Component of Statistics Canada Catalogue no. 11-522-X
Statistics Canada's International Symposium Series: Proceedings

Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Is There Really Any Benefit in Sending Out Introductory Letters in Random Digit Dialling (RDD) Surveys?

by Jimmy Baulne and Robert Courtemanche

2009

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2008
Data Collection: Challenges, Achievements and New Directions

Is There Really Any Benefit in Sending Out Introductory Letters in Random Digit Dialling (RDD) Surveys?

Jimmy Baulne and Robert Courtemanche[1]

Abstract

Currently underway, the Québec Population Health Survey (EQSP), for which collection will wrap up in February 2009, provides an opportunity, because of the size of its sample, to assess the impact that sending out introductory letters to respondents has on the response rate in a controlled environment. Since this regional telephone survey is expected to have more than 38,000 respondents, it was possible to use part of its sample for this study without having too great an impact on its overall response rate. In random digit dialling (RDD) surveys such as the EQSP, one of the main challenges in sending out introductory letters is reaching the survey units. Doing so depends largely on our capacity to associate an address with the sample units and on the quality of that information.

This article describes the controlled study proposed by the Institut de la statistique du Québec to measure the effect that sending out introductory letters to respondents had on the survey's response rate.

Key Words: Letter, RDD, Response rate

1. Introduction

1.1 Description

The Institut de la statistique du Québec (ISQ) is the Quebec government's official statistics agency. In the 1990s, the ISQ, known then as the Bureau de la statistique du Québec, was asked to conduct a series of regional health surveys on behalf of the Ministère de la Santé et des Services sociaux du Québec. After a period in the early 2000s when there were no regional surveys, the Ministère recently decided to launch a new program of regional health surveys to meet its requirements to monitor the health of Quebecers. One of those surveys is the Québec Population Health Survey 2008 (EQSP).

With the substantial changes in the survey business over the last decade – growing numbers of polling and marketing firms, more frequent canvassing of the public, and declining cooperation by respondents – the difficulties and advantages of personal interviewing (the 1990s program) have given way to the challenges of telephone collection (new survey program). Virtually unavoidable, primarily because of the high costs of personal interviewing, this shift has a direct impact on survey response rates. Additional methods are needed to maintain the current response rates. Mailing introductory letters to respondents is part of a set of telephone surveying procedures developed by the ISQ.

To assess the benefits of sending introductory letters to respondents, a controlled study was conducted using a portion of the EQSP's sample. The EQSP was selected for the study because it is a general population survey and is large enough that part of its sample can be used to control the study's findings without having too great an impact on the survey's overall response rate.

[1]Jimmy Baulne and Robert Courtemanche, Direction de la méthodologie et de la qualité, Institut de la statistique du Québec, 200 chemin Sainte-Foy, 3e étage, Québec, Canada, G1R 5T4.

2. Québec Population Health Survey

2.1 Sample design

The EQSP's target population consists of all persons aged 15 and over living in private households in Quebec. Institutional residents (collective households) and people living on Indian reserves are excluded. The survey covers 16 of Quebec's 18 heath regions. Only the Cree and Inuit regions (health regions 17 and 18) are excluded.

The survey's sample was selected through random digit dialling (RDD) of telephone numbers belonging to eligible private households. For each household reached, one person was selected at random from the group of household members aged 15 and over.

To increase the sample's productivity, non-working banks of 100 telephone numbers are eliminated under the RDD method. A bank is considered to be working if it contains at least one residential telephone number listed in a telephone directory; non-working banks are eliminated. The sample frame is the set of all working banks of 100 telephone numbers. A random sample of telephone numbers is then selected according to a stratified sample design (the strata consist of the health regions). For each bank selected, a number between 00 and 99 was generated at random and used to form a telephone number. Only one person per household was selected to complete the survey.

Within the group of numbers generated, those that did not belong to private residences were identified through matching with computerized directories of business numbers. Also at this stage, residential addresses were associated with sample telephone numbers that had one.

The target for the core sample of numbers generated for the EQSP was just over 32,000 completed interviews.[2] The sample was allocated on a non-proportional basis so that each health region would have just over 2,000 respondents.

To facilitate management and field follow-up, the EQSP's sample was split into five separate waves; an eighth of the total sample was used in each of the first two waves, and a quarter of the sample was used in each of the last three waves. On average, the questionnaire took 24 minutes to complete. It covered 12 different topics, including a number of questions on very sensitive subjects, such as suicide, drugs and sexuality. The EQSP's collection period ran from February 2008 to February 2009.

