Comparing Alternative Methods for the Random Selection of a Respondent within a Household for Online Surveys

Geneviève Vézina and Pierre Caron

Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, Ontario, Canada, K1A0T6

Abstract

Online self-reported surveys present an alternative to traditional collection modes. As such, Statistics Canada is gradually moving to online surveys for the majority of its household surveys. This new collection mode however does present challenges in the need to randomly select a household member to complete the survey. Historically, an initial paper invitation is sent by mail and selection of a household member is done through rostering. If the selected member is not the person who completed the roster, a handing-off of the survey to the selected person is needed, which may however increase the survey's non-response rate. Alternative methods for individual selection before accessing the online application are proposed to avoid this hand-off request. We present two such methods that were tested: the last birthday method and an age-order method. For these methods, the selection is done using instructions on the paper invitation so that only the selected household member will have to access the online application. Response rates and selection inaccuracy rates were compared between the two alternative methods as well as with the traditional roster method. Results of the comparisons will be presented and discussed.

Key Words: age-order, last birthday, full roster, full enumeration, rostering, online survey, within-household selection

1. Introduction

Household surveys have historically rostered all eligible members of the household before randomly selecting one person to participate in the survey. The issue with this method in the context of an online self-administrated questionnaire is that the selected person cannot be contacted directly if they are different than the person who completed the roster. In this case, the e-mail address of the selected person is typically asked in order to contact him or her. In other words, the selection of a secondary respondent leads to two contacts and contributes to lower response rates. In the past, methods such as the "last birthday" approach have been suggested as an alternative to a complete roster in order to reach the selected person as quickly as possible. As household surveys continue to move towards electronic questionnaire (EQ) administered via internet as the main mode of collection and response rates continue to decrease, the need to identify the respondent quickly and without interviewer interaction becomes even more important.

In March 2016, as part of the pilot study for the National Travel Survey (NTS), a field test was conducted to compare three potential methods to randomly select a person within a household in an EQ environment; the full roster method, the last birthday method and an age-order method. The focus of this paper is on comparing response rates and selection inaccuracy rates between the three methods. Section 2 of the report explains details of the field test that was conducted. Selection methods will be described in section 3. Results will be shown in section 4 and the conclusion will wrap up this document in section 5.

2. Field test

The National Travel Survey is an address-based survey conducted in the ten Canadian provinces. It provides statistics on Canadian residents and activities related to domestic and international tourism. It was developed to measure the volume and characteristics of travellers and their trips as well as any associated economic impact. To be eligible, an adult must be 18 years of age or older.

2.1 Field Test

The three selection methods described in the next section were tested as part of the pilot for the NTS that was conducted in March 2016. A sample of 22,500 households, evenly split across the three methods, was drawn from Statistics Canada's Dwelling Universe Files (DUF) stratified into 180 strata. This sample size was determined based on the criteria to be able to detect absolute differences of 5% between the methods using a 5% level of significance. In addition to comparing the response rates of each of the methods, efforts were made to collect auxiliary information on each member of the household in order to attempt to measure the selection bias associated with each method. Using this auxiliary information (demographics of each member of the household), it was then possible to determine if the person who filled in the questionnaire was actually the person that had been identified by the selection method.

2.2 Embedded Experiment

The NTS experiment could be considered a randomized block design (RBD). The reader is encouraged to refer to Van den Brakel and Renssen (1998) which explains the comparisons and parallels between sampling theory and experimental design. In order to assess that the results (section 4) show statistical differences between the selection methods, a Wald test was performed as described in Van den Brakel and Renssen (2005). The methodology presented in this paper was programmed using Xper, a SAS-based macro developed at Statistics Canada. Xper was used to compute different Wald statistics with associated p-values.

3. Selection Methods

Each household in the field test was randomly assigned to one of the three respondent selection methods. Then, households were sent a letter in the mail inviting them to complete the National Travel Survey online.

