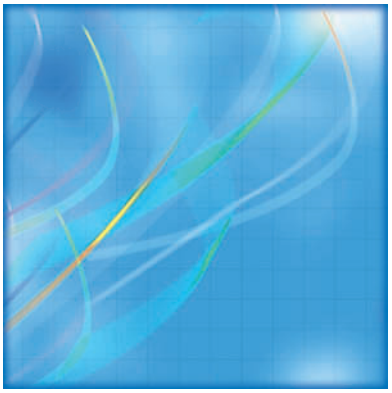
Catalogue no. 92-567-X

2006 Census Technical Report: Coverage



Census year 2006

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website at www.statcan.gc.ca, e-mail us at infostats@statcan.gc.ca, or telephone us, Monday to Friday from 8:30 a.m. to 4:30 p.m., at the following numbers:

Statistics Canada's National Contact Centre

Toll-free telephone (Canada and United States):

Inquiries line	1-800-263-1136
National telecommunications device for the hearing impaired	1-800-363-7629
Fax line	1-877-287-4369

Local or international calls:

Inquiries line	1-613-951-8116
Fax line	1-613-951-0581

Depository Services Program

Inquiries line	1-800-635-7943
Fax line	1-800-565-7757

To access this product

This product, Catalogue no. 92-567-X, is available free in electronic format. To obtain a single issue, visit our website at www.statcan.gc.ca and select "Publications."

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "About us" > "Providing services to Canadians."

Statistics Canada
Social Survey Methods Division

2006 Census Technical Report: Coverage

Census year 2006

Published by authority of the Minister responsible for Statistics Canada

March 2010

Catalogue no. 92-567-X

ISBN 978-1-100-13672-1

Frequency: Occasional

Ottawa

Cette publication est également disponible en français.

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

2006 Census Technical Report

Coverage

Table of contents

Table of contents (continued)

Tables and Charts

Page

Tables and Charts (continued)

Page

Introduction

The 2006 Census required the participation of the entire population of Canada, some 32.5 million people distributed over a territory of nine million square kilometres. Although there are high quality standards governing the collection and processing of the data, it is not possible to eliminate all errors. In order to help users assess the usefulness of census data for their purposes, the 2006 Census technical reports detail the conceptual framework and definitions used in conducting the census, as well as the data collection and processing procedures employed. The principal sources of error, including where possible the size of these errors, are described, as are any unusual circumstances which might limit the usefulness or interpretation of census data. With this information, users can determine the risks involved in basing conclusions or decisions on census data.

This technical report addresses 2006 Census population coverage errors. There are two types of coverage error. **Population undercoverage** refers to the error of excluding someone who should have been enumerated. **Population overcoverage** refers to the error of either enumerating someone more than once or including someone who should not have been enumerated. The latter error is considered negligible. Undercoverage is more common than overcoverage. The net impact of undercoverage and overcoverage on the size of a population of interest is **population net undercoverage**. Net undercoverage is calculated as the number of persons excluded who should have been enumerated (undercoverage) less the number of excess enumerations of persons enumerated more than once (overcoverage). Coverage errors are one of the most important types of error since they affect not only the accuracy of the counts of the various census universes, but also the accuracy of all of the census data describing the characteristics of these universes.

Users of census data should be aware that the presence of coverage error in the 2006 Census means that census products may present the results of a less than complete enumeration and/or include duplicate enumerations. Undercoverage, for example, is highest for young adult males. Users are directed to Section 1 to obtain estimates of 2006 Census population coverage error for a variety of demographic and geographic levels and groupings.

Section 2 gives the 2006 Census conceptual framework and definitions for the population universe, the dwelling universe, and the usual place of residence. This is what the census is trying to measure. Section 3 discusses coverage error, sources of coverage error, census practices that minimize coverage error, the conceptual framework for measuring coverage error, and an introduction to the census coverage studies. Section 4 and Section 5 describe how the 2006 Census was done from the frames to data collection to editing, coding, imputation, and weighting.

Census coverage error is measured by three studies. The 2006 Dwelling Classification Survey (DCS) addressed coverage error resulting from dwelling occupancy classification error. Census data were adjusted for this type of coverage error. The 2006 Reverse Record Check (RRC) measured population undercoverage. The 2006 Census Overcoverage Study (COS) measured population overcoverage. Census data are not adjusted for the population coverage error measured by the RRC and the COS. Rather, estimates of net undercoverage are used in the production of Statistics Canada's demographic estimates of population.

The 2006 Census coverage studies differ from the 2001 Census coverage studies:

1. The 2006 Census was the first time that the names of persons listed on all of the census forms were available in electronic format. This change greatly increased the efficiency of coverage studies since matching could include the name and not be restricted to demographic characteristics.

2. A new coverage study, the COS, was designed to exploit the use of an individual's name for identifying overcoverage. The COS was able to evaluate overcoverage resulting from persons being enumerated more than once with a high degree of accuracy.
3. The measurement of population overcoverage was dropped from the RRC. Consequently, much less field collection was required since only those persons that could not be easily found on the census database were sent to the field.
4. There is a change in terminology. What used to be called 'gross undercoverage' is now 'undercoverage.' The more complete label is '2006 Census population undercoverage.'
5. The Automated Match Study (AMS) was carried out for the 2001 Census and has been in place as a coverage study since the 1991 Census. The AMS was repeated for the 2006 Census but the results were primarily used for evaluating the COS.

Section 6 describes the methodology and results of the 2006 Dwelling Classification Survey (DCS). This survey, carried out after census non-response follow-up, provides information used in the census to account for persons living in non-response dwellings and in occupied dwellings misclassified as unoccupied. This is done by imputing persons onto the census database via the whole household imputation (WHI) procedure. The number of persons added in WHI is a key input for the estimates of population coverage error.

Estimates of coverage error are produced only for the census population universe. Section 7 describes the methodology and results of the 2006 RRC. Section 8 describes the methodology and results of the 2006 COS. Section 9 shows how the results of the RRC and the COS are combined with census data to produce estimates of population coverage error and their associated estimated standard errors. Given the extensive use made of estimates of net undercoverage, it is important to undertake critical and detailed evaluations. Section 10 presents the results of evaluations done for the RRC and the COS as well as an evaluation of the error of closure. The error of closure is the difference between demographic estimates and census counts adjusted for net undercoverage.

Statistics Canada has conducted census population coverage studies since the Reverse Record Check methodology was first applied to the 1961 Census.[1] The historical perspective from the 1971 Census to the 2006 Census is given in Section 11.

Section 12 presents additional topics. We examine the degree to which all persons who should have been enumerated were not and population coverage error for the Aboriginal identity population.

Appendix A contains the 2006 RRC Survey questionnaires and Appendix B lists all of the acronyms used in this report.

This report has been prepared by Mark Armstrong, Karen Bruce, Colleen Clark, Peter Dick, Heather Farr, Josée Morel, Karen Switzer, Alain Théberge and Christian Thibault, members of the Social Survey Methods Division, and Denis Morissette from the Demography Division. Normand Laniel and David Dolson from the Social Survey Methods Division contributed valuable comments on earlier drafts that improved the content and readability of the final report. The support of members of the Census Operations Division, the Demography Division, and the Social Survey Methods Division is noted with appreciation.

1. The first RRC was conducted in 1961 but there was no frame of persons missed in the previous census. The 1966 RRC used the results of the 1961 RRC to construct the frame of persons missed in the 1961 Census.

You can find additional information on census concepts, variables and geography in the *2006 Census Dictionary*. You can find additional information on the complete census process in the *2006 Census reference materials*.

1. Estimates of population coverage error

1.1 Introduction

The census defines the population to be counted and the rules by which the population is to be counted (see Section 4). Coverage errors occur when errors are made relative to these definitions and rules. Important sources of coverage error include the failure to include a dwelling and hence missing the people living there, respondent error by not including all persons who should be included or by including persons who should not be included. This section presents estimates of 2006 Census population net undercoverage, undercoverage and overcoverage. Both undercoverage and overcoverage may result in bias for census counts and estimates because the characteristics of those not included may differ from those who are included, and the characteristics of the duplicates may differ from those who were included only once. Net undercoverage states the extent to which the number of persons included in census data is higher or lower than complete enumeration.

1.2 Net undercoverage

The 2006 Census population net undercoverage rate was estimated at 2.67%.[2] This means that, on a net basis, the census missed 2.67% of those persons who should have been enumerated. The population undercoverage rate was estimated at 4.26% (1,384,372 persons) while the population overcoverage rate was estimated at 1.59% (515,715 persons). An undercoverage rate of 4.26% indicates that persons not included represent 4.26% of the census target population. An overcoverage rate of 1.59% indicates that duplicate enumerations represent 1.59% of the census target population.

Compared to the 2001 Census, coverage error has increased. The rate of undercoverage and the rate of overcoverage increased. Since the overcoverage rate increased more than the undercoverage rate, the estimated net undercoverage rate decreased.

2. This is different from the rate of 2.8% released on September 29, 2008 because incompletely enumerated Indian reserves and settlements are excluded. All of the estimates of coverage error in this document exclude coverage error for this group.

Table 1.2.1 Estimated rates of population coverage error and standard errors for Canada, 2001 and 2006 censuses

	2001 Census		2006 Census	
	Estimated rate	Standard error	Estimated rate	Standard error
Coverage error		**%**		
Undercoverage	3.95	0.13	4.26	0.17
Overcoverage	0.96	0.05	1.59	0.01
Net undercoverage	2.99	0.14	2.67	0.17

Sources: Statistics Canada, 2001 and 2006 Census Coverage Studies.

This section presents estimates of net undercoverage for a variety of geographic and demographic variables:

- Province or territory of current residence at the time of the census
- Age and Sex
- Legal marital status and Sex
- Marital status and Sex
- Mother tongue
- Census metropolitan area (CMA) of Census Day usual residence.

Table 1.2.2 gives the estimated net undercoverage in terms of the estimated net number of persons missed, the standard error of the estimate, and the corresponding estimated net undercoverage rate along with its standard error for each of these variables. Negative estimates of net undercoverage indicate that overcoverage was larger than undercoverage. See Section 9 to understand how this can occur.

Table 1.2.2 Estimated population net undercoverage and standard errors for various characteristics, 2006 Census

Characteristics	Population net undercoverage		Population net undercoverage rate	
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	**868,657**	**53,926**	**2.67**	**0.17**
Provinces and territories				
Newfoundland and Labrador	5,046	2,710	0.99	0.54
Prince Edward Island	1,903	701	1.38	0.52
Nova Scotia	24,558	4,885	2.62	0.53
New Brunswick	16,059	3,105	2.15	0.43
Quebec	60,751	24,077	0.80	0.32
Ontario	465,824	41,363	3.69	0.34
Manitoba	34,330	6,469	2.90	0.56
Saskatchewan	22,594	4,805	2.28	0.50
Alberta	111,353	16,091	3.27	0.49
British Columbia	121,551	16,591	2.87	0.40
Yukon Territory	1,805	194	5.61	0.64
Northwest Territories	1,620	236	3.76	0.57
Nunavut	1,264	176	4.11	0.59
Sex and age group				
Both sexes	**868,657**	**53,926**	**2.67**	**0.17**
0 to 4 years	47,213	10,962	2.72	0.65
5 to 14 years	33,881	17,649	0.86	0.45
15 to 17 years	-9,961	7,949	-0.76	0.61
18 to 19 years	54,842	12,879	6.22	1.55
20 to 24 years	171,783	15,355	7.63	0.73
25 to 34 years	348,205	22,541	8.00	0.56
35 to 44 years	217,060	23,796	4.31	0.49
45 to 54 years	75,931	21,137	1.50	0.42
55 to 64 years	-10,667	19,326	-0.29	0.53
65 years and over	-59,632	17,100	-1.39	0.39
Males	**626,591**	**40,416**	**3.89**	**0.26**
0 to 4 years	25,723	8,123	2.89	0.94
5 to 14 years	15,851	12,577	0.79	0.63
15 to 17 years	-3,303	5,912	-0.49	0.88
18 to 19 years	35,577	10,227	7.78	2.41
20 to 24 years	109,533	11,693	9.46	1.11
25 to 34 years	215,942	16,971	9.91	0.86
35 to 44 years	169,133	18,604	6.66	0.78
45 to 54 years	75,289	16,811	2.98	0.69
55 to 64 years	15,128	13,756	0.83	0.76
65 years and over	-32,284	10,498	-1.74	0.56