The EQSP's target response rate was 65% for all health regions except the Montréal region, for which the target was 60%. At the time the data for the introductory letter impact study were being analyzed, the actual weighted response rate was 54%.

3. Introductory letter impact study

3.1 Impact reported in the literature

The impact of sending introductory letters to respondents in telephone surveys is known and has been reported in the literature. Although the findings of studies that did not produce conclusive results were probably not published, a number of articles have shown that the use of introductory letters in telephone surveys has a positive effect on individuals and households. Among others, De Leeuw et al. (2007) carried out a meta-analysis of 29 independent studies[3] of the impact that introductory letters have on response rates in telephone surveys. The authors reported an average increase of 8% in response rates for telephone surveys in general. When they looked at the impact in relation to the type of sample frame used, De Leeuw et al. (2007) found that the increase averaged 9% for surveys based on a frame of listed telephone numbers, compared with only 4% for surveys based on a frame containing both

[2] An additional sample was generated for six regions that requested it, which raised the target number of respondents to more than 38,000. However, the additional sample was excluded from the introductory letter impact study and will not be discussed in this paper.

[3] All 29 studies used a control group to estimate the letter's impact.

listed and unlisted telephone numbers (RDD). They noted, however, that the letter's content and style might influence the response rate gains.

The authors also reported an increase in cooperation when introductory letters were used, which translates into fewer refusals. Other studies noted that the interviewers had more confidence when letters had been sent to respondents (Groves and Snowden, 1987), while still others mentioned that letters were helpful in the initial contact with households (Collins et al., 1988) or that they prevented the surprise of an unexpected call (Dillman, Gallegos and Frey, 1976). According to De Leeuw et al. (2007), however, sending introductory letters to respondents does not reduce the number of no-contact cases; it simply helps win over potential respondents once the initial contact is made.

3.2 Applicability to Quebec

The beneficial impact that letters have on the response and refusal rates according to the literature has to be weighed against the additional expense of sending out thousands of letters. Indeed, one has to consider whether a gain of only 4% in the response rate is sufficient to justify the higher cost. The magnitude of the gain could be affected by the prominence of the ISQ which is a provincial statistical agency and does not have the stature of a federal agency that is responsible for, among other things, the national census.

Moreover, Quebec may differ in a number of ways from the countries in which the 29 studies analyzed by De Leeuw et al. (2007) were carried out. First there is the proportion of the sample frame for which we had an address. De Leeuw et al. (2007) reported that, in the United States, an average of 40% of the telephone numbers were listed in telephone directories and thus had an address. The corresponding proportion in Quebec is nearly 50%, which could improve the ISQ's results. Second, there is the quality of the addresses obtained. According to the Census of Canada, there are proportionally more renters in Quebec than in other Canadian provinces, which means that a larger proportion of people live in apartments. We know from our previous observations that our success in reaching households by mail is limited in the case of apartments, since apartment numbers are often omitted from addresses. Third, Quebecers' response to small incentives such as introductory letters is unknown and remains to be determined. In particular, a letter sent to the dwelling located at the address (with a general addressee such as "The Baulne family") and not to a specific individual is less likely to reach the right person. Indeed, there is no guarantee that the person who reads the mail will be the person selected to complete the survey or will convey the information in the letter to other household members who could be selected.

3.3 Introductory letter impact study

With the aim of acquiring the tools to maintain our current response rates and stop, or at least slow, the steady decline observed by a number of statistics agencies, the ISQ decided to use part of the EQSP's sample to conduct a controlled study to assess the impact that sending introductory letters to respondents would have on response rates. The ultimate goal, of course, was to be able to extend the practice to all of the ISQ's individual/household surveys if the results were conclusive.

The ISQ's study used about 1/8 of the EQSP's base sample,[4] just over 12,600 telephone numbers. The set of numbers was divided into two groups, a test group and a control group, using the split ballot method. In each group, only some of the units were listed in telephone directories and therefore were associated with an address. That portion was nearly half of the group (49.8%).

In short, two groups were created at random. In each group, we had an address for some units and no address for the rest.

It is very important to use a control group in this kind of study, because it helps control for extraneous factors that may affect the response rate. If we do not control for those other factors, we may not be able to say for sure that the letter was responsible for the response rate increase. The study described in this paper involves two separate groups

[4] Units in the supplementary sample were excluded from the study.

that were subjected to the same conditions, except that, for one group, letters were mailed to respondents before data collection, and for the other, no letters were sent.

