

Using MRP to assess the impact of low voter turnout on the outcome of the 2019 Canadian Federal Election

Cathy Yang

December 22nd, 2020

Abstract

Voting is the most fundamental form of political participation that ensures democratic ideals, however, in many advanced democracies, voter participation has been on the decline. This analysis examines whether low voter participation rates can influence the outcome of the 2019 Canadian federal election. We fit multi-level models of party support to demographic and geographic variables of age, sex, education level, province, and electoral district, then apply these models on a post-stratified population of eligible Canadian voters. Findings show that with universal voter turnout, the Liberals win narrowly over the Conservative party, with 157 and 145 seats, respectively, suggesting that non-voters and abstainers bear minimal impact in shifting the electoral outcome in Canada. Keywords: MRP, elections, voter turnout, surveys.

Code supporting this analysis can be found at: <https://github.com/c-myang/sta304-final>

Introduction

Voting is the most fundamental and basic form of political participation that ensures democratic ideals and equal representation of all citizens. However, in many advanced democracies across the world, voter participation has been on the decline, and Canada is no exception to this trend (Lijphart, 1997). In Canada, the average voter turnout rate is 70.5% for national elections since 1867, however, this rate has been decreasing over the past 50 years (Rubenson et al., 2007). Voter turnout in the 3 most recent elections risen from 61% in 2011 to 68.3% and 67% in 2015 and 2019 respectively, however the voter turnout rate has not been above 70% since 1992 (Elections Canada, 2019b).

In recent elections, most particularly in 2011, low political participation among the younger demographic was believed to drive party support towards the Conservative party (Grenier, 2019). However, the impact of low voter turnout has been minimally investigated, with mixed results. Fowler (2011) demonstrates that the introduction of compulsory voting in Australia, led to an increase in Labor Party support and increased pension spending policy. Meanwhile, Rubenson et al. (2007) found that simulating increased voter turnout for the 2000 Canadian federal election did not significantly impact its outcome.

This analysis will examine whether an increased level of voter turnout has implications in the 2019 Canadian federal election. To do so, we employ a multi-level regression and post-stratification (MRP) strategy. MRP has been widely used to estimate national opinion on a wide range of issues, from climate change (Mildenberger et al., 2016), to COVID-19 reopening policy (Briggs, 2020), and has also been used to accurately estimate the result of federal elections (Wang et al., 2015). The advantage of MRP is the ability to impute estimates from a non-representative sample to a population level by weighting estimates based on each demographic subgroup's relative share in the population. In this case, we use data gathered from the 2019 Canadian Election Study (CES) (Stephenson et al., 2020) to model support for each of the 6 major federal parties: Liberal, Conservative, New Democratic Party (NDP), Bloc Québécois (BQ), Green, and People's

Party of Canada. We then impute estimates of party support to the population of eligible voters in Canada, derived from the 2016 census (Statistics Canada, 2017).

This investigation is organized as follows. The Methods section describes the 2019 CES dataset, variable justification and data cleaning procedures, model specifics, and post-stratification calculations. Then, descriptive statistics and findings of the analysis are presented in the Results section. The Discussion section goes over the interpretations and implications of our findings, and finally, we conclude the paper with weaknesses and next steps.

Methods

MRP consists of a two-stage estimation procedure. In the first stage, multi-level regression models are fitted to estimate the relationship between individual and group-level covariates and the outcome of interest (Little, 1993). In the second stage, the predictions of each model are applied for each subgroup belonging to a different combination of individual and group-level variables, and re-weighted based on the subgroup’s share of the population. Together, this approach can predict an outcome variable to a large population based on demographic or geographic factors.

All data cleaning, modeling, and calculations were carried out using R statistical language (R Core Team, 2020) and the tidyverse and lme4 packages (Wickham et al., 2019; Bates et al., 2015). The packages ggplot, sjPlot, and sf, were used for data visualization (Wickham, 2016; Lüdtke, 2020; Pebesma, 2018). The Canadian Election Survey data was accessed from the cesR package (Hodgetts and Alexander, 2020). The data, data cleaning procedures, model specifics and post-stratification calculations are described in subsequent sections.

Data

The multi-level models build from the 2019 Canadian Election Study (Stephenson et al., 2020), a national, large-scale survey conducted during the campaign period prior to the 2019 Canadian federal election, aiming to gather data on Canadians’ voting behaviours, attitudes, and opinions about key political issues. Respondents were sampled in a modified rolling-cross section during from September 13th to October 21st, 2019. The survey and sampling procedures were administered through the Qualtrics online platform. The primary outcome of interest in the CES data, is based on the respondent’s answer to “cps19_votechoice”, which asks what party the respondent thinks they will vote for in the 2019 federal election. The answer choices included one option for each major federal political party in Canada (Conservative, Liberal, NDP, Green, Bloc Québécois, and People’s Party of Canada). Additional answer choices included “Don’t know/Prefer not to answer”, and an option to specify another party. We disregard entries that contain the latter two responses, as they make up a small fraction of the sample and are not outcomes of interest for our models of party preference.

