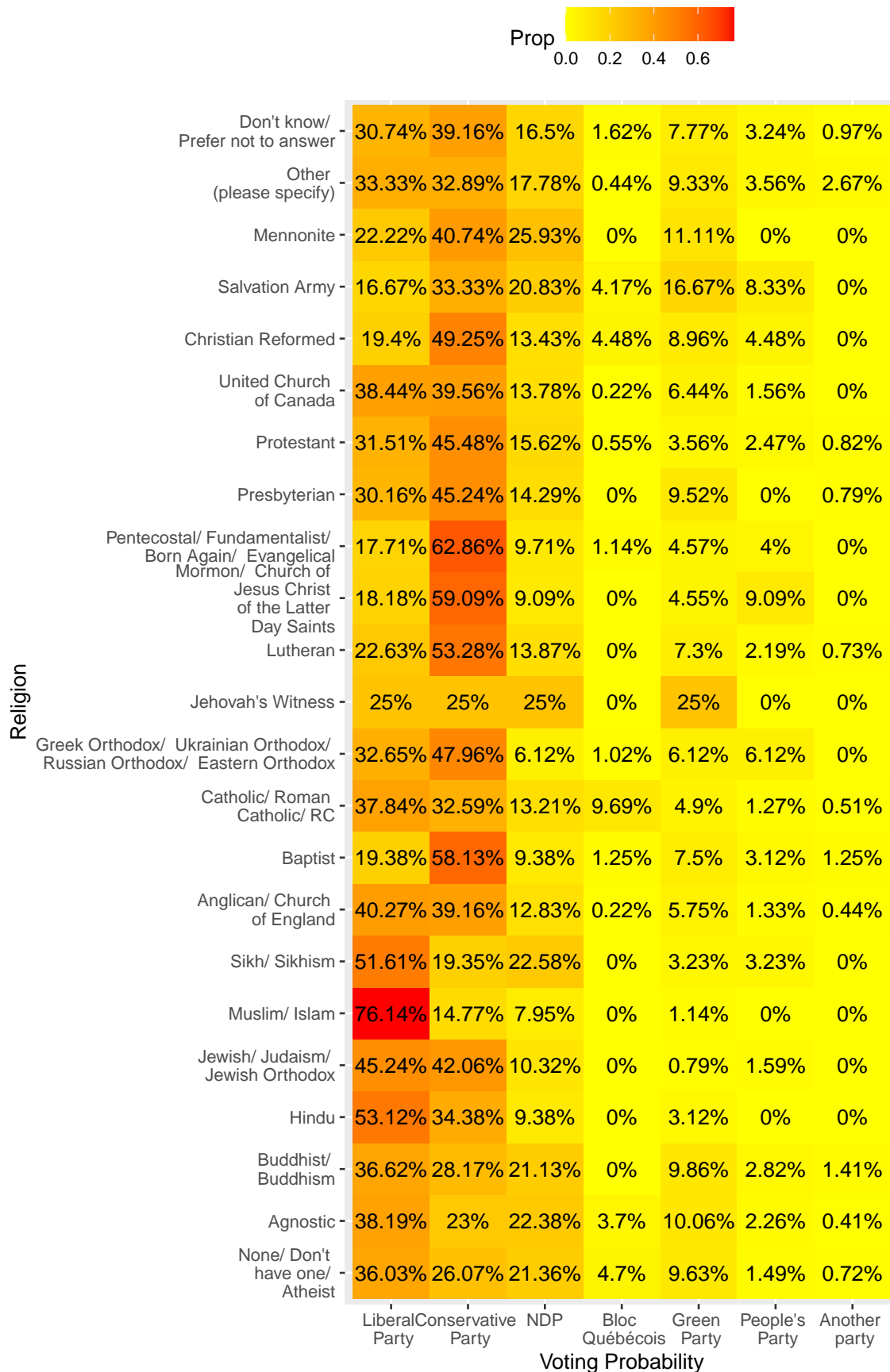


Figure 10:

As shown in Figure 10 above, we can see that the majority of the religions are more supportive for the Liberal, while people with religions of Pentecostal, Jesus Christ, Lutheran, Greek Orthodox and Baptist are more likely to vote for the Conservative.