The main comparison in our study was between the two groups as a whole. However, a stated interest was to compare the impact that the letter had on the two subgroups of units for which we had an address (the listed portions of the two groups). That interest was due to the fact that the listed portion made up not only 49.8% of the sample but also 82.0% of the sample eligible for the survey (once the non-working and out-of-scope numbers were removed). In addition, because the letter was mentioned in the introduction to the telephone interview (see text below) for persons in the test group but not for persons in the control group, comparison of the subgroups with no addresses was also of interest. The object of this third comparison was to assess the "placebo" effect of mentioning the mailing of a letter at the beginning of the interview. It is important to note that the reference to the letter in the introduction implies that introductory letters were sent to some respondents, but not necessarily to all respondents. We were well aware that we had no addresses for part of the sample. The wording of that part of the introduction was as follows:

> [TRANSLATION] We are currently conducting a large study of the population's health on behalf of the Ministère de la Santé et des Services sociaux du Québec. Over the last few weeks, we mailed out introductory letters on this subject. Did you receive this letter?

It should be noted that, to simplify the questionnaire, the same introduction was used for all units in a group. However, different introductions were used for each group: the one for the test group units mentioned the letter, while the one for the control group units did not.

Mentioning at the beginning of the interview that introductory letters had been sent out may compensate in some way for any lack of communication between the household member who opened the mail and the member who took the telephone call. In addition, ISQ's interviewers told us that mentioning the letter made the initial contact with respondents much easier.

In summary, the control group units did not receive a letter, and the mailing of letters was not mentioned at the beginning of the interview. Test group units were told at the beginning of the interview that introductory letters had been mailed out, and those for which we had an address were sent a letter before data collection began.

3.3 Weighting

Since we want to use the results to make inferences about the entire target population, the sample units were weighted using the inverse probability of being selected into the sample. The household's initial weight was important because it allowed us to take into account the regional non-proportionality of the EQSP's sample. Indeed, because every health region in Quebec was to have the same number of respondents, a household in a region with a small population had a greater chance of being selected than a household in a region with a larger population.

In addition, the weights of eligible units were adjusted to compensate for the fact that a certain proportion of the households whose eligibility was unknown were ineligible. On the other hand, since we wanted to compare response rates but did not want to evaluate data quality at the group level, there was no non-response adjustment or post-stratification. The response and refusal rates used in the analyses are the RR3 and REF2 rates defined by the AAPOR (2008).

3.4 Analyses

As mentioned in section 3.2, three comparisons were examined: the main comparison of the test group and the control group, the comparison of the listed portions of the two groups, and the comparison of the unlisted portions of the two groups.

The following measures were analyzed: the response rate, the refusal rate and the number of respondents. For the response and refusal rates, we looked at the overall rates but also the rates at the household level and person level.

In other words, we calculated the proportion of households that answered the questionnaire's first few items, which provided information about the household's eligibility and composition (used to select a household member). At the person level, we calculated the proportion of persons who responded to the survey relative to the persons who were selected. It should be noted that the overall response rate is the product of the household response rate and the person response rate. As for the refusal rate, we decided to present it in a different form: the overall refusal rate is the sum of the household refusal rate and the person refusal rate. For this reason, they both have the same denominator. The advantage of this additive approach is that it shows how many percentage points of the overall refusal rate are attributable to household refusal and how many are attributable to person refusal. Please note that the response and refusal rates used in this study are weighted rates.

In an effort to learn more about the letter's impact on the response rate and the subpopulations most affected, we examined certain socio-demographic characteristics of respondents. For example, for respondent households, i.e. households for which a member was selected, we examined the person response rate in relation to the persons' age and sex. Only significant comparisons at the overall level were examined in relation to age and sex. We also looked at the letter's impact on the person response rate in one-person households and multiple-person households.

Lastly, the responses to certain items in the survey were compared. This comparison was done as a rough guide only, since the weighting applied for the purposes of testing the letter's impact ignored the sample design. Hence the results are not weighted, and no statistical tests were carried out.