For the last birthday and the age-order selection methods, instructions on the invitation letter instructed the reader on how to select the household member who would complete the survey online. For the full enumeration method, the letter directed the reader to visit the survey online where they would be asked to complete a roster of the eligible household members.

3.1 Method 1: Last Birthday Method

A letter is mailed to the selected household and the adult member with the most recent birthday is selected via the letter to complete the electronic questionnaire. This person goes to the internet and accesses the online questionnaire by typing the secure access code (SAC) provided in the letter and completes the survey. The last birthday method is not truly random as the distribution is skewed towards eligible members born in months immediately preceding the interview. However, if the birth month is not related to the topic of interest then this is less of an issue. The last birthday method is appealing since it is

quick to administer, non-intrusive and has lower refusal rates than other methods such as the full enumeration methods. The drawback is that the inaccuracy of the selection, i.e. people not following the instructions and deciding for themselves who will complete the questionnaire. This inaccuracy likely increases with the number of eligible household members and also for households with lower levels of education. It is also important to mention that inaccuracy could lead to a selection bias in the estimate. Studies have shown that the correct respondent is selected approximately 80% of the time for telephone surveys and less than 70% of the time for mail surveys (Lavrakas et al., 1993 and Lavrakas et al., 2000).

The following text box presents the wording used in the invitation letter sent to the selected households for this method.

Who should complete this survey?

The person in your household who had the **most recent birthday**, and is **18 years of age or older**, has been selected to participate.

3.2 Method 2: Age-Order Method

A letter is mailed to the selected household and an adult member is selected via the letter to complete the electronic questionnaire based on the age of all the adult household members. For this method, we restricted the person selection in households with three or more eligible members to six possible versions of the letter randomly assigned to the sample. This means that everyone in households of six or fewer adults will have a chance of being selected. For households exceeding six eligible members, some members will have a zero probability of selection. Note that households of more than six adults represents less than 0.5% of households in Canada. Each version of the letter selects one of the following:

- The first, second or third oldest adult of the household
- The first, second or third youngest adult of the household

The selected person is invited to complete the questionnaire on the internet by accessing the online questionnaire and typing the SAC provided in the letter. Before the test, some concerns were expressed regarding the possible selection inaccuracy, which could again lead to a selection bias. Moreover, producing six variations of the letter means more possibilities of making mistake operationally.

Below, examples of the wording used in the invitation letters are presented. They can be grouped in two different cases even though the wording was different for the six letters.

Case 1: when the oldest or the youngest person is selected from among the adults, 18 years of age or older, in the household.

Who should complete this survey?

- If you are the **only person** in your household who is 18 years of age or older, **you** have been selected to participate in the survey.
- If your household has **two or more members** 18 years of age or older, the **oldest member** among them has been selected.

Case 2: when the 2nd or the 3rd youngest (or oldest) adult is selected (for households with three or more persons 18 years of age or older).

Who should complete this survey?

- If you are the **only person** in your household who is 18 years of age or older, **you** have been selected to participate in the survey.
- If your household has **two members** 18 years of age or older, the **younger member** of them has been selected.
- If your household has **three or more members** 18 years of age or older, list those members in order of youngest to oldest.

1._____ 2._____ 3._____

The **third person on the list** has been selected.

3.3 Method 3: Roster Method

This method can be considered the control group. For this method, a letter is mailed to the household, which invites any household member go to the internet and access the electronic questionnaire by typing in the SAC provided in the letter. The person who starts completing the questionnaire will have to enumerate (or roster) all eligible household members. After the enumeration, one adult household member is randomly selected by the electronic application to complete the rest of the survey. If the selected person is the same as the person who first logged in, it is transparent for the respondent and the survey continues. However, if a different person is selected, then the member who has started the questionnaire is asked to provide the e-mail address of the selected person. Finally, an e-mail will be sent to the selected person with a new SAC and the hyperlink of the electronic questionnaire. With the EQ application doing the selection, the selection bias will likely be lower for this method. However, in some cases two different people are required to go online and complete their part. Therefore the response rate might be lower, which could suggest that the non-response bias is higher for this method.

