[INSERT RESULT] If All Eligible Voters Had Voted in the 2019 Canadian Federal Election

Zakir Chaudry

12/20/2020

Abstract

In this paper we will be using MRP Modeling on the CES data using the GSS 2017 data as a post-stratification dataset to see what the results of the election would be if everyone had voted. We then relay the results (currently unknown, add later) and discuss the potential implications of these findings, as well as any weaknesses/opportunities for future improvement.

Keywords: MRP; 2019 Canadian Federal Election; Eligible Voter; [Winner of Election]

Introduction

Elections are one of the most vital events of the modern age - a submission of votes representing the will of the people, allowing members of the population to make major decisions from local, national, and even international scales. The importance of these events cannot be understated as they directly affect our lives, sometimes in small ways, sometimes much more so. As such, it's incredibly imperative that the "will of the people" mentioned above comes to fruition. In our current voting system, this is what's highlighted, making the majority of people happy with the result. However, can we really say the will of the people is truly being heard? There are many different issues contradicting this assumption, but perhaps is one of the most impactful is simply the lack of voting. In the 2019 Canadian Federal Election, roughly 66% (Turnout Reference) of Canadians reported voting. This is quite a lot at first glance, but that's still 34% of people that are not being accounted for. Frankly speaking it's hard to support the notion that the election was the will of the people given that so many people's votes weren't even heard. However, what if this wasn't the case? What if every single person that could vote did so? What then? Would the result change? Or are the final results truly the will of the people all along. In this report, we will be using MRP (MRP Reference) Modeling to show what could have happened, in aims to answer these questions and go beyond.

Data

CES Data

The data used to generate the model was taken from the 2019 Canadian Election Study. The Canadian Election Study is a large-scale survey conducted on each election year, going back to as early as 1965 (CES site). Its purpose is to give a greater insight into the electoral democracy in Canada by measuring many sorts of details. The data for this report was taken from the 2019 Study, in particular the data gathered from the online survey component of the study. The questionnaire for this study was hosted and conducted via the Qualtrics Platform (Qualtrics Platform) (Codebook), with the population being eligible voters in Canada.

We made use of 5 different variables in order to create our model. The independent variable we looked at was cps19_votechoice, which described which party the respondent was going to vote for. There were 8 choices

in response: “Green Party”, “Conservative Party”, “Liberal Party”, “People’s Party”, “Bloc Québécois”, “ndp”, “Another party (please specify)”, and “Don’t know/ Prefer not to answer”. For the model, we ended up removing last two options, as on the grand scale of things, neither of those two groups would impact the election as a whole. As such, we were left with 6 different parties to choose from. There was another option in terms of finding a variable describing the choice of party labeled pes19_votingchoice, which described who the respondents said they voted for after the election. This may have been a better choice in terms of what we wanted from the variable, but there simply was not enough data for this variable for this switch to have been worth it.

The first dependent variable we chose was cps19_age, which described the age of the respondent. We found that splitting the age into ranges would be better to see from an interpretation standpoint, so we split the variable into the following ranges: “18 to 24”, “25 to 34”, “35 to 44”, “45 to 54”, “55 to 64”, “65 to 74”, and “75 and above”.

The next dependent variable we chose was cps19_education, which described the level of education the respondent had attained up to that point. There was a lot of reshuffling of data values for this variable, mainly due to the fact that the post-stratification dataset’s education had a wildly different set of potential values. The initial values were [Insert Initial Values], but we ended up reorganizing them into the following values: “Less than High School”, “High School”, “Less than Bachelor’s Degree”, “Bachelor’s Degree”, and “Above Bachelor’s Degree”. While some information was lost, the ability to match up with the post-stratification dataset’s education values and the simplicity from an observation standpoint made this a worthy endeavor.

The third dependent variable we chose was cps19_gender, which described the gender of the respondent. Again, there was an issue in terms of matching the post-stratification set, specifically in terms of non-cisgender respondents. The post-stratification set has only male or female options available, as it uses sex rather than gender. Ultimately, we dropped the non-cisgender respondents, as like the voting choices, their impact was negligible relatively speaking (Cis Population).

