The 2008 Canadian Election Survey

Technical Documentation

The 2008 Canadian Election Surveys

Technical Documentation

David Northrup

Institute for Social Research
York University
January, 2010

Conditions of Release

All research based upon these data must include an acknowledgement such as the following:

> Data from the 2008 Canadian Election Surveys were provided by the Institute for Social Research, York University. The survey was funded by Elections Canada, and was completed for the Canadian Election Team of Elisabeth Gidengil (McGill University), Joanna Everitt, University of New Brunswick, Patrick Fournier (Université de Montréal), and Neil Nevitte (University of Toronto). Neither the Institute for Social Research, Elections Canada, or the Canadian Election Survey Team are responsible for the analyses and interpretations presented here.

Researchers are requested to forward a copy of any publications or scholarly papers to the Associate Director, Institute for Social Research, The TEL Building, 88 The Pond Road, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3 and to Dr. Elisabeth Gidengil, Department of Political Science, McGill University, 855 Sherbrooke Street West, Montreal Quebec, H3A 2T7.

Data acquired from the Institute for Social Research may not be re-disseminated outside the recipient institution.

Table of Contents

Conditions of Release. i

Table of Contents. ii

List of Tables. iv

1. Study Description. 1

2. Sample Design. 3

2.1 Introduction. 3
2.2 Selection of Households. 3
2.3 Selection of Respondents. 4
2.4 Household Weights. 4
2.5 Provincial Sample Distribution. 6
2.6 National Estimates. 7
2.7 Post-Election Sample. 8
2.9 Mail-back Sample. 8

3. Data Collection. 9

3.1 Introduction. 9
3.2 Data Collection Procedures. 9
3.21 Number of Call Attempts. 9
3.22 Refusal Conversions 10
3.23 Mail-back Survey. 10
3.3 Response Rates and Reinterview Rates. 11
3.31 Campaign-Period Survey Response Rate. 11
3.32 Post-Election Survey Reinterview Rate. 12
3.33 Mail-back Survey Return Rate. 12

4. Questionnaire Issues and Data Processing. 13
4.1 Introduction. 13
4.2 Assigning Missing Values. 13
4.3 Province Specific Questions 14
4.4 Date Specific Questions. 14
4.5 Randomization of Question Order and Question Wording. 14
4.51 Leader Familiarity (CPS) 15
4.52 Party Leader and Party Ratings (CPS and PES). 15

4.53 Issue Importance (CPS) 16
4.54 Value and Party Images (CPS) 17
4.55 Spending Cuts and Taxes (PES) 17
4.56 Use of Terms "homosexuels" and "les gais et lesbiennes (PES) 18
4.6 Coding of Open-Ended Questions and "Other Specify" Options. 18
4.61 Other Specifies. 18
4.62 Most Important Issue (CPS, PES, MBS) 18
4.63 Occupational Coding. 20
4.64 Geographic Information. 20

References. 21

List of Tables

2.1 Household (HH)Weights 6

2.2 Provincial Sample Distribution and Provincial Weight. 7

3.1 Number of Call Attempts: Campaign-Period and Post-Election Survey. 9

3.2 Final Sample Disposition and Response Rate: CPS. 11

1. Study Description

The January 23, 2006 Canadian Federal Election resulted in a Conservative minority government under the leadership of Stephen Harper. Almost 33 months later, on October 14, 2008 the Conservatives won a second consecutive minority government in Canada's 40th General Election. Telephone surveys, conducted during the campaign (CPS) and after the election (PES), with a sample of Canadians, provide data that can be used to help explain the election outcome. In addition to the telephone surveys, a Mail-back Survey (MBS) that collected more data about the telephone respondent's political and social values and beliefs was completed after the PES. The surveys were completed under the direction of the Canadian Election Team (Gidengil, Everitt, Fournier, and Nevitte). This Technical Report outlines the design and conduct of the 2008 Canadian Election Surveys (CES).

A modified random digit dialling (RDD) sample along with the birthday selection method within households was used for 2008 CPS. Respondents to the 2008 PES included both 2008 CPS respondents, who were willing to complete a second interview, as well as respondents from the 2006 PES who could be located and were willing to complete the survey. Thus, the 2008 PES has both a new (to 2008) sample and a panel sample. PES respondents were asked to provide their name and address for the MBS.

Unlike previous election studies, dating from 1988, a rolling cross sectional sample release was *not* employed for the 2008 CPS. Sample was released on the first five days of calling and respondents were called two (or more) time per day up until midnight the evening before the October 14, 2008 vote. Interviews were completed in the 18 days before the vote and, excepting the first day of calling and Thanksgiving Day (October 10), between 130 to 260 interviews were completed each day. Thus, for the 2008 CPS the interviewers were concentrated more closely to the eve of the election than in previous studies. In total 3,257 CPS interviews were completed.

Calling for the PES started on October 15 the day after vote and all of the respondents were called back within three days of the vote. Of course, not all respondents were available when first contacted and after 10 days of calling 50 percent of the PES interviews were completed. By day 20 almost 70 percent of the interviews were completed but small numbers of interviews were completed until late December. In total 3,689 PES interviews were completed of which 2,451 were from the new 2008 CPS and the remaining 1,238 interviews are 'panel' cases, that is these respondents were interviewed in the 2006 CPS.[1] The total number of respondents, who answered one or both of the telephone components of the 2008 Canadian Election Study was 4,495 of which 1,238 were PES only (i.e. panel respondents).