4. Results

Van den Brakel and Renssen (1998) explain how to determine if two treatments are statistically different for a completely randomized experimental design or a randomized block design. This was generalized to more than two treatments by Van den Brakel and Renssen (2005). All the hypothesis tests done in Sections 4.1, 4.2 and 4.3 are based on the methodology of Van den Brakel and Renssen (2005). In other words, to assess if the response rates and selection inaccuracy rates have differences that are statistically

significant among the three selection methods, Wald tests are performed on the weighted rates at the level α=0.05.

4.1 Response Rates

For this analysis, a household is considered a respondent if we have a completed questionnaire from someone in the household.

Table 4.1 gives weighted and unweighted response rates and the results of Wald tests that were performed at the level α=0.05. The p-value 2.57 E-13 indicates that differences, in terms of weighted response rates, are statistically significant at level α=0.05 (because 2.57 E-13 < α for at least one of the first two methods compared to the control method (roster method). The age-order method seems to perform better than the last birthday method but the difference is not statistically significant (p-value 0.15 > α). Finally, the differences between the last birthday method and the roster method, and between the age-order method and the roster method are statistically significant. The roster method leads to a lower weighted response rate than the last birthday method and the age-order method.

Table 4.1 Response Rates by Method

Method	Unweighted Response Rate	Weighted Response Rate	p-Value (Wald test on weighted results)	Result
Last Birthday	24.2%	19.2%	2.57 E-13 (Roster method as the reference)	Significant
Age-Order	26.1%	20.7%		
Roster	16.4%	13.6%		
Last Birthday	24.2%	19.2%	0.15	Not Significant
Age-Order	26.1%	20.7%		
Last Birthday	24.2%	19.2%	2.23 E-8	Significant
Roster	16.4%	13.6%		
Age-Order	26.1%	20.7%	1.75 E-12	Significant
Roster	16.4%	13.6%		

4.2 Selection Inaccuracy Rate

A household is said to have selection inaccuracy if the respondent who completed the survey is not the person who was selected, i.e. the person who was supposed to answer. It is important to mention that a high selection inaccuracy rate suggests a potential bias in the estimates. This potential bias will be referred to as the selection bias. In order to determine if there was selection inaccuracy for a household, the age and sex of the respondent as provided in the demographic module was compared with the age and sex of all household members in the roster provided at the beginning of the questionnaire to determine if the correct person completed the questionnaire. Depending on the selection method, we identified from the roster who should have completed the questionnaire. If they were different, there was a selection inaccuracy.

Table 4.2 gives weighted and unweighted selection inaccuracy rates and results of the Wald tests that were performed at the level α=0.05.

