[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).

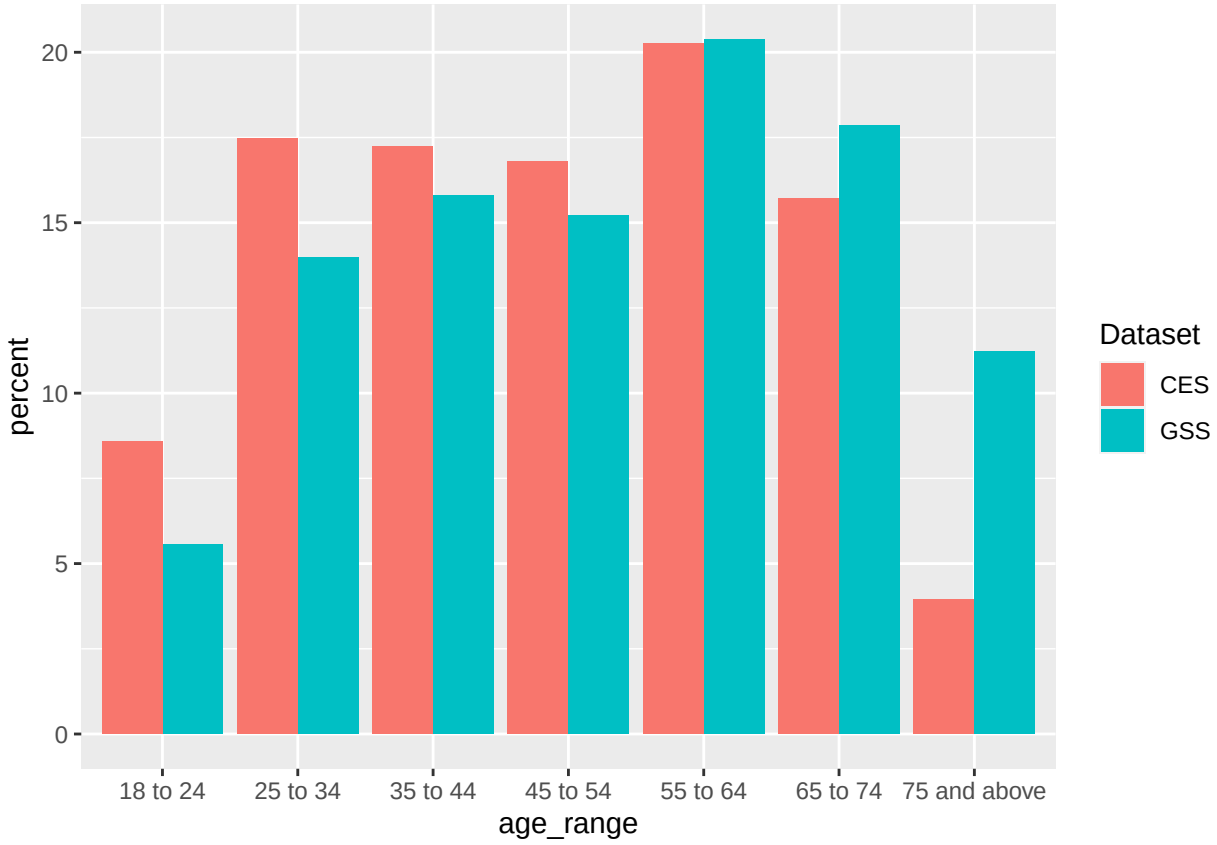
GSS Data

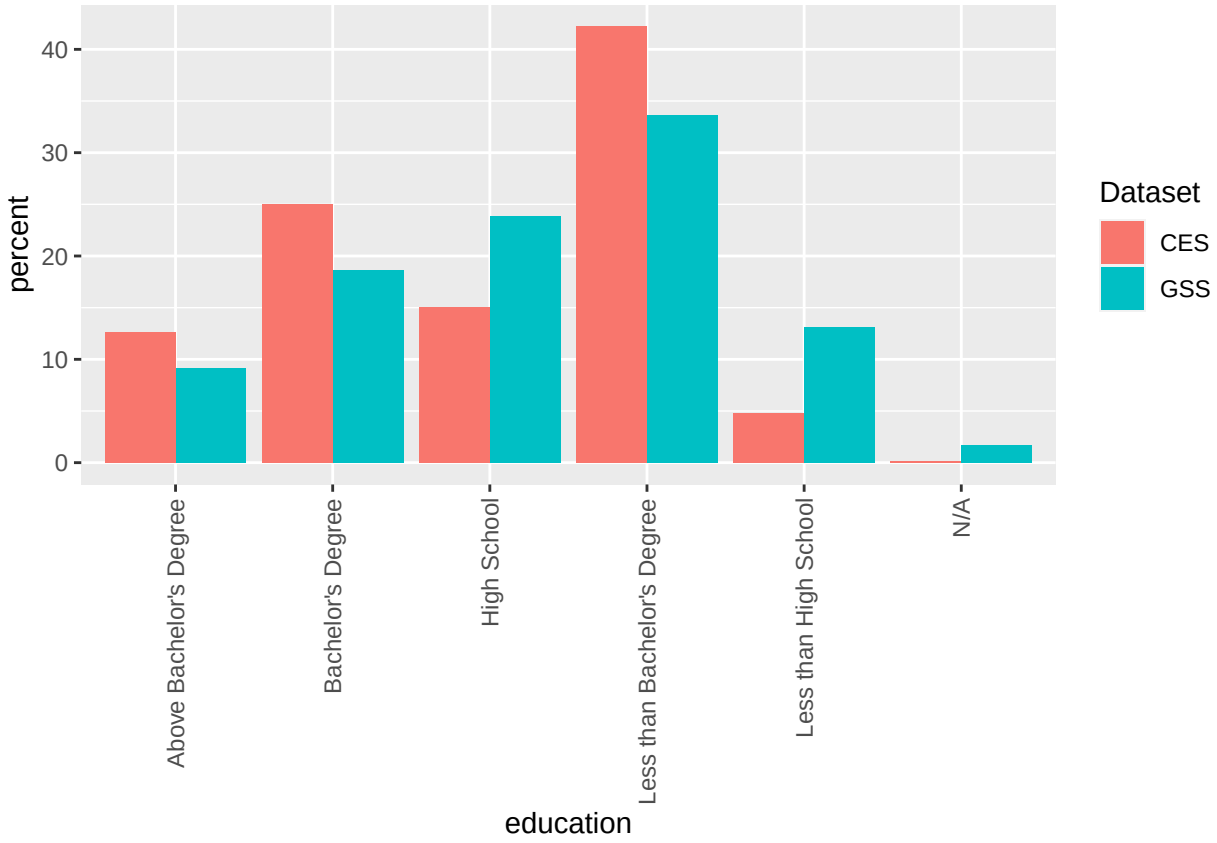
The data used in this report was curated from the General Social Survey (GSS) program., specifically the results of the “General social survey on Family” from 2017. Unfortunately, we cannot provide the data itself, as it is restricted to only those that are allowed to view it. However, if you are a UofT 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.