The sampling frame included all Canadian citizens or permanent residents aged 18 or over living in the provinces. To obtain a representative sample of the Canadian population, sampling targets stratified by region and balanced by gender (50% men and 50% women) and age (28% aged 18-34, 33% aged 35-54, 39% aged 55 or higher) were established. While respondents in the territories were not intentionally sampled, they were still eligible for participation. This resulted in a small, non-representative sample from the territories. In total, 37,822 unique responses were gathered from the campaign period survey.

The sampling design of the CES presents some weaknesses. Firstly, restricting the method of survey administration to a purely online format can introduce non-coverage bias, as individuals in the population without computer or internet access are excluded from participation. Secondly, although the sampling design addresses representativeness in terms of age and gender estimates, there may be under-representation of other sociodemographic variables, such as education. Figure 1 compares the demographic composition of the CES respondents against the general electorate, derived from the Statistics Canada 2016 census. Here, we can see similar distributions in terms of age and province, however, there is an over-representation of female

and some postsecondary-educated respondents, and an under-representation of less than high school or high school-educated respondents in the CES survey. Through MRP, we attempt to correct for some of the biases introduced through non-representative sampling.

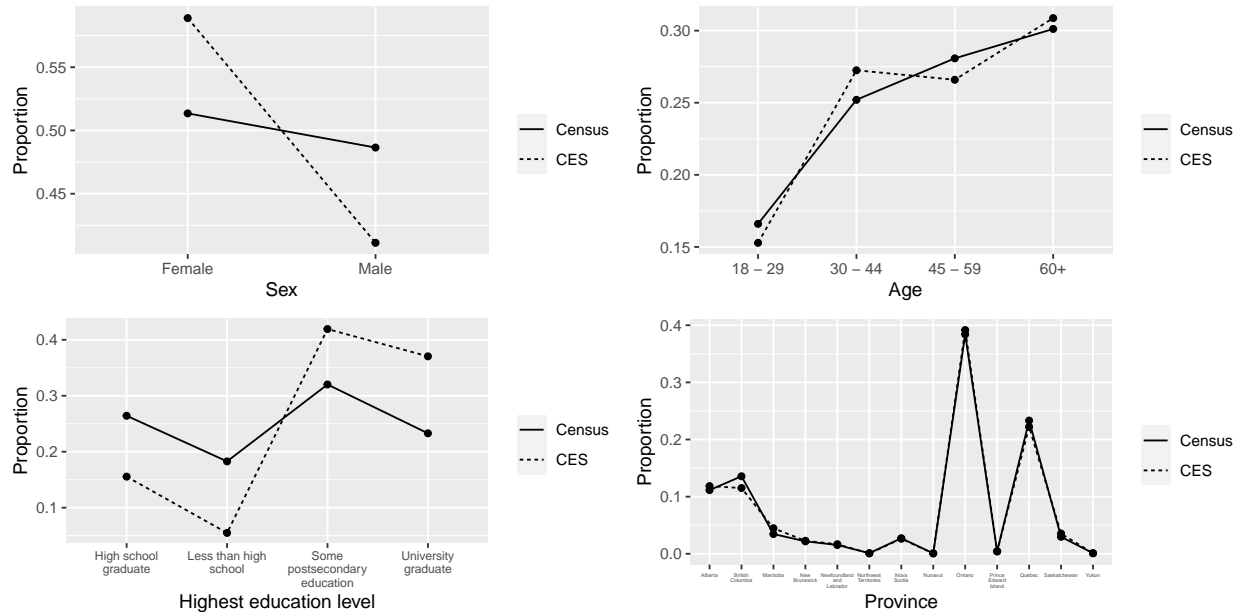


Figure 1: A comparison of the demographic distributions in the CES dataset and the 2016 Canadian census

Variables and data cleaning

Many studies have observed the correlation between age, sex, and education levels on voter preferences (Weakliem, 2002; Wang et al., 2015). As such, we use these as individual-level covariates in our voter preference models. Similarly, the CES data provides sociodemographic data for each respondent. This allows the variables of age, gender, and highest education level to be aggregated into meaningful subgroups that provide the basis for opinion modeling. Age was aggregated into 4 groups: 18-29, 30-44, 45-60, and 60+ and subjects under 18 were filtered out. Additionally, education level was categorized into 4 groups: less than high school, high school graduate, some postsecondary education, and university graduate.