[1] Many of the panel respondents were also interviewed in the 2006 PES and a small number were interviewed in 2004 but not 2006.

Quebec interviews were completed by Jolicoeur et Associés. Data collection was completed with Computer Assisted Telephone Interviewing (CATI). The Institute's CATI software is from the Computer-Assisted Survey Methods Program (CSM) at the University of California, Berkeley.

The naming conventions for the variables in the data file indicate the survey source (CPS, PES or MBS). For example, variables in the campaign-period survey include the prefix CPS, thus CPS_INTDATE indicates the date of interview for the campaign period completions. The suffixes PES and MBS are used to indicate that the variable is from the post-election, and mail-back survey (respectively).

2. SAMPLE DESIGN

2.1 Introduction

The RDD sample for the 2008 CPS was designed to represent the adult population of Canada: Canadian citizens 18 years of age or older who speak one of Canada's official languages, English or French, and reside in private homes[2] in the ten Canadian provinces (thus excluding the territories). Because the survey was conducted by telephone, the small proportion of households in Canada without telephones were excluded from the sample population.[3]

2.2 Selection of Households

To select individual survey respondents for the CPS, a two-stage probability selection process was utilized. The first stage involved the selection of households by randomly selecting telephone numbers. The ideal sampling frame for the campaign-period survey would have been a complete listing of all residential telephone numbers in Canada (excluding the territories). Unfortunately, such a listing does not exist. To select numbers ISR employs a modified form of random digit dialling. All telephone numbers in Canada consist of an area code, a "central office code" or exchange (the first three digits of the telephone number), and a suffix or "bank" (the last four digits of the number). A list of most telephone numbers in Canada can be constructed from CD-ROM versions of telephone books and other commercially available lists of telephone numbers. Numbers from these sources, as well as telephone numbers between or on either side of listed numbers are included in the sampling frame. For example, if the following telephone number was found in a directory, (416) 651-8513 then all numbers from (416) 651-8512 and (416) 651-8514 would be included in the sample. A computer is then used to generate a random sample of telephone numbers from this list. Since unlisted numbers and numbers too new to be included in directories are interspersed among valid numbers, this strategy provides a much better sample than one based on listed numbers alone.

As well as household telephone numbers, RDD samples include "not-in-service" and "non-residential" telephone numbers. Typically, non-household numbers are identified the first time

---

[2] Interviews were not completed with respondents who could not speak English or French well enough to complete the survey and residents of old age homes, group homes, educational and penal institutions were excluded from the sample.

[3] Statistics Canada (/www.statcan.gc.ca/daily-quotidien/070504/dq070504a-eng.htm) estimates that 1.2 percent of the households in Canada do not have no telephone service, that 90.5% of households have landlines, almost 70% have cell phones and 8.3% of households percent only have cell phones. Some cell phones are included in RDD samples.

the interviewer calls. Most of the interviewer's subsequent efforts are then directed at encouraging an informant from the household to provide information about the number of adults living in the home, and after randomly selecting a respondent, completing the interview.

2.3 Selection of Respondents

The second stage of the sample selection process was the random selection of a respondent from the selected household. To be eligible for the interview the household member had to be an adult (18 years of age or older) and a Canadian citizen. If there was more than one eligible person in the household, the person with the next birthday was selected as the survey respondent.[4] The birthday selection method is used as it ensures a random selection of respondents and is a much less intrusive way to start an interviewer than asking about the number of people in the household, thus making it easier for the interviewer to secure the respondent's cooperation. Of course, for the panel component of the survey, the respondent was the person who was initially randomly selected to complete the 2006 CPS (or, in a small number of cases, the 2004 survey). Interviewers, when they asked to speak to the person who did the 2006 survey typically had a first name or initial as well as the respondent's gender and age to maximize their ability to secure the re-interview with the correct respondent.

2.4 Household Weights

The probability of an adult member of the household being selected for an interview varies inversely with the number of people living in that household. In a household with only one adult, this person has a 100 percent chance of selection, in a two adult household each adult has a 50 percent chance of selection, and so on. Analyses based on unweighted estimates are therefore biased: members of one adult households are over represented, and larger households with two or more adults are under represented. Most practitioners of survey research "weight the data" in order to compensate for the unequal probabilities of selection (one adult households are given a weight of one, two adult households are given a weight of two, three adult households are given a weight of three, etc.).[5]

---

[4] See O'Rourke and Blair, 1983; for a review of the birthday selection method.

[5] Weighting to correct for unequal probabilities of selection, stratification, and other factors in order to improve sample estimates is common in survey research. See, for example: Lessler and Kalsbeek, 1992 Chapter 8; Kalton, 1983 Chapter 10; Babbie, 1992 Chapter 5; and Kish, 1965; specifically addresses the issue of weighting to correct for unequal probability of selection at the household level (p. 400) and suggests, unlike most survey researchers, that household weighting may not be necessary.

Survey data can be weighted to the population or to the original sample size. Of course percent figures and all other point estimates are the same whether the data is weighted to the population or the sample size. Until the advent of easy-to-use techniques for weighting complex samples in SPSS or STADA users of survey data tended to weight to sample size so as to have the same number of observations in the weighted and unweighted data set. The population distribution is based on the survey results, thus it is an estimate. The use of adjustment is made by determining the number of cases in each household size category that would have been in the sample, if an interview had been completed with each adult member of the household, and then dividing the population percentage by the sample percentage in each household size category.