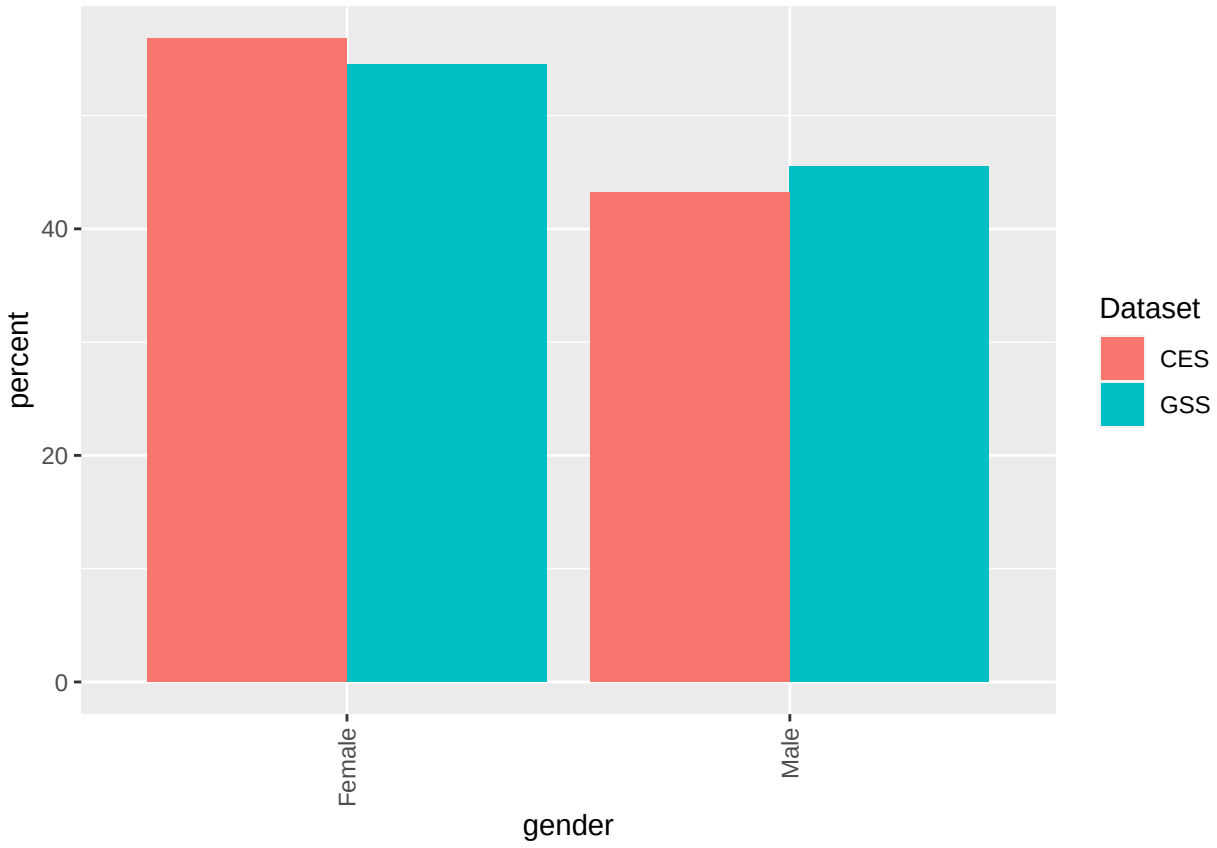
As stated in the overview for the GSS, the target population consisted of those aged 15 and over across all the provinces in Canada, apart from full-time residents of institutions. For this survey, the sample size was approximately 20,000 people, done via Stratified Random Sampling, where the strata 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.

