The 2004 and 2006
Canadian Election Surveys

Technical Documentation

The 2004 and 2006 Canadian Election Surveys

Technical Documentation

David A. Northrup

Institute for Social Research
York University
2007

Conditions of Release

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

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 André Blais, Départment de Politique Science, Université de Montréal, CP6128 Succ. Centreville, Montréal, H3C 3J7.

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

Table of Contents

List of Tables

1. Study Description

1.1   Introduction

As a result of the Canadian Federal Election held on Monday June 28, 2004, the Liberal Party formed a minority government. Telephone surveys, conducted during the campaign and after the election with a random sample of Canadians, provide data that can be used to help explain the election outcome. Within 17 months Canadians again went to the polls and the result of the January 23, 2006 vote was a second minority government, but under the leadership of the Conservative Party. Efforts were made to interview the respondents to the 2004 survey again for the 2006 survey and about one-half of the 2004 survey respondents also answered the 2006 survey. These re-interviews provide a *panel* or longitudinal sample component where a respondent's answers to questions in the first survey can be compared to those of the second survey. Many of the questions asked in 2004 were repeated in the 2006 survey. The remaining respondents to the 2006 survey were a random sample of Canadians

As was the case in the 1988, 1993, 1997 and 2000 Canadian Federal Elections, the Institute for Social Research (ISR) at York University conducted pre- and post-election telephone surveys and, except for the 2006 study, a mail-back survey. The surveys were completed under the direction of the Canadian Election Team (Blais, Everitt, Fournier, Gidengil and Nevitte). This Technical Report outlines the design and conduct of the 2004 and 2006 Canadian Election Surveys (CES). Other than the panel component, the design of the 2004 and 2006 studies largely replicate those of the previous Canadian Election Studies completed at ISR.

For the non-panel respondents, in both 2004 and 2006, modified random digit dialling (RDD) procedures were utilized to select households, and, within households, the birthday selection method was used to select respondents. Data collection was completed with Computer Assisted Telephone Interviewing (CATI).

The 2004 CES included three survey components: the Campaign-Period Survey (CPS), the Post-Election Survey (PES) and the Mail-back Survey (MBS). Interviews in the CPS were collected over the 36 days of the campaign, from Sunday May 23 (the day the writ was dropped) to Sunday June 27, 2004, the day before the election on Monday June 28. The PES survey commenced one week after the election on July 5, and ran to Sunday September 19, 2004. The gap between the election and the start of the PES was not typical of earlier CES where the PES started one or two days after the election. The delay was caused by uncertainty about who would form the government, depending on recounts, and the desire not to start the PES on the Canada Day holiday weekend. The MBS started about one week after the start of the PES, and questionnaires returned to ISR before the end of November, 2004 were added to the data file.

The 2006 CES included two survey components: the Campaign-Period Survey (CPS)[1] and the Post-Election Survey (PES). Interviews in the CPS were collected between November 30, 2005, the day after the Liberal government lost a confidence motion over the budget, and January 22, 2006 the day before the vote. Interviews were not completed between December 22 and December 27, 2005 and December 30, 2005 to January 1, 2006. The PES survey commenced the day after the vote, January 24, 2006 and ran to March 27, 2006.

The number of interviews and sample source for each survey is summarized in Table 1.1. In the 2004 survey there are 4,323 CPS interviews (all from the RDD sample) and 3,138 of these respondents completed the PES and 1,674 the MBS. In the 2006 survey we have 1,999 CPS interviews with respondents from the 2004 survey and 2,059 CPS interviews with respondents from new RDD sample (generated for the 2006 survey). The PES completions were about equally divided between the panel respondents and the new RDD respondents.

Table 1.1   Completed Interviews by Survey and Year

	CPS			PES			MBS
	RDD	panel	total	RDD	panel	total	
2004	4,323	-	4,323	3,138	-	3,138	1,674
2006	2,059	1,999	4,058	3,250	1,566	1,684	-

A rolling cross sectional sample release was employed for both the 2004 and the 2006 CPS. Essentially, the list of randomly generated telephone numbers was equally divided among each of the days of calling. Thus, after a few days of calling and for the rest of the survey interviews were completed with a mix of easy and hard to reach people. The respondents from the 2004 survey who were called as part of the 2006 CPS -- the panel respondents -- were also randomly allocated to each day of calling. In order to increase the chances of obtaining interviews with 2004 respondents however, the panel sample was distributed in the first 45 of the 50 days in the CPS calling period.

Twenty percent of the CPS interviews in the 2004 CES were completed by the Hitachi Survey Research Centre at the University of Toronto. All of the remaining surveys (including all of the

---

[1]   Although the CPS for the 2006 election started in 2005, most of the interviewing was completed in 2006. While the 2005/06 survey would be a more complete name for the survey, using the single year of 2006 is less cumbersome.

2006 interviews) were conducted at ISR. Because the Hitachi centre uses the same Computer Assisted Telephone Interviewing (CATI) software as ISR it was possible to send them an exact copy of the questionnaire. ISR also provided the sample used by Hitachi for the proportion of the CPS they completed. The selection of respondents, number of call attempts, the response rates and the daily distribution of the sample, was the same for the two survey houses throughout the conduct of the CPS.

The naming conventions for the variables in the data file indicate the survey source (CPS, PES or MBS) as well as the survey year (2004 or 2006). Thus 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). In both the CPS and PES surveys, and in both 2004 and 2006, respondents were asked what they thought was the most important issue in the election. The variables are named CES04_CPS_A7 and CES06_CPS_A2, CES04_PES_A1 and CES04_PES_A6 (where CES04_PES_A7 stands for **C**anadian **E**lection **S**tudy 20**04** post-election survey, question **A7**, etc.).

The Institute's CATI software is from the Computer-Assisted Survey Methods Program (CSM) at the University of California, Berkeley.

Easy-to-read copies of the questionnaire accompany this documentation. Complete copies of the CATI questionnaires are available at ISR upon request. (Contact Anne Oram oram@yorku.ca.) An explanation of the way in which CATI was used in the survey is included in Section 4 of this documentation.

2.   SAMPLE DESIGN

2.1   Introduction

The RDD sample for the CPS in both 2004 and 2006 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. 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 blocks of 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) 651-8510 to (416) 651-8519 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 (Residential Telephone Service Survey, Catalogue 56M0001X) estimates that 3.7 percent of the private households in Canada do not have a residential telephone number. Almost equal numbers have either no telephone or a cell phone only. Some cell phones are included in RDD samples but we do not know the overall rate of coverage for cell phones.