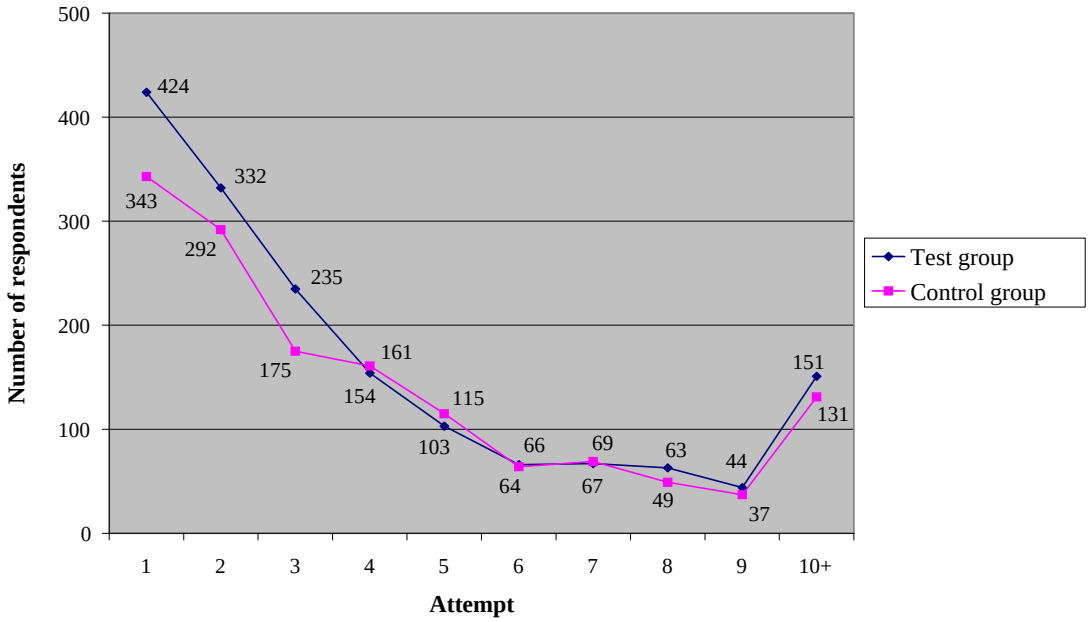
As noted previously, while the target was an overall response rate of 65%, the unweighted response rate at the time the study's results were analyzed was 56.1%. After weighting, i.e. adjustment for the sample's non-proportionality, the rate decreased to 54.0%. The tests conducted for this study were at the 5% significance level.

4. Findings of the study

4.1 Comparison of numbers of respondents

The first finding concerns the number of respondents for each group at each contact attempt (regardless of whether the household was contacted). As mentioned previously, letters were mailed to test group units for which we had an address, and the mailing was mentioned in the introduction to the interview for that group. As shown in Figure 4.1-1, there was a difference between the number of test group respondents and the number of control group respondents. The difference, which favoured the test group, was very clear for the first three contact attempts. However, by the fourth attempt, i.e. some three to five weeks after the letter was mailed, its effects appear to fade. This result is not surprising since, by this time, what is left is the sample's hard core of more recalcitrant respondents.

Figure 4.1-1 Number of respondents by attempt



4.2 Comparison of weighted response rates

Comparison of the response rates of the test group and the control group shows that the letter had a positive effect (Table 4.2-1). Indeed, with regard to the overall response rate, the gap is a significant 7.8 percentage points in favour of the test group. This means that sending out letters tends to significantly improve Quebecers' response. When we look at the household response and the person response, our study shows that the letter had a greater impact at the person level than at the household level, as the differences are 8.1 percentage points (significant) and 2.6 percentage points (non-significant), respectively.

Table 4.2-1
Weighted response rates for the main comparison

	Overall response rate	Household response rate	Person response rate
Test group	57.9%	73.2%	79.1%
Control group	50.1%	70.6%	71.0%
Difference	7.8%	2.6%	8.1%
P-value	< 0.01	0.14	< 0.01

Of course, the results are very different for the listed and unlisted portions of the sample (Table 4.2-2). Aside from the fact that the listed portion has higher response rates (63.8% versus 38.0%), we see that the differences between the two groups are larger for the portion of the sample with differences in the vicinity of 10 percentage points in overall response rate (9.7 percentage points) and person response rate (10.6 percentage points). The overall response rate of the unlisted portion is comparable to the rates observed in other studies for this type of sample. For example, Parsons and Owens (2002) reported a rate of 36.2% for the unlisted portion of the control group in their study.

Table 4.2-2
Weighted response rates for the listed and unlisted portions of the two groups

	Listed portion			Unlisted portion		
	Overall response rate	Household response rate	Person response rate	Overall response rate	Household response rate	Person response rate
Test group	63.8%	78.3%	81.6%	38.0%	56.2%	67.6%
Control group	54.1%	76.2%	71.0%	35.1%	49.8%	70.5%
Difference	9.7%	2.1%	10.6%	2.9%	6.4%	- 2.9%
P-value	< 0.01	0.26	< 0.01	0.46	0.13	0.59

As mentioned in section 3.4, for respondent households we know the age and sex of the person selected to complete the survey. We are therefore able to compare the two groups' person response rates for those characteristics. The results are presented in Table 4.2-3 below.

