Turning *Up* The Vote: The 2019 Federal Election at Full Voter Turnout

Nicholas Callow

December 22, 2020

Abstract

Voter turnout is significant to both democracy and electoral chances. Yet, every federal election cycle, millions of eligible Canadians do not cast a ballot. This paper considers the 2019 Federal Election at full voter turnout. Using several logistic regression models with post-stratification, we estimate changes in vote shares under the assumption that everyone participates in the political process. Our results show changes in vote share for every political party. However, centerist and regional parties, such as the Liberals and Bloc Quebecois, would capture fewer votes at this level of engagement, compared to their more partisan counterparts. Code and data supporting this analysis are available at: https://github.com/ncallow/turningupthevote (https://github.com/ncallow/turningupthevote).

Keywords: Canadian Politics, Federal Election 2019, Voter Turnout, Logistic Regression, Post-Stratification

Introduction

The significance of voter turnout to democracy cannot be understated. Declines in this figure raise questions about the representativeness of government bodies and the integrity of the electoral process (Achen, 2019; Birch, 2010). However, this aggregated parameter does not tell us the full story. Discrepancies in voter turnout are not distributed equally. Age is a particularly salient variable, with young people less likely to vote than their middle-age and older counterparts (Statistics Canada, 2020; Uppal & LaRochelle-Cote, 2012). However, other sociodemographic factors, including education, family status, economic well-being, and immigrant status, are also significant predictors of voter participation (Uppal & LaRochelle-Cote, 2012). Moreover, research shows that these groups have relatively stable voting preferences. For instance, individuals with undergraduate and post-graduate degrees show increasing preferences for leftist parties and policies (Pew Research Center, 2018). Likewise, young and middle-aged voters are more likely to vote for liberal parties than their older counterparts (Pew Research Center, 2018). As such, who votes and in what numbers can significantly impact electoral outcomes.

A surprising number of Canadians do not vote. In the last federal election, there were approximately twenty-seven million eligible voters (Elections Canada, 2020a). Only sixty-seven percent of these individuals cast their ballot, leaving over nine million Canadian voices unheard (Elections Canada, 2020a). Given the significance of voter turnout to election results and democracy, a question remains about how the most recent election would have differed if all eligible Canadians participated in the vote? This paper will use several logistic regression models with post-stratification to answer this question. More specifically, it will compare every major political parties' vote shares from the most recent federal election with their predicted support at full voter turnout.

This analysis will involve two data sets; the 2019 Canadian Election Study (CES) and the 2017 General Social Survey (GSS). The methodology section will explain each data set's purpose, the predictor and outcome variables of the logistic regression model, and the process of post-stratification. The following results section will outline the post-stratified model predictions for propensity to vote for each major political party. The paper will then conclude with a discussion of relevant findings, acknowledgement of limitations, future research directions, and larger societal implications.

Methodology

Data: Individual-Level Survey

The individual-level data comes from two waves of the 2019 Canadian Election Study (CES). These polls considers structural elements of Canadian politics and society while examining contemporary political behaviour, beliefs, and actions (Stephenson et al., 2020a). Data collection for the campaign-period survey took place between September 13 and October 21, in the lead-up to the most recent federal election (Stephenson et al., 2020b). The post-election survey took place between October 24 and November 11 of 2019, immediately following the election (Stephenson et al., 2020b). The target population[1] consists of eligible electors, Canadian citizens and permanent residents over eighteen years of age who live in the ten provinces (Stephenson et al., 2020b). As a result, the political preferences and behaviours of individuals from the territories are excluded from this analysis, and this represents a fundamental weakness in the data. There is little detail on the construction of the CES sampling frame[2]. Researchers hired Qualtrics to construct the sample, who drew from several proprietary market research panels[3], providing no details on their process (Stephenson et al., 2020b).