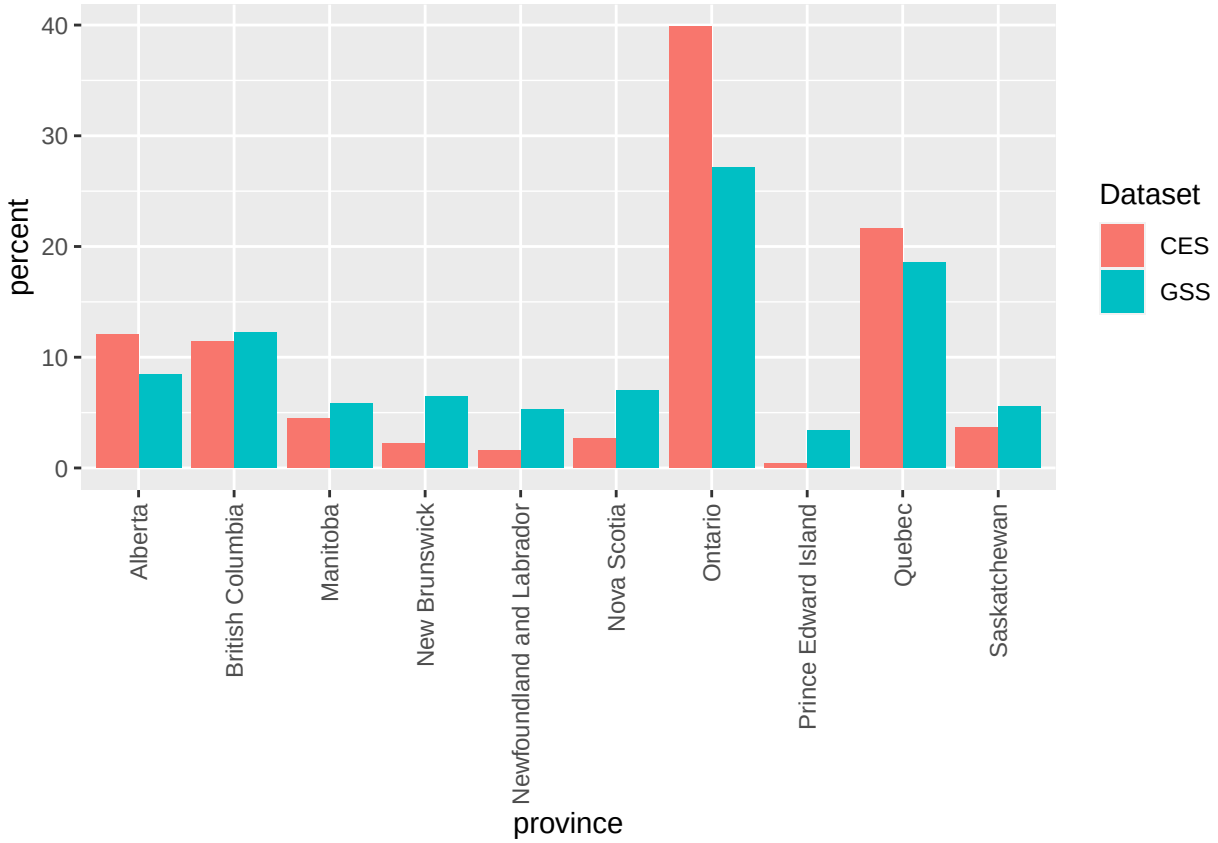
Respondents were contacted via computer assisted telephone interviews. There were many measures for combating non-response. One of which was the resilience shown by the interviewers, including calling up to twice more to a respondent who initially refused, outlining the importance of the survey. In addition, 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