Table 1.2.2 Estimated population net undercoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population net undercoverage		Population net undercoverage rate	
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Females	**242,066**	**36,360**	**1.48**	**0.23**
0 to 4 years	21,490	7,578	2.54	0.92
5 to 14 years	18,030	12,398	0.94	0.65
15 to 17 years	-6,658	5,314	-1.05	0.83
18 to 19 years	19,265	7,830	4.54	1.93
20 to 24 years	62,251	10,001	5.69	0.97
25 to 34 years	132,263	14,947	6.08	0.73
35 to 44 years	47,927	14,930	1.92	0.61
45 to 54 years	642	12,826	0.03	0.51
55 to 64 years	-25,794	13,580	-1.40	0.73
65 years and over	-27,349	13,508	-1.13	0.55
Marital status and sex for persons 15 years and over				
Both sexes	**787,563**	**50,136**	**2.94**	**0.19**
Never married	515,925	30,850	6.70	0.43
Married or common-law	153,998	32,329	1.00	0.21
Separated	69,669	14,710	9.75	2.26
Divorced	67,055	16,630	4.39	1.14
Widowed	-19,449	11,268	-1.28	0.73
Unknown	364	229	…	…
Males	**585,016**	**37,750**	**4.43**	**0.30**
Never married	374,462	24,264	8.82	0.62
Married or common-law	105,454	23,903	1.37	0.31
Separated	55,716	12,956	16.84	4.58
Divorced	50,104	10,986	7.86	1.86
Widowed	-1,055	5,193	-0.38	1.85
Unknown	335	228	…	…
Females	**202,547**	**33,531**	**1.49**	**0.25**
Never married	141,464	19,245	4.09	0.58
Married or common-law	48,544	21,934	0.63	0.29
Separated	13,953	6,967	3.63	1.88
Divorced	16,951	12,492	1.91	1.43
Widowed	-18,394	10,001	-1.48	0.80
Unknown	29	21	…	…
Legal marital status and sex for persons 15 years and over				
Both sexes	**787,563**	**50,136**	**2.94**	**0.19**
Never married	595,420	33,517	6.15	0.37
Married	59,420	28,143	0.47	0.23
Separated	79,435	15,523	9.29	1.98
Divorced	72,700	17,981	3.37	0.86
Widowed	-20,187	11,416	-1.27	0.71
Unknown	775	471	100.00	121.47

Table 1.2.2 Estimated population net undercoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population net undercoverage		Population net undercoverage rate	
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Males	**585,016**	**37,750**	**4.43**	**0.30**
Never married	426,685	26,303	8.13	0.54
Married	35,817	20,319	0.57	0.33
Separated	64,579	13,727	15.82	3.89
Divorced	57,493	12,171	5.97	1.34
Widowed	108	5,484	0.03	1.76
Unknown	335	228	…	…
Females	**202,547**	**33,531**	**1.49**	**0.25**
Never married	168,735	21,021	3.81	0.49
Married	23,604	19,601	0.38	0.31
Separated	14,855	7,251	3.33	1.68
Divorced	15,207	13,245	1.27	1.12
Widowed	-20,295	10,014	-1.58	0.77
Unknown	440	412	…	…
Common-law status and sex for persons 15 years and over				
Both sexes	**787,563**	**50,136**	**2.94**	**0.19**
In a common-law relationship	113,148	16,757	3.98	0.61
Not in a common-law relationship	674,415	47,641	2.81	0.20
Males	585,016	37,750	4.43	0.30
In a common-law relationship	78,539	13,016	5.42	0.95
Not in a common-law relationship	506,477	35,714	4.31	0.32
Females	202,547	33,531	1.49	0.25
In a common-law relationship	34,608	10,574	2.48	0.78
Not in a common-law relationship	167,938	31,989	1.37	0.27
Mother tongue				
Both sexes	**868,657**	**53,926**	**2.67**	**0.17**
English	439,185	41,156	2.31	0.22
French	36,020	21,665	0.52	0.32
Other	385,432	27,086	6.89	0.52
English and French	6,913	7,896	2.81	3.30
English and Other	-3,865	6,160	-0.66	1.04
French and Other	-656	2,027	-0.81	2.49
English, French and Other	-193	1,699	-0.30	2.67
Unknown	5,821	1,761	…	…
Census metropolitan area (CMA)				
St. John's	-799	1,399	-0.44	0.77
Halifax	10,905	3,355	2.84	0.90
Moncton	4,403	1,565	3.37	1.24
Saint John	2,941	1,236	2.35	1.01
Saguenay	-2,636	2,785	-1.77	1.84
Québec	-8,195	6,378	-1.16	0.89

Table 1.2.2 Estimated population net undercoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population net undercoverage		Population net undercoverage rate	
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Sherbrooke	4,354	3,751	2.28	2.01
Trois-Rivières	2,181	3,656	1.52	2.58
Montréal	47,647	16,754	1.29	0.46
Ottawa–Gatineau	31,748	11,747	2.73	1.04
Kingston	10,355	8,371	6.36	5.47
Peterborough	2,352	3,875	1.98	3.32
Oshawa	9,763	5,721	2.87	1.73
Toronto	265,681	28,852	4.94	0.56
Hamilton	8,894	7,655	1.27	1.10
St. Catherines–Niagara	6,996	7,004	1.76	1.79
Kitchener	29,651	9,175	6.17	2.03
Brantford	1,731	3,356	1.37	2.69
Guelph	3,638	3,472	2.78	2.73
London	13,664	7,262	2.90	1.59
Windsor	7,140	5,808	2.16	1.80
Barrie	8,030	4,997	4.34	2.82
Greater Sudbury / Grand Sudbury	12,195	7,911	7.15	4.97
Thunder Bay	-745	2,829	-0.61	2.30
Winnipeg	12,975	4,656	1.83	0.67
Regina	2,203	2,231	1.12	1.14
Saskatoon	4,595	2,199	1.93	0.94
Calgary	29,682	7,953	2.68	0.74
Edmonton	23,425	9,276	2.21	0.90
Kelowna	9,031	3,968	5.27	2.44
Abbotsford	3,528	3,161	2.17	1.99
Vancouver	81,275	12,514	3.70	0.59
Victoria	-2,787	4,038	-0.85	1.22
All CMAs	635,822	46,428	2.87	0.22
Outside a CMA	232,836	28,889	2.25	0.29

… not applicable
Sources: Statistics Canada, 2006 Census, 2006 Reverse Record Check and 2006 Census Overcoverage Study.

The standard error provides a measure of the accuracy of the estimates resulting from sampling. The estimates are considered accurate to within plus or minus two standard errors 19 times out of 20. This means, for example, that there are approximately 19 chances in 20 (95%) that the real 2006 Census population net undercoverage rate falls within the range 2.34% to 3.00% (2.67% + or - two standard errors). Or, there are approximately two chances in three (68%) that the real rate falls within the range 2.50% to 2.84% (2.67% + or - one standard error).

Since net undercoverage is a reflection of both undercoverage and overcoverage, the reader is encouraged to also consult the estimates of undercoverage and overcoverage presented in

Table 1.3. A low rate of net undercoverage, for example, may reflect low undercoverage or high undercoverage along with high overcoverage.

Population net undercoverage was highest in Ontario and the western provinces. Among the provinces, the population net undercoverage rate was highest in Ontario at 3.69%, followed by Alberta at 3.27% and Manitoba at 2.90%. Quebec had the lowest provincial rate of population net undercoverage, 0.80%. Net undercoverage was higher in the territories, from 3.76% for the Northwest Territories to 5.61% for the Yukon Territory.

Population net undercoverage was usually higher for men, and highest for young adults. The rate of net undercoverage for males was slightly more than two and a half times the rate for women, 3.89% versus 1.48%. Net undercoverage was highest for those aged 20 to 34 for both males and females. Males aged 25 to 34 had the highest net undercoverage rate at 9.91% versus 9.46% for the younger 20 to 24 group. Among women, the rate was 5.69% for the 20 to 24 group and 6.08% for those aged 25 to 34. Net undercoverage was negative, indicating more persons counted more than once than considered as undercoverage, for males and females aged 15 to 17, for older women (55+) and for older men (65+).

Population net undercoverage for the 15+ population was higher for single persons. Considering marital status, about two thirds of net undercoverage for the 15+ population was from persons who had never been married and were not in a common-law relationship. The rate of net undercoverage for this group was 6.70%. Net undercoverage for persons who were separated and not in a common-law relationship was high, 9.75%, especially for males, 16.84%.

Population net undercoverage was higher for allophones. The net undercoverage rate for those whose mother tongue is English was larger than for those who reported French, 2.31% versus 0.52%. This reflects lower net undercoverage for Quebec. The rate of net undercoverage for allophones, those whose mother tongue is neither of the official languages, was higher, 6.89%. Net undercoverage for allophones approached the level of net undercoverage for those whose mother tongue is English, 385,432 persons versus 439,185 persons.

Population net undercoverage was slightly more common in urban areas. Among those who should have been enumerated in any of Canada's census metropolitan areas (CMAs), net undercoverage was 2.87%. This is slightly higher than the net undercoverage rate of 2.25% for those not living in CMAs.

1.3 Undercoverage

Persons counted as undercoverage are likely to not have been included as a usual resident on the questionnaire that was completed for their usual residence. Persons, for example, who consider their residence as temporary may not have been included as a usual resident elsewhere. Persons without a usual residence, who were otherwise not enumerated, e.g., the 'homeless population,' are also part of undercoverage.

This section presents estimates of undercoverage for a variety of geographic and demographic variables:

- Province or territory of current residence
- Age and Sex
- Legal marital status and Sex
- Marital status and Sex
- Mother tongue
- Census metropolitan area (CMA) of Census Day usual residence.

Table 1.3 gives the estimated undercoverage as the number of persons missed, the standard error of the estimate, the corresponding estimated undercoverage rate, and its standard error. There are some occurrences of negative estimates of undercoverage such as -4,127 persons for women aged 55 to 64. See Section 9 to understand how this can occur.

Table 1.3 Estimated population undercoverage and overcoverage and standard errors for various characteristics, 2006 Census

Characteristics	Population undercoverage				Population overcoverage			
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	**1,384,372**	**53,831**	**4.26**	**0.17**	**515,715**	**3,207**	**1.59**	**0.01**
Provinces and territories								
Newfoundland and Labrador	13,355	2,698	2.62	0.54	8,309	254	1.63	0.05
Prince Edward Island	4,185	697	3.04	0.52	2,282	74	1.66	0.06
Nova Scotia	37,711	4,875	4.02	0.54	13,153	301	1.40	0.03
New Brunswick	26,543	3,099	3.56	0.43	10,485	204	1.41	0.03
Quebec	187,047	24,014	2.46	0.32	126,296	1,749	1.66	0.02
Ontario	654,118	41,310	5.18	0.34	188,294	2,091	1.49	0.02
Manitoba	51,113	6,453	4.32	0.57	16,783	444	1.42	0.04
Saskatchewan	37,734	4,791	3.81	0.50	15,140	365	1.53	0.04
Alberta	161,337	16,072	4.74	0.49	49,984	784	1.47	0.02
British Columbia	204,722	16,539	4.83	0.41	83,171	1,306	1.96	0.03
Yukon Territory	2,325	193	7.23	0.64	521	24	1.62	0.08
Northwest Territories	2,475	233	5.74	0.57	854	34	1.98	0.08
Nunavut	1,706	174	5.55	0.60	442	20	1.44	0.07
Sex and age group								
Both sexes	**1,384,372**	**53,831**	**4.26**	**0.17**	**515,715**	**3,207**	**1.59**	**0.01**
0 to 4 years	70,670	10,902	4.07	0.65	23,457	1,142	1.35	0.07
5 to 14 years	121,698	17,434	3.10	0.46	87,816	2,745	2.24	0.07
15 to 17 years	20,368	7,751	1.56	0.60	30,329	1,764	2.33	0.14
18 to 19 years	78,170	12,797	8.86	1.58	23,327	1,451	2.65	0.17
20 to 24 years	236,589	15,156	10.50	0.74	64,806	2,463	2.88	0.11
25 to 34 years	410,458	22,405	9.43	0.56	62,253	2,478	1.43	0.06
35 to 44 years	269,695	23,689	5.36	0.50	52,635	2,249	1.05	0.05
45 to 54 years	133,175	21,014	2.64	0.43	57,244	2,274	1.13	0.05
55 to 64 years	34,708	19,220	0.95	0.53	45,375	2,029	1.24	0.06
65 years and over	8,840	16,903	0.21	0.40	68,472	2,583	1.60	0.06
Males	**887,171**	**40,339**	**5.51**	**0.26**	**260,581**	**2,494**	**1.62**	**0.02**
0 to 4 years	37,760	8,084	4.24	0.95	12,036	788	1.35	0.09
5 to 14 years	61,012	12,427	3.04	0.64	45,161	1,938	2.25	0.10

Table 1.3 Estimated population undercoverage and overcoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population undercoverage				Population overcoverage			
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
15 to 17 years	12,583	5,770	1.88	0.88	15,885	1,289	2.37	0.20
18 to 19 years	45,985	10,186	10.06	2.45	10,407	918	2.28	0.21
20 to 24 years	141,357	11,572	12.21	1.12	31,824	1,677	2.75	0.15
25 to 34 years	248,935	16,880	11.42	0.86	32,993	1,755	1.51	0.08
35 to 44 years	197,141	18,535	7.77	0.79	28,007	1,602	1.10	0.06
45 to 54 years	104,490	16,732	4.14	0.69	29,201	1,625	1.16	0.07
55 to 64 years	38,835	13,672	2.13	0.77	23,708	1,520	1.30	0.09
65 years and over	-927	10,338	-0.05	0.56	31,357	1,826	1.69	0.10
Females	**497,200**	**36,300**	**3.04**	**0.23**	**255,134**	**2,101**	**1.56**	**0.01**
0 to 4 years	32,911	7,533	3.88	0.92	11,421	825	1.35	0.10
5 to 14 years	60,686	12,246	3.17	0.66	42,656	1,935	2.23	0.10
15 to 17 years	7,786	5,177	1.23	0.83	14,444	1,200	2.28	0.19
18 to 19 years	32,185	7,750	7.58	1.96	12,920	1,122	3.04	0.28
20 to 24 years	95,232	9,837	8.70	0.98	32,982	1,803	3.01	0.17
25 to 34 years	161,522	14,845	7.43	0.73	29,259	1,747	1.35	0.08
35 to 44 years	72,554	14,846	2.90	0.61	24,627	1,578	0.99	0.06
45 to 54 years	28,685	12,728	1.13	0.51	28,043	1,586	1.11	0.06
55 to 64 years	-4,127	13,514	-0.22	0.73	21,667	1,340	1.18	0.07
65 years and over	9,767	13,384	0.40	0.56	37,115	1,825	1.53	0.08
Marital status and sex for persons 15 years and over								
Both sexes	**1,192,004**	**49,990**	**4.44**	**0.19**	**404,441**	**3,825**	**1.51**	**0.01**
Never married	682,631	30,654	8.86	0.43	166,705	3,479	2.16	0.05
Married or common-law	336,734	32,140	2.19	0.21	182,736	3,496	1.19	0.02
Separated	80,142	14,673	11.21	2.28	10,473	1,043	1.46	0.15
Divorced	85,719	16,585	5.61	1.15	18,664	1,225	1.22	0.08
Widowed	6,414	11,169	0.42	0.74	25,863	1,492	1.70	0.10
Unknown	364	229	…	…	0	0	…	…
Males	**788,400**	**37,638**	**5.97**	**0.30**	**203,384**	**2,909**	**1.54**	**0.02**
Never married	463,054	24,134	10.91	0.63	88,592	2,505	2.09	0.06
Married or common-law	201,536	23,764	2.61	0.32	96,082	2,576	1.25	0.03