An issue in the data involved the measurement of gender and sex characteristics. While the CES data records gender for all its respondents, the post-stratification data only records sex variables. Conflation of gender and sex measurement in survey and census data is commonly observed and brings about important structural and statistical considerations (Kennedy et al., 2020). For this analysis, we impute all non-male genders as female. While this option will mis-gender a proportion of non-binary respondents, it ultimately avoids the ethical concern of removing all of these individuals from the data (Kennedy et al., 2020). As a result, male gender respondents were recategorized as male sex, while female and non-binary genders were grouped into female sex. From a sociological perspective, this option groups non-binary respondents under the sex that has traditionally faced more structural exclusion and oppression (Kennedy et al., 2020).

We use data from the 2016 Canadian census (Statistics Canada, 2017) as the post-stratification dataset. These data provide cross-tabulations for age and sex counts for each federal electoral district. Age and sex variables were coded under the same classifications as the survey data. Education level counts by age and sex group were not provided in the raw data, however, by multiplying the proportion of individuals of each education group in a federal district population by the count values of each age-sex cell, we could obtain an estimate of the number of people in each education level for each age and sex combination.

Along with demographic data, geographic variables at the level of the provinces and federal electoral districts

were matched to both the census and post-stratification datasets. This is important for predicting the distribution of Parliamentary seats among federal electoral districts, which serve as primary geographic unit that determines the outcome of Canadian federal elections (Elections Canada, 2019b). As such, each response of the is associated with two geographic-level variables: federal electoral district ID and province.

Model Specifics

In the modeling stage, we fit multi-level logistic regression models of party support based on individual and group-level covariates. Individual-level covariates include age, sex and highest education level, while group-level covariates include the province and federal electoral district. A logistic model serves the purpose of this analysis, as support for a federal party can be modelled on a binary basis (yes or no). Since there are 6 major federal parties in Canada, we fit six models of voter preference for each party.

The varying-intercept model also allows us to consider the influence that geography can play in determining party support. Inclusion of provincial and federal electoral districts as geographic covariates in MRP modeling has been widely used to estimate national opinion in Canada (Mildenberger et al., 2016; Briggs, 2020). Following the random intercept notation adapted from Sonderegger et al. (2018) and Mildenberger et al. (2016), the model is expressed as follows:

$$(1) P(Y_{ij} = 1) = \text{logit}^{-1}(\alpha_{0,j[i]}^{province} + \delta_{0,k[i]}^{fed} + \beta_{l[i]}^{age} + \beta_{m[i]}^{sex} + \beta_{o[i]}^{edu})$$

$$\alpha_{0,j[i]}^{province} = r_{00} + r_{01} * W_j$$

$$\delta_{k[i]}^{fed} = r_{00} + r_{01} * W_k$$

Y_{ij} represents support for each major Canadian political party. $\alpha_{0,j[i]}^{province}$ and $\delta_{0,k[i]}^{fed}$ are the random intercepts for each province or territory ($J = 13$) and federal electoral district ($K = 338$), respectively. The remaining terms, denoted by $\beta_{l[i]}^{var}$ correspond to the coefficients associated with each fixed predictor variable, which are age, sex, and education. Subscripts $l[i]$, $m[i]$, $o[i]$ indicate the demographic subgroup for which the i^{th} respondent belongs. For instance, depending on the the i^{th} respondent's age, they can belong in any of the 4 age group cells: β_{18-29}^{age} , β_{30-44}^{age} , β_{45-59}^{age} , β_{60+}^{age} .

Under each random intercept term, r_{00} is the intercept, r_{01} is the slope, and W_j and W_k indicate the value of the province and federal electoral district, respectively.

This model was fitted for each major Canadian party, with the exception of the Bloc Québécois. In Canada, the Bloc Québécois only campaigns for federal seats within the province of Quebec. As a result, support for the party is only observed by respondents located within Quebec. This skew prevented the BQ model from converging. As a compromise, the random intercept term for province, $\alpha_{0,j[i]}^{province}$ was dropped. This allows estimation of support for the BQ party within the province of Quebec, while also keeping a geographic variable in the model.

Post-Stratification

In the post-stratification stage, the census data was split into cells containing population counts for each combination of age, sex and education within each federal electoral district. With 338 federal electoral districts, 2 sex groups, 4 age groups, and 5 education groups, 40,560 unique cells were created. We then use the models of party support to estimate the probability for voting for each of the 6 parties within each cell. Estimation of party support for BQ was only conducted on cells within Quebec.