Likewise, there is little explicit information on the sampling approach[4] undertaken by Qualtrics on behalf of the CES research team. However, the information provided by both this firm and the researchers seems to suggest a probability sampling approach[5]. Qualtrics (n.d.) explains that their research panels are often “selected carefully so that they represent a sample of your target population” (para 2). Moreover, Stephenson et al. (2020b) explain that the target population was stratified by gender, geographic location, and language. These factors taken together point towards a stratified random sampling method[6]. If this was the strategy employed, there are several benefits and drawbacks. Probability sampling allows for inferences about the larger population, a crucial requirement for this paper (Caetano, 2020a). Additionally, this approach emphasizes the validity and reliability of measures, as sampling errors can be calculated and results reproduced by other researchers (Caetano, 2020a). However, probability sampling is a costly and time-consuming endeavour, heavily reliant on the sampling frame’s accuracy and subject to non-participation (Caetano, 2020a). These conclusions about the sampling approach should be interpreted very cautiously, given significant amounts of missing information.

In total, over seventy-thousand and thirteen-thousand individuals interacted with the CES campaign period and post election surveys, respectively (Stephenson et al., 2020b). Amongst these groups, some elected not to participate or were screened out by the researchers. Several determining factors for excluding responses included ineligibility to vote, could not match to campaign period survey, postal code mismatch, and speeding through the survey (Stephenson et al., 2020b). After these data cleaning procedures, a total of $n = 10,337$ respondents made up the sampled population[7] as they completed both the campaign period and post election surveys. We reduced this value to $n = 8,197$ after further cleaning and exclusions (see Data: Inclusions, Exclusions, and Modifications). The presence of missing data is a persistent concern for survey research. However, this is not a significant issue in this data set, as each of the sociodemographic predictor contain no missing data, and any ommissions in the vote choice variable are minimal.

Data: Census-Level Survey

The census-level data comes from the 2017 General Social Survey (GSS). This survey took place between February 2 and November 30 of 2017, and considers the changing role of families in Canadians' lives (Beaupré, 2020). The target population (see Note 1) of the GSS includes all individuals older than fifteen years of age of the ten provinces. Like the CES, this desired demographic excludes the territories in addition to residents of institutions (Beaupré, 2020). Statistics Canada constructs the sampling frame (see Note 2) using a two-stage approach. First, personnel pull landline and cellular numbers from several sources, including the Census and phone companies (Beaupré, 2020). Statistics Canada then associates these phone numbers with a specific household using the Address Register, a list of all the provinces' residences (Beaupré, 2020). This strategy does produce some discrepancies between the target and frame populations, introducing coverage error (see Note 2) into the survey data. Expressly, individuals without a known phone number are excluded from the analysis. However, the inclusion of households with only cellular phone numbers has improved coverage since this group "constitutes a constantly growing portion of the population" (Beaupré, 2020, p. 10).

The GSS employs a stratified random sampling approach (see Note 6). Statistics Canada created strata by geographic area, relying on information from other sources, including the Census (Beaupré, 2020). A simple random sample without replacement was then procured in each group (Beaupré, 2020). There are numerous benefits and drawbacks of this probability sampling approach, discussed above (see Data: Individual-Level Survey). The sampled population (see Note 7) came to a total of $n = 20,602$, yielding a response rate of 52.4% (Beaupré, 2020). This figure was further reduced to $n = 19,841$ after additional cleaning and exclusions (see Data: Inclusions, Exclusions, Modifications). Missing data on the variables of interest is minimal or non-existent. Therefore, the GSS is considered appropriate census-level data for this analysis.

Data: Inclusions, Exclusions, Modifications

We elected to exclude several additional respondents from the data, in addition to the cleaning processes completed by the CES research team. First, we removed all individuals ($n = 4$) who reported no formal schooling as their level of education. The likelihood that a person has never attended school in Canada, yet is completing this election survey is very rare. Moreover, the CES contains a category for individuals who attended some elementary school. Therefore, we excluded these respondents from the survey given the chance this option was selected by accident and the small subset size. Second, individuals who responded "Don't know / prefer not to answer" ($n = 12$) on the CES education variable were removed. Including this group in our model offered little additional benefit or insight. Therefore, we filtered them out of the data set. Third, we excluded participants who left the vote choice question blank. Since the goal of this paper is to predict political party support, including these individuals in the sample would have been meaningless. Fourth, age was recoded from a continuous to ordinal-level variable. We made this change because at the continuous level, there were so many combinations of variables that the bins used to calculate the post-stratification estimates were too small to make a reasonable prediction. Finally, we recoded gender to match the sex variable in the GSS (see Discussion: Weaknesses).