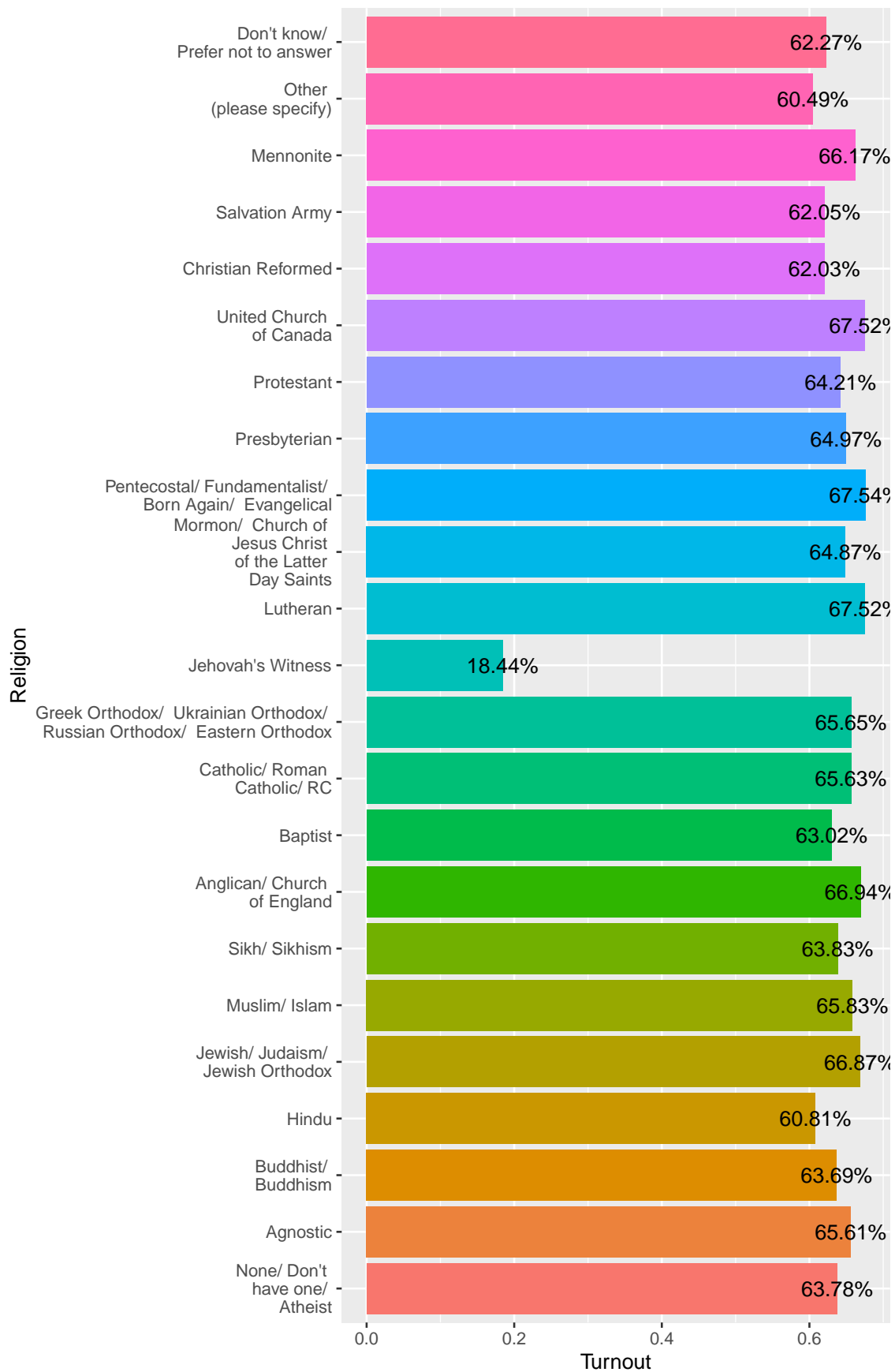


Figure 11:

Turnout by Religion

From the Figure 10 above, we can see that the turnout of individuals with religions of the United Church of Canada, Anglican, Jewish and Menonite are relatively higher. On the contrary, the turnout of individuals with the religion of Jehovah's Witness have the lowest turnout compare to those with other religions.