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 2006 survey, the respondent was the person who was randomly selected to complete the 2004 CPS. Interviewers, when they asked to speak to the person who did the 2004 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.

Conventionally, most users of survey data wish to have the same number of observations in the weighted and unweighted data set.[6] This 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 sample among each household size category according to the proportion of interviews completed in each household size category.

In the 2004 campaign-period survey there were 4,323 households in the sample and 1,267 were one-adult households, 2,271 were two-adult households, 508 were three-adult households, etc. (Table 2.1 and variable CPS_NADULTS in the 2004 CPS data). 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: 508 times 3 which gives 1,524 three-adult households in the weighted sample. In the campaign-period survey there are 8,513 weighted cases. Second, the 8,513 weighted cases are adjusted down to the original sample size of 4,323 (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: (1,524/8,513) * 4,323 = 773.90. 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: 773.90/508 = 1.523435.

Table 2.1   2004 Calculation of Household (HH) Weights (CES04_CESHHWGT)

HH Size	# of HHs	Weighted Cases	Adjustment	Weight
1 adult	1,267	1,267	643.3973	0.507812
2 adults	2,271	4,542	2,306.4800	1.015623
3 adults	508	1,524	773.9049	1.523435
4 adults	219	876	444.8429	2.031246
5 adults	47	235	119.3357	2.539058
6 adults	8	48	24.3750	3.046869
7 adults	3	21	10.6640	3.554681
Totals	4,323	8,513	4,323.0000	

---

[6] While such weights are common they do not include a downward adjustment in sample size to compensate for design effects. Another option used by some researchers is to "weight up" to the population.

Using the same procedures the households weights for the 2006 survey are detailed in Table 2.2.

Table 2.2   2006 Calculation of Household Weights (CES 2006 HHWGT)

Household Size	No. of Households	Weighted Cases	Adjustment	Weight
1 adult	1,550	1,550	841.0082	0.542586
2 adults	1,913	3,826	2,075.9340	1.085172
3 adults	371	1,113	603.8981	1.627758
4 adults	171	684	371.1288	2.170344
5 adults	39	195	105.8043	2.712930
6 adults	7	42	22.7886	3.255515
7 adults	2	14	7.5962	3.798101
11 adults	5	55	29.8422	5.968445
Totals	4,058	7,479	4,058.0000	

For the 2006 data file there is an issue with respect to whether or not the CPS interviews should be treated as a cross section, that is combining the RDD and panel sample, and weighting it to produce national estimates. Given the different methods of sample selection caution should be used when combining the two sample components in the 2006 CPS. The extent to which the panel is the same as, or different from the cross section can be examined by using the variable RECALL, which identifies panel and RDD observations in the 2006 data file.

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 ®, 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.

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 are detailed in Tables 2.2 and 2.3. In terms of the percentage of sample per province, the design called for a slight over representation of the eight smaller provinces and a corresponding under representation in Ontario and Quebec. For example, in the 2004 CPS survey Alberta has 9.02 percent of the households in the country, but 10.11 percent of the households in the sample. Conversely, Ontario which has 36.17 percent of Canada's population has only 30.49 percent of the sample.

Table 2.3   2004 Sample Distribution and Provincial Weight Variable (CES04_CESPWGT)

Province	# of HHs*	% of HHs	# HHs Sample	% HHs Sample	Weight
Nfld	185,495	1.71	120	2.78	0.6159
PEI	47,960	0.44	115	2.66	0.1662
Nova Scotia	342,590	3.16	112	2.59	1.2188
NB	271,155	2.50	116	2.68	0.9314
Quebec	2,882,030	26.56	1,048	24.24	1.0957
Ontario	3,924,515	36.17	1,318	30.49	1.1864
Manitoba	419,385	3.87	217	5.02	0.7700
Saskatchewan	372,820	3.44	214	4.95	0.6941
Alberta	979,175	9.02	437	10.11	0.8928
BC	1,424,640	13.13	626	14.48	0.9068
Totals	10,849,765	100.00	4,323	100.00	

*   Statistics Canada, 1997. Dwellings and Households: The Nation. Ministry of Industry, Science and Technology, Catalogue No. 93-111, pp 78-89.

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 1.1864 (36.17/30.49). In preparing national estimates, each

Ontario case counts for 1.1864 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 Alberta (.8928), are "weighted down." Caveats about the effect of weighting on the variance estimates noted above apply here as well.

Table 2.4 2006 Sample Distribution Provincial Weight Variable (CES06_CES PWGT)

Province	# of HHs*	% of HHs	# HHs Sample	% HHs Sample	Weight
Nfld	185,495	1.71	102	2.51	0.6522
PEI	47,960	0.44	96	2.37	0.1862
Nova Scotia	342,590	3.16	109	2.69	1.1623
NB	271,155	2.50	101	2.49	0.9889
Quebec	2,882,030	26.56	1,013	24.96	1.0345
Ontario	3,924,515	36.17	1,231	30.34	1.2062
Manitoba	419,385	3.87	203	5.00	0.7498
Saskatchewan	372,820	3.44	206	5.08	0.6486
Alberta	979,175	9.02	408	10.05	0.9523
BC	1,424,640	13.13	589	14.52	0.9167
Totals	10,849,765	100.00	4,058	100.00	

* Statistics Canada, 1997. Dwellings and Households: The Nation. Ministry of Industry, Science and Technology, Catalogue No. 93-111, pp 78-89.

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. ROCNWGT (**r**est **o**f **C**anada) allows for "national" estimates excluding the province of Quebec. As in the case for the National Weight, this weight is the product of a household weight and provincial weight, but these weights are computed by excluding the province of Quebec cases for the computations.

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 Daily Sample Distribution for the Campaign-Period Survey

The importance of campaign dynamics in understanding election results has been documented by a number of researchers (Nevitte, Blais, Gidengil, and Nadeau, 2000; Holbrook, 1996; Blais and Boyer, 1996; Johnston, Blais, Gidengil, and Nevitte, 1996; Johnston, Blais, Brady and Crête, 1992; Bartels, 1988; and Brady and Johnston, 1987). By interviewing a cross section of Canadians each day (and including date of interview as a variable in the data set), it is possible to determine the impact of events during a campaign. Using data from the election survey, the analyst can determine if support for specific policy issues, predictions of the results of the election, or ratings of the Prime Minister or the opposition leaders varied, or remained constant, over the course of the election campaign. Similarly, utilization of a rolling cross section sample facilitates division of the campaign-period data sets into temporal components. For example, analysts can divide the campaign-period data into before and after the leaders' debates, four nine day periods, the beginning, middle and end of the campaign, etc.