Table 1.3 Estimated population undercoverage and overcoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population undercoverage				Population overcoverage			
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Separated	61,157	12,935	18.49	4.63	5,441	750	1.64	0.24
Divorced	58,308	10,962	9.15	1.88	8,203	723	1.29	0.12
Widowed	4,010	5,142	1.43	1.86	5,065	727	1.81	0.27
Unknown	335	228	…	…	0	0	…	…
Females	**403,604**	**33,436**	**2.96**	**0.25**	**201,058**	**2,526**	**1.48**	**0.02**
Never married	219,577	19,092	6.35	0.59	78,114	2,418	2.26	0.07
Married or common-law	135,198	21,806	1.77	0.29	86,654	2,370	1.13	0.03
Separated	18,985	6,929	4.94	1.89	5,031	723	1.31	0.19
Divorced	27,411	12,453	3.08	1.44	10,460	988	1.18	0.11
Widowed	2,404	9,916	0.19	0.80	20,798	1,303	1.68	0.11
Unknown	29	21	…	…	0	0	…	…
Legal marital status and sex for persons 15 years and over								
Both sexes	**1,192,004**	**49,990**	**4.44**	**0.19**	**404,441**	**3,825**	**1.51**	**0.01**
Never married	789,350	33,316	8.15	0.37	193,931	3,664	2.00	0.04
Married	204,900	27,946	1.64	0.23	145,479	3,325	1.16	0.03
Separated	91,642	15,472	10.72	2.00	12,207	1,256	1.43	0.15
Divorced	98,542	17,918	4.56	0.87	25,842	1,502	1.20	0.07
Widowed	6,794	11,313	0.43	0.71	26,982	1,525	1.69	0.10
Unknown	775	471	…	…	0	0	…	…
Males	**788,400**	**37,638**	**5.97**	**0.30**	**203,384**	**2,909**	**1.54**	**0.02**
Never married	528,715	26,172	10.07	0.55	102,030	2,623	1.94	0.05
Married	113,187	20,169	1.80	0.33	77,371	2,462	1.23	0.04
Separated	71,078	13,694	17.41	3.94	6,499	941	1.59	0.24
Divorced	69,390	12,130	7.20	1.35	11,897	1,000	1.24	0.11
Widowed	5695	5,435	1.83	1.78	5587	733	1.79	0.24
Unknown	335	228	…	…	0	0	…	…
Females	**403,604**	**33,436**	**2.96**	**0.25**	**201,058**	**2,526**	**1.48**	**0.02**
Never married	260,636	20,864	5.88	0.50	91,901	2,563	2.07	0.06
Married	91,712	19,473	1.47	0.32	68,109	2,236	1.09	0.04
Separated	20,564	7,203	4.60	1.69	5,709	831	1.28	0.19

Table 1.3 Estimated population undercoverage and overcoverage and standard errors for various characteristics, 2006 Census (continued)

Characteristics	Population undercoverage				Population overcoverage			
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Divorced	29,152	13,198	2.44	1.13	13,945	1,119	1.17	0.10
Widowed	1,100	9,924	0.09	0.78	21,394	1,336	1.67	0.11
Unknown	440	412	…	…	0	0	…	…
Common-law status and sex for 15 and over								
Both sexes	**1,192,004**	**49,990**	**4.44**	**0.19**	**404,441**	**3,825**	**1.51**	**0.01**
In a common-law relationship	150,404	16,669	5.29	0.62	37,257	1,710	1.31	0.06
Not in a common-law relationship	1,041,600	47,480	4.34	0.21	367,184	3,904	1.53	0.02
Males	**788,400**	**37,638**	**5.97**	**0.30**	**203,384**	**2,909**	**1.54**	**0.02**
In a common-law relationship	97,251	12,960	6.71	0.95	18,712	1,205	1.29	0.09
Not in a common-law relationship	691,149	35,593	5.88	0.32	184,672	2,933	1.57	0.03
Females	**403,604**	**33,436**	**2.96**	**0.25**	**201,058**	**2,526**	**1.48**	**0.02**
In a common-law relationship	53,153	10,504	3.81	0.78	18,545	1,213	1.33	0.09
Not in a common-law relationship	350,451	31,883	2.87	0.27	182,512	2,610	1.49	0.02
Mother tongue								
Total	**1,384,372**	**53,831**	**4.26**	**0.17**	**515,715**	**3,207**	**1.59**	**0.01**
English	736,800	41,018	3.87	0.22	297,615	3,374	1.56	0.02
French	153,017	21,559	2.22	0.32	116,997	2,139	1.70	0.03
Other	470,155	26,984	8.41	0.52	84,723	2,353	1.52	0.04
English and French	11,540	7,883	4.69	3.35	4,626	452	1.88	0.20
English and Other	5,661	6,124	0.96	1.05	9,526	664	1.62	0.12
French and Other	571	2,012	0.71	2.51	1,227	244	1.52	0.31
English, French and Other	807	1,695	1.27	2.70	1,000	119	1.58	0.20
Unknown	5,821	1,761	100.00	60.49	0	0	0.00	0.00
Census metropolitan area (CMA)								
All CMAs	**982,592**	**46,278**	**4.44**	**0.22**	**346,770**	**3,733**	**1.57**	**0.02**
St. John's	2,460	1,379	1.36	0.78	3,259	236	1.81	0.13
Halifax	16,320	3,337	4.25	0.91	5,415	354	1.41	0.09
Moncton	6,358	1,555	4.86	1.25	1,954	181	1.49	0.14
Saint John	4,440	1,215	3.54	1.00	1,499	227	1.20	0.18
Saguenay	239	2,741	0.16	1.84	2,875	495	1.93	0.34

Table 1.3 Estimated population undercoverage and overcoverage and standard errors for various characteristics, 2006 Census (continued)

	Population undercoverage				Population overcoverage			
Characteristics	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Québec	4,488	6,309	0.63	0.90	12,683	931	1.79	0.13
Sherbrooke	8,085	3,716	4.23	2.02	3,731	510	1.95	0.27
Trois-Rivières	4,994	3,609	3.48	2.60	2,813	582	1.96	0.42
Montréal	101,779	16,663	2.76	0.46	54,133	1,741	1.47	0.05
Ottawa–Gatineau	49,831	11,662	4.29	1.05	18,083	1,406	1.56	0.12
Kingston	12,936	8,356	7.95	5.54	2,580	493	1.59	0.32
Peterborough	4,138	3,845	3.48	3.35	1,785	482	1.50	0.41
Oshawa	14,044	5,678	4.13	1.74	4,282	705	1.26	0.21
Toronto	347,238	28,739	6.46	0.57	81,558	2,554	1.52	0.05
Hamilton	17,815	7,583	2.54	1.11	8,921	1,047	1.27	0.15
St. Catharines–Niagara	14,254	6,929	3.59	1.81	7,258	1,018	1.83	0.26
Kitchener	35,526	9,134	7.39	2.04	5,875	858	1.22	0.18
Brantford	3,525	3,329	2.79	2.71	1,794	426	1.42	0.34
Guelph	5,459	3,453	4.18	2.75	1,821	364	1.39	0.29
London	19,702	7,217	4.18	1.60	6,038	812	1.28	0.18
Windsor	11,089	5,752	3.36	1.80	3,949	808	1.19	0.25
Barrie	10,523	4,967	5.69	2.84	2,492	544	1.35	0.30
Greater Sudbury / Grand Sudbury	15,308	7,883	8.98	5.04	3,113	659	1.83	0.40
Thunder Bay	1,460	2,762	1.19	2.29	2,205	611	1.80	0.51
Winnipeg	22,093	4,631	3.12	0.67	9,118	477	1.29	0.07
Regina	4,737	2,214	2.40	1.15	2,534	272	1.29	0.14
Saskatoon	8,432	2,173	3.54	0.94	3,837	336	1.61	0.14
Calgary	45,521	7,899	4.10	0.74	15,839	932	1.43	0.09
Edmonton	37,740	9,238	3.57	0.90	14,315	844	1.35	0.08
Kelowna	11,734	3,933	6.85	2.45	2,703	523	1.58	0.31
Abbotsford	7,223	3,118	4.44	2.00	3,695	522	2.27	0.33
Vancouver	129,827	12,406	5.91	0.60	48,552	1,646	2.21	0.08
Victoria	3,275	3,983	1.00	1.23	6,063	663	1.85	0.21
Outside a CMA	401,780	28,722	3.89	0.29	168,944	3,099	1.63	0.03

… not applicable
Sources: Statistics Canada, 2006 Census, 2006 Reverse Record Check and 2006 Census Overcoverage Study.

There are some demographic trends in undercoverage:

- The rate of undercoverage for men was almost twice the rate for women, 5.51% versus 3.04%.
- For both men and women, undercoverage was highest for young adults aged 18 to 34.
- Among young adult males, undercoverage was 10.06% for the 18 to 19 age group, 12.21% for the 20 to 24 age group and 11.42% for the 25 to 34 age group.
- Considering marital status, undercoverage was highest in the population aged 15 or more for those who were not married nor in a common-law relationship. The rate for this group was 8.86%. Undercoverage was also high for separated persons who were not common-law. The rate for this group was 9.75%. In both cases, rates were higher for men than women.

A profile of the person most likely to have been missed in the 2006 Census emerges from Table 1.3 as male, between 18 and 34 years of age, and single (never married nor common-law). Mother tongue other than English or French is also important.

1.4 Overcoverage

Population overcoverage is the number of enumerations in excess of persons who are included in census tabulations more than once, usually twice. This is an error resulting in bias for census counts and estimates because they should only have been included once. Examples of overcoverage include children whose parents have separate residences and each parent includes the children on their census form, persons who need to reside away from their family for reasons of work who are listed on their family's form and also on the form for the dwelling they live in while working, and students away at school who are listed both by their roommates and their parents.

This section presents estimates of overcoverage for a variety of geographic and demographic variables:

- Province or territory of current residence
- Age and Sex
- Legal marital status and Sex
- Marital status and Sex
- Mother tongue
- Census metropolitan area (CMA) of Census Day usual residence.

Table 1.3 also gives the estimated overcoverage in terms of the number of persons included more than once, the standard error of the estimate, the corresponding estimated overcoverage rate and its standard error.

The estimates of overcoverage in Table 1.3 are subject to less sampling error than the estimates of undercoverage. There are some demographic trends in overcoverage:

- Across the provinces and territories, overcoverage varied much less than undercoverage did.
- Overcoverage was only slightly higher for females than males, 1.62% versus 1.56%.
- Overcoverage is highest for children and young adults from age 5 to 34. Rates are highest for young adults aged 20 to 24 at 2.88%.
- As for undercoverage, overcoverage was highest neither for persons who were never married nor in a common-law relationship at 1.96%.

A profile of persons most likely to be counted twice emerges from Table 1.3 as equally likely to be male as female, equally likely to be a child or a young adult, and, if 15+, single or widowed. Analysis of overcoverage revealed that about half of the persons counted more than once were children with parents in separate households, or young adults away from their family home, or families who moved.

2. Census universes

2.1 Introduction

The 2006 Census involved the enumeration of the following five universes:

1. the population universe
2. the dwelling universe
3. the household universe
4. the family universe
5. the geography universe.

The 2006 Census Coverage Error Measurement Program estimates coverage error for the population universe only. This section provides a description of the population universe and the dwelling universe. Since coverage error can be a result of misinterpretation of the concept of usual residence as presented on census forms, this section gives the information provided on the census form itself, and the 2006 Census definition of usual residence. Information on the variables associated with each universe can be found in the *2006 Census Dictionary*.

2.2 Population universe

The population universe of the 2006 Census includes the following groups:

- Canadian citizens (by birth or by naturalization) and landed immigrants with a usual place of residence in Canada.
- Canadian citizens (by birth or by naturalization) and landed immigrants who are abroad, either on a military base or attached to a diplomatic mission.
- Canadian citizens (by birth or by naturalization) and landed immigrants at sea or in port aboard merchant vessels under Canadian registry.
- Non-permanent residents:
  - persons with a usual place of residence in Canada who are claiming refugee status and members of their families living with them
  - persons with a usual place of residence in Canada who hold study permits (covering Census Day) and members of their families living with them
  - persons with a usual place of residence in Canada who hold work permits (covering Census Day) and members of their families living with them.