In the 2008 campaign-period survey there were 3,257 households in the sample and 982 were one-adult households, 1,752 were two-adult households, 359 were three-adult households, etc. (Table 2.1 and variable CPS_NADULTS). The weights for each household are calculated as follows. First, the total number of weighted cases is calculated (number of cases times the number of adults in the household). For three-adult households the calculation is: 359 times 3 which gives 1,077 three-adult households in the weighted sample. In the campaign-period survey there are 6,278 weighted cases. Second, the 6,278 weighted cases are adjusted down to the original sample size of 3,257 (calculated as weighted cases for each household size divided by the weighted sample size times the original sample size). For three-adult households the calculation is: (359/6,278) * 3,257 = 558.74. Third, the weight for each household size is calculated (by dividing the adjustment to original sample size by the number of cases). For three-adult households the calculation is: 558.74/359 = 1.556387. (In essence, weights are obtained by dividing the proportion of households in the population (as estimated from the data) by the proportion of the households in the sample.)

Note that in the calculation of the household weights the total number of observations in the sample –the "weighted sample size" – is based on the original sample size, but we do not have a true random sample (as households were used to locate adults) and there is no accounting for sample design effects. Weighting in this manner, so that the weighted sample size is equal to the actual number of interviews, provides researchers with a very good approximation of the precision of their sample for point estimates (such as percentages, means, correlation coefficients (r), coefficients of determination (r^2), and so on). But, treating the sample as if it was a simple random sample of equal size results in *in*correct estimates of standard errors and, of course, incorrect significance tests. Worse, the errors are *down*wardly biased and so give a false sense of the precision of estimates as well as significance tests with too many false positives. Researchers should consider the use of the complex sample module in SPSS or use a statistical package that takes proper account of weights (such as STATA) when analysing the data.

Table 2.1 Household (HH) Weights

# of adults	# of HHs	weighted sample	adjusted to original sample size	weight
1 adult	982	982	509.4575	0.518796
2 adults	1,752	3,504	1,617.8605	1.037592
3 adults	359	1,077	558.7431	1.556387
4 adults	125	500	259.3979	2.075183
5 adults	32	160	83.0073	2.593979
6 adults	3	18	9.3383	3.112775
7 adults	1	7	3.6316	3.631571
10 adults	3	30	15.5639	5.187958
Totals	3,257	6,278	3,257.0000	

2.5 Provincial Sample Distribution

The distribution of Canadian households and survey sample among the provinces, as well as the Provincial Weights (CESPWGT) for each survey is detailed in Tables 2.2. In terms of the percentage of sample per province, the design called for a slight over representation of the seven smaller provinces and a corresponding under representation in Ontario and Alberta. For example, in the 2008 CPS survey Newfoundland and Labrador has 1.64 percent of the households in the country, but 3.15 percent of the households in the sample. Conversely, Ontario which has 36.59 percent of Canada's population has only 31.61 percent of the sample.

Because the sample distribution is not proportional to the population size (pps) of the provinces, the data must be weighted before national estimates are derived. (No province weight is required in comparisons between provinces.) Weights are obtained by dividing the proportion of households in the province by the proportion of the households in the sample for that province. For example, Ontario has a weight of 1.1577 (36.59/31.61 plus rounding error). In preparing national estimates, each Ontario case counts for 1.1577 observations in the weighted data set; in other words, Ontario cases are "weighted up" so that the impact of the Ontario sample on national estimates is an accurate reflection of Ontario's proportion of the total number of

households in Canada. Conversely, provinces where the weights are less than one, for example Newfoundland and Labrador (.5207), are "weighted down." Caveats about the effect of weighting on the variance estimates noted above apply here as well.

Table 2.2 Provincial Sample Distribution and Provincial Weight

Province	# of HHs	% of HHs	# HHs Sample	% HHs Sample	Weight
Nfld	189,045	1.64	102	3.15	0.5207
PEI	50,800	0.04	101	3.12	0.1413
Nova Scotia	360,025	3.12	102	3.15	0.9917
NB	283,825	2.50	109	3.36	0.7316
Quebec	2,978,115	25.58	880	27.16	0.9508
Ontario	4,219,410	36.59	1,024	31.61	1.1577
Manitoba	432,555	3.75	150	4.63	0.8102
Saskatchewan	379,680	3.29	140	4.32	0.7620
Alberta	1,104,100	9.57	143	4.41	2.1693
BC	1,534,335	13.31	506	15.62	0.8520
Totals	11,531,890	100.00	3,257	100.00	

2.6 National Estimates

In order to produce national estimates it is advisable to correct for both the unequal probabilities of selection at the household stage and the unequal probabilities of selection based on province of residence. CESNWGT (National Weight) is the product of the household weight and the provincial weight and should be used with the National Sample when national estimates are required..

Although the weights are provided as part of the data set, users must specify the weights they wish to use in the appropriate programming language before analysing the data. If weights are not invoked the tabulations produced will be for unweighted data. Because the weights include fractions that are rounded and missing values vary by item, there will be minor variations in the number of cases for different analytical procedures and subsets of the data.