It is critical to any analysis which includes date of interview as a continuous or contingent variable, that the sociodemographic characteristics of the survey respondents do not systematically vary over time. Because easy-to-reach respondents (people who are more often home and willing to do the interview when first contacted) have different characteristics than hard-to-reach respondents (Groves, 1989; Hawkins, 1975; and Dunkleberg and Day, 1973), it is important that each day of interviewing includes a mix of easy and hard-to-reach people. Assume, for example, that educational achievement is found to covary with attitudes about a specific election issue such as the importance of creating jobs. If more of the interviews at the beginning of data collection were completed with respondents with lower levels of education (and if they were more supportive of job creation efforts as compared to paying down the debt), and if more of the interviews at the end of data collection were completed with respondents with high levels of education (and they were less supportive of job creation efforts), it would be possible to mistake a change in respondent characteristics for a change in attitudes.

Given the small sample for any one day, the daily variation in the number of completed interviews is expected. However this variation is less pronounced when the number of completed interviews is averaged over a three or five day period. Variation in the number of interviews per day varies in part, because some days, for example Friday tend to have lower co-operation rates and other days, such as Sunday, have higher rates. Other factors such as the weather ("nice" days have lower co-operation rates), the compliment of interviewers working each shift (there is variation among interviewers in the response rates they obtain) and the number of days before the vote (all things being equal the co-operation increases the closer to the vote the interview attempt is made).

There is an attempt to minimize the variation by controlling the amount of sample released each day for calling. Each day of sample release was, within provinces, divided into "sample replicates." Each sample replicate was a random sample of the day's release. Three to five replicates were released each day depending on the factors noted above.

In the 2006 survey, the panel respondents were also randomly allocated to each day of calling. The chances of getting an interview from the sample released on any one day decreases as the date of the election approaches. Telephone numbers from the RDD sample component, and names and numbers of the 2004 respondents released on the first days of calling, were called over a 14 day time period. Numbers released on the day before the election could only be called on one day, numbers released on the second last day before the election could be called on two days, etc. In order to increase the likelihood of getting interviews from the 2004 respondents in the 2006 survey, the release of panel sample was distributed over the first 45 of the 50 days of calling.

2.8   Post-Election Samples

The sample for the post-election surveys was comprised of respondents to the CPS. For the 2006 survey this included not only the new 2006 RDD sample but also the panel sample. At the end of the CPS, interviewers 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." At the start of the PES, the cover sheets were put into a random order (shuffled) so that the time of the first call for the PES was not related to the date of interview, or the day of sample release during the CPS. 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.

2.9   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 75 percent of the PES respondents. This number is down from the 79 percent obtained in the 2000 MBS.

Separate weights were not prepared for the PES and MBS data sets. The re-interview rates are reasonably high and sample attrition between the surveys was not associated with household size or province of interview.

3.   Data Collection

3.1   Introduction

A description of the data collection procedures is 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   Campaign-Period Survey

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 two and four call attempts were made each day (split between day and evening hours) during the first four days that a sample was released. Although over half of the interviews completed in the CPSs took three or fewer call attempts, 10 percent of the completed interviews required ten or more calls (Table 3.1).

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

| | CPS | | | | PES | | | |
| | 2004 | | 2006 | | 2004 | | 2006 | |
Calls	#	%	#	%	#	%	#	%
1	953	22	839	21	581	19	726	22
2	867	20	828	20	507	16	646	20
3	617	14	628	16	351	11	465	14
4	429	10	424	10	302	10	326	10
5	342	8	302	7	240	7	230	7
6-9	693	16	624	15	585	19	477	15
10-14	287	7	265	7	292	9	227	7
15/15+	135	3	148	4	280	9	153	5
Totals	4,323	100	4,058	100	3,138	100	3,250	100

The number of call attempts to complete an interview was about the same in the 2006 survey as in the 2004 survey (Table 3.1). Indeed, the pattern is fairly consistent with that for previous campaign-period surveys completed at ISR. The survey data files and accompanying documentation for these studies are also available at ISR's web site (http://www.isr.yorku. ca/index.html).

3.22   Post-Election Survey

Traditionally, when it comes to the timing for data collection of the PES the goal is to start interviewing one or two days after the election and to complete the PES as quickly as possible after the election. This was the case for the 2006 PES. Reinterviewing for the 2004 PES survey however, commenced seven days after the election on July 5, and ran to Sunday September 19. The delay in starting the PES was caused by the need for recounts in several ridings which had implications about who would form the government, thus affecting the wording of several key questions in the PES and the desire not to start the PES on the Canada Day holiday weekend. In 2006 the PES survey did start the day after the election. The calling for the 2006 PES concluded on March 27, 2006.

On average, the number of call attempts to complete an interview for the PES is about the same as for the CPS. The variables CPS_ATEMPTS and PES_ATEMPTS identify the number of calls required to obtain a completion.

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 in both the 2004 and 2006 surveys. Respondents and/or households who refused to participate when initially contacted by the interviewer were called a second time in both surveys. In the CPS refusal conversion attempts had to be made within the 10 day calling period whereas in the PES the conversion attempts were typically made two or three weeks after the initial refusal. It is not surprising, therefore, that interviewers had somewhat more success in converting refusals in the PES (11.1 and 24.6 percent for 2004 and 2006 respectively) than in the CPS (8.7 and 13.0 percent for 2004 and 2006 respectively). However, given the much larger number of refusals to the CPS than the PES there are more "converted refusals" in the CPS data files (233 and 267 observations in 2004 and 2006 respectively) than in the PES file (66 and 104 for 2004 and 2006 respectively). It is not clear why the conversions rate was better in 2006 than 2004. In general, there were somewhat fewer experienced interviewers in the 2006 study than in 2004, thus conversions may have been easier to complete by the more senior interviewers in 2006. The variables CPS_REFUSALS and PES_REFUSALS, for each study year, 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

13

that are somewhat different from typical survey responders (Dunkelberg and Day, 1973; Fitzgerald and Fuller, 1982; and McDonald, 1979).