Additional changes to the GSS were fairly minimal. First, we excluded individuals under the age of eighteen since they are ineligible to vote. Including this group in our model would have compromised the results, given research that demonstrates young people are more likely to vote for left-leaning parties (Pew Research Center, 2018). Lastly, we removed individuals who left their level of education blank so as to match the CES data set.

Model: Logistic Regression

This paper will employ several logistic regression models with post-stratification to estimate each parties vote share at full turnout. In the CES data set, we created eight binary variables, coded 1 if an individuals plans to vote for a given political party, and 0 otherwise. Logistic regression was selected due to the binary nature of the outcome variables, voting for the one political party over all others (Caetano, 2020c). Likewise, because we have no prior information on the variables of interest, we chose a Frequentist approach over comparable Bayesian methods (Caetano, 2020d).

First, we will use the individual-level survey data from the Canadian Election Study to construct a logistic regression model. This model estimates the likelihood that an individual with a given sociodemographic background will vote for a specified political party over the competing options. For comparability, the predictor variables remain the same across all eight models. The logistic regression models take the following mathematical form. Please refer to Table 1 for further detail on variable categories and reference levels.

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \underbrace{\left(\hat{\beta}_2 x_{2,1} + \cdots + \hat{\beta}_4 x_{2,3}\right)}_{\text{Age Categorical Variables}} + \underbrace{\left(\hat{\beta}_5 x_{3,1} + \cdots + \hat{\beta}_8 x_{3,4}\right)}_{\text{Education Categorical Variables}}$$

Table 1: Breakdown of Variable Categories

Notation	**Variable**	**Category**	**Reference Level**
x_1	Sex	Male	Female
$x_{2,1}$	Age	30-49 years old	18-29 years old.
$x_{2,2}$	Age	50-64 years old	18-29 years old.
$x_{2,3}$	Age	65 years and older	18-29 years old.
$x_{3,1}$	Education	Below Bachelors	Bachelors
$x_{3,2}$	Education	Below High School	Bachelors
$x_{3,3}$	Education	Graduate Degree	Bachelors
$x_{3,4}$	Education	High School	Bachelors

The left side of this equation produces the log odds of a binary result, where p is the probability that the event will take place (Caetano, 2020c). In this paper, this translates to the log-odds that an individual will vote for a specific political party (e.g. Liberals) instead of someone else. The slope coefficients $\hat{\beta}_1, \ldots, \hat{\beta}_8$, represent the change in log-odds associated with a given predictor variable $x_1, \ldots, x_{3,4}$. We have included three such variables, age, sex, and education. Lastly, $\hat{\beta}_0$ is the logistic regression intercept. In our model, this corresponds to the likelihood that an eighteen to twenty-nine-year-old female with a bachelor's degree will vote for a given party.

We selected these three predictor variables given available research on voter preferences. The Pew Research Center (2018) validates that these sociodemographic variables are consistently significant predictors of political behaviour. For instance, there is a growing gender gap, where women are significantly more likely to vote for leftist parties than their male counterparts (Pew Research Center, 2018). Similarly, educational attainment continues to inform partisan behaviour in meaningful ways. While individuals with a post-secondary education gravitate towards liberal policies and groups, those without university education are split between the left and the right (Pew Research Center, 2018). Finally, age is a fundamental element of political support. The Pew Research Center (2018) concludes that the "generational gap in partisanship is now more pronounced than in the past" and continues to widen (para 28). Younger individuals are consistently more likely to vote for leftist parties than older generations (Pew Research Center, 2018). These well-established voting patterns lend support to their inclusion in our model.