Next, we determine the party supported by the majority of the population within each federal electoral district. Within each cell, the multi-level regression estimates for each party was multiplied by its relative population proportion within their electoral district. The result produces estimated proportions of the population that support each of the major Canadian parties. The calculation is described as follows:

$$(2) \hat{y}_s^{PS} = \frac{\sum_{j \in J_s} N_j \hat{y}_j}{\sum_{j \in J_s} N_j}$$

where \hat{y}_s^{PS} is the estimate of the proportion of voters for a given party in electoral district s , and J_s is the set of all cells that comprise district s . We then assign a win to the party that has the highest \hat{y}_s^{PS} in the federal electoral district. To determine the outcome of the election, we sum for each party the number of districts where they have a higher proportion of votes to determine the national-level distribution of federal Parliamentary seats.

Results

Descriptive statistics

Figure 2 shows the distribution of party support across sex, age groups, education levels, and province. This shows the results of the voter support before multi-level modeling and post-stratification. The distribution of party support is split between Liberals and Conservatives, followed by the NDP, Green, and Bloc Québécois. The Liberal, NDP and Green parties are more favoured by female voters, while Conservative, BQ, and People's party support are balanced between the sexes. Additionally, there is a larger share of university graduates for Liberal supporters compared to Conservative supporters. Finally, we observe a higher share of voters in Alberta, Saskatchewan and Manitoba in support of the Conservative party, while there is a relatively higher share of supporters for the Liberal party in Quebec and Ontario.

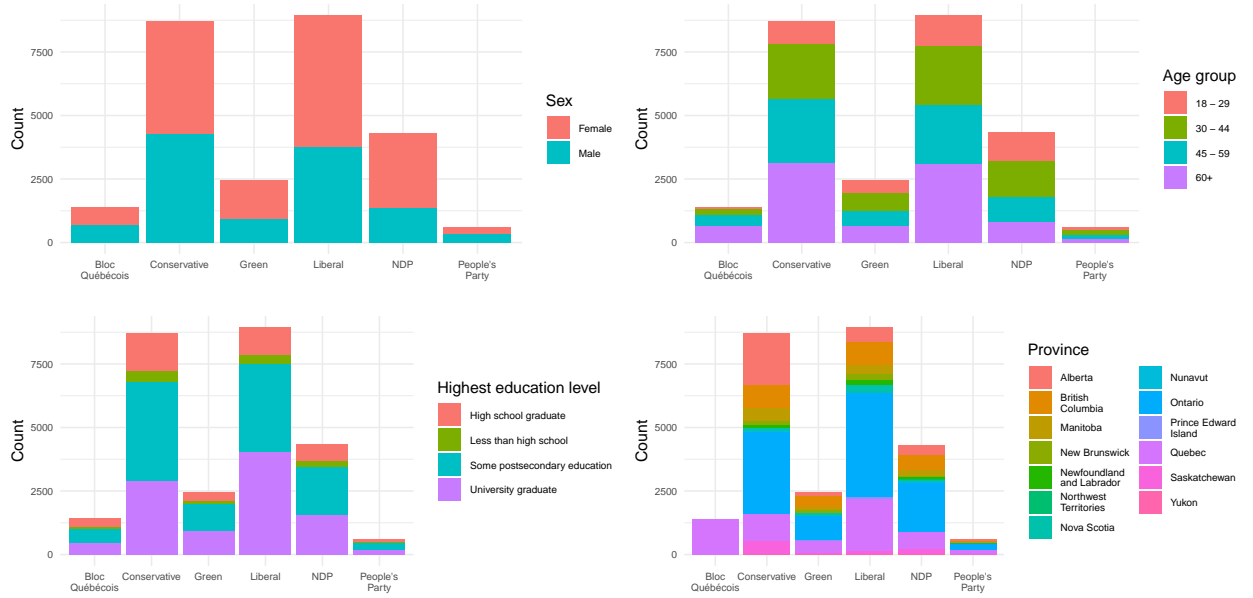


Figure 2: Distribution of demographic variables by vote preference in the Canadian Election Survey, 2019

Model results

Figure 3 shows the results of the multi-level models, comparing the odds ratio estimates of voter support for each party based on fixed effect variables of age, sex, and education level. These values are summarized in a table in Appendix A. In terms of individual-level effects, the models indicate greater support for the Conservative and Green parties and a decrease in support for the NDP among older age groups. Additionally, male sex increases the odds ratio for Conservative, Green and People's party support, and decreases support

for NDP. For education, models suggest some postsecondary education positively influences support for the Liberal party and decreases support for the People’s and Conservative parties.

