

Minimum Response Rate is a Crucial Component of the Polling Procedure*

Traffic-Calming-related polls are still consistently invalidated by the amended minimum response rate in the City of Toronto

Tian Yi Zhang

06 February 2022

Abstract

Polls are a widely used and effective way for administrative divisions to determine the general opinion of residents and businesses on various topics. We obtain and analyze the polls conducted by the city of Toronto broken down based on the type of application and analyze it using two graphs and a table. The data shows a consistent trend for parking-related polls to meet the required engagement rate, while the opposite is seen for traffic calming applications. The findings suggest a change in percentage for the number of returned ballots required for polls regarding traffic calming.

1 Introduction

What is the best way to find out what people think about a particular topic? Ask them. This sentiment is shared by Joshua Clinton from the Vanderbilt University (Clinton, Abby, and Winkelried 2021), who believes that “public opinion polling, done right, remains the best way of obtaining citizens’ opinions.” While he does raise the point that polls should not be the only thing that public policy is based on, it is nevertheless a good way to gauge the public’s opinions and expectations. This is the crux of the concept of polling, and is why it is a widely employed tactic. As an administrative body, the City of Toronto is no stranger to utilizing polls to determine the opinions of its residents and business owners. These polls could be for any topic, ranging from changes to by-laws, to a simple addition of a speed hump on a particular road, and have no doubt aided the City of Toronto in constructing improvements to the city.

A important factor in conducting a poll is the idea of the **minimum response rate**, which could be described as the percentage of ballots that need to be returned in order for the poll to be considered valid. The City of Toronto initially adopted a minimum response rate of 50% in 2006 (City of Toronto 2011), and while many polls were considered valid under this 50% rate, it is also true that a significant portion of polls were invalidated because of it. This would not be an issue on its own, but when the invalidated polls often have an overwhelmingly positive response, as is illustrated by Councilor Frances Nunziata (Nunziata 2011), there is a case to be made for the argument that perhaps the minimum response rate of 50% has invalidated many useful polls and is therefore too high. The minimum response rate was eventually lowered down to 25% after many reviews (City of Toronto 2011), and it is the primary goal of this paper to examine whether this change was a meaningful one.

This paper will be using the dataset for polls conducted by the City of Toronto from 2015 to 2021 to analyse the relationship between the minimum response rate and the number of ballots cast by the voters to see if the change made to the minimum response rate had any impact on the validity of polls, particularly the

*Code and data are available at: https://github.com/zhan7818/sta304_paper1_repo.

ones concerned with Front Yard Parking and Traffic Calming. The results from this analysis will have clear implications for further changes towards the minimum response rate for polls conducted by the City.

The remainder of this paper contains two sections: Section 2 explains the sources of the dataset, how it is structured, its strengths, and weakness. The findings from observations on the graph and table are also included in this section. Section 3 contains a link to the repo where the code used to construct this paper can be found. Section 3 contains the references.

2 Data

The dataset used in this paper is obtained from the City of Toronto open data portal, using the `opendatatoronto` package (Friendly et al. 2020) and the statistical programming language R (R Core Team 2020). A new variable, `ballots_needed_to_proceed_double`, will be constructed by doubling the number of ballots needed to proceed, `ballots_needed_to_proceed`, to simulate the minimum response rate before the amendment from 50% to 25% as discussed in Development and Reviews of Front Yard Parking Poll Requirements (City of Toronto 2011). This newly constructed variable is used in Figure 1.

In addition to using R (R Core Team 2020) to construct this paper, several R packages are used as well: The feature names in the dataset are cleaned using the `janitor` package (Firke 2021). Subsequent graphs and tables shown in this paper also utilize the `tidyverse` (Wickham et al. 2019) and `knitr` (Xie 2021) package respectively.

The dataset is collected by the Toronto City Clerk’s Office, whom are responsible for administering polls and collecting the results. The data collection started from as early as April 1, 2015, and is updated daily after each closing and certification of a poll. The polls included in this dataset are used to determine whether property owners and businesses have a favorable attitude towards a change in their neighbourhood that could affect them. The types of polls are divided into the following five categories:

- Boulevard Cafes
- Off-Street Parking (such as front yard parking and commercial boulevard parking)
- Permit Parking
- Traffic Calming
- Business Improvement Area

The majority of the polls conducted in this dataset are parking and traffic-calming related, with the rest of the polls dealing with public opinion concerning business improvement areas and boulevard cafes.