2.7 Post-Election Sample

The sample for the post-election surveys was comprised of respondents to the CPS. For the 2008 survey this included not only the new 2008 RDD sample (n = 2,451) but also the 2006 panel sample (n = 1,238). At the end of the CPS in 2008, as in 2006, the interviewer ensured that they had a first name or some other identifier (such as the respondent's initials or position in the household, e.g., mother). This information, as well as the sex and year of birth of the CPS respondent, and the respondent's telephone number, was recorded on a "cover sheet." The interviewer called and asked for the person by name or identifier. If there was any concern about reaching the correct person the interviewer also checked age and gender.

2.8 2004 Mail-back Sample

At the end of the post-election survey, respondents were asked to provide their address so they could be sent the mail-back survey. Mail-back information was provided by 76 percent of the PES respondents.

3. Data Collection

3.1 Introduction

A description of the data collection procedures is briefly outlined in this section of the technical documentation. Supervisors monitored (listened to) about 10 percent of interviewers' calls to verify that the interviewers were reading questions and recording answers correctly.

3.2 Data Collection Procedures

3.21 Number of Call Attempts

In order to maximize the chances of getting a completed interview from each telephone number in the CPS sample, call attempts were made during the day and the evening - for both week and weekend days. Typically, between each number was called twice each day, once before and once after 6:00pm. (split between day and evening hours) but the calling was limited to the 18 day period immediately before voting day (October 14, 2008). Although a bit more than one-half of the interviews completed in the CPS took three or fewer call attempts, and almost 10 percent required ten or more calls (Table 3.1).

Table 3.1 Number of Call Attempts:
Campaign-Period and Post-Election Survey

	CPS		PES	
Calls	#	%	#	%
1	614	18.85	712	19.30
2	616	18.91	649	17.59
3	456	14.00	451	12.23
4	383	11.76	344	9.33
5	294	9.03	268	7.26
6-9	586	17.99	593	16.07
10-14	236	7.25	377	10.22
15/15+	72	2.21	295	8.00
Totals	3,257	100.00	3,689	100.00

The number of call attempts to complete an interview was about the same in previous campaign-period surveys despite the shorter calling time period (Table 3.1). Indeed, the pattern is fairly consistent with that for all previous campaign-period surveys completed at ISR.

The number of call attempts to complete an interview for the PES is about the same as for the CPS except almost twice as many of the interviews were completed on the tenth or subsequent call. The variables CPS_ATEMPTS and PES_ATEMPTS identify the number of calls required to obtain a completion.

3.22 Refusal Conversions

In addition to making numerous call attempts and spreading these attempts over day, evening and weekend time slots, efforts were made to "convert" refusers on both the CPS and PES. Respondents and/or households who refused to participate when initially contacted by the interviewer were called a second time. In the CPS refusal conversion attempts were made in the last several days calling period whereas in the PES the conversion attempts were typically made over the last three weeks of the 69 day calling period. It is not surprising, therefore, that interviewers had somewhat more success in converting refusals in the PES than in the CPS. However, given the much larger number of refusals to the CPS than the PES there are more "converted refusals" in the CPS data files (374 observations) than in the PES file (210). The variables CPS_REFUSALS and PES_REFUSALS identify whether the interview was a "standard" completion or a "converted" refusal.

The careful attention to the number and timing of callbacks and refusal conversions is designed to increase the response rate, thereby improving sample representativeness. Many researchers have found that respondents who are "hard-to-reach" and those who "refused" have characteristics that are somewhat different from typical survey responders (Dunkelberg and Day, 1973; Fitzgerald and Fuller, 1982; and McDonald, 1979).

3.23 2008 Mail-back Survey

The PES respondents who provided mailing addresses received up to five contacts encouraging them to complete and return the mail-back questionnaire. The first contact included the questionnaire, a covering letter, and a postage-paid pre-addressed return envelope. The second was a reminder/thank you card (physically like a post card) sent one week after the first questionnaire package was sent. The first and second mail contacts were sent to all respondents. The mailings were staggered and sent every week at the start of the PES calling and somewhat less often near the end of calling. A second questionnaire (covering letter and return envelope) were sent only to non-responders and typically were mailed about three weeks after the first reminder card. One week later the second reminder card was sent. Finally, telephone calls were made to all non-responders about two weeks after the last reminder card was sent.

3.3 Response Rate and Reinterview Rates

3.31 Campaign-Period Survey Response Rate

There are numerous ways to calculate response rates in survey research (Dillman, 2000; Smith, 1995; Groves, 1989; and Groves and Lyberg, 1988). The method used in this project is conservative; most other ways of calculating the response rate would produce inflated values. The response rate was defined as the number of completed interviews divided by the estimated number of eligible households times 100 percent.

Details on the calculation of the response rate for the CPS are as follows. Of the 9,422 telephone numbers included in the sample, 6,652 were identified as being eligible households (completions [n=3,257] + refusals [n=2,634] + callbacks [n=761], see Table 3.2). Not eligible households (respondent was unable to speak English or French, was not physically or mentally healthy enough to complete the interview, was not a Canadian citizen (858), and nonresidential and not in service numbers (1,498) accounted for 2,356 of the telephone numbers. It was not possible to determine the eligibility status for 414 of the sample telephone numbers. For response rate calculations, it was assumed that the proportion of these 414 numbers which were eligible household numbers was the same as it was in the rest of the sample.