Model: Post-Stratification

Now, we will use the census data to post-stratify our individual-level results. This technique allows us to extrapolate beyond the CES respondents and consider the likelihood of voting for a given political party amongst the larger Canadian electorate. We calculate our post-stratified estimates using the following formula (Caetano, 2020e).

$$\hat{y}^{ps} = \frac{\sum N_j \hat{y}_j}{\sum N_j}$$

The post-stratification process consists of the following steps. First, we divide our CES sample data into j subsets by enumerating every combination of our three predictor variables (Caetano, 2020e). Therefore, one bin will include all eighteen-year-old females with a bachelor's degree. Another cell will contain fourty-five-year-old males with a graduate degree. These examples could continue indefinitely as we consider every possible combination of responses to these sociodemographic variables. Then, we compute an estimate of our response variable y_j, propensity to vote for specified political party, for each of these subsets using our logistic regression model (Caetano, 2020e). Finally, we weigh these estimates according to the proportion of these individuals in the census-level data (Caetano, 2020e). We sum over all these weighted estimates to compute the final post-stratified estimate $\hat{y}^{ps}$ (Caetano, 2020e).

Results

Table 2: Summary of Post-Stratified Model Results

Political Party	**Actual Vote Share**	**Predicted Vote Share**	**Difference**
Liberal	33.1%	32.73%	-0.37%
Conservative	34.3%	35.01%	$+0.71\%$
Bloc Quebecois	7.6%	5.59%	-2.01%
New Democratic	16.0%	16.43%	$+0.43\%$
Green	6.5%	7.00%	$+0.50\%$
People's	1.6%	2.22%	$+0.62\%$
Other	0.6%	0.52%	-0.08%
Spoil	$\mathrm{N/A}$	0.32%	$\mathrm{N/A}$

Table 2 offers a comparison of our estimated results against the actual vote share totals from the 2019 Federal Election (Elections Canada, n.d.). We find that each political party would experience a change in support at full voter turnout, using our logistic regression model with post-stratification that accounts for age, sex, and education level. Two of the six major parties, the Liberals and Bloc Quebecois, would capture less of the vote at this level of political participation, losing 0.37% and 2.01%, respectively. In contrast, the remaining four serious parties, the Conservatives, NDP, Greens, and PPC, would benefit from full turnout, increasing their vote shares by $0.71\%, 0.43\%, 0.50\%$ and 0.62%, respectively. Finally, fringe parties, such as the Communist Party of Canada, would see their vote share decrease, and spoiling ballots would become less common.

Discussion

Summary

This paper highlights the significance of voting to democracy. Every federal election cycle, millions of eligible Canadians do not cast a ballot. However, the distribution of voter turnout is complex. Research consistently validates that some individuals are more likely to vote than others, and that these groups of people have established political preferences. As such, who votes and in what numbers can significantly impact electoral outcomes. Building on this, we have employed several logistic regression models with post-stratification to estimate vote shares at full turnout. Using the 2019 Canadian Election Study, we constructed eight logistic regression models, each of which predict the probability of voting for one political party based on age, sex, and education sociodemographic. Our results (see Table 1) indicate a shift in support for every political party, and that some of these vote shifts are significant.

Conclusions

This model yields several significant conclusions. First, the data indicates that non-centerist parties, such as the Greens, NDP, PPC, and Conservatives, would see increases in vote share at full turnout. However, the same cannot be said of the Liberals, as well as regional parties like the Bloc Quebecois. This trend may be an indicator of increasing political polarization in federal politics. Elections Canada (2020b) surveyed a group of youth non-voters, asking them their reasons for not participating in the federal election process. This demographic maintains that a principal reason they do not vote is because the government does not reflect their interests or values (Elections Canada, 2020b). These findings, coupled with our data, demonstrate that Canadian voters are becoming increasingly unwilling to compromise on their political beliefs and values. As a result, they are throwing their support behind parties further along both sides of the political spectrum. Moreover, at a higher degree of turnout, party supports becomes increasingly dispersed. Therefore, if this level of political participation occurs, we may see a shift away from sole-party minority and majority governments to coalition structures common in other constitutional monarchies.