3.23   2004 Mail-back Survey

At the end of the 2004 PES, respondents were asked if they would be willing to provide an address so that a mail-back questionnaire could be sent to them. A quarter of the PES respondents declined to provide an address and could not be included in the MBS component of the CES. (This is five percent higher than the 2000 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.

As a result of limited funding, there was no mail-back survey in the 2006 Canadian Election Survey.

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 2004 CPS are as follows. Of the 11,495 telephone numbers included in the sample, 7,531 were identified as being eligible households (completions [n=4,323] + refusals [n=2,439] + callbacks [n=769], 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, etc. [n=1,316], and nonresidential and not in service numbers [n=2,300]) accounted for 3,616 of the telephone numbers. It was not possible to determine the eligibility status for 518 of the sample telephone numbers. For response rate calculations, it was assumed that the proportion of these 518 numbers which were eligible household numbers was the same as it was in the rest of the sample.

This proportion, or "household eligibility rate" was .676 (eligibles [7,531]/(eligibles [7,531] + not eligibles [3,616]) = .67561). The estimated total number of eligibles was then computed as 6,081 (5,777 + [.73 x 395] = 7,881). Dividing the number of completions (4,323) by the estimated number of eligibles (7,881) gives a final response rate of 55 percent (54.854)

Many organizations would not include "eligibility not determined" numbers in the denominator for the response rate calculations on the argument that few of these numbers would be eligible households. (See: Groves and Lyberg, 1988 for a debate on this issue.)   This version of the response rate, sometimes called a completion rate, calculated as completions/known eligibles is 57.4 percent (4,323/7,531).   Other organizations calculate response rates as the number of completions over the number of completions plus refusals. This version of the response rate, which is sometimes known as the participation rate, is 63.9 percent (4,323/4,232+2,439).

Table 3.2   Final Sample Disposition and Response Rate: 2004 and 2006 CPS

Results	number	percent	number	percent
completions	4,323	38	4,058	43
refusals	2,439	20	1,792	19
callbacks	769	7	646	7
not able to locate	-	-	755	8
ill/aged/language problem/ absent/not a citizen	1,316	11	732	8
not-in-service & nonresidential	2,300	20	1,213	13
eligibility not determined	518	4	225	2
total	11,495	100	9,421	100
household eligibility rate	-	.676		.770
estimated number of eligibles	7,881	-		6,669
response rate	-	55		61

Because of the rolling cross section sample design, numbers released for calling on the last days of the campaign survey get fewer calls and no refusal conversion attempts thus a lower response rate is achieved. The response rate for the sample released for the last three days of the campaign was almost 20 points lower than that obtained for the sample released earlier. While this sample release strategy helps to explain why the election survey has a lower response rate than most other

surveys completed at ISR it does not explain the general decline in response rates at ISR and elsewhere. The response rate to the 2004 CPS is five points lower than the 2000 survey. The lower response rate is likely indicative of a similar 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.

While the overall response rate, of 61%, for the 2006 CPS is higher than the 53% obtained in 2004, this results partly from the higher reinterview rate with the 2006 CPS panel respondents. The 1,999 panel respondents represent 64% of the 2004 PES respondents called back by ISR. This 64 percent reinterview rate is higher than the response rate for the new RDD sample (57%).

3.32   Post-Election Survey Re-Interview Rate

The PES reinterview rate in 2004, 73 percent, was lower than any previous CES. (The figures for 2000, 1997 and 1993 and 1988 were 78, 80, 88 and 81 percent, respectively). In 2006, for the RDD component the reinterview rate was 76 percent, somewhat of an improvement over 2004. The reinterview rate for the panel was higher at 84 percent and the overall reinterview rate for the 2006 PES was 80%.

For both years, about three-quarters of the non-response to the PES was accounted for by refusals and callbacks. Illness or death of CPS respondents, never answered telephones (typically 15 or more calls), and changes in telephone numbers (PES respondents had their number changed and the new number was unlisted; the number was changed and the new number listed by the telephone company reached the wrong household; respondent left the household and those remaining in the household either could not or would not provide a new number) account for the remaining non-response to the PES.

3.33   Mail-back Survey Return Rate: 2004

Of those who provided an address at the end of the PES and were sent the MBS, 71 percent returned a completed questionnaire. This represents 53 percent of the respondents to the PES (the same as in 2000) and 39 percent of the CPS respondents, as compared to 42 percent in 2000. (This three percent decrease in the percent of CPS respondents completing the MBS in comparison to the 2000 results from the lower PES reinterview rate (73 percent in 2004 as compared to 78 percent in 2000) as well as the lower percentage of respondents willing to provide an address in 2004 (75 percent in 2004 as compared to 79 percent 2000).

There was limited provincial variation in the return rate, but New Brunswick (48 percent) was somewhat below the average (53 percent) and Newfoundland (61 percent) was somewhat above the average.

4.   Questionnaire Issues and Data Processing

4.1   Introduction

With CATI, interviewers read questions from a computer screen and enter answers directly into a series of computer files 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. (The 2004 CPS, for example, is 120 pages long.) To facilitate use of the data, easy-to-read copies of the CPS and PES questionnaire (as well as an exact copy of the 2004 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 for the 2004 and 2006 Canadian Election Surveys are available from ISR upon request (contact Anne Oram oram@yorku.ca).

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). In addition, in the combined 2004 - 2006 data file the year of the survey is included in the variable names. In both the CPS and PES surveys, and in both 2004 and 2006, respondents were asked what they thought was the most important issue in the election. The variables are named CES04_CPS_A7 and CES06_CPS_A2, CES04_PES_A1 and CES06_PES_A6. So, for example, CES04_CPS_A7 stands for **C**anadian **E**lection **S**tudy 20**04** **c**ampaign-**p**eriod **s**urvey, question **A7**, etc.).

4.2   Assigning Missing Values

Frequently, in both 2004 and 2006 and both the CPS and PES surveys, whether or not respondents are asked questions is conditional on answers to previous questions. For example, respondents who said they were certain not to vote (CES04_CPS_B3@3) were not asked questions about their vote intentions (CES04_CPS_B4@3). These respondents have "missing data" for the questions they skipped. Unlike all other respondents however, these respondents were asked if there was any particular reason why they would not be voting (CES04_CPS_B3A and B3B). Thus for (CES04_CPS_B3B and B3B) we have substantive responses for the non-voters and missing data values for all other respondents. The most complex conditional logic used in the surveys was utilized in the vote intention section of the questionnaires. The logic used in the 2004 and 2006 questionnaires was the same. In the 2006 questionnaire, however, there were two versions of the vote intention question. The first version of the question, CES06_CPS_B4, was the same as the version used in 2004 but the second verison, CES06_CPS_B4N, included the Green Party in the list of parties read out to respondents.