Table 4.2-3
Weighted response rated by sex and age for the main comparison

	Sex		Age				
	Male	Female	15-24	25-39	40-49	50-64	65 +
Difference between the two groups	8.2%	8.0%	5.3%	5.2%	6.0%	11.1%	11.3%
P-value	< 0.01	< 0.01	0.36	0.23	0.18	< 0.01	< 0.01

An analysis of this table shows that the letter affected both men and women, with differences between the two groups of 8.2 and 8.0 percentage points respectively. With regard to age, our study shows that the letter had a greater positive impact on people aged 50 and over, as the differences were over 11 percentage points for the 50-64 and 65-and-over age groups (11.1 and 11.3 respectively). It is also worth noting the differences observed for the other age groups, which were greater than 5 percentage points, though they were not significant.

4.3 Comparison of weighted refusal rates

Our analysis of refusals also reflected the positive impact of the introductory letter sent to test group units. Table 4.3-1 shows a 5.0 percentage-point difference between the test group's refusal rate and the control group's refusal rate (18.3% versus 23.3%). As noted above, the refusal rate is computed additively, i.e. the household refusal rate and the person refusal rate are added together to produce the overall refusal rate. In contrast to the response rate, where most of the letter's effect was seen at the person level, its effect on the refusal rate was more evenly split between the household and the person levels. Specifically, 2.3 percentage points (non-significant) of the total difference between the groups was due to household refusal, while 2.7 percentage points (significant) were due to person refusal.

Table 4.3-1
Weighted refusal rates for the main comparison

	Overall refusal rate	Household refusal rate	Person refusal rate
Test group	18.3%	11.2%	7.0%
Control group	23.3%	13.5%	9.7%
Difference	- 5.0%	- 2.3%	- 2.7%
P-value	< 0.01	0.07	< 0.01

4.4 Comparison of questionnaire responses

Despite the lack of a significance test, we decided to compare, for reference purposes only, the two groups' responses to a few of the survey's questions: household size, language spoken at home, type of household (person

living alone, family with or without children, etc.), education, main occupation, household income, propensity to consent to matching of the data collected with data from another source, and propensity to provide personal information required for matching, such as the Quebec health insurance number (NAM).[5] The results in Table 4.4-1 are not weighted, and no statistical tests were carried out.

Table 4.4-1
Propensity to consent to matching and to provide personal information by group (unweighted data)

	Permission to match	Willingness to provide NAM	Match potential
Test group	84.2%	79.9%	67.3%
Control group	80.5%	77.6%	62.5%
Difference	3.7%	2.3%	4.8%

Only results that appeared to distinguish between the two groups are shown in Table 4.4-1. The data in this table suggest that mailing letters to respondents might produce a 5-percentage-point increase in the number of individuals willing to have their information matched to other data sources such as health care databases.

5. Conclusion

In conclusion, our study shows that the letter was effective for Quebecers. In particular, we observed an increase of nearly 8 percentage points in the response rate. That is a substantial gain, as the average reported in other studies of the same kind was 4 percentage points. Our study also shows that the letter had a positive impact on both men and women, and that its effect was greater on people aged 50 and over. Young people are not very susceptible to small incentives such as the letter. Its positive impact also extends to refusals. Our study found a significant 5-percentage-point drop in the refusal rate when introductory letters were sent out.

Two other important findings emerged from our study. First, it is important to take advantage of the letter's impact within the first few contact attempts, since the effect appears to fade after the third contact attempt (after about four weeks). The letter also seems to boost respondents' level of trust, as evidenced by the increase in consent to allow matching and provide personal information.

References

American Association for Public Opinion Research (2008). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, 5th edition, Lenexa KS: AAPOR.

Collins, M., Sykes, W., Wilson, P. and Blackshaw, N. (1988). Nonresponse: The UK Experience, *Telephone Survey Methodology* (Eds. R.M. Groves et al.), New York: Wiley, 213-231.

De Leeuw, E., Callegaro, M., Hox, J., Korendijk, E. and Lensvelt-Mulders, G. (2007). The Influence of Advance Letters on Response in Telephone Surveys, *Public Opinion Quarterly*, 71, 413-443.

Dillman, D., Gallegos, J.G. and Frey, J.H. (1976). Reducing Refusal Rates for Telephone Interviews, *Public Opinion Quarterly*, 40, 66-78.

Groves, R. and Snowden, C. (1987). The Effects of Advance Letters on Response Rates in Linked Telephone Surveys, *Proceedings of the Survey Research Methods Section*, American Statistical Association 633-638.

Parsons, J., Owens, L. and Skogan, W. (2002). Using Advance Letters in RDD Surveys: Results of Two Experiments, *Survey Research newsletter*, 33, 1-2.

[5] The health insurance number is the sole identification number for the Quebec health care system.