The population universe of the 2006 Census does not include foreign residents but, since 1991, non-permanent residents are included in the population universe.

The definition of the population universe indicates which persons should be included in the census, but not where these persons should be enumerated. The Canadian census uses the modified *de jure* method of enumeration, whereby persons are to be enumerated at their usual place of residence, even if they are temporarily away at the time of the census. Persons away from their usual place of residence and residing elsewhere in Canada are to be enumerated at their usual place of residence and are considered 'temporary residents' at the other location ('Temporary residents' should not be confused with 'non-permanent residents' which refers to the legal status of the person while in Canada). Persons without a usual place of residence are to be enumerated wherever they happen to be on Census Day. Some countries use the *de facto* method, whereby all persons are to be enumerated wherever they are on Census Day, regardless of their usual place of residence.

2.3 Dwelling universe

A dwelling is defined as a set of living quarters in which a person or group of persons resides or could reside. Only dwellings in Canada are included. There are two types of dwellings:

- A private dwelling is a separate set of living quarters with a private entrance either from outside or from a common hall, lobby, vestibule or stairway inside the building. The entrance to the dwelling must be one which can be used without passing through the living quarters of someone else. In addition, a private dwelling must have a source of heat or power and must be an enclosed space that provides shelter from the elements, as evidenced by complete and enclosed walls and roof and by doors and windows that provide protection from wind, rain and snow.

- A collective dwelling is a dwelling of a commercial, institutional or communal nature. It may be identified by a sign on the premises or by an enumerator speaking with the person in charge, a resident, a neighbour, etc. Included are lodging or rooming houses, hotels, motels, tourist homes, nursing homes, hospitals, staff residences, communal quarters (military bases), work camps, jails, missions, group homes, and so on. Collective dwellings may be occupied by usual residents or solely by foreign and/or temporary residents.

These two main types of dwellings are subject to more detailed classifications:

- Private dwellings can be regular private dwellings or occupied marginal dwellings. Regular private dwellings are further classified into three major groups: occupied dwellings (occupied by usual residents), dwellings occupied solely by foreign and/or temporary residents and unoccupied dwellings. An occupied marginal dwelling is an occupied private dwelling which, because it was not built, maintained or converted for year-round use, does not meet the two conditions for year-round occupancy (a source of heat or power and shelter from the elements). To be included in the census, the marginal dwelling must be permanently occupied by a person or a group of persons who have no other usual place of residence. Examples of occupied marginal dwellings are non-winterized cottages or cabins and unconverted barns or garages. Occupied marginal dwellings are classified as either occupied by usual residents or occupied solely by foreign and/or temporary residents. Marginal dwellings that were unoccupied on Census Day are excluded.

- Collective dwellings are classified into occupied dwellings and unoccupied dwellings. Occupied dwellings are either occupied by usual residents or occupied solely by foreign or temporary residents. In the case of unoccupied collective dwellings, data were collected but are not included in census products.

In summary, the dwelling universe includes:

- regular private dwellings occupied by usual residents
- regular private dwellings occupied solely by foreign or temporary residents
- regular private dwellings that are unoccupied
- marginal dwellings provided they were occupied on Census Day
- collective dwellings occupied by usual residents
- collective dwellings occupied solely by foreign or temporary residents.

The dwelling universe does not include:

- marginal dwellings that were unoccupied on Census Day
- collective dwellings that were unoccupied on Census Day
- dwellings outside Canada.

2.4 Usual place of residence

In most cases, people have only one residence and enumerating them at their usual place of residence is easily done. Enumeration consists of listing all usual residents of the dwelling on Census Day by following the step-by-step instructions at the beginning of the census questionnaire: 'Including yourself, how many persons usually live here, at this address, as of May 16, 2006. *Include all persons who usually live here, even if they are temporarily away.* Consult the instructions on page 3 if needed.'

The Page 3 instructions were:

1. **WHOM TO INCLUDE IN STEP B**
   - **EVERYONE WHO USUALLY LIVES HERE, AT THIS ADDRESS,** including newborn babies and room-mates;
   - **STUDENTS** who return to live with their parents during the year should be included at their parents' address, even if they live elsewhere while attending school or working at a summer job;
   - **CHILDREN IN JOINT CUSTODY** who live here most of the time. Children who spend equal time with each parent should be included in the home of the parent where they are staying on May 16, 2006;
   - **SPOUSES OR COMMON-LAW PARTNERS WHO LIVE ELSEWHERE** while working or studying, but who return here periodically;
   - **LANDED IMMIGRANTS** who usually live at this address;
   - Persons claiming **REFUGEE STATUS** and family members living here with them;
   - **PERSONS FROM ANOTHER COUNTRY WITH A WORK OR STUDY PERMIT** and family members living here with them;
   - **PERSONS** who usually live here but are now **IN AN INSTITUTION** (such as a home for the aged, a hospital or a prison), **IF THEY HAVE BEEN THERE LESS THAN SIX MONTHS**;
   - **PERSONS** staying here on May 16, 2006, **WHO HAVE NO USUAL HOME ELSEWHERE**.

2. **DO NOT INCLUDE IN STEP B**
   - Persons who have their **usual home at another address in Canada** and who are staying here temporarily (for example, persons visiting or persons who have their secondary residence here, at this address).
   - Residents of another country **visiting** Canada (for example, on a business trip or on vacation).
   - Government representatives of another country and their family members.

There are a number of situations where the determination of an individual's usual place of residence is not elementary and special rules have been created to define the usual place of residence:

1. Persons with more than one residence.

   This category includes all persons who have more than one dwelling in Canada that could be considered by them as their usual place of residence. In this situation, the usual place of residence is the place where a person spends the major part of the year. If the time spent at each residence is equal or the person is not sure which one to choose, the residence where he or she stayed overnight on Census Day (between May 15 and 16, 2006) should be considered as his or her usual place of residence. There are two exceptions to this rule:

   (a) Sons or daughters who live somewhere else while attending school, but return to live with their parents part of the year, should consider the residence they share with their parents as their usual place of residence, even if they spend most of the year elsewhere.

(b) Husbands, wives or common-law partners who live away from their families while working, but return to their families regularly (for example, on weekends), should consider the residence they share with their spouse or partner as their usual place of residence, even if they spend most of the year elsewhere.

2. Persons in institutions such as a hospital, a home for the aged, a prison or a correctional centre.

   Persons with no other usual place of residence elsewhere in Canada, or persons who have been in one or more institutions for a continuous period of six months or longer, are to be considered to be usual residents of the institution.

3. Residents with no usual place of residence.

   Residents who do not have a usual place of residence should be enumerated in the dwelling where they stayed overnight between May 15 and May 16, 2006.

4. Persons residing outside Canada.

   Canadian citizens and landed immigrants residing outside Canada on Census Day including:

   - persons aboard Canadian government or merchant vessels
   - Canadian federal and provincial government employees and their families
   - members of the Canadian Armed Forces and their families who do not have a permanent place of residence within Canada occupied by one or more members of their family

   were asked to provide on the census questionnaire the address they use for election purposes or their last permanent address within Canada. This information is then used to determine a geographic location for defining usual place of residence.

3. Population coverage error

3.1 Sources

Although there are high quality standards governing the collection and processing of census data, it is not possible to eliminate all errors. There are two kinds of population coverage error. Population undercoverage is the extent to which persons who should have been enumerated are not included in census data while population overcoverage is the degree to which census data include persons who were enumerated more than once, usually twice.

Undercoverage can occur in the first stage of the census if the list of dwellings constructed to cover the census dwelling universe is incomplete. This risk is higher, for example, if a dwelling is under construction. Conversely, overcoverage can occur if a dwelling is listed twice.

Coverage error is also likely to occur during the field data collection stage. Respondent error is responsible for coverage error when the person completing the census form either excludes or omits someone whose usual place of residence, according to the census rules, is the dwelling; this is undercoverage. Or, he/she includes someone whose usual place of residence is not the dwelling; this will become overcoverage if this person is also included at his/her usual place of residence or somewhere else. In most cases, it is easy to determine someone's usual place of residence. However, there are a number of situations where the process, as stated in the previous section, is not elementary and special rules have been created in order to define an individual's usual place of residence. Although the rules are set out in the census form, the list is long and there may be comprehension challenges. Coverage error may result when the rules are not consulted or when they are incorrectly applied. The notion of Census Day as the reference

date, for example, for determining usual residence is also critical to the potential for coverage error.

Coverage errors can also be introduced during the processing stage at any point where records for persons or households are added or removed from the census database. Records can be erroneously cancelled or lost. Questionnaires may be linked to the wrong record or returned too late to be included.

Although efforts are made to enumerate the homeless population, the risk of undercoverage is high for this population. Some other living arrangements are particularly vulnerable to coverage error. Young adults newly away from home, for example, can be either undercovered because neither the roommates nor a parent lists them, or overcovered because the person is listed on both census forms. Similarly, persons who maintain a second residence because of their employment may be at risk of coverage error.

Users should also be aware of the extent to which Indian reserves and Indian settlements participated in the 2006 Census. In some cases enumeration was not permitted or was interrupted before it could be completed. In other cases the quality of the enumeration was considered inadequate. These geographic areas, a total of 22, are called incompletely enumerated Indian reserves and Indian settlements. Data for 2006 are therefore not available for the incompletely enumerated reserves and settlements, and are not included in tabulations. Similar problems have occurred in previous censuses. In the 2001 Census there were 30 Indian reserves and Indian settlements that were declared incompletely enumerated. Among these, 14 became participating reserves in the 2006 Census.

In order to produce population estimates covering persons living on the 22 incompletely enumerated Indian reserves and Indian settlements, model-based estimates are produced. Since no reliable source exists to verify the assumptions used in the model, the estimates must be used with caution. You can find more information in Section 12.2.

3.2 Control

Potential sources of coverage error were recognized during the planning of the 2006 Census, and the following measures were taken to minimize them:

(a) Collection unit (CU) boundaries were carefully defined and mapped in order to ensure that no geographic areas were left out or included twice.

(b) List/leave areas: The enumerator's (EN) manual contained instructions on how to canvass a CU so as to minimize the risk of missing dwellings. The total number of dwellings from the 2001 Census was provided to the field manager to enable him/her to identify notable change. Also, when the listing operation resulted in a significant difference in the number of dwellings relative to the 2001 Census, the listing was checked. Finally, specific quality control procedures were applied to the EN work to assess and eventually to correct the changes done to the listing. Census frames including the definition of list/leave and mail-out areas are described in Section 4.2.

(c) Mail-out areas: Mail-out was based upon a list of addresses taken from Statistics Canada's Address Register. This list was verified and updated in the fall of 2005 via a block canvassing field operation. The work of the enumerator was subject to quality control procedures.

(d) Collective dwellings: Collectives dwellings are identified before collection. Field staff verify that these dwellings are indeed collectives and, if so, determine whether or not they are occupied.

(e) Special procedures were developed to enumerate persons who have difficulty responding (e.g., difficulty in English and French or literacy problems) and to enumerate persons who are located in special core areas of major cities.

(f) Special procedures were developed to enumerate the population on Indian reserves.

(g) Publicity messages informed Canadians about the census and indicated what to do if they did not receive a questionnaire.

(h) The Census Help Line was available to answer any questions about the census including questions related to coverage.

(i) The questionnaire contained instructions on 'Whom to include' to inform respondents of whom should be included.

(j) The questionnaire included questions asking if there were any persons the respondent was not sure whether or not to list. A telephone follow-up was then done with the respondent to determine if the person(s) in question should or should not be listed on the questionnaire.

(k) Telephone follow-up was done after questionnaire editing when inconsistencies were found on coverage issues or to verify status of households including only foreign or temporary residents.

(l) Non-response follow-up included some dwelling coverage checks.

These procedures, along with appropriate training, supervisory checks, and quality control systems during census collection and processing, helped to reduce the number of coverage errors.

3.3 Definitions

Algebraic definitions of coverage errors are given in this section. Let T represent the total or 'true' number of persons in the census target population. Then, let C be the published census count of the number of persons in the census target population. The error in using C instead of T is then:

$$N = T - C$$

This error, denoted as N, is the **net population coverage error**.

Let U denote **population undercoverage**. U is the number of persons not included in C who should have been.

Let O denote **population overcoverage** where O is the number of persons included in C who should not have been. There are two components to O. The first is persons who were enumerated more than once. These duplicate enumerations should not have been included in C. The census coverage studies focus on duplicate enumerations. The second component of O is persons who were included in C who are not in the census target population. Foreign residents visiting Canada, for example, who are listed on a census form as usual residents of a dwelling should not be included in C. Fictitious persons are another example. The number of persons included that are not in the census target population has been seen by previous studies to be negligibly small. Therefore, the 2006 Census coverage studies did not measure this component of coverage error.

Since U refers to persons who should be included in C and O refers to persons who should not be included in C, the difference between T and C is U less O. That is:

$$N = U - O$$

The true number of persons in the census target population is then:

$$T = C + N = C + U - O$$

An estimate of T is given by $\hat{T}$ where:

$$\hat{T} = C + \hat{N} = C + \hat{U} - \hat{O}$$

$\hat{U}$ is an estimate of the number of persons not included in C that should have been; and $\hat{O}$ is an estimate of the number of persons included in C who should not have been. Let us assume that overcoverage from persons included in C who are not in the census target population is zero. Therefore, $\hat{O}$ is restricted to an estimate of the number of duplicate enumerations. It is the goal of the census coverage studies to produce $\hat{U}$ and $\hat{O}$.