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 both years (2004 and 2006), 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.) Similarly, respondents who say they do not know anything about any one party are not asked about the remaining parities.

In the CPS, (in both 2004 and 2006) respondents were not asked the leadership traits questions about honesty and competency (H1A to H2D), if they indicated they did not know any of the leaders or if they answered they did not know about the leader being asked about in the question. The same convention was used in the 2004 PES. Respondents who could not provide leader ratings were not asked which leader they would describe as arrogant, intelligent, caring, cannot be trusted (I1 to I4). These questions (arrogant, intelligent, etc.) were not asked in the 2006 PES.

In both years, respondents who indicated in the campaign-period survey that they were not working (S4), were not asked to describe their occupation in the post-election survey (SD3).

In addition, there are a few questionnaire design decisions that account for large amounts of missing data. These include questions that were province specific and those asked only before or after set dates.

4.21   Province Specific Questions

In both 2004 and 2006 and 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 rating the leader of the Bloc Québécois. A small number of questions were only asked of residents of the other nine provinces (for example the 0-100 rating about feelings towards Quebec). In the 2004 PES, one question was asked only of respondents in Ontario (about the 2004 provincial budget, see K9) and two questions (about the citizens assembly, see S4A and S4B) were only asked of residents of British Columbia.

4.22   Date Specific Questions

Some questions were only asked as the campaign moved forward, such as voting in the advance poll or seeing the leadership debates. In the 2004 CPS questions about the French and English television debates (variables R1 to R6B) were asked starting on June 15 and 16 respectively. A frequency count of these questions will produce missing data for all interviews completed before these dates. In the 2004 post-election survey, the debate questions (R1 to R6B) were only asked of those respondents who were interviewed in the CPS before the debates (i.e., interviewed before June 15 or 16) so all other PES respondents have missing data for these questions. In the 2006

18

campaign-period survey there were two sets of leadership debates: the first before Christmas and the second after Christmas. In the 2006 CPS respondents were asked about both the English and French debate starting on December 17, 2005. CPS respondents interviewed before December 17, 2005 have missing values for the debate questions. In the 2006 survey, respondents interviewed after January 9 (English) and January 10 (French), were asked about the second debate. Questions about the debates were not asked in the 2006 PES.

One question was added after the start of the 2004 CPS, about the likelihood of there being a minority government (P0). Many more questions were added after the start of the 2006 CPS. Two factors account for these additions. First, a number of questions were added as a result of the passage of time during the campaign. So for example, questions about advance polls were started after the polls were held and questions about Elections Canada advertisements were asked after the ads ran on televison for a few days. Second, given the somewhat unexpected nature of the election call, the questionnaire was still in development when the writ was dropped. Questions (other than the debate questions), added to the 2006 CPS and the date added are:

fin_1	Nov. 30	The maximum amount a person can give to each political party, its candidates, and its riding association in any one year is $5,000. Is this too much, about right, too little, or do you have no opinion?
cand_1	Dec. 5	Are you very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied with the way parties choose their local candidates, or do you have no opinion on this?
i14	Dec. 6	Does the possibility of Quebec separating worry you a lot, somewhat, a little, or not at all?
know_8@1 - know_8@5	Dec. 6	Which party is promising to cut the GST from 7 to 5 percent?
know_7a	Dec. 9	During the campaign have you seen or heard an ad saying "Why not speak up when everyone is listening?"
know_7b	Dec. 9	Did you see it on TV?
know_7c	Dec. 9	Did you hear it on the radio?
know_7d	Dec. 9	Do you know who put out the ad?
b2	Jan. 13	Did you vote in the advance poll?

Respondents interviewed before these dates will have missing values for these questions.

19

4.3   Randomization of Question Order and Response Options

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 RANDOMX, where X is the specific random number used for the question(s) of interest, along with the year and survey identifiers). 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.

4.31   Question and Response Order Experiments in the 2004 and 2006 Campaign-Period Questionnaire

4.311   Most Important Issue (close ended version), 2004 CPS

In question CPS_D1 respondents were asked which of five issues was most important to them personally in the election. The response categories were: a, health care; b, taxes; c, social welfare programs; d, the environment; and e, corruption in government. When CPS_RANDOM1 equalled "1", the order in which the response categories were read was: a, b, c, d, e (where the letters represent the categories as indicated). When RANDOM1 equalled "2", the order was b, c, d, e, a and when it was 3 the order was c, d, e, a, b, etc. The data file has the frequency count for all five orders of the response categories combined. Running a cross tabulation of CPS_D1 by RANDOM1 will provide the frequency count for each of the five orders in which the response categories were read to respondents.

The order of presentation used in D1 was maintained for the follow-up question (D2) which asked the respondents the second most important issue to them in the election.

4.312   Leader Familiarity: 2004 and 2006 CPS

Respondents were asked if they could "recall the name of"each party leader (E1-E4) in both the 2004 and 2006 campaign-period surveys. 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)

In both 2004 and 2006 *all* respondents were first asked about the Liberal leader. The order in which they were asked the name of the three remaining leaders (Conservative, NDP, and Bloc Québécois) was randomized. Each case (respondent) was randomly assigned a three digit string of numbers by CATI's internal random generator (see variables: CPS_1ST, CPS_2ND, and CPS_3RD). For example, for the first respondent the order could be 123, for the second the order could be 321, etc. During the execution of the survey CATI "read" the first digit of the three digit string and then followed the code as constructed. For example, if the first number in the first three digit string of numbers (CPS_1ST) was a "1" he respondent was asked Layton first, if it was a '2' Harper was asked about first and if a "3" Duceppe. CATI then went to the second of the three digit string of numbers (CPS_2ND) and checked the first number, if it was for a party where the interviewer had not asked for the leader CATI then took the interviewer to that question. If the leader pointed to by the random number already had been rated CATI went back to CPS_2ND and took the next number until it pointed to a party for which the interviewer had not asked about. CATI then proceeded to the third string of numbers (CPS_3RD) until all the leader familiarity questions were asked.