Second, many discussions of federal politics typically focus on the Liberals and Conservatives, who effectively run the country on a two-party system. In every modern federal election, one of these two parties has formed the government. On this topic, our findings suggest that the Conservative Party would fair better at higher levels of political participation. Although our estimated increase of 0.71% is modest at best, it may be enough to win increasingly competitive and close elections. Likewise, as parties further from the center gain more of the vote share, the balance of power between the Liberals, Conservatives, and the remainder of the legislature may change. However, our data does dispell any notion that there is a block of non-voters who, if they participated, could significantly change electoral outcomes. Even at full participation, it remains a race between the Liberal and Conservative parties.

Finally, our results suggest that the political preferences of non-voters are complex. Prior research indicates that among individuals who do not participate in the political process, a great number are youth (Elections Canada, 2020b). There is a well-established association between age and voting behaviour, with young voters more likely to select leftist parties (Pew Research Center, 2018). As such, we would expect that at full voter turnout, where these demographics increase in size, that left-leaning parties would take votes away from right-spectrum parties. However, our data shows that the parties with the greatest increases in vote share are the Conservatives and People’s Party. We do not have an explanation for why this is the case. However, future research may further probe the political preferences of non-voters.

Weaknesses and Next Steps

This study is not without its limitations. First, a mismtach exists between the individual-level survey and census data. The Canadian Election Study took place in 2019, meanwhile the General Social Survey was completed in 2017. Unfortunately, this discrepancy was largely unavoidable. At the time of writing this paper, more recent versions of the GSS were not available for analysis. Likewise, readers may wonder why we did not select the 2016 Census of Population. Our choice was informed by two reasons. First, compared to the 2017 GSS, the 2016 Census was much less accessible. Second, the Census of Population was even more outdated than the GSS, which could have further compromised the results. This limitations has several implications. Primarily, the 2016 Census of Population, or a similarily large and comprehensive data set, would have produced more accurarte population proportions. As such, our weighted estimates would better reflect the Canadian population. Likewise, the Census and similar data sets contain more sociodemographic variables, which could have strengthened the logistic regression models.

Second, our selection of logistic regression to compute the individual-level estimates is inefficient for a study of this nature. As explained earlier, logistic regression is used to model relationships where the response variable is binary (Caetano, 2020c). However, our outcome variable of interest is traditionally categorical, as there are numerous political parties to choose from when voting. To overcome this problem, we created eight separate logistic regression models to model support for each individual political party. A more efficient approach would create a singular model using multi-nomial logistic regression. Time limitations and knowledge gaps prevented such an approach. However, future research could pair multi-nomial logistic regression with post-stratification for a less time-consuming and more efficient model.

Third, discrepancies exist between our sample and census data on the issue of sex and gender. Although sometimes interpreted as such, sex and gender are not synonymous. Kennedy et al. (2020) explain that sex is “a set of biological attributes in humans and animals … including chromosones, gene expression hormone levels and function, and reproductive/sexual anatomy” (p.4). In contrast, gender is a socially constructed role that reflects an individuals identity (Kennedy et al., 2019). While the 2019 CES records gender at three levels, male, female, and other, the GSS 2017 measures sex on a strict male-female binary. To complete the post-stratification process, these variables must be recoded to be identical. Building on the work of Kennedy et al. (2019), we elected to impute non-binary individuals as female. This choice reflects an understanding that non-binary respondents experience the same, if not more, marginalization and discrimination as those who identify as female (Kennedy et al., 2019). The benefit of this approach is that these individuals remain in the data set. However, the concern is that this further confuses the unique and significant differences between sex and gender (Kennedy et al., 2019).