Last Vote

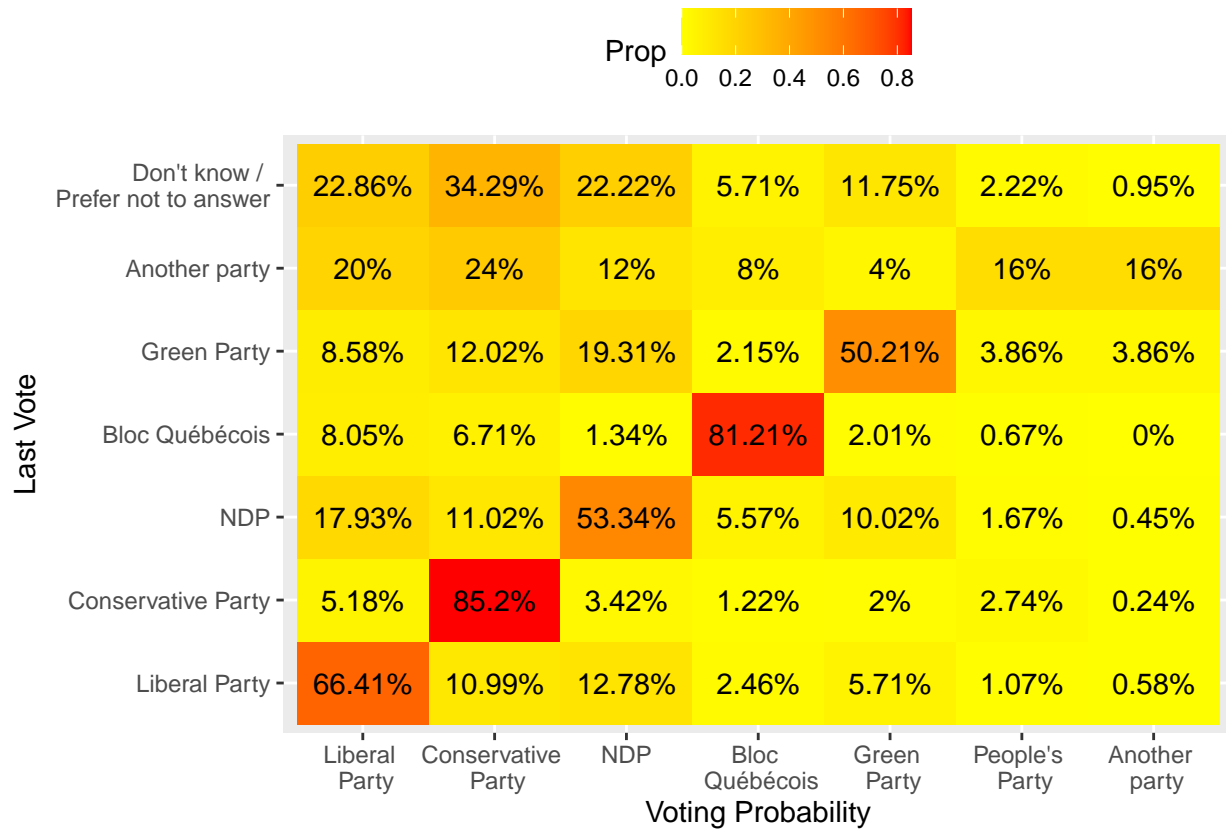


Figure 12: Voting Probability by Last Vote

The Figure 12 above has shown the relationship between the last vote and the actual vote choice of respondents. We can see clearly that people tend to stick to their last vote choice.