4.313   Party Leader Ratings

Respondents were asked to rate the (mainline) party leaders on a 0 to 100 scale where 0 meant they really disliked and 100 they really liked the leader (G1 -G4) for both the 2004 and 2006 campaign-period survey. Respondents who volunteered they did not know *any* of the leaders well enough to rate them were not asked for ratings and non-Quebec respondents were not asked to rate Duceppe, thus accounting for the missing values at the leader ratings questions. As in previous versions of the CPS, the order in which a respondent was asked to rate the leaders was randomized.

Each respondent (case/observation) was randomly assigned a four digit string of numbers (see variables 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 ratings.

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

21

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.

4.314   Party Ratings

As was the case for the ratings of party leaders, the 0-100 ratings for parties were randomized, in both 2004 and 2006 campaign-period survey (G7 to G10). The same logic was used for the randomization, however, the four digit string variables used for the party ratings were: CPS_1ST_PARTY, CPS_2ND_PARTY, etc.). Again there were 24 orders for respondents in Quebec and 12 for respondents from other provinces who were not asked to rate the Bloc Québécois party.

Note, while respondents were asked to rate how much they liked or disliked the Green Party on the 0 to 100 scale this question always followed the rating of the other main line parties.

4.315   Leader Traits and Leader Rankings on These Traits

In both the 2004 and 2006 CPS, respondents were asked to rate the party leaders (four in Quebec and three elsewhere) on two characteristics: honesty (H1A-H1D) and competence (H2A -H2D). In 2004, when RANDOM2 was "1", respondents were asked to give a honesty rating for each leader then a competency rating for each leader. When RANDOM2 was "2" they were asked for the competency ratings first and the honesty ratings second. The order in which the leaders were presented to respondents was the same for each trait, but the 'leader' order was determined by RANDOM3. When RANDOM3 was "1" the order of leaders (for both the honesty and competency ratings) was: Martin, Layton, Duceppe, Harper. When RANDOM3 was 2 the order was: Layton, Duceppe, Harper, Martin; etc.

If, for example RANDOM2 was "2" and RANDOM3 was "3", then the respondent was asked to rate the leaders on competency first and the order the leaders were presented to respondents for the ratings was Duceppe, Harper, Martin and Layton. Respondents were then asked the honesty ratings for the leaders in the same order (Duceppe, Harper, Martin and Layton). Respondents who in the earlier 0 to 100 party leader ratings indicated they did not know enough about any of the leaders to provide a rating skipped the leader traits rankings. In addition, respondents who indicated they did not know or could not rate a particular leader were not asked the leader trait ratings for that leader. Further, when being asked the leader trait ratings, if a respondent indicated twice that they could not rate a leader on a trait, they were not asked to rate the remaining one (or two in Quebec) leader(s) on this trait. Respondents who skipped the leader trait ratings for these reasons have missing data on the leader trait rating questions.

The same randomization was used in 2006, but RANDOM3 was used to determine if respondents were asked about leader honesty or competency first and RANDOM4 was used to determine what order the leaders were presented to the survey respondent.

4.32   Question and Response Order Experiments in the Post-Election Questionnaire

Most of the randomization of questions used in the CPS were duplicated in the PES. This design feature facilitates comparison of answers on questions in both the CPS and PES.

4.321   Leader Familiarity

As in the CPS, PES respondents, in both 2004 and 2006 were asked if they could "recall the name of" each party leader (E1- E4). The same logic, was used in the PES as used in the CPS Each case was randomly assigned a four digit string of numbers by CATI's internal random generator (PES_1ST, PES_2ND, PES_3RD, and PES_4TH). The order of the four digits is randomized for each variable for each respondent. 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).

4.322   Party and Party Leader Ratings

In both the 22004 and 2006 PES survey, respondents were again asked to rate, on the 0 to 100 scale, both the party (C1A - C1D) 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, PES_3RD_PARTY and PES_4TH_PARTY; for leaders see variables: PES_1ST_LEADER, PES_2ND_ LEADER, PES_3RD_LEADER, and PES_4TH_LEADER).

4.323   Spending Cuts and Taxes

In the 2004 PES, respondents were asked if the government should "spend more, spend less or spend about the same as now" on seven different areas. The question identifiers and areas were as follows: D1A = defence and the military; D1B = welfare; D1C = health care; D1D = education; D1E = aid to developing countries; D1F = the environment and D1G = social housing.[7] 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.

When RANDOM1 was "1" respondents were asked the spending questions first and the tax questions second and, when it was "2" the order was reversed (taxes then spending). In addition, the order in which the respondents were asked about seven different areas of spending (defence, welfare, health care, etc.) was also randomized according to the value of RANDOM3. When it was

---

[7]   Note there are no questions PES_D1H and PES_D1I. The list of possible spending cuts was reduced from 9 to 7 during the development of the PES questionnaire.

"1", the order of the spending questions was A to G, when it was "2", the order was B to G followed by A, when it was "3", the order was C to G followed by A then B, etc.

The spending cuts and taxes questions were also asked in the 2006 PES and the same style of randomization was used but with random number 2 (determining the order of taxes and spending) and random number 3 (determining the order in which the spending areas were asked) were used. One of the spending items -- aid to developing countries (D1E in 2004) -- was dropped in the 2006 survey and the wording of the defence and military spending question (D1A in 2004 and 2006) was modified. Half the time respondents were asked about 'spending on defence' and half the time 'military spending.' Which version of the military spending question was asked was determined by the value of RANDOM5.

4.4 Question Wording Experiments: CPS and PES

4.41 Use of Terms "homosexuels" and "les gais et lesbiennes"

In the 2006 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" and the other half were asked about "les gais et lesbiennes." Which version of the question Quebec respondents were asked was determined by RANDOM1. When it was "1" respondents were asked the first version of the question (homosexuels), otherwise they were asked the second version (les gais et lesbiennes).

4.42 Federal Party Identification

The same wording experiment was used in asking federal party identification in both the campaign and post-election survey. In the first version of the questions respondents were asked if they usually thought of themselves "as a Liberal, Conservative, NDP, or none of these." In the second version of the question, "none of these" was replaced with "another party or no party." In the 2004 CPS see Q1A@3 and Q1B@3 and random number 4, and Q1A and Q1B and random number 5 in 2006. For the 2004 PES see K1A@3 and K1B@3 and random number 6, and K1A and K1B and random number 6 in 2006.

4.43 Vote in Last Provincial Election

The last provincial vote question was asked in the PES. In 2004 respondents were randomly assigned (RANDOM7) to one of two versions of the question. The first (K8A), when the random number had a value of "1" asked: "Which party did you vote for?" In the second version of the question (K8B) the respondents were asked the same question, but were also read out the names of the main parties for the respondent's province.