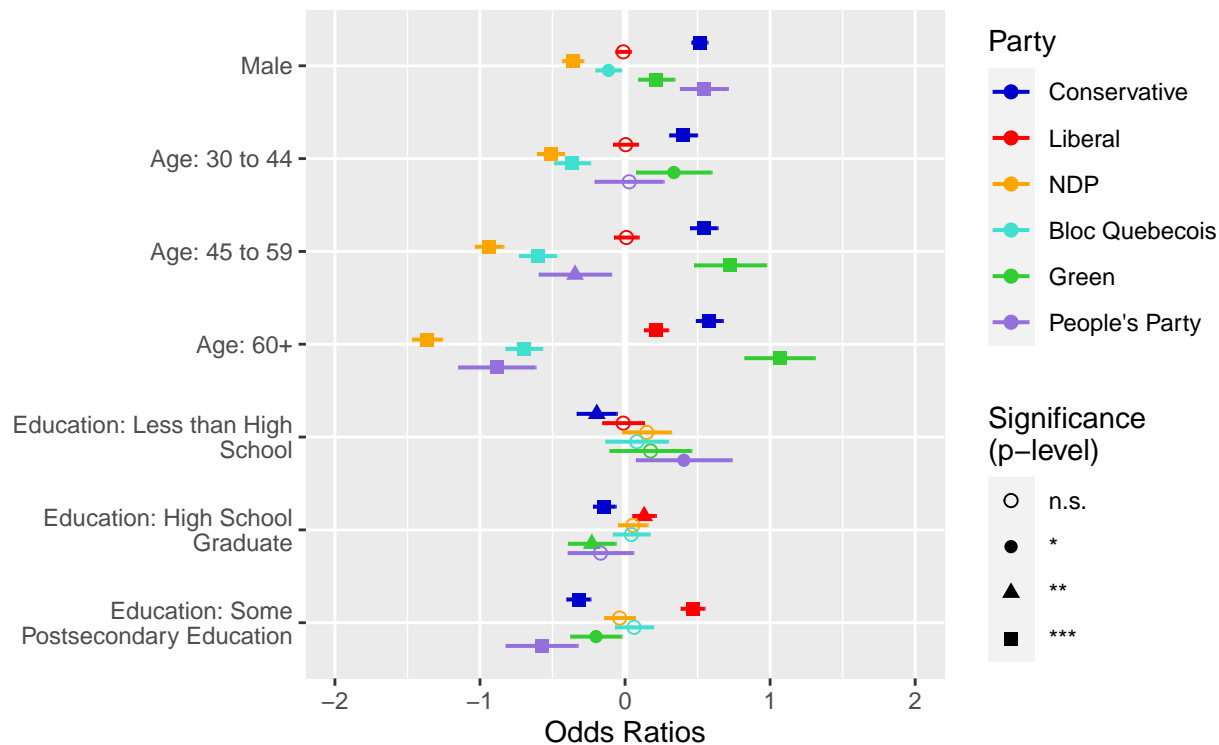


Figure 3: Fixed effects estimates for the multi-level model of voter preference for each major Canadian federal party

Post-stratification estimates

Table 1 shows the estimated distribution of federal Parliamentary seats following the poststratification calculation at the scale of each electoral district. Figure 4 shows the geographic distribution of votes across Canada. This estimation suggests that if everyone was to vote in the 2019 Canadian federal election, the Liberal party would have a narrow margin of victory with 157 Parliamentary seats. Conservatives would follow with 145 seats, Bloq Québécois with 25, NDP with 10, and Green with 1. In comparison to the 2019 election results, there is no change in seats for the Liberal party, and a significant increase in the share of Conservative seats at the expense of NDP and Bloc Québécois seats. This result suggests that there is no large change in the election outcome if everyone voted in the 2019 federal election.

Table 1: Estimated distribution of Parliamentary seats compared against the 2019 Canadian federal election results

Party	MRP estimate	Official results
Liberal	157	157
Conservative	145	121
NDP	10	24
BQ	25	32
Green	1	3
People's	0	0