According to information gathered from chapter 190 of City of Toronto bylaws (City of Toronto 2006) and the City of Toronto webpage regarding polls (City of Toronto 2021), the procedure of polling can be summarized as follows: First, the City reviews a proposal for an application and decides whether to approve it. If approved, the City conducts a poll of the population in the area affected by the application. Every owner, resident and tenant at least 18 years old in the affected area will be mailed the materials necessary to engage in the poll, including but not limited to: a notice regarding the poll and its deadline, the ballot distributed, and a return envelope. Then, the results from the poll are collected. Note that for a poll to be considered positive, it must first reach a certain percentage of ballots returned so as to ensure the poll is an accurate depiction of the general consensus in the area of interest. This threshold differs for each type of poll as determined by specific by-laws and/or city policies. Lastly, should the result of the poll be positive, the application can then proceed through the approval process (Final approval by City Council may be required based on the type of poll).

The dataset contains several features that are variations of other features with additional details that may otherwise be useful for a different analysis. For example, the dataset records both the number of ballots distributed, `BALLOTS_DISTRIBUTED`, and the number of ballots returned to City Clerk’s Office, `BALLOTS_RECEIVED_BY_VOTERS`. The difference in the data recorded between these two variables are minor or nonexistent and would perhaps be useful in another analysis regarding the element of human error

Table 1: First ten rows of a dataset that shows poll engagement in Toronto from 2015 to 2021

Address	Potential Voters	Ballots Distributed	Ballots Cast
8 Brumell Ave	41	34	18
90 Belsize Dr	40	36	30
376 Roehampton Ave	135	97	43
889 Queen St E	120	106	30
1747 Eglinton Ave W	333	235	26
24 Harlton Cres	144	109	35
291 Deloraine Ave	100	80	48
353 St Germain Ave	93	73	36
2152 Danforth Ave	23	16	3
866 Bloor St W	26	24	1

present in the execution of polling procedures. Other minor but potentially crucial details, such as the number of ballots received but not clearly marked as “in favour” or “opposed,” recorded as `BALLOTS_SPOILED`, are also included in the dataset. These small details included in the dataset allow for more accurate data analysis depending on the topic presented in a paper, and speaks volumes about how comprehensive the dataset is.

That is not to say that the dataset is flawless, however. Certain issues arise with the aforementioned polling methods, too: Despite stating that every eligible voter in the affected area will be able to participate in the poll, it is not guaranteed that this will be the case. Table 1 showcases this discrepancy:

Note that the number of potential voters differs from the number of ballots distributed when they should ideally be equal. This difference could be due to human errors or is intentional, which means that even if all ballots distributed are properly returned, the results are not representative of the true population. Alternatively, this discrepancy could be due to the way the data is collected and labelled: The City of Toronto’s (Friendly et al. 2020) official description of `POTENTIAL_VOTERS`, the feature in the dataset which records the number of potential voters, states that this feature represents the “number of people residing within poll boundary range.” However, as we previously stated, there are certain requirements that must be fulfilled for one to be considered eligible for the poll (for example, all participants must be at least 18 years of age). The vagueness in the description and naming of the `POTENTIAL_VOTERS` variable means that we are not sure if residents under the age of 18 in the affected area are also included in the dataset as a “potential voter,” and regrettably cannot accurately assess the reasoning behind the discrepancy between the number of potential voters and the number of ballots distributed.

Assuming that the first reasoning is true, then the dataset may be biased in several possible ways depending on how the ballot is distributed: Recall from chapter 190 of the City of Toronto bylaws (City of Toronto 2006) that the City Clerk’s Office is responsible for compiling the polling list which includes the names of the participants and their addresses. Is it possible that a racial bias may be present in the dataset as the City Clerk’s Office intentionally omits affected residents of a certain minority group? Or perhaps the City Clerk’s Office is prone to contextual bias as they get to decide which residents are considered affected by the application proposal and which ones are not? There is no way to analyze these possibility because relevant data such as the ethnicity of affected residents are not included in the dataset, a flaw that we should not ignore when examining the reliability of the dataset.