Vote Choice Campaign Period

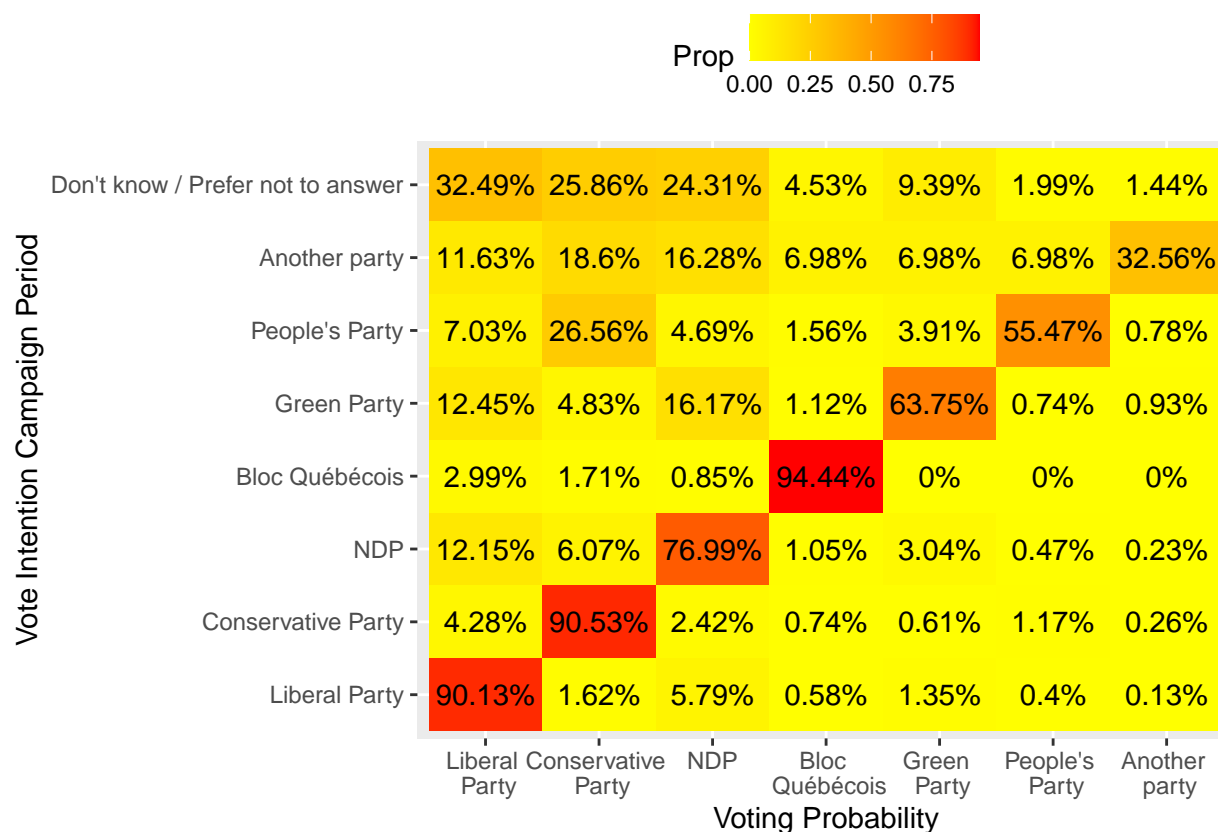


Figure 13: Voting Probability by Vote Intention

The Figure 13 above exhibits the relationship between respondents' vote intention during the campaign period and their actual vote choice during the election. It is obvious that people tend to follow their initial vote intention in the election.

Discussion

In this section, I will firstly explain and discuss the results from the Results Section (Section 3) to make conclusions. Furthermore, I will discuss the weaknesses of this analysis, and propose next steps which can be carried out later.

Conclusion

The Election Result If “Everyone” Had Voted

To begin with, I will compare the estimated election result using poststratification and the actual election outcome. The model is based on the 7,244 individuals in the CES data, who actually had voted in the 2019 Canadian Election. The poststratification is based on the CPS data with 37,822 respondents, and the result from the Results Section (Section 3) has shown the estimated voting proportions for each party if all of these respondents had voted. Since the CPS data has already excluded the respondents not eligible to vote, directly estimating the voting proportion of these 37,822 individuals is reasonable.

As shown in Figure 1, the estimated election result if “everyone” had voted using poststratification based on the multilevel logistic regression model from the Model Section (Section 2) is not significantly different from the actual election result. If everyone of those 37,822 individuals in the CPS had voted in the election, the

Liberal Party, Justin Trudeau would still be the winner of the 2019 Canadian Election with Andrew Scheer closely following behind. Although the overall results would not have significantly changed if “everyone” had voted, the voting proportions for the Liberal and Conservative would have changed by a crucial amount, especially for the Conservative party. The voting proportion for the Liberal will increase by 0.89% if “everyone” had voted, while the voting proportion for the Conservative will increase by 2.89% if “everyone” had voted. This increase would very likely to inverse the situation and make Andrew Scheer the winner, since the difference of voting proportions between Trudeau and Scheer is only 1.47%.

This result may have implications for the party leaders that the focus of campaign should also include encouraging your people to really participate, turn out and vote in the election. Having more people voted in the election may help the leader beat the competitor and win, especially in the situation where there is small difference in the voting proportion between 2 leaders.