Census population coverage error can be usefully expressed as rates relative to the true population: The undercoverage rate R_U is U expressed as a percentage of T. The overcoverage rate R_O is O expressed as a percentage of T. The net undercoverage rate R_N is the difference between U and O expressed as a percentage of the census target population. These three rates can be estimated by $\hat{R}_U$, $\hat{R}_O$, and $\hat{R}_N$ as follows:

$$\hat{R}_U = 100 * \frac{\hat{U}}{\hat{T}} = 100 * \frac{\hat{U}}{C + \hat{N}}$$

$$\hat{R}_O = 100 * \frac{\hat{O}}{\hat{T}} = 100 * \frac{\hat{O}}{C + \hat{N}}$$

$$\hat{R}_N = 100 * \frac{\hat{N}}{\hat{T}} = 100 * \left(\frac{\hat{U} - \hat{O}}{C + \hat{N}} \right)$$

A positive net undercoverage rate indicates that undercoverage is larger than overcoverage. That is, there are more people not included in the published census count C than the number of duplicated enumerations. This has been, and continues to be, the experience of the Canadian census. For some domains of interest, however, negative net undercoverage has recently been observed.

As defined above, U is the number of persons not included in C who should have been. The census count C is composed of two elements:

$$C = E + I$$

where:

$E =$ the number of enumerations. This is the number of people who were listed on a census form.

$I =$ the number of imputed persons. This is an estimate of the number of persons missed in non-response dwellings and in occupied dwellings erroneously classified as unoccupied. More information on whole household imputation (WHI) can be found in Section 5.7.

Undercoverage, therefore, is a subset of all persons who were not listed on a census form but should have been. It does not include those who were not enumerated either because no completed census form was returned for the dwelling (non-response dwelling) or the dwelling did not receive a form because they were erroneously classified as unoccupied (misclassified occupied dwelling).

In summary, the true population T is composed of the census count C and net undercoverage N. C consists of E plus the number of persons added in WHI I where the imputations are for persons living in non-response dwellings or in misclassified occupied dwellings. N is undercoverage U less overcoverage O.

3.4 Measurement

Two postcensal studies were carried out to estimate 2006 Census population coverage error. The Reverse Record Check (RRC) provided estimates for population undercoverage while the Census Overcoverage Study (COS) estimated population overcoverage.

The RRC and the COS were conducted after census field collection and processing were complete. Preliminary estimates of 2006 Census population coverage error were released March 27, 2008. Following a lengthy and detailed validation exercise with the Demography Division and the provincial and territorial statistical focal points, final estimates were released on September 29, 2008. This release was concurrent with the release of new official population estimates reflecting the update of the base population to the 2006 Census. Census population counts adjusted for net population undercoverage formed the updated base population.

The methodology of the two census coverage studies can be briefly described as follows:

Reverse Record Check (RRC)

In the RRC, a random sample of individuals representing the 2006 Census target population was taken from frames independent of the census such as a list of persons enumerated in the 2001 Census and a list of intercensal births according to provincial birth registries. The 2006 RRC sample consisted of 67,813 persons in the provinces and 1,938 persons in the territories. In addition, 84,522 enumerated persons with a weight of one contributed to the territorial estimates.[3] The 2006 Census database was searched to determine if the persons selected in the sample had indeed been enumerated.

When required, a telephone interview via computer-assisted telephone interviewing (CATI) out of the regional offices (ROs) was conducted to collect further information to declare the individual as in scope or not in scope for the census, and when in scope, to provide further data for searching. An interview was achieved for 84.2% of the 20,114 cases sent to the ROs. Sampling weights were adjusted to account for non-response whereby the total sampling weights of the non-respondents was shared among a group respondents most like the non-respondents in their propensity to respond.

3. The large sample size in the territories is because a different methodology is used. The sample frames were first matched to the entire census database. Matches were classified as enumerated if they were found in the same territory or out of scope if they were found elsewhere. All of the matched persons from the sample frames were included in the RRC sample with a weight of one. An additional sample of 1,938 persons was selected from the non-matches.

Estimates of population undercoverage are based on persons in the RRC sample who were classified 'missed.' These persons have been found to be in scope for the 2006 Census but no evidence of enumeration in the 2006 Census could be found in the 2006 Census Response Database. Nationally, there were 5,431 persons selected by the RRC classified as missed in the provinces and 676 in the territories.

Census Overcoverage Study (COS)
Overcoverage was measured by matching the final 2006 Census database to a partial list of persons who should have been enumerated constructed from administrative data sources, and by matching the 2006 Census database to itself. The COS applied automated exact matching to the administrative sources and probabilistic matching to the census database. Probabilistic matching identifies matches that are close but not exact. Pairs of these potential duplicates were sampled and name and demographic characteristics were examined to identify overcoverage.

4. Census data collection

4.1 General

The data collection stage of the 2006 Census ensured that each of the 13.5 million dwellings in Canada received a census questionnaire. The census aims to enumerate the entire population of Canada, which consists of Canadian citizens (by birth and by naturalization), landed immigrants, and non-permanent residents together with family members living with them. Non-permanent residents are persons living in Canada who have a work or study permit covering the Census Day, or who are claiming refugee status, and family members living with them.

The census also counted Canadian citizens and landed immigrants who were temporarily outside the country on Census Day. This included federal and provincial government employees working outside Canada, Canadian embassy staff posted to other countries, members of the Canadian Forces stationed abroad and all Canadian crew members of merchant vessels.

The Census of Canada uses different forms and questionnaires to collect data. The basic short questionnaire is called the 2A. It is distributed to four in five private dwellings. The 2B is a longer questionnaire that collects the same information as the 2A as well as additional information on a variety of topics. The 2B questionnaire is distributed to one in five private dwellings. Each household that receives a 2A or 2B census questionnaire is asked to enumerate and provide information on all household members who fall into the census target population.

A Form 2C is mainly used to enumerate people posted outside Canada, including Canadian government employees (federal and provincial) and their families, and members of the Canadian Forces and their families. The 2C contains the same questions as the 2B with the exception of housing questions. A Form 2D contains the same questions as the 2B but is used to enumerate northern areas and most Indian reserves, Indian settlements, Indian government districts and 'terres réservées.' In canvasser areas, it is also used to enumerate usual residents of Hutterite colonies.

A Form 3 is an individual census questionnaire used to enumerate persons in collective dwellings. Each person in the collective dwelling completes a separate Form 3. It can also be used to enumerate usual residents in a private household who prefer to be enumerated on their own census questionnaire rather than be included on a 2A or 2B questionnaire. Form 3A is the short version of the questionnaire, and Form 3B is the long version.

4.2 Frames

To ensure the best possible coverage, the country was divided into small geographic areas called collection units (CU). In the 2006 Census, there were approximately 50,000 CUs. A small number

of CUs were designated as canvasser areas where a census enumerator completed a Form 2D for each dwelling. The remaining CUs were assigned a questionnaire delivery method of either mail-out (MO) or list/leave (L/L). The MO method was used for all large and medium-sized cities. Smaller municipalities with a population under about 25,000 were designated L/L along with the remaining rural areas.

A dwelling list frame was used for the MO areas. The list provided the means of identifying and contacting every dwelling in the dwelling universe as of May 16, 2006. The number of dwellings and their addresses were taken from the Address Register (AR), a list of dwellings maintained by Statistics Canada which covers most of the country. Various administrative data are used to update the AR such as telephone billing files and the GST New Housing Rebate File.

The quality and coverage of the dwelling frame in MO areas was enhanced by block canvassing, which took place about eight months prior to Census Day. This was a process whereby field staff surveyed MO areas to update AR dwellings lists to ensure the accuracy of the addresses and dwelling classifications, and ensure that dwellings were correctly coded to the block they belong to. Block canvassing not only identified dwellings that were in existence at the time, but also attempted to identify new planned dwellings and dwellings under construction that were likely to exist by Census Day. There was also a late block canvass operation three to four months prior to Census Day to recanvass high-growth areas.

The frame for the L/L and canvasser areas contains a list of all potential dwellings in these areas. In L/L and canvasser areas enumerators constructed lists of dwellings in a control register known as a Visitation Record. As they did this, questionnaires were dropped off (L/L) or interviews were conducted (canvasser). The AR was also a coverage improvement tool in some L/L areas.

4.3 Collection methods

About 98% of households were enumerated using self-enumeration. For the 70% of dwellings located in MO areas, Canada Post delivered a census questionnaire starting May 2. In the L/L areas covering 28% of the dwellings, the households received their questionnaire from a census enumerator. Householders were asked to complete the questionnaire for themselves and for members of their household and return it either online or in the postage-paid yellow envelope by May 16, Census Day.

About 2% of households were enumerated using the canvasser method. An enumerator visited a household and completed a questionnaire for the household by a personal interview. This method was normally used in remote and northern areas of the country and on most Indian reserves. It was also used in large urban downtown areas where many residents are transient.

For the first time, the 2006 Census offered all households in Canada the option of completing their questionnaire online. Each paper questionnaire had a unique Internet access code printed on the front along with the 2006 Census website address (www.census2006.ca). Respondents needed this access code to complete their questionnaire online. If a questionnaire was completed and returned online, the information was directly submitted into the Data Processing Centre system and was verified for completeness. Approximately 18% of households responded via the Internet. Some households were enumerated through the Census Help Line (CHL), a free, nationwide, multilingual service that was available to all respondents. The CHL collected census information through a computer-assisted telephone interviewing (CATI) system.

During the non-response and follow-up stage of collection (NRFU), households from which a questionnaire had not been received within an acceptable time frame were contacted individually by enumerators (EN) in order to collect their information.

5. Census data processing

5.1 Introduction

Census data processing encompasses everything from the capture of questionnaire data from the completed questionnaires through to the creation of an accurate and complete census database: questionnaire registration, data capture, questionnaire imaging, editing, error correction, coding, imputation and weighting. This section describes each operation.

Automated processes, implemented for the 2006 Census, had to be monitored to ensure that all Canadian residences were enumerated once and only once. The Master Control System was built to control and monitor the process flow. The Master Control System held a master list of all the dwellings included in the census. Each dwelling had a unique identifier providing the link to its questionnaire. This system was updated on a daily basis with information of each dwelling's status in the census process flow (i.e., delivered, received, processed, etc.). Reports were generated and accessible online to the census managers to ensure that operations were efficient and effective.

5.2 Receipt and registration

Respondents completing paper questionnaires mailed them back to a centralized data processing centre (DPC). Canada Post scanned the barcode on the front of the questionnaire through the transparent portion of the return envelope. The envelopes were then transported to the DPC along with a compact disc containing the list of all of the identifiers for the scanned questionnaires. The returned questionnaires were then registered on the Master Control System at Statistics Canada. About ten days after Census Day, a list of all of the dwellings for which a questionnaire had not been received was generated by the Master Control System and then transmitted to field operations for follow-up. Afterwards, registration updates were sent to field operations on a daily basis to prevent follow-up on households whose questionnaires (either paper or electronic) were received after that point in time.

5.3 Imaging and data capture

The 2006 Census was Canada's first census to capture data using automated capture technologies rather than manual keying. There were five steps in the imaging process:

- **Document preparation:** mailed-back questionnaires were removed from envelopes and foreign objects, such as paper clips and staples were detached in preparation for scanning. Forms that were in a booklet format were separated into single sheets by cutting off the spine.
- **Scanning:** scanning, using 18 high-speed scanners converted the paper to digital images (pictures).
- **Automated image quality assurance:** an automated system verified the quality of the scanning. Images failing this process were flagged for rescanning or keying from paper.
- **Automated data capture:** optical mark recognition and optical character recognition technologies were used to extract respondents' data from the images. Where the systems could not recognize the handwriting with sufficient accuracy, data repair was done by an operator.
- **Check-out:** as soon as the questionnaires were processed successfully through all of the above steps, the paper questionnaires were checked out of the system. Check-out is a quality assurance process that ensures the images and captured data are of sufficient quality that the paper questionnaires are no longer required for subsequent processing. Questionnaires that had been flagged as containing errors were pulled at check-out and reprocessed as required.

5.4 Coverage edits

At this stage, a number of automated edits were performed on the respondent data. These edits were designed to detect cases where invalid persons may have been created either due to respondent error or data capture error. Examples include data erroneously entered in a blank person column, crossed off data that was captured in error, or data provided for the same person more than once, usually due to the receipt of duplicate forms (e.g., a husband completed the Internet version and his wife filled in the paper form and mailed it back). The edits were also designed to detect the possible absence of usual residents, when data are not provided for every household member listed at the beginning of the questionnaire. There was also some telephone follow-up for these edit failures.

Data from questionnaires that failed the edits were forwarded to processing clerks for verification. An interactive system enabled the clerks to examine the captured data and compare it with the image if available (online questionnaires would not have an image). Edit failures were resolved by manually deleting invalid/duplicate persons and adding missing ones (i.e., creating blank person records), as necessary and appropriate.