The next primary subject of interest is the relationship between the number of ballots cast, the number of ballots needed for the poll to be considered valid, and the type of application. This relationship could be used as a measure of the public’s willingness to engage in polls depending on the type of application and affect future choices for setting new percentages of returned ballots needed to consider a poll valid. Figure 1 shows the relationship between the number of ballots returned and the number of ballots needed to proceed **before the reduction of minimum response rate from 50% to 25%** as discussed in Development and Reviews of Front Yard Parking Poll Requirements (City of Toronto 2011).

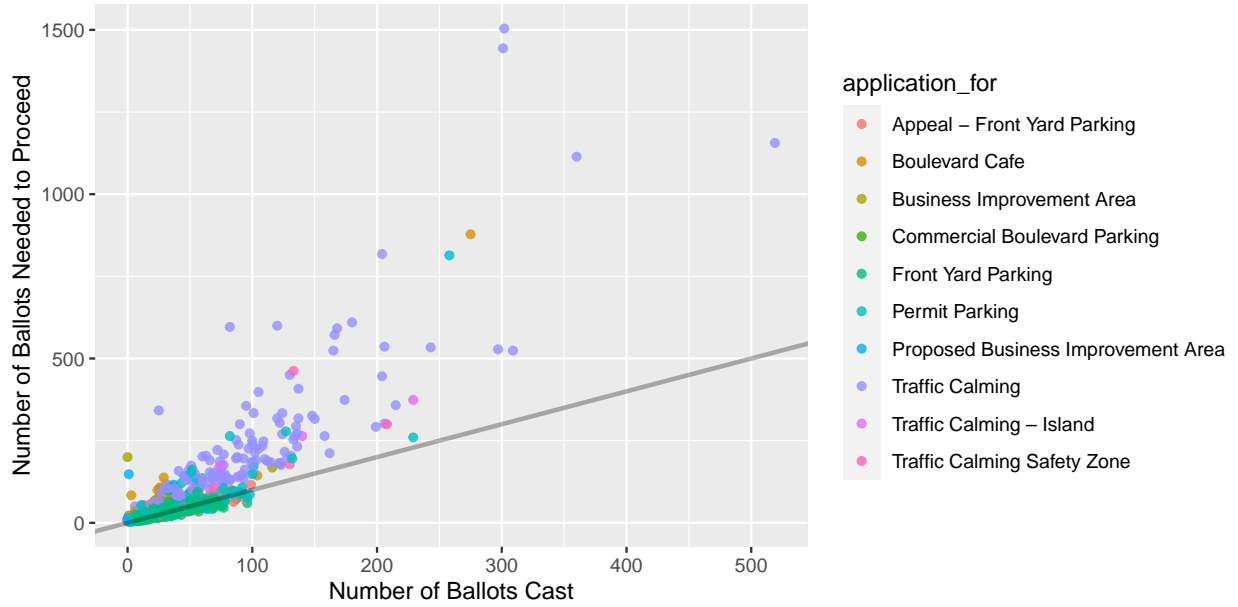


Figure 1: Ballots distributed versus ballots cast for each type of application, 50% return requirement

We can see from Figure 1 that the majority of the polls from this dataset are for traffic calming and parking applications, as mentioned earlier in the data segment of this paper. Note that the graph shows the majority of parking-related polls having less than 100 ballots cast. This makes sense because applications concerning parking only affects the residents of the property being approved for parking. In Contrast, polls regarding traffic calming vary much more in terms of number of ballots cast, which makes sense because improvements to certain road sections would affect more people than the rest.

The black line represents the line where the ballots cast is equal to the required number of ballots needed to proceed; the points above the black line represents polls that did not meet the required number of returned ballots to be considered valid. Note that the number of required returned ballots for a poll to be considered valid was 50% of the number of ballots distributed. We can see that Councillor Frances Nunziata’s (Nunziata 2011) concerns are supported by this dataset; There was not just a few, but a significant number of Front Yard Parking polls considered invalid polls under the 50% minimum response rate policy.

The previously mentioned percentage of 50 was subject to many discussions and revisions, as noted in Development and Reviews of Front Yard Parking Poll Requirements (City of Toronto 2011), to the point where it is now reduced from 50% to 25%, the latter shown in Figure 2.

Observations from Figure 2 tells that that most of the applications for parking-related applications meet the amended minimum response rate of 25%. This, in conjunction with our observation from Figure 1, shows that the issues brought up and reviewed by City Council (City of Toronto 2011), particularly Councillor Frances Nunziata (Nunziata 2011), were valid concerns regarding the polling procedures, and that the reduction of minimum response rate from 50% to 25% was a step in the right direction.