Table 4.2 Selection Inaccuracy Rates by Method

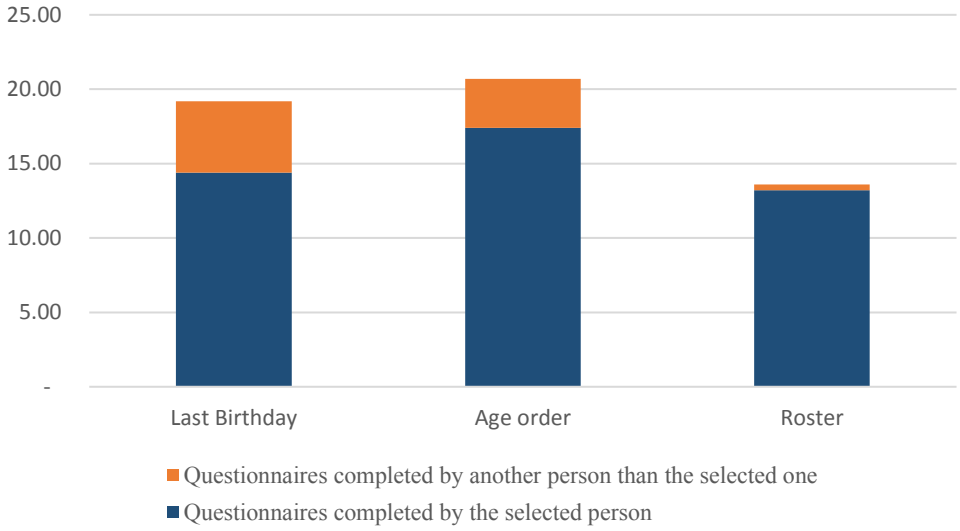
Method	**Unweighted Selection Inaccuracy Rate**	**Weighted Selection Inaccuracy Rate**	**p-Value** (Wald test on weighted results)	**Result**
Last Birthday	26.0%	23.0%	0 (Roster method as the reference)	Significant
Age-Order	18.3%	13.4%		
Roster	2.4%	2.4%		
Last Birthday	26.0%	23.0%	1.00E-4	Significant
Age-Order	18.3%	13.4%		
Last Birthday	26.0%	23.0%	0	Significant
Roster	2.4%	2.4%		
Age-Order	18.3%	13.4%	1.08 E-9	Significant
Roster	2.4%	2.4%		

In Table 4.2, we can observe that all the comparisons between methods lead to differences that are statistically significant at significance level α=0.05. In other words, all of the three methods lead to different inaccuracy rates. In light of the previous results, roster method is the most accurate method in terms of the selection of the person and the last birthday method is the least accurate.

4.3 Combined Rate: Response and Selection Bias Rates

At this point, we know that the response rate is significantly lower for roster method, but so is the selection inaccuracy rate. In the Graph 4.1, response rates and inaccuracy rates are displayed together to give an overall picture of the situation. The orange (upper) portion of the bars represent the questionnaires filled in by a person other than the selected person. In other words, the orange portion represents the contributor to the inaccuracy rate. The blue portion of the bars represent the proportion of questionnaires completed by the correct (i.e. selected) respondent.

Graph 4.1 Weighted Combined Rates by Method



In order to remove the potential selection bias, the questionnaires completed by a person other than the selected one are considered as non-response for the following results. In Graph 4.1, it means that the orange portion is removed and 'a new response rate' is calculated that will be referred to as the *accurate response rate*.

Table 4. 3 gives weighted and unweighted accurate response rates and results of the Wald tests that were performed at the level α=0.05.

Table 4.3 Accurate Response Rates by Method

Method	**Unweighted Combined Rate**	**Weighted Combined Rate**	**p-Value** (Wald test on weighted results)	**Result**
Last Birthday	17.9%	14.4%	0.000 (Roster method as the reference)	Significant
Age-Order	21.3%	17.4%		
Roster	16.0%	13.2%		
Last Birthday	17.9%	14.4%	0.002	Significant
Age-Order	21.3%	17.4%		
Last Birthday	17.9%	14.4%	0.213	Not significant
Roster	16.0%	13.2%		
Age-Order	21.3%	17.4%	0. 000	Significant
Roster	16.0%	13.2%		

Based on Table 4.3, the age-order method has a significantly higher combined rate than last birthday and roster methods. Despite last birthday method having a higher combined rate than the roster method, it is not significantly different at the level α=0.05.

From all the tests done in Sections 4.1, 4.2 and 4.3, we conclude that the age-order method outperformed the last birthday method, especially in terms of the inaccuracy rate. However, the roster method shows a lower inaccuracy rate, but at a cost of a significantly lower response rate. Therefore, if we do not use the selection biased cases, roster method still has a significantly lower response rate than age-order method.