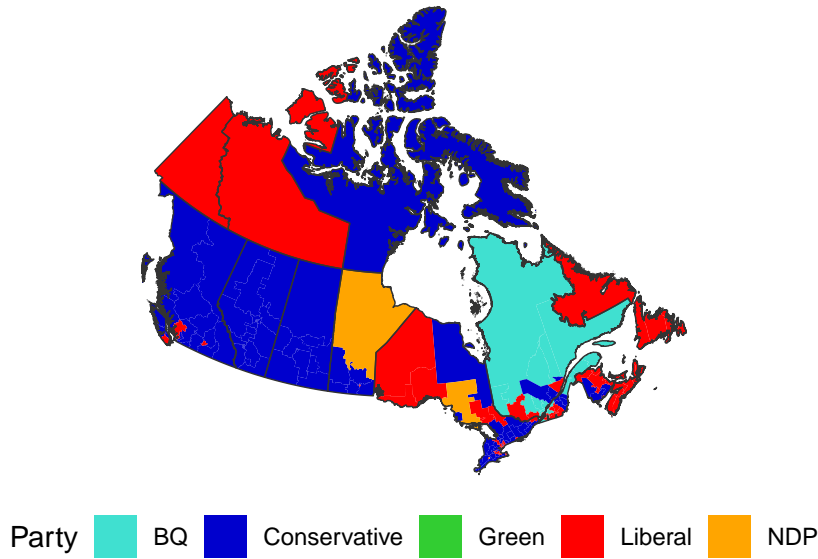


Figure 4: Map of estimated electoral vote distribution if everyone voted in the 2019 Canadian federal election

Discussion

Conventional wisdom generally assumes that low turnout can bias election outcomes toward right wing parties at the expense of left-leaning parties (Lijphart, 1997). The reasoning behind this is that citizens who are better off may be more likely to vote and have marked differences in political affiliation than those who do not vote (Rubenson et al., 2007). We tested this assumption for the 2019 federal election, using Canadian Election Study data to fit a multi-level logistic models of voter for support for each political party based on demographic and geographic variables, then imputing these estimates to the Canadian electorate population through post-stratification. Findings suggest that if everyone voted in the election, the Liberal Party would still hold on to a minority government, keeping hold of the same number of Parliamentary seats.

We found limited evidence that non-voters and abstainers hold significant ability to shift the electoral outcome towards left-leaning parties. Rather, we find a significant increase in support for the Conservative party, gaining 24 seats at the expense of left-of-centre parties such as the NDP and the Bloc Québécois (Table 1). This may be due to a number of factors. Firstly, our multi-level models suggest that support for Conservative party was greatly favoured among older age groups of 30 to 44, 45 to 59, and 60 and above. Meanwhile, support for the NDP and Bloc Québécois significantly decreases with age (Figure 3). Despite

this polarity in party support between age groups, in Canada, the proportion of the younger demographic is relatively low compared to older age groups, suggesting that even with higher voter turnout rates, there would not be any shift towards left-leaning parties. Additionally, a higher-educated demographic also tends to favour Liberal party support in the multi-level models. However, with a higher proportion of less than or high school-educated individuals in the electorate compared to the CES survey, (Figure 1), the post-stratification calculations suggest increased support in the population for the Conservative party.

Geography also plays a role in influencing political opinion, as Figure 4 displays overwhelming support for the Conservative party in the prairie provinces of Alberta, Saskatchewan, Manitoba, and in rural Ontario. Meanwhile, Liberal and NDP support is evident in large city centres, such as Toronto, Vancouver, and Montreal, and also in the Maritime provinces. The geographic distribution from our estimation overall does not change much from the actual election outcome (Appendix B), with the exception of a few electoral districts in west British Columbia and Quebec.

Weaknesses and next steps

There were several limitations in the analysis which could be improved upon in future investigations. Firstly, the CES did not intentionally sample from residents residing in the territories, however, respondents were still able to participate. This resulted in a very sparse sample size, which may have skewed support towards parties that may not be accurate of true party support in these populations. Given that the aim of this analysis was to determine the outcome of the election if the entire eligible population voted, the territories were kept in the survey. However, they could have been removed and the analysis could have been restricted to the provinces to obtain more accurate modeling results.

Additionally, our modeling and post-stratification calculations could have been conducted on a wider variety of demographic variables, such as income/class, or race. These variables have been used previously in MRP-based simulations of electoral outcomes (Wang et al., 2015; Rubenson et al., 2007). However, in the CES survey, the coding of race was on a “choose all that apply” basis, presenting challenges in coding respondents to mutually exclusive race categories. Additionally, the response rates for income in the survey were very low, which significantly decreases the sample size for modeling stage of MRP. Future work could use political opinion surveys with more robust race and income data to determine whether varying political affiliations along race and class lines would impact the outcome of future elections.

Finally, this analysis does not assess how the political opinions of non-voters or abstainers may differ from active voters. Some political science scholars argue that if abstainers were to become mobilized to vote, their political preferences would also change (Lijphart, 1997). As such, an alternative method could have fitted party preference models based solely on the respondents in the CES who specify they are likely to vote. This analysis would then follow the assumption that if everyone in the eligible population actively participated in elections, that they would also have a strong shift in voter opinion that may shift the political climate of the country towards a different direction.