\ ## Model We used bayesian regression using the brms package (Brms) in R (R). However, we have 6 parties to deal with, which can be hard to deal with given all of our categorical variables. So, we actually created 6 different models, with each one's independent variable being a vector of binary values, with a value of 1 for the party in question and 0 for the rest.

The formula for each was (Party) ~ age_range + education + gender + province.

Results

Our final results were the following:

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: Green ~ age_range + education + gender + cps19_province
##    Data: ces (Number of observations: 26146)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                    Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                             -2.23      0.12    -2.47    -2.00 1.01
## age_range25to34                       -0.40      0.08    -0.55    -0.26 1.00
## age_range35to44                       -0.62      0.08    -0.77    -0.46 1.00
## age_range45to54                       -0.64      0.08    -0.80    -0.48 1.00
## age_range55to64                       -0.75      0.08    -0.90    -0.60 1.00
## age_range65to74                       -0.74      0.08    -0.91    -0.57 1.00
```

```
## age_range75andabove                       -1.03      0.14    -1.32    -0.76 1.00
## educationBachelorsDegree                  -0.21      0.07    -0.35    -0.07 1.00
## educationHighSchool                       -0.21      0.08    -0.37    -0.05 1.00
## educationLessthanBachelorsDegree          -0.12      0.07    -0.25     0.01 1.00
## educationLessthanHighSchool               -0.16      0.11    -0.39     0.06 1.00
## genderMale                                -0.21      0.05    -0.30    -0.12 1.00
## cps19_provinceBritishColumbia              1.32      0.10     1.12     1.53 1.00
## cps19_provinceManitoba                     0.70      0.14     0.43     0.97 1.00
## cps19_provinceNewBrunswick                 1.72      0.13     1.46     1.97 1.00
## cps19_provinceNewfoundlandandLabrador     -0.40      0.29    -1.01     0.12 1.00
## cps19_provinceNovaScotia                   1.23      0.14     0.96     1.51 1.00
## cps19_provinceOntario                      0.71      0.10     0.53     0.90 1.01
## cps19_provincePrinceEdwardIsland           2.01      0.25     1.52     2.49 1.00
## cps19_provinceQuebec                       0.62      0.10     0.43     0.82 1.00
## cps19_provinceSaskatchewan                 0.16      0.16    -0.15     0.48 1.00
##                                       Bulk_ESS Tail_ESS
## Intercept                                 1086     1989
## age_range25to34                           2062     2795
## age_range35to44                           1815     2746
## age_range45to54                           1882     2545
## age_range55to64                           2000     2937
## age_range65to74                           1859     2426
## age_range75andabove                       3176     3111
## educationBachelorsDegree                  2508     2662
## educationHighSchool                       2595     2998
## educationLessthanBachelorsDegree          2231     2718
## educationLessthanHighSchool               3124     2946
## genderMale                                4816     2868
## cps19_provinceBritishColumbia             1421     2118
## cps19_provinceManitoba                    2202     2715
## cps19_provinceNewBrunswick                1868     2622
## cps19_provinceNewfoundlandandLabrador     3501     2817
## cps19_provinceNovaScotia                  2065     2699
## cps19_provinceOntario                     1361     2123
## cps19_provincePrinceEdwardIsland          3378     2913
## cps19_provinceQuebec                      1390     2027
## cps19_provinceSaskatchewan                2525     2985
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).