5.5 Completion edits

Following the coverage edits, another set of automated edits was run to detect cases where there were either too many missing responses, or there were indications that data may not have been provided for all usual residents in the household. Households failing these edits were subject to follow-up whereby an interviewer used a computer-assisted telephone interview (CATI) application to telephone the respondent to resolve any coverage issues and to fill in the missing information. The data were then sent back to the Data Processing Center for reintegration into the system for subsequent processing.

5.6 Coding

The long-form questionnaires (2B, 2C, 2D and 3B) contained questions where answers could be checked off against a list, as well as questions requiring a written response from the respondent in the boxes provided. These written responses underwent automated coding to assign each one a numerical code, using Statistics Canada reference files, code sets and standard classifications. Reference files for the automated match process were built using actual responses from past censuses. Specially trained coders and subject matter experts resolved cases where a code could not be automatically assigned. The variables for which coding applied were: Relationship to Person 1, Place of birth, Citizenship, Non-official languages, Home language, Mother tongue, Ethnic origin, Population group, Indian band/First Nation, Place of residence 1 year ago, Place of residence 5 years ago, Major Field of Study, Location of study, Place of birth of parents, Language at work, Industry, Occupation and Place of work.

About 37 million write-ins were coded from the 2006 long-form questionnaires. An average of about 82% of these was coded automatically.

As the responses for a particular variable were coded, the data for that variable were sent to the edit and imputation phase.

5.7 Adjustments for non-response and misclassified occupied dwellings

The Dwelling Classification Survey (DCS) was carried out during processing after non-response follow-up to estimate the error rates in classifying dwellings in the self-enumerated collection areas as occupied or unoccupied in the field. Based on this information, adjustments were made to the census database. The DCS selected a random sample of 1,405 self-enumerated CUs that were revisited in July and August 2006 to reassess the occupancy status as of Census Day for each dwelling for which no questionnaire had been received. The DCS found that 17.4% of the

934,564 dwellings classified as unoccupied were actually occupied and that 29.1% of the 366,527 dwellings with no responses that were classified as occupied or with occupancy status classified as unknown were actually unoccupied. Estimates based on the DCS samples were used to adjust the occupancy status for individual dwellings so as to change (impute) appropriate proportions of unoccupied dwellings to occupied and of occupied non-responding dwellings to unoccupied. This resulted in an increase of 3.6% in the number of occupied dwellings (relative to the number of dwellings originally classified as occupied) and a decrease of 5.2% in the number of unoccupied dwellings at the Canada level (relative to the number of dwellings originally classified as unoccupied). More information on the DCS can be found in Section 6.

After this adjustment of the occupancy status on the basis of the DCS results, occupied dwellings with total non-response had the number of usual residents (if not known) and all the responses to the census questions imputed by borrowing the unimputed responses from another household within the same CU that had its type of questionnaire (long or short). This process, called whole household imputation (WHI), imputed 96% of the total non-response households. The other 4% of the total non-response households where no donor household was found under the WHI process were imputed as part of the main edit and imputation (E&I) process. Utilizing a single donor under WHI was more efficient computationally and was less likely to produce implausible results than using several donors as part of the main E&I process, as was done in 2001. More information on WHI can be found in Section 6.2.4.

5.8 Edit and imputation

The data collected in any survey or census contains some omissions or inconsistencies. For example, a respondent might be unwilling to answer a question, fail to remember the right answer, or misunderstand the question. Also, census staff may code responses incorrectly or make other mistakes during processing.

The final clean-up of the data was done in edit and imputation and was, for the most part, fully automated. Two types of imputation were applied. The first type, called 'deterministic imputation,' involved assigning specific values under certain conditions. Detailed edit rules were applied to identify these conditions, and then the variables involved in the rules would be assigned a pre-determined value. The second type of imputation, called 'minimum-change donor imputation,' applied a series of detailed edit rules that identified any missing or inconsistent responses. These missing or inconsistent responses were corrected by changing as few variables as possible. For minimum-change donor imputation, a record with a number of characteristics in common with the record in error was selected. Data from this 'donor' record were borrowed and used to change the minimum number of variables necessary to resolve all missing or inconsistent responses. The Canadian Census Edit and Imputation System (CANCEIS) was used for nearly all deterministic and minimum-change donor imputation in 2006.

5.9 Weighting

Questions on age, sex, marital status, mother tongue and relationship to Person 1 were asked of 100% of the population, as in previous censuses. However, the bulk of census information was acquired on a 20% sample basis, using the additional questions on the 2B questionnaire. Weighting was used to project the information gathered from the 20% sample to the entire population.

For the 2006 Census, weighting employed the same methodology used in the 2001 Census, known as calibration estimation. This began by first assigning initial weights of approximately 5 to the sampled households. These weights were then adjusted by the smallest possible amount needed to ensure closer agreement between the sample estimates and the population counts for a number of characteristics related to age, sex, marital status, common-law status and household size (e.g., number of males, number of people aged 15 to 19). More information on sampling and weighting can be found in the *2006 Census Technical Report on Sampling and Weighting*.

6. Dwelling Classification Survey

6.1 Introduction

As described in Section 5.7, census data are adjusted for occupied non-respondent dwellings. The number of people living in these dwellings is estimated by the Dwelling Classification Survey (DCS). These estimates are used in census processing to specify how many people should be imputed during whole household imputation (WHI). The second objective of the DCS is to measure three types of dwelling classification error.

One of the potential sources of error in a census is the misclassification of dwellings. When a questionnaire is not returned from a household, the enumerator has to determine if the dwelling is occupied or not. Two types of errors can occur. **First, an occupied dwelling can be incorrectly classified as unoccupied**. Census dwelling and population undercoverage are the result of this classification error because the dwelling is excluded from the census database. **Second, an unoccupied dwelling can be incorrectly classified as occupied**. When this error occurs, no questionnaire will be received for this dwelling and it will be subject to non-response follow-up (NRFU) as described in Section 4.3. The dwelling will be considered as a non-respondent dwelling and therefore subject to imputation. This would add persons to the census database when, in fact, no one is living at that dwelling. That is, this classification error results in population overcoverage. Estimates from the DCS are used to adjust census data for both of these coverage errors.

The third component of dwelling classification error measured by the DCS is the error incurred when marginal dwellings or dwellings under construction are classified in error as dwellings. Since the dwelling would be classified as unoccupied, no population overcoverage results as only occupied dwellings can be classified as non-respondent dwellings and therefore be subject to imputation. However, there is dwelling overcoverage. Census data is not adjusted for these dwellings so census estimates of the housing stock include some degree of overcoverage.

6.2 Methodology

6.2.1 Stratification and sample selection

The DCS target population was all non-response dwellings and all unoccupied dwellings excluding dwellings in collective collection units (CU), canvasser CUs and Indian reserves CUs. These areas were excluded because of cost and operational considerations.

The sample size for the DCS was set at 1,405 CUs. The sampling frame consisted of all self-enumeration CUs with the exception of Indian reserves. Consequently, the sampling frame for the territories included only Whitehorse, Yellowknife, Hay River and Fort Smith. The sample design was as follows. First, CUs in Whitehorse, Yellowknife, and Hay River and Fort Smith in the Northwest Territories formed one stratum. All of these CUs were selected for the DCS sample. All of the CUs in Prince Edward Island formed a second stratum from which a simple random sample of 44 CUs was selected.

The remaining CUs were grouped into urban and rural strata. A CU was considered urban if it initially had been part of a census metropolitan area (CMA) or a census agglomeration (CA) that had 40,000 or more occupied dwellings. Further, all of the CUs within a crew leader district (CLD) were considered urban if more than 50% of the CUs in the CLD were urban. All of the remaining CUs formed the rural strata. Urban CUs were stratified by CMA and CA. A simple random sample of at least five CUs was selected within each stratum. From past census data, it was determined that five CUs was an appropriate workload for an interviewer. There were 812 urban CUs in the sample. In order to control field costs, the rural sample was chosen to be geographically close. This was done via a two-stage stratified simple random sampling design. In the first stage, CLDs

were selected within each province. In the second stage, five CUs were selected from each of the selected CLDs. There were 593 rural CUs in the sample.

All of the unoccupied dwellings and non-response dwellings in the sampled CUs formed the DCS sample of dwellings, a total of 32,345 unoccupied and 6,788 non-response dwellings. Table 6.2.1 shows the distribution of the sample by province and territory.

Table 6.2.1 Sample size for Canada, provinces and territories

Provinces and territories	Number of census units	Number of unoccupied dwellings	Number of non-response dwellings
Canada	**1,405**	**32,345**	**6,788**
Newfoundland and Labrador	84	2,653	219
Prince Edward Island	44	2,184	166
Nova Scotia	94	3,384	701
New Brunswick	88	2,116	321
Quebec	299	5,130	1,982
Ontario	272	7,400	1,040
Manitoba	84	1,486	223
Saskatchewan	85	1,663	156
Alberta	156	2,331	819
British Columbia	152	2,258	1,007
Yukon Territory	29	307	64
Northwest Territories	18	433	90
Nunavut	0	0	0

Source: Statistics Canada, 2006 Dwelling Classification Survey.

6.2.2 Field interviews

All dwellings in the sampled CUs that were classified as unoccupied on Census Day or classified as occupied but for which no census form had been returned, were to be checked again in late June or early July 2006 to determine the true occupancy status of the dwellings on Census Day. A DCS questionnaire was used for this purpose.

The timing of this operation was left to the discretion of each regional office (RO). In order to determine occupancy status and collect other information, enumerators were instructed to contact current occupants, neighbours, landlords, or any other person with some knowledge about the dwelling. Up to three contact attempts were made for each dwelling. If the dwelling was found to have been occupied on Census Day, the number of occupants on Census Day was obtained.

6.2.3 Processing, coding, and editing

All completed questionnaires were sent to Ottawa for processing.

Some preliminary edits and general grooming were then performed before the questionnaires were sent for data capture (key entry). Once data capture was completed, the questionnaires were subjected to an extensive set of consistency edits. The questionnaires failing edits were examined manually in order to resolve the inconsistencies.

For each dwelling in the DCS sample found to have been occupied on Census Day, the DCS questionnaire was consulted to determine whether another address was listed where the household members may have been enumerated. If an alternate address was given, then the Visitation Record (VR) and the census questionnaires for the alternate addresses were checked to see if the household members were indeed enumerated elsewhere. If they were found to have been enumerated elsewhere, they were considered as already having been enumerated and therefore they need not add to overcoverage by being included again. The dwelling itself, however was added to the occupied dwelling count.

At this point in processing, the unoccupied dwellings and the non-response dwellings in the sample were separated and the classification of these dwellings was confirmed against final census listing. The questionnaires completed for each sampled CU were matched to the final census listing of unoccupied dwellings. If a match could not be found, the sampled dwelling was discarded and no further processing was required. Dwellings listed as unoccupied on the census list for which no DCS questionnaire was received were considered as total non-response and went onto the next step of processing. Similarly, the final census listing of all dwellings for which a census questionnaire was not received was used to establish which of the DCS dwellings for which a DCS questionnaire was not received would be considered as total non-response.

Total non-response was addressed by a weighting adjustment while item imputation was used for item non-response. The procedure was the same for the unoccupied dwellings and non-response dwellings. When there was no information for a dwelling, the design weights of the respondents were adjusted to account for the design weight of the non-respondents. The adjustment was done separately for each of the Montréal, Toronto and Vancouver CMAs, for the remaining urban areas in each province and territory, and for the rural areas for each province and territory. Item non-response for occupancy status, number of usual residents, and dwelling type was addressed by imputation. Occupancy status was imputed first and then used in the imputation of the other variables. Design weights were then adjusted so that the sum of the adjusted weights for each subprovincial area equaled the number of unoccupied/non-response dwellings.

6.2.4 Census whole household imputation

Once the DCS estimates were produced, census data were adjusted for non-respondent dwellings and for occupied dwellings classified in error as unoccupied. This was done in the whole household imputation (WHI) step of census processing as follows for the non-respondent dwellings; unoccupied were handled in a similar, but simpler, fashion. First, within a DCS post-stratum all the non-respondent dwellings were identified. Second, any non-respondent dwelling for which field collection had obtained the number of usual residents was deemed to be occupied and assigned the recorded household size. Finally, an additional number of non-respondent dwellings were randomly selected and declared to be occupied. The selection was done such that the final number of non-respondent dwellings converted to occupied dwellings in the post-strata equaled the DCS estimate of occupied dwellings in the non-respondent dwelling universe. This process resulted in all private dwellings on the database being classified as either occupied or unoccupied.

A second procedure was used to impute the household dwelling size and other variables for the selected non-respondent dwelling. Household size was determined by randomly selecting a dwelling from all dwellings that had completed a census questionnaire in the same CU. The complete record from this donor household was then assigned to the non-respondent dwelling. If no donor was found, then only a household size was assigned.

More information on WHI can be found in Dick (2007).

6.3 Estimates

Census data are adjusted for non-respondent dwellings and for occupied dwellings that are classified in error as unoccupied using DCS estimates. The estimates are given in Section 6.3.1.1 and 6.3.2. Census data are not adjusted for marginal dwellings or dwellings under construction that are classified in error as dwellings. Section 6.3.1.2 presents estimates of the number of marginal dwellings and dwellings under construction that are classified in error as dwellings and therefore erroneously included in the housing stock.