Another important piece of information presented by Figure 2 is that despite the reduction in the required number of ballots returned, a significant portion of polls regarding traffic calming still lies above the black line, failings to meet the required amount of 25% of distributed ballots returned to be considered a valid poll. It would be bizarre to attribute consistent invalidation of polls to a simple lack of engagement from the affected population. Therefore, this observation implies that there may be a need for further reduction in percentage for the poll validity requirement.

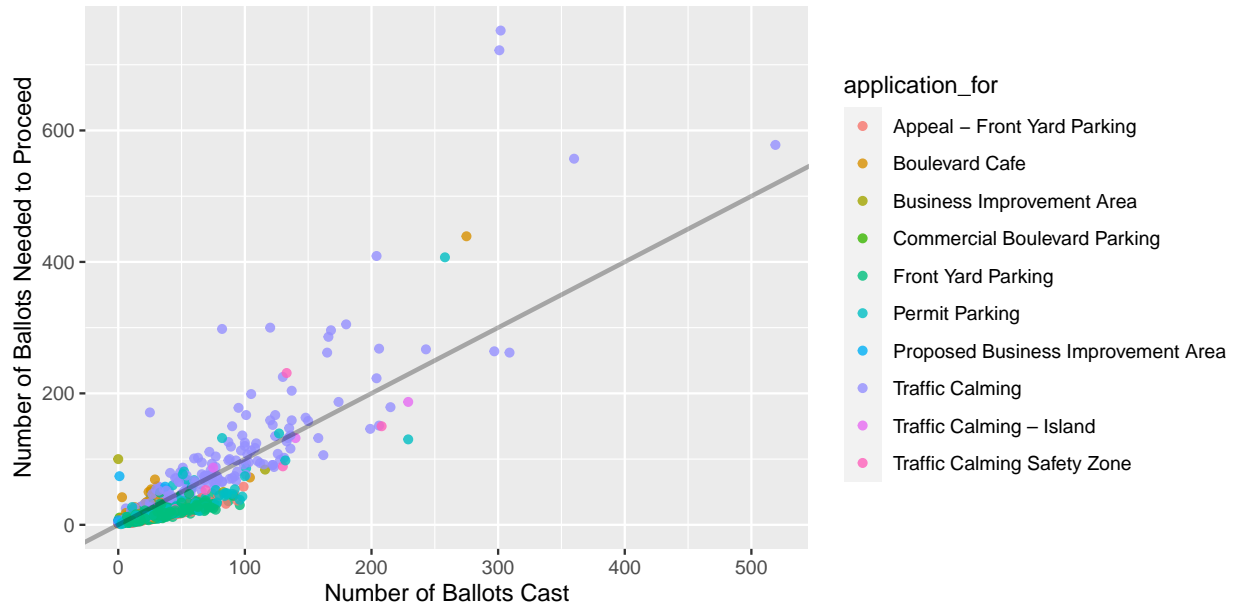


Figure 2: Ballots distributed versus ballots cast for each type of application, 25% return requirement

3 Footnote

References

- City of Toronto. 2006. *Chapter 190 Polling and Notification - Toronto*. https://www.toronto.ca/legdocs/municode/1184_190.pdf.
- . 2011. *Pw3.-Polling for Front Yard Parking - Attachment 1 - Toronto*. <http://www.toronto.ca/legdocs/mmis/2011/pw/bgrd/backgroundfile-37202.pdf>.
- . 2021. *Polls Regarding Changes in a Neighbourhood*. <https://www.toronto.ca/city-government/planning-development/polls-regarding-changes-in-a-neighbourhood/>.
- Clinton, Joshua, Abby, and Jon Winkelried. 2021. *Polling Problems and Why We Should Still Trust (Some) Polls*. <https://www.vanderbilt.edu/unity/2021/01/11/polling-problems-and-why-we-should-still-trust-some-polls/>.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean ‘Lahman’ Baseball Database*. <https://CRAN.R-project.org/package=Lahman>.
- Nunziata, Frances. 2011. *Front Yard Parking - Polling Requirement - Toronto*. <https://www.toronto.ca/legdocs/mmis/2011/pw/bgrd/backgroundfile-35072.pdf>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.