Table 3.2 Final Sample Disposition and Response Rate: CPS

Results	number	percent
completions	3,257	35
refusals	2,634	28
callbacks	761	8
ill/aged/language problem/ absent/not a citizen	858	9
not-in-service & nonresidential	1,498	16
eligibility not determined	414	4
total	9,422	100
household eligibility rate	-	.738455
estimated number of eligibles	6,958	-
response rate	-	47

This proportion, or "household eligibility rate" was .7385 (eligibles [6,652]/(eligibles [6,652] + not eligibles [2,356]) = .7385). The estimated total number of eligibles was then computed as 6,958 (6652 + [.7385 x 414] = 6,958). Dividing the number of completions (3,257) by the estimated number of eligibles (6,958) gives a final response rate of 47 percent. This is a lower response rate than any previous CPS. In 2006, for new RDD sample the response rate was 57% and in 2004 it was 55%.

The short data collection time period for the CPS, 18 days, had a negative effect on response rates. All things being equal, spreading call attempts over a longer period of time and maximizing time between initial refusals and conversions attempts improve response rates. In addition, there has been a general decline in response rates at ISR and elsewhere over the last two decades. The lower response rate for the Canadian Election Survey is part of a trend in declining response rates reported by American survey researchers in the 1990s. See reports by: Curtin, Presser and Singer, 2005 and Groves, Dillman, Eltinge and Little, 2002.

3.32 Post-Election Survey Reinterview Rate

The PES reinterview rate for the new 2008 sample was 78%, a bit higher the 2004 and 2006 survey. The reinterview rate, at 60%, was lower for the panel sample. The overall reinterview rate was 69%. For the new RDD sample about two thirds of the non-response was accounted for by refusals. Illness or death of CPS respondents, never answered telephones (typically 15 or more calls), and changes in telephone numbers accounted for the remaining non response. A small number of PES respondents had their number changed and the new number was unlisted. In addition, neither the telephone provider nor other household members either could not or would not provide a new number. For the panel sample more than half of the non-response resulted from an inability to locate the 2006 respondent and most of the remaining non-response resulted from refusals.

3.33 Mail-back Survey Return Rate

Of those who provided an address at the end of the PES (76%) and were sent the MBS, 70 percent returned a completed questionnaire. This represents 66 of the respondents to the PES, a considerable higher rate than the approximately 55 percent rate obtained in the last few MBS.

4. Questionnaire Issues and Data Processing

4.1 Introduction

With CATI, interviewers read questions from a computer screen and enter answers directly into a computer file for processing. CATI software automates skip patterns so that interviewers do not have to determine the next appropriate question to be asked, allows questions to be date stamped so they are asked on set days, and provides a mechanism for systematically varying the order in which respondents receive questions or deliberate variations in question wording. CATI code, while relatively easy to follow is cumbersome and requires considerable space as each question (almost always), no matter how small, requires a separate computer screen (a page in CATI language) for viewing. To facilitate use of the data, easy-to-read copies of the CPS and PES questionnaires (as well as an exact copy of the MBS) are provided by ISR. In the easy-to-read versions of the questionnaire CATI code has been replaced with an abbreviated description of how the questionnaire was delivered to respondents. Copies of the CATI questionnaires are available from ISR upon request. Note that most variables in the campaign-period survey include the prefix CPS. The prefixes PES and MBS are used to indicate that the variable is from the post-election, and mail-back survey (respectively).

4.2 Assigning Missing Values

The data file has 4,495 cases but the 1,238 panel respondents did not complete the campaign-period survey thus have "missing data" for all CPS questions. If the user selects for all respondents who completed the CPS before running tabulations (see variable SURVEY_TYPE_08) these cases will be excluded there will be no "missing data" (unless the question was skipped based on the answer to previous CPS questions). Similarly the 806 respondents who only completed the CPS do not have responses to the PES questions, that is they have "missing data" for all PES questions. Respondents who did not complete the MBS have "missing data" in all mail back survey questions.

In the PES a substantial number of questions were not asked of panel respondents, some on the assumption that they provided this information in the 2006 survey and others to keep the overall length of the survey a bit shorter. For these questions, for respondents who completed the PES will have missing data. The skipped questions were: A11, B8C, B8F, party ratings, M1A-M1B, F1N-F2B, G8-G10, G14, K5-K7, S2, and S7-S9B.

Frequently, in both the CPS and PES surveys, whether or not respondents are asked questions is conditional on answers to previous questions. For example, respondents who did not vote in the advance poll and said they were certain not to vote (CPS_B2 and CPS_B3A) were not asked questions about their vote intention (CPS_B4N) and other vote intention related questions. These respondents have "missing data" for the questions they skipped. The most complex conditional logic used in the surveys was utilized in the vote intention section of the questionnaires.

By and large, the reasons for having missing data are self-evident. There are times, however, when the reason for skipping questions is not quite as obvious. In both surveys (CPS and PES) and respondents are asked to rate the leaders on a 100 point scale G1 to G4 in the CPS and F1 to F4 in the PES. Respondents who indicated, when asked about any one leader, that they do not know anything about any of the leaders are not asked to rate the other leaders. (Because the order in which the leaders was asked was random the number of cases assigned a missing value on these questions varies somewhat between questions.)

The same conditional logic is used for in the party ratings CPS_G7 to CPS_G11 and PES_C1A-PES_C1C and PES_C1E and PES_C1F. Respondents who say they do not know anything about any one party are not asked about the remaining parities.