Finally, our last dependent variable was cps19_province. We found that province was a very important variable as there can be great variation across provinces regarding political choice. Perhaps the largest change was that the post-stratification dataset did not contain data on the Northwest Territories, Nunavut, and Yukon. While this at first may seem like a reason to not use this variable, we found that the province of the respondent to be very important and stripped away the 3 options not found. Similarly, we’re not losing much (Province Population).

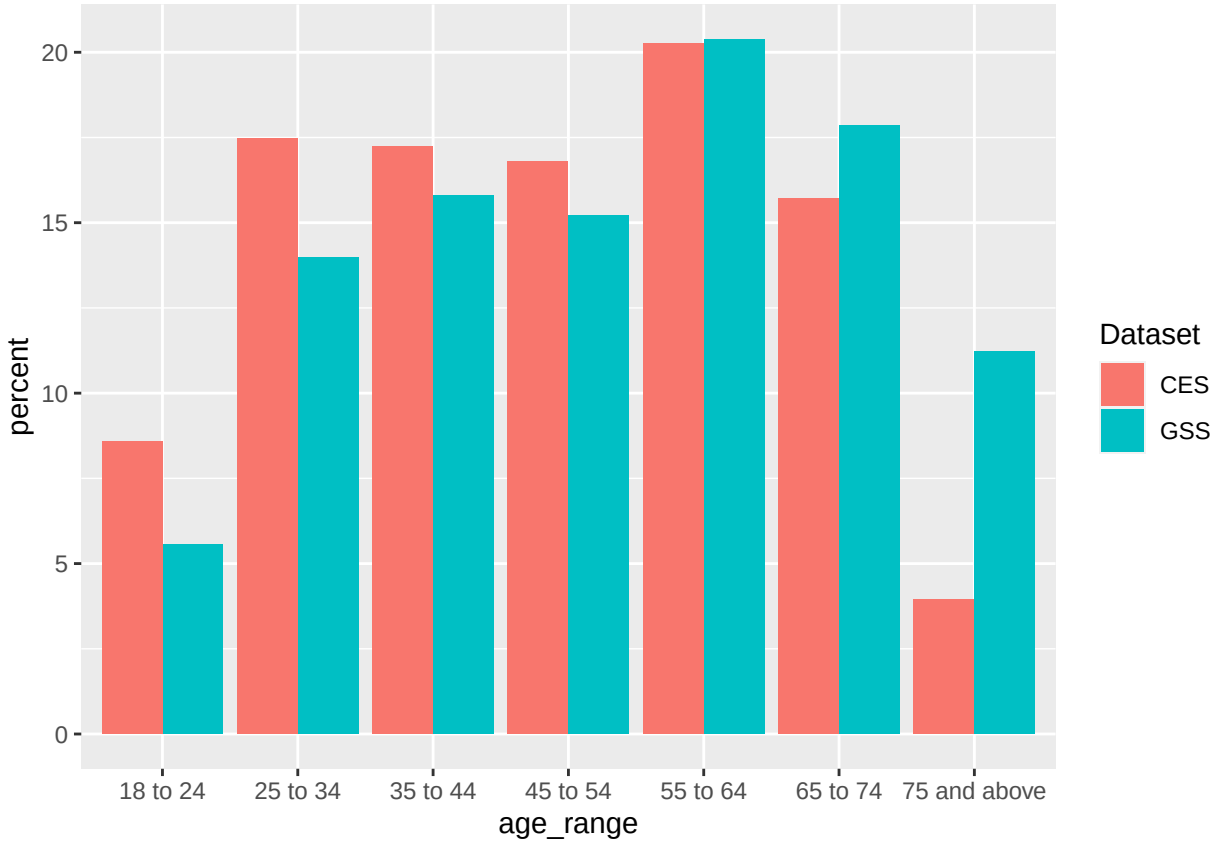
GSS Data

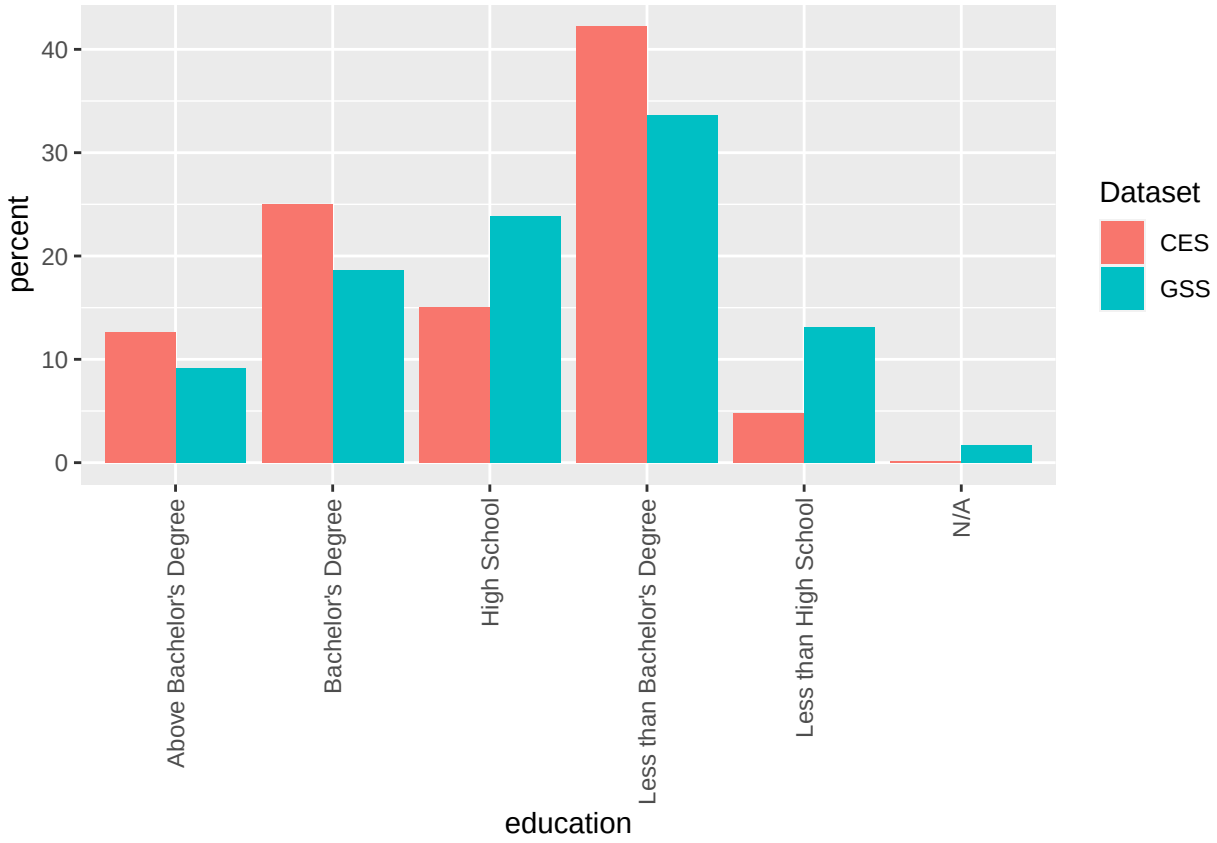
The data used in this report was curated from the General Social Survey (GSS) program., specifically the results of the “General social survey on Family” from 2017. Unfortunately, we cannot provide the data itself, as it is restricted to only those that are allowed to view it. However, if you are a UofT student or have a UofT login, you can follow the steps outlined in “gss_cleaning-1.R” located in the git repository to acquire and clean the data for use.