The Influence of the Turnout on the Election Result

To discuss the impact of the turnout on the actual election result, the analysis of turnouts of different demographic groups will be compared and explained.

First of all, as we can see from the graphs showing turnouts of different demographic groups, the turnouts approximately fall between 60% - 70%, which is very close to the actual turnout of 67% of the 2019 Canadian Election. This result has confirmed that the dataset as a whole are representative of the target population.

Moreover, from the result and Figure 3, we can see there is no obvious gap of turnouts between different gender groups. As for the education levels, the result and Figure 5 has shown that the more educated are more likely to participate and vote in the election, while individuals with lower education levels have lower turnout. And from Figure 4, the less educated are more supportive for the Conservative, therefore, if Andrew Scheer could successfully increase the turnout of those people by encouraging them to vote, the final voting proportion for the Conservative would have increased by a sufficient amount. Regarding different age groups, as we can see from the Figure 7, the turnout of youth is the lowest. Meanwhile, as shown in the Figure 6, those young-aged are more likely to vote for the Liberal, which confirms the poststratification result that the voting proportion for the Liberal would have increased if “everyone” had voted. In this situation, the increase of turnout of youth would have increased the voting proportion for the Liberal.

Furthermore, as for the predictor last vote, from Figure 12, we can see there is a strong relationship between this predictor and the actual vote in the election. People tend to vote for the party same as their last vote. This result has confirmed that the last vote should be a significant predictor for the actual vote choice. Last but not least, as shown in Figure 13, we can see that respondents’ vote intention during the campaign period will well reflect their actual vote. This result has further verified that the vote intention is statistically significant for predicting the actual election result.

Weaknesses

As mentioned in the Data part of the Methodology Section (Section 2) and the Table 1 in the Appendix, the young-aged respondents only account for 1.9%. This bias in the dataset will lower the representativeness of the result for this age group. The sampling approach is not thoroughly explained in the CES report. As for the survey itself, though offering flexible options for questionnaires is more considerate to respondents, this increase the complexity of the structure of the data, thus making the data cleaning and analysis more difficult. Besides, there are more missing values due to the flexibility of the questionnaire. Furthermore, for the PES, the aim was to collect 50% of the CPS respondents, which is 18,911. However, it turned out that the sample size of the PES only reached 10,340, which did not meet the expectation. The useful sample size of the CES is not sufficiently large.

Next Steps

Since the sample is not sufficiently representative for all age groups, especially limited for the young-aged group, the CES could adjust and improve the age proportions of the respondents to reach a balance between different age groups. Regarding the model, I have employed multilevel logistic regression model with random intercept, and the group-level variable is province. Future research and analysis could add more province-level predictors such as the vote share, population proportion, area of each province to further improve the predictiveness of the model.

Appendix

1. The overview of the data with the variables of interest.

n	Overall
age (%)	7244
18-24	138 (1.9)
25-39	1369 (18.9)
40-59	2469 (34.1)
60+	3268 (45.1)
gender (%)	
Man	3477 (48.0)
Woman	3727 (51.4)
Other	40 (0.6)
last_vote (%)	
Liberal Party	3084 (45.7)
Conservative Party	2047 (30.3)
NDP	898 (13.3)
Bloc Québécois	149 (2.2)
Green Party	233 (3.5)
Another party	25 (0.4)
Don't know / \n Prefer not to answer	315 (4.7)
education (%)	
No schooling	2 (0.0)
Some elementary school	8 (0.1)
Completed elementary school	33 (0.5)
Some secondary/ high school	201 (2.8)
Completed secondary/ high school	908 (12.5)
Some technical, community college, CEGEP, College Classique	732 (10.1)
Completed technical, community college, CEGEP, College Classique	1577 (21.8)
Some university	726 (10.0)
Bachelor's degree	1980 (27.3)
Master's degree	696 (9.6)
Professional degree or doctorate	372 (5.1)
Don't know/ Prefer not to answer	9 (0.1)
province (%)	
Alberta	928 (12.8)
British Columbia	998 (13.8)
Manitoba	350 (4.8)
New Brunswick	177 (2.4)
Newfoundland and Labrador	144 (2.0)
Northwest Territories	5 (0.1)