4.3 Province Specific Questions

In both the CPS and the PES, a number of questions were only asked of respondents from Quebec. Typically, these were questions judged not germane to respondents in the rest of Canada, for example, in the CPS only Quebec respondents were asked if they knew the name of Bloc Québécois leader (C3), to rate Duceppe (G4), to rate the Bloc Québécois as a party (G10) and views on Quebec sovereignty (I5). In the Post Election Survey only Quebec respondents were asked to rate the Bloc (C1E) and the leader (F4) of the party. Two Quebec only questions (C10 and G9) were asked about sovereignty.

4.4 Date Specific Questions

In the CPS the question about voting in the advance poll (B2) and seeing the leadership debates (variables R1 to R6B) were asked starting on October 3. A frequency count of these questions will produce missing data for all interviews completed before these dates.

4.5 Randomization of Question Order and Question Wording

The logical operators resident in CATI were used to randomize the order in which respondents received items in several sections of the questionnaire. Given that order effects have been identified in surveys, but are not always easy to predict (Schuman and Presser, 1981), the order randomization was designed primarily as a precautionary measure to limit the impact question order had on overall response. CATI was also used to vary the wording questions. The importance of the way in which issues are framed in question wording has been recognized by survey researchers (Converse and Presser, 1986; and Schuman and Presser, 1981).

The software used at ISR makes it easy for users of the data to determine what effect, if any, the random order and variation in question wording had on response. To examine the effect of randomization the user must run cross tabulations of the questions of interest by the random number variables (in the data set random numbers have the name RANDOM*X*, where *X* is the specific random number used for the question(s) of interest. The random numbers were created

before interviewing commenced and were added to the data set as part of the sample record (along with telephone number, ID number, etc.). The range and value of each random number (i.e., a range of 2 with values 1 and 2 each of which was used about one-half of the time, or a range of 3 with values of 1, 2 and 3 with each used one-third of the time, etc.) can be determined by running a frequency count on the random number, as each random number is a variable in the data set. As noted below, for a small number of questions randomization variables are called CPS_1ST, CPS_2ND, etc, or CPS_1ST_LEADER, CPS_2ND_LEADER, or CPS_1ST_PARTY, CPS_2ND_PARTY, etc.

4.51 Leader Familiarity (CPS)

Respondents were asked if they could "recall the name of"each party leader (E1-E4) in the campaign-period survey. Respondents outside of Quebec were not asked if they knew the name of the Bloc Québécois leader (and, as a result, they have missing data for this question). Answers were considered correct if the respondents gave the first name, last name (or both first and last). The order in which they were asked the name of the leaders was randomized. Each case (respondent) was randomly assigned a four digit string of numbers by CATI's internal random generator (see variables: CPS_1ST, CPS_2ND, CPS_3RD and CPS_4TH). For example, for the first respondent the order could be 1234, for the second the order could be 2341, the third, 3412, etc. During the execution of the survey CATI "read" the first digit of the four digit string and then followed the code as constructed. For example, if the first number in the string of numbers (CPS_1ST) was a "1" he respondent was asked if they knew the name of the leader of the Liberal Party first, the name of the leader of NDP second, the name of the Conservative leader third, and Bloc leader fourth. When the first number in the string was 2 the respondents was asked if he/she knew the leaders in the following order: NDP, Conservative, Bloc and Liberal. Each leader was asked first about one-third of the time outside of Quebec and one-fourth of time in Quebec (as respondents from Quebec were asked to rate all four leaders but those outside of Quebec were not asked to rate Duceppe).

Respondents were asked if they knew the leader of the Green Party but this question was always asked after respondents answered for the other four leaders.

4.52 Party Leader and Party Ratings (CPS and PES)

The same procedures and CATI used logic for rating party leaders and parties as were used for leader familiarity. Respondents in Quebec were asked to rate all four leaders and all four parties and those in the rest of Canada were asked about three leaders and parties. In the CPS respondents were asked to rate the party leaders on a 0 to 100 scale where 0 meant they really disliked and 100 they really liked the leader (G7-G10). Respondents who volunteered they did not know *any* of the leaders well enough to rate them were not asked for ratings. The variable used for the randomization was CPS_1ST_LEADER, CPS_2ND_LEADER, etc). The order of the four digits in the string is randomized for each variable for each respondent. During the execution of the survey CATI went to look at the first digit of the four digit string and then followed the code as constructed in the same way as the code was used for leader familiarity questions.

Rating of party leaders (G7 TO G10) were asked after all party ratings were completed and the variables CPS_1ST_PARTY, CPS_2ND_PARTY, etc., were used in the randomization.

If all four leaders were rated there were 24 different possible orders (the product of 4*3*2*1) for the four leader ratings questions. Given the small number of respondents receiving each of the possible sequences of questions the randomization is precautionary. There is some evidence that ratings on a scale are relative to the first rating given by the respondent. For example if leader "A" is given a 40 and leader "B" is liked more they well be given a number higher than 40, or liked less, a number lower than 40. Thus the first rating acts as an anchor point that respondents adjust up and down as they are asked their ratings for other leaders. Because the exact placement of the first rating may have more variance than the ratings that follow, randomizing the order in which the leaders are ranked will minimize this effect.

In the PES respondents were again asked to rate, on the 0 to 100 scale, both the party (C1A - C1C and C1E) and party leaders (F1-F4). The same logic, employing a four digit string of numbers was used for parties see variables: PES_1ST_PARTY, PES_2ND_PARTY, etc., and for leaders see variables: PES_1ST_LEADER, PES_2ND_ LEADER, etc.