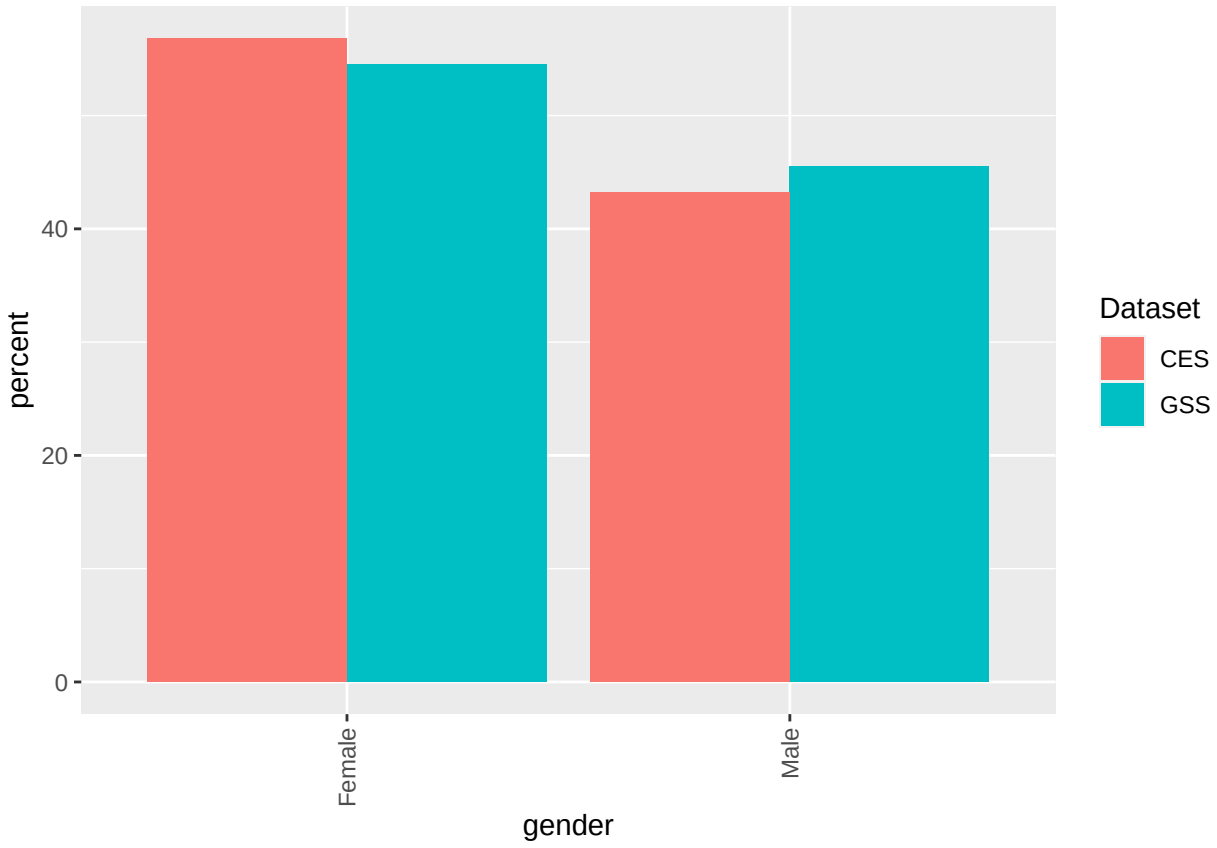
As stated in the overview for the GSS, the target population consisted of those aged 15 and over across all the provinces in Canada, apart from full-time residents of institutions. For this survey, the sample size was approximately 20,000 people, done via Stratified Random Sampling, where the strata were geographic areas within each province. The frame consisted of two aspects: the list of telephone numbers in use available to Statistics Canada and the list of all dwellings within the ten provinces.