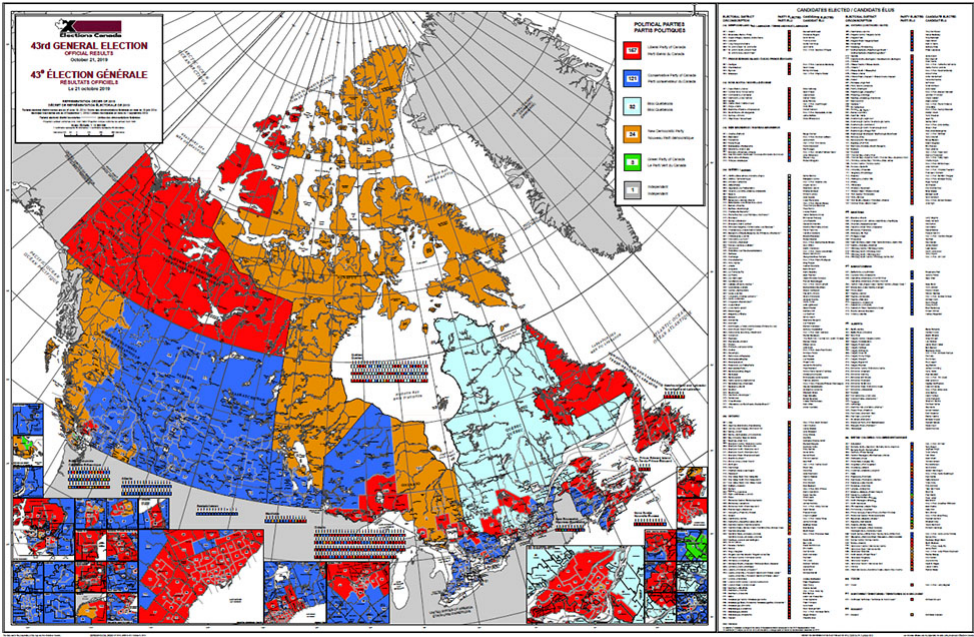
In summary, while there are many consequences of decreased voter participation that have yet to be explored, in our analysis, we do not observe a change in the electoral outcome for the 2019 Canadian federal election. This is mainly attributed to the demographic make-up in Canada, in which the older demographic outnumbers the younger demographic. Although we do not find any significant shift in Canada, this analysis could be of use to examine whether electoral outcomes may change for countries with a more significant share of eligible young voters.

Appendix

Appendix A: Table of model estimates of voter preference for each major federal Canadian party

Model estimates of voter preference for each major federal Canadian party																		
Predictors	Conservative			Liberal			NDP			Bloc Quebecois			Green			People's Party		
	Odds Ratios	CI	p	Odds Ratios	CI	p	Odds Ratios	CI	p	Odds Ratios	CI	p	Odds Ratios	CI	p	Odds Ratios	CI	p
Intercept	0.22	0.15 – 0.32	<0.001	0.26	0.20 – 0.35	<0.001	0.33	0.26 – 0.42	<0.001	0.12	0.09 – 0.16	<0.001	0.14	0.09 – 0.20	<0.001	0.02	0.02 – 0.03	<0.001
Male	1.67	1.58 – 1.77	<0.001	0.99	0.94 – 1.04	0.641	0.70	0.65 – 0.75	<0.001	1.24	1.10 – 1.41	0.001	0.89	0.82 – 0.97	0.011	1.72	1.46 – 2.03	<0.001
Age: 30 to 44	1.49	1.36 – 1.64	<0.001	1.00	0.92 – 1.09	0.925	0.60	0.55 – 0.66	<0.001	1.40	1.08 – 1.82	0.011	0.69	0.61 – 0.79	<0.001	1.03	0.81 – 1.30	0.813
Age: 45 to 59	1.72	1.57 – 1.89	<0.001	1.01	0.93 – 1.10	0.831	0.39	0.36 – 0.43	<0.001	2.06	1.61 – 2.65	<0.001	0.55	0.48 – 0.62	<0.001	0.71	0.55 – 0.91	0.007
Age: 60+	1.79	1.63 – 1.96	<0.001	1.24	1.14 – 1.35	<0.001	0.26	0.23 – 0.28	<0.001	2.90	2.28 – 3.70	<0.001	0.50	0.44 – 0.57	<0.001	0.41	0.32 – 0.54	<0.001
Education: Less than High School	0.82	0.72 – 0.94	0.006	0.99	0.86 – 1.14	0.863	1.16	0.98 – 1.37	0.082	1.19	0.90 – 1.58	0.222	1.08	0.87 – 1.35	0.461	1.50	1.08 – 2.09	0.016
Education: Some Postsecondary Education	0.87	0.80 – 0.94	<0.001	1.14	1.05 – 1.24	0.001	1.06	0.95 – 1.17	0.297	0.80	0.67 – 0.94	0.006	1.04	0.92 – 1.19	0.503	0.84	0.67 – 1.06	0.142
Education: University Graduate	0.73	0.67 – 0.79	<0.001	1.59	1.47 – 1.73	<0.001	0.96	0.87 – 1.07	0.489	0.82	0.69 – 0.98	0.026	1.06	0.93 – 1.21	0.352	0.56	0.44 – 0.72	<0.001
Random Effects																		
σ ²	3.29			3.29			3.29			3.29			3.29			3.29		
τ ₀₀	0.20 constituencynumber			0.14 constituencynumber			0.24 constituencynumber			0.33 constituencynumber			0.14 constituencynumber			0.20 constituencynumber		
	0.37 province			0.21 province			0.11 province						0.34 province			0.00 province		
ICC	0.15			0.10			0.10			0.09			0.13			0.06		
N	13 province			13 province			13 province			78 constituencynumber			13 province			13 province		
	338 constituencynumber			338 constituencynumber			338 constituencynumber						338 constituencynumber			338 constituencynumber		
Observations	31054			31054			31054			6961			31054			31054		
Marginal R ² / Conditional R ²	0.031 / 0.174			0.012 / 0.106			0.071 / 0.161			0.047 / 0.135			0.016 / 0.141			0.065 / 0.120		