Note, while respondents were asked to rate the Green Party, and the Green Party Leader, on the 0 to 100 scale, these questions, in both the campaign-period and the post-election surveys, always followed the rating of the other parties and leaders.

4.53 Issue Importance (CPS)

Respondents were asked to indicate the level of importance (very, somewhat or not very important) they assigned to a number of issues. The issues and question numbers were:

- I1 fighting crime,
- I2 improving health care / improving social welfare programs,
- I3 creating jobs / dealing with the economy,
- I4 protecting the environment, and
- I5 defending the interests of Quebec (only asked of respondents residing in Quebec).

The order they were asked the issues was determined by RANDOM3. When it was 1 the questions were asked in sequence above, that is: I1, I2, I3, I4, I5. When RANDOM3 was 2 respondents started with the second question and ended with the first, as follows: I2, I3, I4, I5.I1, and when it was three they started at I3 and ended with I1, followed by I2, etc.

Further respondents were assigned to one of the two issues in I2, That is they were asked about *either* improving health care *or* improving social welfare programs. When RANDOM1 was 1 they were asked about health care and when it was 2 they were asked about social welfare. Similarly at I3, when RANDOM2 was 1 they were asked about creating jobs and when it was 2 they were asked about dealing with the economy.

4.54 Values and Party Images (CPS)

Respondents were asked four questions to gage their view of a major policy position for the Liberal and Conservative Parties as well as a their Leaders. Respondents were asked one of two policy questions for the Liberals: the first (JO1) was on "the Green Shift impact on the economy" and the second (JO2) on "the impact of a Carbon Tax on the economy." When RANDOM8 was 1 they were asked about the "Green Shift" and when it was 2 they were asked about the "Carbon Tax." The Conservative question was: "the party was a threat to Canada's social programs" (J1_1) when RANDOM6 was 1 or "if the party had hidden agenda" (J1_2) when RANDOM6 was 2).

For Dion the leadership question was that he either "only cared about is the environment" (J31) or "was a weak leader" (J3_2). For Harper the question asked if he "was just too extreme" (J41) or "had a hidden agenda" (J421). RANDOM7 (Dion) and RANDOM6 (Harper) determined which question was asked where the first statement was read wen the random number was 1 and second one when it was 2.

4.55 Spending Cuts and Taxes (PES)

In the PES, respondents were asked if the government should "spend more, spend less or spend about the same as now" on six different areas.

D1A, defence and the military,
D1B, welfare,
D1C, health care,
D1D, education,
D1F, the environment (there was no question D1E), and
D1G, arts and culture

The order in which respondents in these areas were asked was determined by random number 3. When it was 1 the order was: defence and the military, welfare, health care, education, the environment, and arts and culture. When RANDOM3 was 2, the respondents started with welfare and ended with defence and the military, when it was 3 they started with health care and ended with welfare, etc. Respondents were asked one of the two versions of D1A. When RANDOM5 was 1 respondents were asked about speeding on defence and when it was 2 they were asked about spending on the military.

Respondents were also asked if they thought their personal taxes (D1K) and corporate taxes (D1L) should be increased, decreased, or stay about the same as now. One-half the time respondents were asked the spending questions before the tax questions first (when RANDOM2 was "1") and the rest of the time the respondents were asked about taxes before spending.

At the start of the spending cuts and tax questions section there was an introductory statement. One-half of the respondents got the short version and the other half got a long version. The short

version, when the spending questions were asked before the tax questions read: "And now government spending. Should the Federal government spend more, less, or about the same as now on the following areas?" The long version included the following caveat: "Keep in mind that spending more in one area means spending less in another area or increasing taxes." Which version of the introduction that respondents heard depended on the value of RANDOM4. When it was 1 the short introduction was used and when it was 2 the longer introduction was used.

There was also two versions of the introduction to the pair of questions on taxes. Again, the short version: "Now what taxes" was used when RANDOM4 was 1. The long version, read when RANDOM4 was 2 read: "And now taxes. Keep in mind that cutting taxes means spending less in some areas."

4.56 Use of Terms "homosexuels" and "les gais et lesbiennes" (PES)

Respondents were asked to indicate how much they disliked or like a number of groups on the "0 to 100 scale." In all provinces but Quebec, respondents were asked about "gays and lesbians." In Quebec one-half of the respondents were asked about "homosexuels" (when RANDOM1 was 1) and the other half were asked about "les gais et lesbiennes."(when it was 2).

4.6 Coding of Open-Ended Questions and "Other Specify" Options

4.61 Other Specifies

In a number of items, particularly questions about political parties, and in the demographics, interviewers had the option of writing in an "other specify" response. The information provided by interviewers was reviewed and placed into existing categories when appropriate and additional categories were added when appropriate. Observations that remain in the other category in the final data set normally are few in number, or cover such a wide range of possible options that it was not sensible to create specific codes.

4.62 Most Important Issue (CPS, PES, MBS)

In both the campaign-period (CPS_A2), post-election (PES_A11), and mailback (MBS_G1) respondents were asked to identify the issue which was most important to them personally in the election. Most respondents provided a single response and codes were developed for the more common "double answers" for example, health care and education. If a respondent provided more than one response, that could not be coded into a single category, the first response was coded (unless there was no appropriate code and then the second response was used). The same set of codes (listed below) was used to code both the CPS and PES responses. An attempt was made, when possible, to use categories developed for the 1993, 1997 and 2000, 2004 and 2006 Canadian Election Study. For some of the codes listed below, the text definition is more complete than that used in the value label.