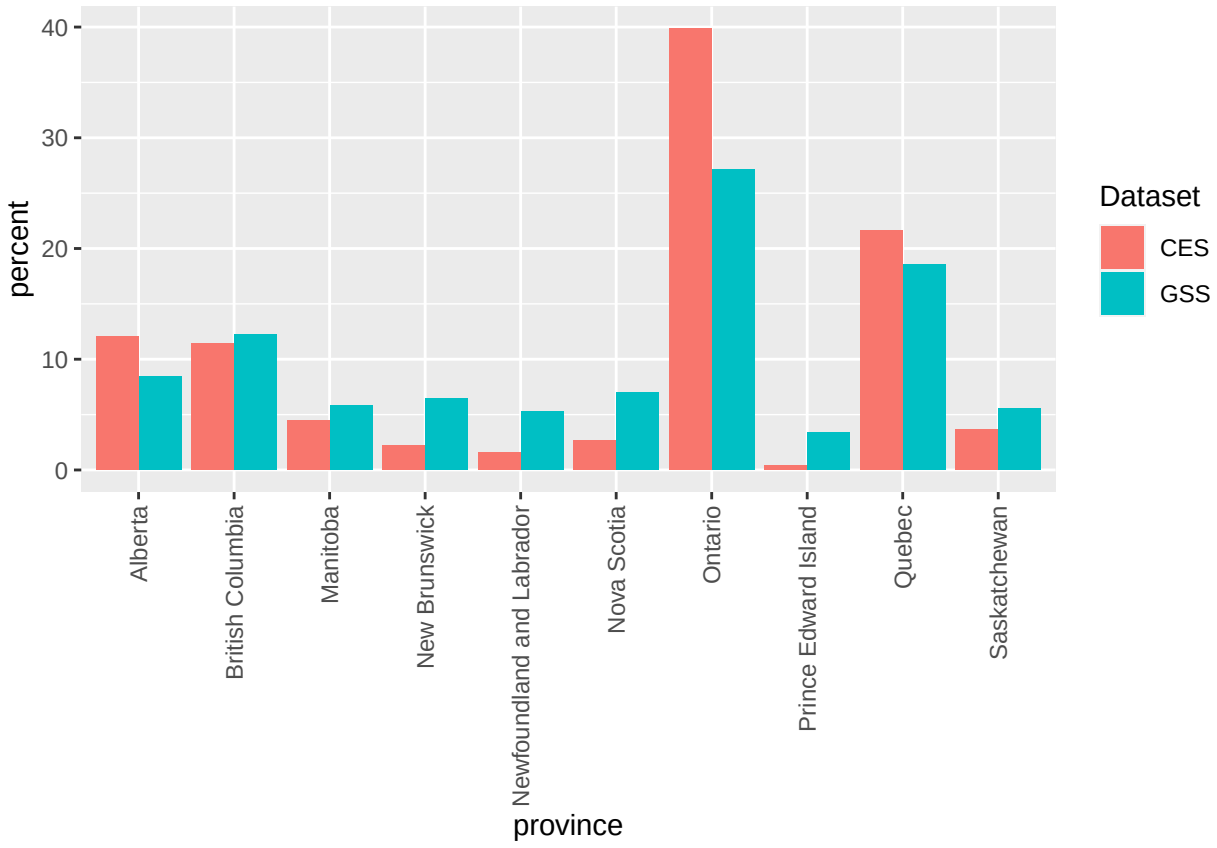
Respondents were contacted via computer assisted telephone interviews. There were many measures for combating non-response. One of which was the resilience shown by the interviewers, including calling up to twice more to a respondent who initially refused, outlining the importance of the survey. In addition, numerous calls were made to those that didn’t pick up. In addition, when dealing with actual non-response, it was not permitted for questions that required weighting. Furthermore, the non-response adjusted accordingly depending on the level of information available.

The dependent variables were the same as from the CES data, with the variable names being age, education, sex, and province. Age and education were the only two variables that had to have refactoring. Age was split into age ranges exactly as before. Education again was a struggle to work through, but was transformed into the same list of values: “Less than High School”, “High School”, “Less than Bachelor’s Degree”, “Bachelor’s Degree”, and “Above Bachelor’s Degree”.




40
30
20
10
0
percent
Above Bachelor's Degree
Bachelor's Degree
High School
Less than Bachelor's Degree
Less than High School
N/A
education
Dataset
CES
GSS


percent
40
20
0
Female
Male
gender
Dataset
CES
GSS



Analysis

Results

Discussion

Appendix (?)

References