Appendix B: Map of the official results for the 2019 Canadian federal election. From Elections Canada (2019a)



References

- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1):1–48.
- Briggs, R. C. (2020). Canadians Think that Nearly All of Us Will Be Allowed Back to Work around August. Canadian Journal of Political Science/Revue canadienne de science politique, 53(2):245–251. Publisher: Cambridge University Press.
- Elections Canada (2019a). 43rd General Election – October 21, 2019. Last Modified: October-26-2020.
- Elections Canada (2019b). Voter Turnout at Federal Elections and Referendums. Last Modified: December-08-2020.
- Fowler, A. (2011). Electoral and Policy Consequences of Voter Turnout: Evidence from Compulsory Voting in Australia. SSRN Scholarly Paper ID 1816649, Social Science Research Network, Rochester, NY.
- Grenier, E. (2019). How voter turnout might affect who wins in October. CBC News.
- Hodgetts, P. A. and Alexander, R. (2020). cesR: Access the CES Datasets a Little Easier. R package version 0.1.0.
- Kennedy, L., Khanna, K., Simpson, D., and Gelman, A. (2020). Using sex and gender in survey adjustment. arXiv:2009.14401 [stat]. arXiv: 2009.14401.
- Lijphart, A. (1997). Unequal Participation: Democracy’s Unresolved Dilemma. The American Political Science Review, 91(1):1–14. Publisher: [American Political Science Association, Cambridge University Press].
- Little, R. (1993). Post-stratification: A modeler’s perspective. Journal of the American Statistical Association, 88(423):1001–1012.
- Lüdtke, D. (2020). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.5.
- Mildenberger, M., Howe, P., Lachapelle, E., Stokes, L., Marlon, J., and Gravelle, T. (2016). The Distribution of Climate Change Public Opinion in Canada. PLOS ONE, 11(8):e0159774. Publisher: Public Library of Science.
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal, 10(1):439–446.
- R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rubenson, D., Blais, A., Fournier, P., Gidengil, E., and Nevitte, N. (2007). Does low turnout matter? Evidence from the 2000 Canadian federal election. Electoral Studies, 26(3):589–597. Publisher: Elsevier Science.
- Sonderegger, M., Wagner, M., and Torreira, F. (2018). Mixed-effects logistic regression. In Quantitative Methods for Linguistic Data. McGill Department of Linguistics.
- Statistics Canada (2017). Age Groups (15), Sex (2), and Selected Education Characteristics (7), for the Total Population of Canada and Federal Electoral Districts (2013 Representation Order), 2016 Census [Table]. Retrieved from <http://dc.chass.utoronto.ca.myaccess.library.utoronto.ca/cgi-bin/census/2016/displayCensus.cgi?year=2016&geo=fed>.
- Stephenson, L. B., Harell, A., Rubenson, D., and Loewen, P. J. (2020). 2019 Canadian Election Study - Online Survey. Publisher: Harvard Dataverse type: dataset.
- Wang, W., Rothschild, D., Goel, S., and Gelman, A. (2015). Forecasting elections with non-representative polls. International Journal of Forecasting, 31(3):980–991.

- Weakliem, D. (2002). The Effects of Education on Political Opinions: An International Study. International Journal for Quality in Health Care - INT J QUAL HEALTH CARE, 14:141–157.
- Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43):1686.