Coding Categories for "Most Important Issue" Questions: 2004 and 2006

1	other & multiple responses [not coded elsewhere]
6	party platform/ what the parties stand for, seeing who stands for what, where they stand on the issues, etc.
10	create jobs/reduce unemployment
20	general mention: debt, deficit, fiscal planning , etc.
25	government spending, government waste, wasting tax dollars, etc.
26	balance the budget/budget
30	economy, fixing the economy
31	economy and health
32	economy and environment
33	health care and environment
35	agriculture, helping farmers
39	oil & gas (fuel) prices: high price of, rising cost of, etc.
45	Afghan war
46	election timing
47	Canadian Wheat Board
48	military/military spending/defense
49	arts & culture
50	taxation
57	health care
58	health & taxes
59	health & jobs
60	social programs, social services, welfare, health & social programs
61	seniors: pensions, ability to retire, other retirement issues, seniors health, etc.
62	family (children's) benefits, childcare funding & funding child care programs
64	health care & education
65	education, programs & funding
71	crime/violence, gun crime, justice system
72	poverty as agenda issue
73	abortion (pro or con)
74	rights & justice issues: aboriginal, women, immigrants, etc.
75	environment & ecological issues (pollution, etc.)
76	moral issues, family values (regardless of direction)
77	gun control/registry, Bill C68
78	immigration as an issue
79	foreign affairs/US relations, security issues
80	Quebec sovereignty/Quebec interests
81	national unity
82	Federal/Provincial relations, "fiscal inequality"
83	Electoral reform & procedural reform issues
84	Canada's future, stability
90	sponsorship issue, corruption, dishonesty, etc.
91	ethics & effectiveness: accountability, transparency, leadership, honesty, integrity, etc., in leaders and government
92	getting a majority government
93	getting a minority government
94	to defeat Conservatives (Harper) / elect Liberals (Dion)
95	to defeat Liberals (Dion / elect Conservatives (Harper)
97	none, no issue important/too many to single out
98	don't know/not sure/not paying attention
99	refused

4.63 Occupation Coding

Respondents to the post-election survey were asked for their occupation. Respondents who only completed the campaign-period survey, panel respondents, and those who were not working or refused to provide their occupation all have missing values for these questions. The text answers provided by respondents were coded in numerical codes. The value numbers and value labels refer to NOC codes (HRDC version). Dictionaries for both sets of codes are government of Canada publications.

4.64 Geographic Information

Using postal codes, provided by the respondents or as determined by using a reverse telephone directory when the respondent did not provide this information, and the Statistics Canada Postal Code Conversion File geographics information was added for survey respondents. The conversion file provides dissemination areas, municipalities, census tracts, urban/rural classification and considerable additional information, including most importantly the federal electoral districts. In the data file these variables are presented after the telephone data and before the mail back data.

References

Babbie, Earl. 1992. *The Practice of Survey Research* (6th Edition). Belmont, California: Wadsworth.

Converse, Jean M. And Stanley Presser. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverley Hills: Sage.

Curtin, Richard, Stanley Presser and Eleanor Singer. 2005. "Changes in Telephone Survey Nonresponse over the Past Quarter Century." *Public Opinion Quarterly*, 69(1), 87-98.

Dillman, Don A. 2000. *Mail and Internet Surveys: The Tailored Design Method* (2nd edition). New York: John Wiley and Sons.

Dunkelberg, William C. and George S. Day. 1973. "Nonresponse Bias and Callbacks in Sample Surveys." *Journal of Marketing Research,* 10, 160-168.

Fitzgerald, Robert and Linda Fuller. 1982. "I Hear You Knocking But You Can't Come In: The Effects of Reluctant Respondents and Refusers on Sample Survey Estimates." *Sociological Methods and Research*, 11 (1), 3-32.

Groves, Robert M. and Lars E. Lyberg. 1988. "An Overview of Nonresponse Issues in Telephone Surveys." In *Telephone Survey Methodology*, Editors: Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II and Joseph Waksberg. New York: John Wiley and Sons, 191-212.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley and Sons.

Groves, Robert M., Don A. Dillman, John L. Eltinge and Roderick J. A. Little. 2002. *Survey Nonresponse*. New York: John Wiley & Sons

Hawkins, Darnell F. 1975. "Estimation of Nonresponse Bias." *Sociological Methods and Research,* 3 (4) 461-488.

Kalton, Graham. 1983. *Introduction to Survey Sampling*. Newbury Park, California: Sage.

Kish, Leslie. 1965. *Survey Sampling*. New York: John Wiley and Sons.

Lessler, Judith T. and William D. Kalsbeek. 1992. *Nonsampling Errors in Surveys*. New York: John Wiley and Sons.

McDonald, P. 1979. "Noncontact and Refusal Rates in Consumer Telephone Surveys." *Journal of Marketing Research*, 16, 478-484.

O'Rourke, D. and J. Blair. 1983. "Improving Random Respondent Selection in Telephone Surveys." *Journal of Marketing Research*, 20, 428-432.

Schuman, Howard and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. Orlando: Academic Press.

Smith, Tom W. 1995. "Trends in Non-Response Rates." *International Journal Of Public Opinion Research,* 7 (2), 157-70.