Work remains to be done in this area. First, future research may examine this problem with alternative data sources. For example, others may use the 2016 Census of Population to post-stratify their individual level estimates. The larger sample size and inclusions of additional sociodemographic variables would produce significant findings. Moreover, a new Census is scheduled to take place in 2021. Given the traditional instability of minority governments, there may be future opportunities to engage with a new Canadian Election Study in light of a more recent census data. Second, research confirms that geographic location is a significant predictor of voting behaviour. Time limitations prevented the addition of region in this paper. However, an appropriate next step to account for these variations would be to employ a random-intercept or random-slope model conditioned on province or territory. This practice would better illuminate change in the Bloc Quebecois vote share at full turnout, who competes only in the province of Quebec.

References

Achen, C. H. (2019). *Understanding Voter Turnout in Canada: What Data Do We Lack?*. Commonwealth Parliamentary Association.

Beaupré, P. (2020). *General Social Survey Cycle 31 : Families*. https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/sdaweb/dli2/gss/gss31/gss31/more_doc/GSS31_User_Guide.pdf (https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/sdaweb/dli2/gss/gss31/gss31/more_doc/GSS31_User_Guide.pdf)

Birch, S. (2010). Perceptions of Electoral Fairness and Voter Turnout. *Comparative Political Studies, 43*(12), 1601-1622. 10.1177/0010414010374021

Caetano, S. (2020a). Probability versus Non-Probability Sampling [PowerPoint Slides]. Quercus. https://q.utoronto.ca/courses/184060/files/8975319?module_item_id=1816946 (https://q.utoronto.ca/courses/184060/files/8975319?module_item_id=1816946) (https://q.utoronto.ca/courses/184060/files/8975319 (https://q.utoronto.ca/courses/184060/files/8975319)?module_item_id=1816946)

Caetano, S. (2020b). Sampling Techniques [PowerPoint Slides]. Quercus. https://q.utoronto.ca/courses/184060/files/9058287?module_item_id=1828122 (https://q.utoronto.ca/courses/184060/files/9058287?module_item_id=1828122)

Caetano, S. (2020c). Introduction to Logistic Regression [PowerPoint Slides]. Quercus. https://q.utoronto.ca/courses/184060/files/9309406?module_item_id=1855302 (https://q.utoronto.ca/courses/184060/files/9309406?module_item_id=1855302)

Caetano, S. (2020d). Introduction to Bayesian Inference [PowerPoint Slides]. Quercus. https://q.utoronto.ca/courses/184060/files/9490196?module_item_id=1872494 (https://q.utoronto.ca/courses/184060/files/9490196?module_item_id=1872494)

Caetano, S. (2020d). Multilevel Regression and Post-Stratification [PowerPoint Slides]. Quercus. https://q.utoronto.ca/courses/184060/files/9632737?module_item_id=1891982 (https://q.utoronto.ca/courses/184060/files/9632737?module_item_id=1891982)

Caetano, S. & Alexander, R. (2020a). gss_cleaning. [R code]. https://q.utoronto.ca/courses/184060/files/9422740?module_item_id=1867317 (https://q.utoronto.ca/courses/184060/files/9422740?module_item_id=1867317)

Caetano, S. & Alexander, R. (2020b). 01-data_cleaning-survey1. [R code]. https://q.utoronto.ca/courses/184060/files/9827183?module_item_id=1913034 (https://q.utoronto.ca/courses/184060/files/9827183?module_item_id=1913034)

Caetano, S. & Alexander, R. (2020c). 01-data_cleaning-post-strat1. [R code]. https://q.utoronto.ca/courses/184060/files/9827182?module_item_id=1913033 (https://q.utoronto.ca/courses/184060/files/9827182?module_item_id=1913033)

Elections Canada. (n.d.) *Official Voting Results: Forty-Third General Election*. Government of Canada. http://www.elections.ca/res/rep/off/ovr2019app/home.html (http://www.elections.ca/res/rep/off/ovr2019app/home.html)