6.3.1 Unoccupied dwellings

6.3.1.1 Occupied dwellings misclassified as unoccupied

Table 6.3.1.1.1 gives the estimated number of dwellings classified as unoccupied that should have been classified as occupied and the corresponding error rate for unoccupied dwellings by urban and rural,[4] by province and territory, for the three largest CMAs, and by type of dwelling. For comparison, Table 6.3.1.1.2 gives the same estimates for the 2001 Census. Table 6.3.1.1.3 gives the estimated number of persons living in occupied dwellings misclassified as unoccupied. Table 6.3.1.1.4 shows the number of households and persons added to the initial 2006 Census counts to adjust for these misclassifications.

Table 6.3.1.1.1 shows that 17.4% of all dwellings classified as unoccupied were actually occupied. This is an increases from 12.7% found in 2001. The misclassification of dwellings was much more prevalent in urban areas (25.7%) than in rural areas (8.1%). Both areas show increases from 2001. Among the three largest CMAs, there was a large decrease in the rate of misclassification in both Toronto and Vancouver while a large increase occurred for Montréal. Increases in the misclassification rates occurred for all provinces except Prince Edward Island where it remained stable. The misclassification rate decreased for apartment buildings of five or more stories between 2001 and 2006 but the rates increased for all other types of dwellings.

Among the provinces and territories, British Columbia had the highest misclassification rate, 25.5%, followed by the Yukon Territory, 23.5%, Alberta, 21.4%, Quebec, 21.1% and Ontario, 16.0%. The rates for the other provinces and territories ranged from 14.9% for Nova Scotia to 7.3% for Newfoundland and Labrador. Among the three largest CMAs, the 2006 rate of misclassification is very high in all three areas with the rate in Montréal (34.0%) being higher than the rates in Vancouver (25.5%) or Toronto (23.3%). Among the types of private dwellings classified in the census, the rate of misclassification is lowest in single-detached houses (17.4%) and highest in apartments in buildings of five or more storeys (39.6%). The rate of misclassification in the 'Other' category, which includes semi-detached houses, row houses, duplexes, apartments in buildings with fewer than five storeys, mobile homes and other movable dwellings, is also high at 38.2%.

Because of error in the initial classification of dwellings, approximately 162,897 households were not enumerated in the 2006 Census. This is the number of households added to the census during WHI. Table 6.3.1.1.4 shows the number of households and persons added to adjust for occupied dwellings misclassified as unoccupied.

4. Urban refers to urban areas with a population of over 50,000 persons. The remaining geographies constitute the rural areas.

Table 6.3.1.1.1 Number of occupied dwellings misclassified as unoccupied dwellings for various characteristics, 2006 Census

Characteristics	Number of unoccupied dwellings	Occupied dwellings misclassified as unoccupied			
		Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	**934,565**	**162,897**	**10,915**	**17.4**	**1.2**
Urban > 50,000	495,781	127,404	10,052	25.7	2.0
Rural	438,784	35,493	3,900	8.1	0.9
Atlantic provinces	**119,899**	**12,937**	**983**	**10.8**	**0.8**
Newfoundland and Labrador	34,267	2,510	344	7.3	1.0
Prince Edward Island	9,077	711	160	7.8	1.8
Nova Scotia	45,298	6,765	825	14.9	1.8
New Brunswick	31,257	2,953	379	9.4	1.2
Quebec	192,297	40,479	3,796	21.1	2.0
Ontario	363,808	58,111	8,837	16.0	2.4
Prairies	**139,653**	**21,078**	**2,308**	**15.1**	**1.7**
Manitoba	35,126	3,755	838	10.7	2.4
Saskatchewan	43,899	4,377	696	10.0	1.6
Alberta	60,628	12,946	2,035	21.4	3.4
British Columbia	118,087	30,154	4,511	25.5	3.8
Territories	**821**	**137**	**24**	**16.7**	**2.9**
Yukon Territory	307	72	9	23.5	2.9
Northwest Territories	514	65	22	12.6	4.3
Selected CMAs					
Montréal	52,642	17,882	2,314	34.0	4.4
Toronto	78,696	18,358	4,021	23.3	5.1
Vancouver	35,457	9,040	1,941	25.5	5.5
Type of private dwelling					
Single-detached	328,775	57,151	3,999	17.4	1.2
Apartment in a building with five or more storeys	66,867	26,492	5,294	39.6	7.9
Other type of dwelling	207,406	79,254	6,623	38.2	3.2
Not in housing stock[1]	331,517	…	…	…	…

... not applicable

Note1: These dwellings were originally classified as unoccupied dwellings. They are actually dwellings that are not in the housing stock. Therefore, none of them are misclassified occupied dwellings.

Source: Statistics Canada, 2006 Dwelling Classification Survey.

Table 6.3.1.1.2 Number of occupied dwellings misclassified as unoccupied dwellings for various characteristics, 2001 Census

Characteristics	Number of unoccupied dwellings	Occupied dwellings misclassified as unoccupied			
		Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	**904,236**	**114,603**	**7,166**	**12.7**	**0.8**
Urban > 50,000	405,427	86,370	5,814	21.3	1.4
Rural	498,809	28,233	4,176	5.7	0.8
Atlantic provinces	**112,357**	**6,961**	**693**	**6.2**	**0.6**
Newfoundland and Labrador	35,896	1,511	291	4.2	0.8
Prince Edward Island	5,121	396	100	7.7	2.0
Nova Scotia	42,931	3,411	519	7.9	1.2
New Brunswick	28,409	1,643	341	5.8	1.2
Quebec	213,062	31,007	4,577	14.6	2.1
Ontario	316,474	36,647	4,676	11.6	1.5
Prairies	**150,294**	**15,303**	**1,729**	**10.2**	**1.2**
Manitoba	37,592	2,737	434	7.3	1.2
Saskatchewan	47,503	3,113	486	6.6	1.0
Alberta	65,199	9,452	1,602	14.5	2.5
British Columbia	110,755	24,545	2,252	22.2	2.0
Territories	**1,294**	**140**	**27**	**10.8**	**2.1**
Yukon Territory	395	72	10	18.2	2.5
Northwest Territories	899	68	25	7.6	2.8
Selected CMAs					
Montréal	42,174	8,435	1,737	20.0	4.1
Toronto	40,020	17,557	4,330	43.9	10.8
Vancouver	29,565	11,034	1,614	37.3	5.5
Type of private dwelling					
Single-detached	423,328	46,616	3,554	11.0	0.8
Apartment in a building with five or more storeys	39,365	18,796	4,427	47.7	11.2
Other type of dwelling	230,899	49,191	4,434	21.3	1.9
Not in housing stock[1]	210,644	…	…	…	…

… not applicable

Note: 1. These dwellings were originally classified as unoccupied dwellings. They are actually dwellings that are not in the housing stock. Therefore, none of them are misclassified occupied dwellings.

Source: Statistics Canada, 2006 Dwelling Classification Survey.

Table 6.3.1.1.3 Estimated number of persons living in misclassified occupied dwellings and standard errors for various characteristics, 2006 Census

Characteristics	Estimated number	Standard error
Canada	**331,145**	**17,663**
Urban > 50,000	259,903	16,333
Rural	71,242	6,372
Atlantic provinces	**24,578**	**1,541**
Newfoundland and Labrador	5,300	663
Prince Edward Island	1,392	282
Nova Scotia	11,952	1,153
New Brunswick	5,934	725
Quebec	78,285	6,212
Ontario	121,389	14,590
Prairies	**44,313**	**3,958**
Manitoba	7,728	1,616
Saskatchewan	8,735	1,411
Alberta	27,851	3,325
British Columbia	62,334	6,517
Territories	**246**	**42**
Yukon Territory	120	15
Northwest Territories	126	39
Selected CMAs		
Montréal	33,808	4,577
Toronto	41,088	7,297
Vancouver	20,526	4,689
Type of private dwelling		
Single-detached	132,363	8,667
Apartment in a building with five or more storeys	44,106	8,555
Other type of dwelling	154,675	11,437

Source: Statistics Canada, 2006 Dwelling Classification Survey.

Table 6.3.1.1.4 Imputed households and persons for various characteristics, 2006 Census

Characteristics	Number of imputed households	Number of imputed persons
Canada	**162,897**	**331,145**
Urban > 50,000	127,404	259,903
Rural	35,493	71,242
Atlantic provinces	**12,937**	**24,578**
Newfoundland and Labrador	2,510	5,300
Prince Edward Island	711	1,392
Nova Scotia	6,765	11,952
New Brunswick	2,953	5,934
Quebec	40,479	78,285
Ontario	58,111	121,389
Prairies	**21,078**	**44,313**
Manitoba	3,755	7,728
Saskatchewan	4,377	8,735
Alberta	12,946	27,851
British Columbia	30,154	62,334
Territories	**137**	**246**
Yukon Territory	72	120
Northwest Territories	65	126
Selected CMAs		
Montréal	17,882	33,808
Toronto	18,358	41,088
Vancouver	9,040	20,526
Type of private dwelling		
Single-detached	57,151	132,363
Apartment in a building with five or more storeys	26,492	44,106
Other type of dwelling	79,254	154,675

Source: Statistics Canada, 2006 Dwelling Classification Survey.

6.3.1.2 Housing stock overcoverage

Table 6.3.1.2 shows the estimated number of unoccupied dwellings not in the housing stock and the corresponding error rate for unoccupied dwellings for various geographic areas. No adjustments are made to the census database to account for dwellings not in the housing stock that were erroneously classified as unoccupied.

The enumeration of unoccupied dwellings that fall outside the housing universe results in overcoverage of dwellings. Dwellings are considered to be outside the housing universe if they are used for commercial purposes, if they are not habitable year round, or if they are double counted in the census. This last situation can happen when the dwelling appears to have two addresses associated with it, or when two questionnaires are mistakenly returned for a dwelling which no longer contains a separate apartment within it.

The Dwelling Classification Survey estimates of the number of unoccupied misclassified as dwellings are not used to adjust the census database because of the degree of subjectivity associated with classifying a dwelling as suitable for year-round occupancy. A dwelling must have a source of heat or power and provide complete shelter from the elements to be considered as suitable. It is sometimes difficult to tell whether or not a dwelling is habitable such as when the dwelling is a cottage, when the dwelling is under construction and almost complete, or when the dwelling has deteriorated.

Dwellings outside the housing stock account for 35.5% of all dwellings classified as unoccupied. Among the provinces and territories, the incidence of dwellings outside the housing stock having been classified as unoccupied ranges from 8.3% in Prince Edward Island to 46.1% in Manitoba. The problem is evenly found in urban areas (35.0%) and the rural areas (36.0%). For the three largest CMAs, the rate ranges from 29.9% in Montréal to 54.4% in Toronto.

Dwellings actually outside the housing stock represent 2.5% of all private dwellings in the 2006 Census. This is an increase from the 2001 error rate of 1.7%. Among the provinces and territories, the error ranges from a low of 0.9% in the Yukon Territory to a high of 3.8% in Newfoundland and Labrador. For the three largest CMAs, the error ranges from 1.0% in Montréal to 2.3% in Toronto.

Table 6.3.1.2 Dwellings not in housing stock misclassified as unoccupied dwellings for various geographic areas, 2006 Census

Geographic areas	Number of unoccupied dwellings	Dwellings not in housing stock misclassified as unoccupied dwellings			
		Estimated number	Standard error	Estimated rate[1] (%)	Standard error (%)
Canada	**934,565**	**331,517**	**29,897**	**35.5**	**3.2**
Urban > 50,000	495,781	173,452	22,543	35.0	4.6
Rural	438,784	158,065	18,884	36.0	4.3
Atlantic provinces	**119,899**	**29,449**	**3,337**	**24.6**	**2.8**
Newfoundland and Labrador	34,267	8,741	1,835	25.5	5.4
Prince Edward Island	9,077	756	202	8.3	2.2
Nova Scotia	45,298	12,890	2,404	28.5	5.3
New Brunswick	31,257	7,062	1,396	22.6	4.5
Quebec	192,297	70,370	10,576	36.6	5.5
Ontario	363,808	141,930	26,882	39.0	7.4
Prairies	**139,653**	**50,434**	**4,444**	**36.1**	**3.2**
Manitoba	35,126	16,195	2,514	46.1	7.2
Saskatchewan	43,899	8,800	1,567	20.1	3.6
Alberta	60,628	25,439	3,312	42.0	5.5
British Columbia	118,087	39,115	5,331	33.1	4.5
Territories	**821**	**219**	**42**	**26.7**	**5.1**
Yukon Territory	307	79	12	25.7	3.9
Northwest Territories	514	140	41	27.2	8.0
Selected CMAs					
Montréal	52,642	15,720	2,314	29.9	4.4
Toronto	78,696	42,808	26,507	54.4	33.7
Vancouver	35,457	14,227	3,824	40.1	10.8

Note: 1. The rate is the estimated number of occupied non-response dwellings as a percent of all non-response dwellings.

Source: Statistics Canada, 2006 Dwelling Classification Survey.

6.3.2 Non-response dwellings

6.3.2.1 Persons added in non-response dwellings

Table 6.3.2.1.1 gives the estimated number and rate of occupied non-response dwellings in the census by urban (> 50,000) and rural, by province and territory, for the three largest CMAs, and by type of dwelling. Table 6.3.2.1.2 gives the number of persons estimated by the DCS to be living in these non-response dwellings while Table 6.3.2.1.3 gives the same information for the 2001 DCS.

Table 6.3.2.1.1 shows that 70.9% of all dwellings classified as non-response were actually occupied. The census did a slightly better job of classifying non-response dwellings in urban areas (72.1%) than it did in rural areas (64.1%). At the province and territory level in 2006, the

Northwest Territories had the highest rate of correctly classified non-response dwellings at 91.0%, while New Brunswick had the lowest rate at 59.5%.