4.44   Lowering the Voting Age

In the 2004 PES respondents were asked about lowering the voting age from either 18 to 17 (G9A) or 18 to 16 (G9B). Which version of the question they were asked was determined by RANDOM4. When it was "1" respondents were asked the A version of the question and when it was "2" they were asked the B version.

4.45   Diversifying Candidate Selection

In the 2004 post-election survey, there were four versions of the question asking respondents about how to insure more women ran in elections. The first version asked about how to get a "minimum percentage of female candidates," the second asked about "having an equal number of male and female candidates." The response options were that the party be "required by law" or this should be "left up to the party to decide." This allowed for four versions of the questions. The first of the four versions, when RANDOM8 was "1" read:

> Which of the following comes closest to your view: 1, political parties should be required by law to have a minimum percentage of female candidates; or 2, each political party should decide for itself how to get more female candidates.

The four versions of the question were:

when PES RANDOM8 was:	1	minimum percentage	by law or up to the party
	2	minimum percentage	up to the party or by law
	3	equal number	by law or up to the party
	4	equal number	up to the party or by law

There was also a question about female candidates in the 2006 PES. It was very similar to the first version of the 2004 questions (detailed above), but it was only asked of half of the respondents. Respondents who were not asked the question about female candidates were asked a question about young candidates. The question read:

> Which of the following comes closest to your view: 1, political parties should be required by law to have a minimum percentage of *young* candidates; or 2, each political party should decide for itself how many *young* candidates it wants.

Which question respondents were asked: female or young candidates was determine by random number 8 (1 for female candidates and 2 for young candidates).

Using the same question format respondents to the 2006 PES were also asked about visible minority candidates (when random number 9 was 1) or aboriginal candidates when random number 9 was 2).

4.5   Imputing Answers for The Chance of Winning Questions

In the 2004 campaign-period survey respondents were asked if each of the Liberals, Conservatives and NDP had a chance to win the most seats in the whole country (see CHANCE1, CHANCE2 and CHANCE3, respectively). For every party that a respondent indicated had a chance, they were asked to assess these chances on a 0 to 100 scale where 0 meant 'no chance at all' and 100 meant 'certain to win.' (See CHANCE4, CHANCE5 and CHANCE6 for the Liberals, Conservatives and NDP, respectively.) Respondents who indicated *only one party*, for example the Liberals had a chance to win the most seats, were not asked to assess the Liberals' chances to win the country on the 0 to 100 scale, nor the chances of the Conservatives or NDP (as they had already answered "no" when asked if these parties had a chance to win the most seats). Rather than leave these "skipped over" assessing the chances on the 0 to 100 scale questions as missing data, the Liberals' chances to win were imputed as 100 and the other two parties chances of winning were imputed as 0. In the survey data the value labels distinguish between 0s and 100s given by respondents and those imputed (values 995 and 996).

The same logic was applied to the questions about each party's chances of winning the respondent's riding (see: CHANCE7 to CHANCE14).

The chances of winning questions were also asked in the 2006 CPS, but respondents were asked which party had the about best chance of winning the most seats, and second best chance, etc. That is, they were not asked about the Liberals chances, etc as was the case in 2004. This wording allowed for a rating of each parties chances, thus there was no imputing of answers in the 2006 chance of winning questions.

4.6   Response Time Measurement

Researchers have explored the relationship between the length of time it takes a respondent to answer a question and how firmly committed they are to their answer (Basal, 1996; Basal, 1993; and Basal and Fletcher, 1991). The 2004 questionnaire was programmed, using the clock resident in the CATI system, to measure how long it took respondents to answer a number of questions. The length of time, in hundredths of a second, was stored in a separate variable. The clock time was set by interviewers hitting a key when they stopped reading the question and again when the respondents started to answer).   Response-time measurements were used for the following questions:

- likelihood of voting CPS_B3@3,
- vote intention CPS_B4@3 and vote choice PES_A3@3

- whether or not Paul Martin knew about the sponsorship scandal before he became the Leader of the Liberal Party CPS_L3@3
- favour/oppose same sex marriage CPS_I1@3 and gays and lesbians allowed to get married (PES_G12@3), and
- federal party identification CPS_G1A@3 and PES_K1A@3 and K1B@3

There was no timing of respondent answers in the 2006 survey.

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

4.71   Open Ended Questions

4.711 Most Important Issue (CPS and PES Questionnaires)

In both the campaign-period (CPSA1) and post-election (PESA1) surveys in 2004 and again in 2006 in both surveys (CPSA2 and PESA6) 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 it was not codeable and then the second response was used). The same set of codes (listed below) was used to code both the CPS and PES responses. The list of categories used is extensive and the number of observations in some categories are quite small. However, the use of a large number of categories makes it easier for the analyst to recode the responses into a smaller set of broader categories. An attempt was made, when possible, to use categories developed for the 1993, 1997 and 2000 Canadian Election Study. Numbers/codes that are *italicised* were only used in 2006, all other codes were used in both 2004 and 2006. 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]	35	agriculture, helping farmers
5	*who wins, general interest in outcome of election*	39	oil & gas (fuel) prices: high price of, rising cost of, etc.
6	*party platform/ what the parties stand for, seeing who stands for what, where they stand on the issues, etc.*	48	military/military spending/defense
		50	taxation
		57	health care
		58	health & taxes
10	create jobs/reduce unemployment	59	health & jobs
20	general mention: debt, deficit, fiscal planning , etc.	60	social programs, social services, welfare, health & social programs
25	government spending, government waste, wasting tax dollars, etc.	61	seniors: pensions, ability to retire, other retirement issues, seniors health, etc.
26	balance the budget/budget		
30	economy, fixing the economy		

62	family (children's) benefits, childcare funding & funding child care programs	*83*	*Electoral reform & procedural reform issues*
63	jobs & social programs	*84*	*Canada's future, stability*
64	health care & education	90	scandal, sponsorship scandal, corruption, dishonesty, etc.
65	education, programs & funding		
70	same sex marriage	91	ethics & effectiveness: accountability, transparency, leadership, honesty, integrity, etc., in leaders and government
71	crime/violence, gun crime		
72	poverty as agenda issue		
73	abortion (pro or con)		
74	rights & justice issues: aboriginal, women, immigrants, etc.	92	getting a majority government
		93	getting a minority government
75	environment & ecological issues (pollution, etc.)	94	to defeat Conservatives (Harper) / elect Liberals (Martin)
76	moral issues (regardless of direction)	95	to defeat Liberals (Martin) / elect Conservatives (Harper)
77	gun control/registry, Bill C68		
78	*Immigration as an issue*		
79	foreign affairs/US relations, security issues	97	none, no issue important/too many to single out
80	Quebec sovereignty/Quebec interests	98	don't know/not sure/not paying attention
81	national unity		
82	*Federal/Provincial relations, "fiscal inequality"*	99	refused