##  Family: bernoulli
##   Links: mu = logit
## Formula: Conservative ~ age_range + education + gender + cps19_province
##    Data: ces (Number of observations: 26146)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                      Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                              -0.69      0.08    -0.85    -0.54 1.00
## age_range25to34                         0.50      0.07     0.37     0.63 1.00
## age_range35to44                         0.78      0.07     0.65     0.91 1.00
```

```
## age_range45to54                             0.91      0.07     0.78     1.04 1.00
## age_range55to64                             0.91      0.07     0.78     1.03 1.00
## age_range65to74                             0.91      0.07     0.77     1.04 1.00
## age_range75andabove                         1.16      0.09     0.99     1.33 1.00
## educationBachelorsDegree                    0.20      0.05     0.10     0.30 1.00
## educationHighSchool                         0.56      0.05     0.46     0.67 1.00
## educationLessthanBachelorsDegree            0.41      0.05     0.33     0.51 1.00
## educationLessthanHighSchool                 0.40      0.08     0.25     0.55 1.00
## genderMale                                  0.41      0.03     0.35     0.47 1.00
## cps19_provinceBritishColumbia              -1.40      0.06    -1.51    -1.29 1.00
## cps19_provinceManitoba                     -0.92      0.07    -1.06    -0.78 1.00
## cps19_provinceNewBrunswick                 -1.54      0.10    -1.74    -1.34 1.00
## cps19_provinceNewfoundlandandLabrador      -1.74      0.12    -1.97    -1.50 1.00
## cps19_provinceNovaScotia                   -1.89      0.10    -2.09    -1.70 1.00
## cps19_provinceOntario                      -1.36      0.04    -1.45    -1.28 1.00
## cps19_provincePrinceEdwardIsland           -1.97      0.26    -2.50    -1.48 1.00
## cps19_provinceQuebec                       -2.16      0.05    -2.26    -2.05 1.00
## cps19_provinceSaskatchewan                 -0.38      0.07    -0.53    -0.24 1.00
##                                       Bulk_ESS Tail_ESS
## Intercept                                 1402     1863
## age_range25to34                           1466     2018
## age_range35to44                           1482     2234
## age_range45to54                           1438     2133
## age_range55to64                           1424     2012
## age_range65to74                           1436     2256
## age_range75andabove                       2001     2181
## educationBachelorsDegree                  2593     2577
## educationHighSchool                       2929     2740
## educationLessthanBachelorsDegree          2593     2557
## educationLessthanHighSchool               3076     2818
## genderMale                                4804     2849
## cps19_provinceBritishColumbia             2301     2521
## cps19_provinceManitoba                    3060     3050
## cps19_provinceNewBrunswick                3455     2320
## cps19_provinceNewfoundlandandLabrador     3824     2992
## cps19_provinceNovaScotia                  3040     3093
## cps19_provinceOntario                     1911     2555
## cps19_provincePrinceEdwardIsland          4760     3041
## cps19_provinceQuebec                      2263     2674
## cps19_provinceSaskatchewan                2817     2990
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: Liberal ~ age_range + education + gender + cps19_province
##    Data: ces (Number of observations: 26146)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                    Estimate Est.Error l-95% CI u-95% CI Rhat
```