Nova Scotia	251 (3.5)
Nunavut	3 (0.0)
Ontario	2968 (41.0)
Prince Edward Island	34 (0.5)
Quebec	1070 (14.8)
Saskatchewan	312 (4.3)
Yukon	4 (0.1)
religion (%)	
None/ Don't have one/ Atheist	1807 (24.9)
Agnostic	487 (6.7)
Buddhist/ Buddhism	71 (1.0)
Hindu	32 (0.4)
Jewish/ Judaism/ Jewish Orthodox	126 (1.7)
Muslim/ Islam	88 (1.2)
Sikh/ Sikhism	31 (0.4)
Anglican/ Church of England	452 (6.2)
Baptist	160 (2.2)
Catholic/ Roman Catholic/ RC	1961 (27.1)
Greek Orthodox/ Ukrainian Orthodox/ Russian Orthodox/ Eastern Orthodox	98 (1.4)
Jehovah's Witness	4 (0.1)
Lutheran	137 (1.9)
Mormon/ Church of Jesus Christ of the Latter Day Saints	22 (0.3)
Pentecostal/ Fundamentalist/ Born Again/ Evangelical	175 (2.4)
Presbyterian	126 (1.7)
Protestant	365 (5.0)
United Church of Canada	450 (6.2)
Christian Reformed	67 (0.9)
Salvation Army	24 (0.3)
Mennonite	27 (0.4)
Other (please specify)	225 (3.1)
Don't know/ Prefer not to answer	309 (4.3)
cps19_votechoice (%)	
Liberal Party	2228 (30.8)
Conservative Party	2312 (31.9)
NDP	856 (11.8)
Bloc Québécois	234 (3.2)
Green Party	538 (7.4)
[reached getOption("max.print") -- omitted 11 rows]	

Table 1: Data Overview

References

Bailey, J. [@PoliSciJack]. (2020, November 15). Ever wondered how the 2017 General Election might have gone if “everyone” had voted? I did, so I worked it out using MRP. Turns out, @jeremycorbyn would probably be the Prime Minister of a minority Labour government. [Tweet]. Twitter. <https://twitter.com/PoliSciJack/status/1327920037198499840>

Chung-hong Chan, Geoffrey CH Chan, Thomas J. Leeper, and Jason Becker (2018). rio: A Swiss-army knife for data file I/O. R package version 0.5.16.

David B. Dahl, David Scott, Charles Roosen, Arni Magnusson and Jonathan Swinton (2019). xtable: Export Tables to LaTeX or HTML. R package version 1.8-4. <https://CRAN.R-project.org/package=xtable>

- Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.
- Fournier, P. (2020, June 28). The biggest divide in Canadian politics? Men vs. Women. Retrieved December 20, 2020, from <https://www.macleans.ca/politics/ottawa/the-biggest-divide-in-canadian-politics-men-vs-women/>
- Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.1. <https://CRAN.R-project.org/package=stargazer>
- Kazuki Yoshida and Alexander Bartel (2020). tableone: Create ‘Table 1’ to Describe Baseline Characteristics with or without Propensity Score Weights. R package version 0.12.0. <https://CRAN.R-project.org/package=tableone>
- Snyder, R. E. (2011, November). The impact of age, education, political knowledge and political context on voter turnout. December 20, 2020, <http://dx.doi.org/10.34917/3310658>
- Statistics Canada. (2020, February 26). Voter turnout by province, 2011, 2015 and 2019 federal elections. Retrieved December 20, 2020, from <https://www150.statcan.gc.ca/n1/daily-quotidien/200226/cg-b002-eng.htm>
- Stephenson, L. B., Harell, A., Rubenson, D., & Loewen, P. J.. (2020). The 2019 Canadian Election Study – Online Collection. [dataset]
- The Economist. (2019, July 6). If everyone had voted, Hillary Clinton would probably be president. Retrieved December 17, 2020, from <https://www.economist.com/graphic-detail/2019/07/06/if-everyone-had-voted-hillary-clinton-would-probably-be-president>
- Wang, W., et al., Forecasting elections with non-representative polls. *International Journal of Forecasting* (2014), <http://dx.doi.org/10.1016/j.ijforecast.2014.06.001>
- Wickham et al., (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>
- Youth Voting Trends - Elections Canada. (2019). Retrieved December 17, 2020, from <https://electionsanddemocracy.ca/canadas-elections/youth-voting-trends>