4.712 Main Reason for Not Voting

Most respondents in the 2004 PES reported that they had voted. The 17 percent (or 479 respondents) who did not vote were asked, in an open ended question, why they did not vote (PES_A2B). The answers were coded into eight categories and the value labels for these categories are mostly self- explanatory. Code 1 (did not know who to vote for/what the issues were) and code 6 (not interested in the election or the issues) are similar. To the extent possible code 1 included respondents who seemed to express some interest but feel they do not have enough knowledge whereas code 6 indicated the respondents did not care about the election or the issues. There is good reason to assume this distinction is fuzzy and it may be sensible to combine these categories.

Codes 3 (could not physically make it to the poll), 4 (did not know where to vote) and 8 (not registered) suggest a more structural reason for not voting. Code 7 (did not like the candidates/ they are all the same) and code 5 (vote does not make a difference) suggests voter cynicism and code 2 (no time/forgot to vote) may, in a less direct way, also suggest a lack of faith in the role of elections and party politics in Canada. Given that most respondents do vote, the small number of responses is somewhat limiting for the data analyst.

4.72   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.

4.73   Occupation Coding

In both 2004 and 2006 respondents to the post-election survey were asked for their occupation. Respondents who only completed the campaign-period survey both years as well as respondents who were not working or refused to provide their occupation all have missing values for these questions. Panel respondents were not asked their occupation a second time when they were reinterviewed in 2006. The text answers provided by respondents were coded in numerical codes. Running frequencies on these variables (CES04_PES_SD3 and CES06_PES_SD3) will provide a long list of numbers, where each number corresponds to an occupation. The 2004 numbers refer to CCDO codes and the 2006 numbers refer to NOC codes (HRDC version). Dictionaries for both sets of codes are government of Canada publications.

For the 2004 occupations two socio-economic index scores have also been provided . These indices are commonly referred to as "Blishen Scores" and are based on the male labour force population who reported an occupation in the 1981 Canadian Census. The development of the scale is reported in Blishen, Carroll and Moore (1987). Pineo, Porter and McRoberts (1977), have also created an index and it is based on the 1971 Canadian Census. This index was updated in 1985 to reflect the 1981 Census and is reported in McMaster University (1985). While these indexes are more 'user friendly' than just having the occupation score, they are very dated. The NOC scores are based on 2001 census data.

4.8   Linking Respondents from Three Surveys and the WAVE and RTYPE Variables

Considerable effort was made to ensure, within each household, that the same person completed each survey. For example, in the post-election survey, interviewers were provided with the first name, initial, or other identifier (mother, only male in household, etc.) of the respondent who completed the campaign-period survey as well as their sex and year of birth. However, in comparing the name (or identifier), sex, and year of birth for respondents across the surveys, it is possible to isolate cases where there are differences in sex, age, or name (identifier). In 2004, there are 16 cases where we have reason to be concerned that the same respondent did not answer the surveys. These cases are identified as probable good links in the variable RLINK. There are five cases of concern in 2006 (see variable BADLINK).

While the same questions may have been asked in both surveys and in both years, the questions are presented as separate data points, that is, questions and data were survey specific. A frequency tabulation (marginal) for an item from the 2004 mail-back survey will include valid cases only for the 1,517 respondents who completed the MBS. A "missing values/system missing" code will be assigned to respondents who were part of the Canadian Election Survey but did not complete the MBS.

An alternative to including the missing cases is to specify that only a subset of the data is to be used in the analysis. In 2004 a series of RTYPE variables have been created. The variable RTYPE3 for example, identifies respondents to the mail-back survey (and RTYPE1 and RTYPE2 identify campaign-period and post-election survey respondents respectively). For the 2006 survey the same information is available but for both the survey components and the years of the survey. See variable CES0406_WAVE.

References

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

Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice*. Princeton: Princeton University Press.

Basal, John N. 1993. "Response Latency versus Certainty." *Public Opinion Quarterly*, 57 (1) 54-61.

Basal, John N. 1996. "The How and Why of Response Latency Measurement in Surveys." In: *Answering Questions,* eds., Norbert Schwarz and Seymour Sudman. San Francisco: Jossey-Bass.

Basal, John N. and Joseph F. Fletcher. 1991. "Response-Time Measurement in Surveys." *Public Opinion Quarterly*, 55 (3) 331-346.

Blais, André and Martin Boyer. 1996. "Assessing the Impact of Televised Debates: The Case of the 1988 Canadian Election." *British Journal of Political Science,* 26:143-64.

Blishen, Bernard, Bill Carroll and Catherine Moore. 1987. "The 1981 Socioeconomic Index for Occupations in Canada." *Canadian Review of Sociology and Anthropology*, 24 (4), 465-488.

Brady, Henry E. and Richard Johnston. 1987. "What's the Primary Message: Horse Race or Issue Journalism?" In Gary R. Orren and Nelson P. Polsby, eds., *The New Hampshire Primary and Nomination Politics*. Chatham, New Jersey: Chatham House.

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.

Holbrook, Thomas M. 1996. *Do Campaigns Matter*? Thousand Oaks, California: Sage.

Johnston, Richard, André Blais, Elisabeth Gidengil, and Neil Nevitte. 1996. *The Challenge of Direct Democracy*. McGill-Queen's University Press: Montreal.

Johnston, Richard, André Blais, Henry E. Brady, and Jean Crête. 1992. *Letting the People Decide: Dynamics of a Canadian Election.* McGill-Queen's University Press: Montreal.

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.

Nevitte, Neil, André Blais, Elisabeth Gidengil, and Richard Nadeau. 2000. *Unsteady State: The 1997 Canadian Federal Election*. Oxford, Toronto.

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

Pineo, P., J. Porter and H. McRoberts. 1985. "Revisions of the Pineo-Porter-McRoberts Socio-economic Classification of Occupations for the 1981 Census", *Program for Quantitative Studies in Economics and Population Report*. No. 125. McMaster University.

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.