```
## Intercept                                -1.27      0.07    -1.42    -1.12 1.00
## age_range25to34                          -0.04      0.06    -0.15     0.07 1.00
## age_range35to44                           0.06      0.06    -0.05     0.17 1.00
## age_range45to54                           0.06      0.06    -0.05     0.17 1.00
## age_range55to64                           0.22      0.05     0.11     0.32 1.00
## age_range65to74                           0.27      0.06     0.16     0.39 1.00
## age_range75andabove                       0.26      0.08     0.11     0.41 1.00
## educationBachelorsDegree                 -0.07      0.04    -0.15     0.02 1.00
## educationHighSchool                      -0.62      0.05    -0.72    -0.52 1.00
## educationLessthanBachelorsDegree         -0.43      0.04    -0.52    -0.35 1.00
## educationLessthanHighSchool              -0.66      0.07    -0.81    -0.52 1.00
## genderMale                               -0.12      0.03    -0.17    -0.06 1.00
## cps19_provinceBritishColumbia             0.72      0.06     0.60     0.84 1.00
## cps19_provinceManitoba                    0.68      0.08     0.53     0.83 1.00
## cps19_provinceNewBrunswick                1.07      0.10     0.88     1.25 1.00
## cps19_provinceNewfoundlandandLabrador     1.51      0.11     1.30     1.72 1.00
## cps19_provinceNovaScotia                  1.35      0.09     1.18     1.53 1.00
## cps19_provinceOntario                     1.07      0.05     0.97     1.17 1.00
## cps19_provincePrinceEdwardIsland          1.22      0.22     0.79     1.64 1.00
## cps19_provinceQuebec                      0.96      0.05     0.85     1.07 1.00
## cps19_provinceSaskatchewan               -0.25      0.11    -0.47    -0.05 1.00
##                                       Bulk_ESS Tail_ESS
## Intercept                                 1473     1980
## age_range25to34                           1816     2494
## age_range35to44                           1942     2518
## age_range45to54                           1991     2595
## age_range55to64                           1921     2318
## age_range65to74                           2007     2256
## age_range75andabove                       2464     3007
## educationBachelorsDegree                  2264     2338
## educationHighSchool                       2354     2499
## educationLessthanBachelorsDegree          1963     1739
## educationLessthanHighSchool               3040     3080
## genderMale                                3721     2953
## cps19_provinceBritishColumbia             1985     2505
## cps19_provinceManitoba                    1932     2799
## cps19_provinceNewBrunswick                2568     2351
## cps19_provinceNewfoundlandandLabrador     2849     2833
## cps19_provinceNovaScotia                  2507     2815
## cps19_provinceOntario                     1463     2121
## cps19_provincePrinceEdwardIsland          4160     2983
## cps19_provinceQuebec                      1717     2443
## cps19_provinceSaskatchewan                3307     3043
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).

## Warning: There were 535 divergent transitions after warmup. Increasing
## adapt_delta above 0.8 may help. See http://mc-stan.org/misc/
## warnings.html#divergent-transitions-after-warmup