4.4 Evaluation of the Potential Bias

In order to evaluate the potential bias, the weighted proportions by age and sex groups were estimated and compared to the known population demographic proportions. In order to get the weighted proportions for each method, design weights were adjusted to compensate for nonresponse. This adjustment was a simple calibration at the stratum level based on the number of in-scope units in the stratum. In other words, the weight for the units in a given stratum is the count of in-scope units divided by the number of respondents for each method. Based on these weights, the weighted proportions were calculated (refer to Table 4.4). Note that the estimates provided in this table used all respondents, including ones where the incorrect person completed the questionnaire (i.e. the orange part of Graph 4.1)

Table 4.4 Comparison of the Weighted Distribution by Method

Sex	Age Group	Demographic proportions	Last Birthday proportions	Age-Order proportions	Roster proportions
Male	18-34	14.5%	11.3%	11.5%	6.6%
Male	35-44	8.2%	8.5%	8.2%	6.8%
Male	45-54	8.9%	11.0%	7.2%	10.2%
Male	55-64	8.5%	9.7%	13.0%	10.1%
Male	65+	9.2%	10.6%	10.7%	16.1%
Female	18-34	14.3%	10.3%	11.2%	8.2%
Female	35-44	8.3%	8.3%	9.1%	10.6%
Female	45-54	8.9%	10.2%	11.3%	10.0%
Female	55-64	8.6%	11.8%	9.6%	11.8%
Female	65+	10.7%	8.4%	8.3%	9.6%
Euclidean distance of methods (1 to 3) from the demographic proportions			7.1	7.5	13.1

By looking at the proportions in this table, it is difficult to determine which method is the best, i.e., the closest to the known population demographic proportions. The last row of the table represents the Euclidean distance of proportions (for each of the three methods) to the known population demographic proportions. The distance is larger, almost double, for the roster method, which suggests a larger bias (selection and nonresponse bias) than the other two methods.

5. Conclusion

Following this analysis, the last birthday method can be dismissed since the performance of the age-order method is superior in terms of response rates and selection accuracy rates. Furthermore, it was shown that the age-order method has much higher response rates (see Table 4.1) than the roster method. On the other hand, the roster method outperformed the age-order method in terms of selection accuracy rates. The decision regarding which method to use between age-order and roster is not obvious and depends on many aspects, such as, for example, the budget for the survey, the resources available for non-response follow-up and the expected response rate of the survey.

The age-order method is recommended for surveys such as the NTS, where the main or the only mode of collection is self-administrated electronic questionnaire and budget for non-response follow-up is very limited. For example, if mail reminders are the only non-response follow-up strategy considered as it was the case for this pilot, then it was shown in Table 4.1 that weighted response rates for the age-order method is 7 percentage points higher than weighted response rates for the roster method. Moreover, Table 4.4 suggests that the bias associated with non-response could be more significant than the bias associated with selection. For surveys with more non-response follow-up resources and where more precise estimates are required, evaluations should be conducted to see if the selection inaccuracy generated from the age-order method could lead to a bias in the estimates. Note that in this case, both selection methods should lead to approximately the same response rates since it is assumed that all non-respondents will be followed-up. If this assumption is true, the roster method might be preferable.

References

Lavrakas, P., Bauman, S. and Merkle, D. (1993). "The Last-Birthday Selection Method and Within-Unit Coverage Problems". Paper presented at the annual meeting of the American Association for Public Opinion Research, St Charles, IL.

Lavrakas, P., Stasny, A. and Harpuder, B. (2000). "A Further Investigation of the Last-Birthday Respondent Selection Method and Within-Unit Coverage Error." Proceedings of the *American Statistical Association*, Survey Research Method Section [CD-ROM], Alexandria, VA: American Statistical Association.

Van den Brakel, J.A., Renssen, R.H. (1998). Design and analysis of experiments embedded in sample surveys. *Journal of Official Statistics*, Vol. 14, No. 3, pp. 277-295.

Van den Brakel, J.A., Renssen, R. H. (2005). Analysis of Experiments Embedded in Complex Sampling Designs. *Survey Methodology*, Vol. 31, No 1, pp. 23-40, Statistics Canada, No 12-001-XPB.