Elections Canada. (2020a). Explaining the Turnout Decline in Canadian Federal Elections: A New Survey of Non-voters. Government of Canada. https://www.elections.ca/content.aspx?section=res&dir=rec/part/tud&document=youth&lang=e (https://www.elections.ca/content.aspx?section=res&dir=rec/part/tud&document=youth&lang=e)

Elections Canada. (2020b). *Voter Turnout at Federal Elections and Referendums*. https://www.elections.ca/content.aspx?section=ele&dir=turn&document=index&lang=e#ftn6 (https://www.elections.ca/content.aspx?section=ele&dir=turn&document=index&lang=e#ftn6)

Hodgetts, P.A. & Alexander, R. (2020). cesR. (Version 0.1.0) [R code]. https://hodgettsp.github.io/cesR/ (https://hodgettsp.github.io/cesR/)

Pew Research Center. (2018). Trends in Party Affiliation Among Demographic Groups. https://www.pewresearch.org/politics/2018/03/20/1-trends-in-party-affiliation-among-demographic-groups/ (https://www.pewresearch.org/politics/2018/03/20/1-trends-in-party-affiliation-among-demographic-groups/)

Qualtrics. (n.d.) *Online Panels: Get Responses for Surveys & Research*. Qualtrics XM. https://www.qualtrics.com/research-services/online-sample/ (https://www.qualtrics.com/research-services/online-sample/)

Statistics Canada. (2020). *Reasons for Not Voting in the Federal Election, October 21, 2019*. Government of Canada. https://www150.statcan.gc.ca/n1/daily-quotidien/200226/dq200226b-eng.htm (https://www150.statcan.gc.ca/n1/daily-quotidien/200226/dq200226b-eng.htm)

Statistics Canada. (2017). *General Social Survey Cycle 31 : Families* [Data set]. https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/cgi-bin/sda/hsda?harcsda4+gss31 (https://sda-artsci-utoronto-ca.myaccess.library.utoronto.ca/cgi-bin/sda/hsda?harcsda4+gss31)

Stephenson, L. B., Harell, A., Rubenson, D. & Loewen, P. J. (2020a). *Welcome to the 2019 Canadian Election Study* Canadian Election Study. http://www.ces-eec.ca/ (http://www.ces-eec.ca/)

Stephenson, L. B., Harell, A., Rubenson, D. & Loewen, P. J. (2020b). *2019 Canadian Election Study - Online Survey Technical Report and Codebook*. V1

Stephenson, L. B., Harell, A., Rubenson, D. & Loewen, P. J. (2020). *Canadian Election Study 2019* [Data set]. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DUS88V (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DUS88V)

Uppal, S., & LaRochelle-Cote, S. (2012). Factors Associated with Voting. *Perspectives on Labour and Income, 24*(1), 4. https://search.proquest.com/docview/1015345477 (https://search.proquest.com/docview/1015345477)

Wu, C., & Thompson, M. E. (2020). *Sampling Theory and Practice.* Cham: Springer International Publishing.

---

1. The target population refers to the set of individuals who meet the study requirements (Wu & Thompson, 2020).↩

2. Under ideal conditions, the sampling frame, or frame population, lists all individuals in the target population (Wu & Thompson, 2020). In practice, there is often a discrepancy between the target and frame populations, leading to issues of coverage error and representativeness (Wu & Thompson, 2020).↩

3. Market research panels refer to “a group of people recruited to respond to a survey” (Qualtrics, n.d.).↩

4. The sampling approach is how individuals are selected from the sampling frame (see Note 2).↩

5. The difference between a probability and non-probability sampling approach is random selection. Information from the CES team and Qualtrics suggests the latter, indicating that respondents were randomly selected.↩

6. Stratified random sampling is one example of a probability sampling approach. It involves dividing the population into homogenous groups, also known as strata, before constructing the sample (Caetano, 2020b). Then, in each stratum, a simple random sample is performed (Caetano, 2020b).↩

7. The sampled population refers to the set of individuals who participate in the survey (Wu & Thompson, 2020).↩