##  Family: bernoulli
##   Links: mu = logit
```

```
## Formula: Peoples ~ age_range + education + gender + cps19_province
##    Data: ces (Number of observations: 26146)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                    Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                             -4.04      0.22    -4.46    -3.61 1.00
## age_range25to34                        0.28      0.16    -0.03     0.60 1.01
## age_range35to44                       -0.02      0.17    -0.34     0.31 1.01
## age_range45to54                       -0.06      0.17    -0.38     0.29 1.00
## age_range55to64                       -0.56      0.18    -0.88    -0.20 1.01
## age_range65to74                       -0.93      0.20    -1.33    -0.54 1.01
## age_range75andabove                   -0.80      0.30    -1.41    -0.24 1.00
## educationBachelorsDegree              -0.19      0.16    -0.51     0.11 1.01
## educationHighSchool                    0.48      0.15     0.18     0.77 1.01
## educationLessthanBachelorsDegree       0.33      0.14     0.06     0.61 1.01
## educationLessthanHighSchool            0.91      0.19     0.53     1.29 1.01
## genderMale                             0.50      0.08     0.33     0.67 1.00
## cps19_provinceBritishColumbia         -0.10      0.17    -0.43     0.24 1.00
## cps19_provinceManitoba                -0.58      0.27    -1.15    -0.07 1.00
## cps19_provinceNewBrunswick             0.41      0.25    -0.09     0.90 1.00
## cps19_provinceNewfoundlandandLabrador -0.83      0.48    -1.84     0.02 1.01
## cps19_provinceNovaScotia               0.09      0.28    -0.47     0.60 1.01
## cps19_provinceOntario                  0.04      0.13    -0.21     0.30 1.01
## cps19_provincePrinceEdwardIsland    -775.36    913.46 -3348.89   -17.32 1.01
## cps19_provinceQuebec                  -0.07      0.15    -0.36     0.22 1.01
## cps19_provinceSaskatchewan             0.07      0.23    -0.43     0.49 1.01
##                                    Bulk_ESS Tail_ESS
## Intercept                               320      653
## age_range25to34                         281      564
## age_range35to44                         258      508
## age_range45to54                         297      511
## age_range55to64                         338      735
## age_range65to74                         362      769
## age_range75andabove                     556      910
## educationBachelorsDegree                552     1026
## educationHighSchool                     528     1380
## educationLessthanBachelorsDegree        543     1160
## educationLessthanHighSchool             514      982
## genderMale                             1226     1714
## cps19_provinceBritishColumbia           660     1053
## cps19_provinceManitoba                  753     1059
## cps19_provinceNewBrunswick              798     1319
## cps19_provinceNewfoundlandandLabrador   773     1085
## cps19_provinceNovaScotia                782     1105
## cps19_provinceOntario                   474      639
## cps19_provincePrinceEdwardIsland        509      460
## cps19_provinceQuebec                    520      846
## cps19_provinceSaskatchewan              719      765
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
##  Family: bernoulli
##   Links: mu = logit
## Formula: NDP ~ age_range + education + gender + cps19_province
##    Data: ces (Number of observations: 26187)
## Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                                    Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                             -1.23      0.09    -1.40    -1.05 1.00
## age_range25to34                       -0.27      0.06    -0.38    -0.15 1.00
## age_range35to44                       -0.57      0.06    -0.69    -0.44 1.00
## age_range45to54                       -0.83      0.06    -0.95    -0.70 1.00
## age_range55to64                       -1.15      0.06    -1.27    -1.02 1.00
## age_range65to74                       -1.28      0.07    -1.42    -1.14 1.00
## age_range75andabove                   -1.57      0.13    -1.85    -1.33 1.00
## educationBachelorsDegree              -0.04      0.06    -0.16     0.08 1.00
## educationHighSchool                    0.07      0.07    -0.06     0.20 1.00
## educationLessthanBachelorsDegree       0.11      0.06     0.00     0.23 1.00
## educationLessthanHighSchool            0.21      0.09     0.04     0.38 1.00
## educationNDA                           0.07      0.42    -0.80     0.87 1.00
## genderMale                            -0.44      0.04    -0.52    -0.37 1.00
## cps19_provinceBritishColumbia          0.76      0.07     0.62     0.90 1.00
## cps19_provinceManitoba                 0.55      0.10     0.36     0.73 1.00
## cps19_provinceNewBrunswick            -0.35      0.16    -0.68    -0.05 1.00
## cps19_provinceNewfoundlandandLabrador  0.78      0.13     0.51     1.04 1.00
## cps19_provinceNovaScotia               0.39      0.12     0.16     0.62 1.00
## cps19_provinceOntario                  0.54      0.06     0.42     0.66 1.00
## cps19_provincePrinceEdwardIsland      -0.13      0.32    -0.79     0.46 1.00
## cps19_provinceQuebec                  -0.02      0.07    -0.16     0.12 1.00
## cps19_provinceSaskatchewan             0.82      0.10     0.64     1.01 1.00
##                                    Bulk_ESS Tail_ESS
## Intercept                              1602     1725
## age_range25to34                        1753     2536
## age_range35to44                        1874     2295
## age_range45to54                        1998     2462
## age_range55to64                        2109     2641
## age_range65to74                        2229     2803
## age_range75andabove                    3263     2986
## educationBachelorsDegree               2181     2687
## educationHighSchool                    2370     2545
## educationLessthanBachelorsDegree       1758     2652
## educationLessthanHighSchool            2627     2706
## educationNDA                           4397     2806
## genderMale                             4202     3121
## cps19_provinceBritishColumbia          1897     2365
## cps19_provinceManitoba                 2821     2848
## cps19_provinceNewBrunswick             3655     2777
## cps19_provinceNewfoundlandandLabrador  3482     2861
## cps19_provinceNovaScotia               3132     3234
## cps19_provinceOntario                  1677     2137
## cps19_provincePrinceEdwardIsland       4336     2498
## cps19_provinceQuebec                   2023     2313
## cps19_provinceSaskatchewan             2583     2605
```

```
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Discussion

Appendix (?)

References