Among the three largest CMAs, the 2006 rates for occupied non-response dwellings ranged from 68.3% in Vancouver to 75.6% in Montréal. Finally, when examining types of private dwellings, the rates for occupied non-response dwellings ranged from 74.3% in apartments in buildings of five or more storeys to 78.9% in the 'Other' category, which includes semi-detached houses, row houses, duplexes, apartments in buildings with fewer than five storeys, mobile homes and other movable dwellings.

Table 6.3.2.1.2 shows the number of non-response dwellings in the 2006 Census, and gives the number of persons added in those dwellings through the DCS. Table 6.3.2.1.3 shows the same data from the 2001 DCS. In 2006, a total of 571,521 persons were added to the census in 259,894 dwellings. The comparable 2001 numbers are 317,587 persons in 143,681 dwellings.

Table 6.3.2.1.1 Occupied non-response dwellings for various characteristics, 2006 Census

Characteristics	Number of non-response dwellings	Occupied non-response dwellings			
		Estimated number	Standard error	Estimated rate[1] (%)	Standard error (%)
Canada	**366,528**	**259,894**	**3,030**	**70.9**	**0.8**
Urban > 50,000	310,218	223,821	2,750	72.1	0.9
Rural	56,310	36,074	1,265	64.1	2.2
Atlantic provinces	**23,435**	**15,578**	**647**	**66.5**	**2.8**
Newfoundland and Labrador	4,426	2,722	284	61.5	6.4
Prince Edward Island	1,216	762	69	62.7	5.7
Nova Scotia	10,900	7,991	472	73.3	4.3
New Brunswick	6,893	4,103	333	59.5	4.8
Quebec	111,467	82,877	1,552	74.4	1.4
Ontario	103,289	72,111	1,594	69.8	1.5
Prairies	**62,679**	**43,457**	**1,200**	**69.3**	**1.9**
Manitoba	10,200	8,104	273	79.5	2.7
Saskatchewan	9,362	6,296	381	67.3	4.1
Alberta	43,117	29,057	1,105	67.4	2.6
British Columbia	65,372	45,627	1,540	69.8	2.4
Territories	**286**	**245**	**9**	**85.7**	**3.1**
Yukon Territory	109	83	7	76.1	6.4
Northwest Territories	177	161	6	91.0	3.4
Selected CMAs					
Montréal	62,045	46,904	3,408	75.6	5.5
Toronto	48,563	33,980	4,257	70.0	8.8
Vancouver	40,370	27,560	3,129	68.3	7.8
Type of private dwelling					
Single-detached	125,028	97,612	5,215	78.1	4.2
Apartment in a building with five or more storeys	50,130	37,265	5,517	74.3	11
Other type of dwelling	158,393	125,017	4,590	78.9	2.9
Not in housing stock[2]	32,977	…	…	…	…

… not applicable

Notes:

1. The rate is the estimated number of occupied non-response dwellings as a percent of all non-response dwellings.
2. These dwellings were originally classified as unoccupied dwellings. They are actually dwellings that are not in the housing stock. Therefore, none of them are occupied non-response dwellings.

Source: Statistics Canada, 2006 Dwelling Classification Survey.

Table 6.3.2.1.2 Persons living in occupied non-response dwellings for various characteristics, 2006 Census

Characteristics	Occupied non-response dwellings		Persons living in occupied non-response dwellings	
	Estimated number	Standard error	Estimated number	Standard error
Canada	**259,894**	**3,030**	**571,521**	**3,918**
Urban > 50,000	223,821	2,750	489,840	3,477
Rural	36,074	1,265	81,681	1,774
Atlantic provinces	**15,578**	**647**	**31,059**	**823**
Newfoundland and Labrador	2,722	284	5,783	360
Prince Edward Island	762	69	1,458	90
Nova Scotia	7,991	472	15,215	593
New Brunswick	4,103	333	8,603	434
Quebec	82,877	1,552	171,274	1,927
Ontario	72,111	1,594	163,184	2,083
Prairies	**43,457**	**1,200**	**97,102**	**1,677**
Manitoba	8,104	273	16,952	373
Saskatchewan	6,296	381	13,587	474
Alberta	29,057	1,105	66,563	1,564
British Columbia	45,627	1,540	108,296	1,952
Territories	**245**	**9**	**607**	**10**
Yukon Territory	83	7	131	7
Northwest Territories	161	6	476	7
Selected CMAs				
Montréal	46,904	3,408	99,246	6,772
Toronto	33,980	4,257	74,171	8,333
Vancouver	27,560	3,129	64,866	6,871
Type of private dwelling				
Single-detached	97,612	5,215	256,758	13,008
Apartment in a building with five or more storeys	37,265	5,517	64,163	10,484
Other type of dwelling	125,017	4,590	250,600	10,442

Source: Statistics Canada, 2006 Dwelling Classification Survey.

Table 6.3.2.1.3 Occupied non-response dwellings and persons living in them for Canada, provinces and territories, 2001 Census

Provinces and territories	Number of non-response dwellings	Occupied non-response dwellings		Persons living in occupied non-response dwellings	
		Estimated number	Standard error	Estimated number	Standard error
Canada	**179,788**	**143,681**	**2,352**	**317,587**	**14,841**
Newfoundland and Labrador	1,431	1,185	67	2,268	179
Prince Edward Island	508	392	83	978	206
Nova Scotia	5,063	3,980	332	8,042	511
New Brunswick	3,303	2,676	60	5,501	66
Quebec	65,787	50,834	1,473	100,741	7,789
Ontario	59,784	48,396	1,686	124,825	12,282
Manitoba	3,798	3,254	142	6,602	161
Saskatchewan	3,246	2,313	144	5,142	172
Alberta	14,197	11,834	370	26,982	423
British Columbia	22,472	18,697	470	36,269	2,862
Yukon Territory	118	90	10	177	12
Northwest Territories	81	30	21	60	29

Source: Statistics Canada, 2001 Dwelling Classification Study.

6.3.2.2 Dwellings not in the housing stock misclassified as non-response

Table 6.3.2.2 shows the 2006 Census dwelling classification error from dwellings erroneously classified as non-response because they should not have been included in the housing stock. Section 6.3.1.2 provides the definition of dwellings outside of the housing universe and comments on the difficulty in determining whether a dwelling should be included in the housing stock.

At the national, dwellings outside the housing stock account for 9.0% of all non-response dwellings. The error rate is slightly higher in the rural areas (10.5%) than in the urban areas (8.7%). For provinces and territories, the incidence of dwellings outside the housing stock having been classified as non-response ranges from 2.4% in Prince Edward Island to 11.2% in both Alberta and British Columbia. For the three largest CMAs, the rate ranges from 8.4% in Montréal to 15.6% in Vancouver. At the national level, non-response dwellings outside the housing stock account for 0.3% of all private dwellings. For provinces and territories, this error ranges from a low of 0% rounded in Prince Edward Island to 0.4% in both Alberta and British Columbia. For the three largest CMAs, the error ranges from 0.2% in Toronto to 0.7% in Vancouver.

Table 6.3.2.2 Dwellings not in housing stock misclassified as non-response dwellings for various geographic areas, 2006 Census

Geographic areas	Number of non-response dwellings	Dwellings not in housing stock misclassified as non-response dwellings			
		Estimated number	Standard error	Estimated rate[1] (%)	Standard error (%)
Canada	**366,528**	**32,977**	**2,291**	**9.0**	**0.6**
Urban > 50,000	310,218	27,091	2,092	8.7	0.7
Rural	56,310	5,887	982	10.5	1.7
Atlantic provinces	**23,435**	**1,336**	**220**	**5.7**	**0.9**
Newfoundland and Labrador	4,426	160	53	3.6	1.2
Prince Edward Island	1,216	29	17	2.4	1.4
Nova Scotia	10,900	640	162	5.9	1.5
New Brunswick	6,893	506	137	7.3	2.0
Quebec	111,467	10,112	974	9.1	0.9
Ontario	103,289	8,427	1,268	8.2	1.2
Prairies	**62,679**	**5,765**	**994**	**9.2**	**1.6**
Manitoba	10,200	550	130	5.4	1.3
Saskatchewan	9,362	406	179	4.3	1.9
Alberta	43,117	4,809	969	11.2	2.2
British Columbia	65,372	7,327	1,288	11.2	2.0
Territories	**286**	**11**	**6**	**3.8**	**2.1**
Yukon Territory	109	5	3	4.6	2.8
Northwest Territories	177	6	5	3.4	2.8
Selected CMAs					
Montréal	62,045	5,214	753	8.4	1.2
Toronto	48,563	4,644	1,129	9.6	2.3
Vancouver	40,370	6,310	1,219	15.6	3.0

Note: 1. The rate is the estimated number of occupied non-response dwellings as a percent of all non-response dwellings.
Source: Statistics Canada, 2006 Dwelling Classification Survey.

7. Reverse Record Check

7.1 Sampling

The target population, which consisted of all persons who should have been enumerated in the 2006 Census, was formed from six sources (sampling frames) presented in Table 7.1.1. The first five frames were used to estimate undercoverage in the ten provinces, whereas estimates for the three territories were calculated based on samples from the last frame only.

Table 7.1.1 Sample size, sampling frames for Canada

Sampling frames	Definition	Sample size
Canada	**...**	**69,751**
Provinces	**...**	**67,813**
2001 Census	All persons enumerated in the 2001 Census.	56,317
Missed	All persons from the 2001 Reverse Record Check (RRC) sample who were classified as missed. Their weight is their 2001 RRC weight.	2,797
Births	All children born between May 15, 2001 and May 15, 2006.	3,874
Immigrants	All landed immigrants who arrived in Canada between May 15, 2001 and May 15, 2006.	2,977
Non-permanent residents	All persons from another country, who held employment or student permits covering May 16, 2006 and persons claiming refugee status on May 16, 2006. Family members living with them in Canada are also in this frame.	1,848
Territories	**...**	**1,938**
Health care files	All persons listed in the health care files of the Yukon Territory[1], the Northwest Territories and Nunavut who were living in these territories on May 16, 2006.	1,938

... not applicable
Note: 1. Some persons from other sources were added.
Source: Statistics Canada, 2006 Reverse Record Check.

One disadvantage of multiple sampling frames is the possibility that someone will be included in more than one frame. For example, a person in the immigrants frame may have been in Canada on a work permit in May 2001, and thus have been enumerable in the 2001 Census. The person would then be in both the immigrants frame and the census frame if he or she was enumerated, or in the immigrants frame and the missed frame if not enumerated. It is important to identify all cases of frame overlap. If this is not done, estimates may be too high because some people have been included twice in the frames. Though such overlap was identified wherever possible when preparing the sampling frames, some was also identified later based on information provided by the respondents.

Another difficulty is that none of the first five sampling frames covered people who had emigrated, or who were outside the country at the time of the 2001 Census and had returned during the intercensal period ('returning Canadians within a province'). According to demographic estimates, this population is estimated to contain 210,406 people. To this number add 12,817 persons returning from a territory to a province, and 4,955 from Indian reserves or Indian settlements that were partially enumerated in 2001 and enumerated in 2006. Coverage error estimates do not include these populations, which are estimated to total some 228,178 people.

Sample allocation was done in two stages. First, the national sample was allocated to the provinces using a combination of proportional allocation to achieve the same variance for all the provincial estimates of the undercoverage rate and optimal allocation to achieve the national estimate of the undercoverage rate with the smallest variance. The second step was to determine the allocation of the provincial samples to the strata. This was also done via optimal allocation based on historical undercoverage rates (overcoverage was also taken into account in 2001, but not in 2006), historical non-response rates, and stratum size. The exception is the missed frame where everyone who was classified missed in the 2001 RRC was selected. It should be noted that the allocations are only approximately optimal because assumptions were made about the size of some populations such as the projected number of intercensal births and immigrants. The total allocated sample was 69,602 people distributed among the frames (67,664 in the provinces, and 1,938 in the territories). Table 7.1.1 presents the sample allocation by sampling frame. Table 7.1.2 gives the allocation by sampling stratum for all provinces.

Table 7.1.2 Sample allocation, sampling frames, strata for Canada

Sampling frames	Strata within each province[1]	Sample allocation
All	**…**	**67,664**
2001 Census	Females, 0 to 14 years	4,950
	Females, 15 to 24 years	5,136
	Females, 25 to 34 years, married	1,908
	Females, 25 to 44 years, not married	3,559
	Females, 35 years and over, married	5,652
	Females, 45 years and over, not married	3,492
	Males, 0 to 14 years	5,680
	Males, 15 to 24 years	6,400
	Males, 25 to 34 years, married	1,936
	Males, 25 to 44 years, not married	5,563
	Males, 35 years and over, married	7,345
	Males, 45 years and over, not married	2,630
	On reserve	2,066
Missed	No further stratification	2,628
Births	No further stratification	3,856
Immigrants	No further stratification	2,971
Non-permanent residents	No further stratification	1,892

… not applicable
Note: 1. In Quebec, persons in common-law relationships are included in the married strata.
Source: Statistics Canada, 2006 Reverse Record Check.

Table 7.1.3 gives the allocation by stratum for all territories.

Table 7.1.3 Sample allocation, strata for territories

Strata	Number of persons
Total	**86,460**
Matched	**84,522**
Unmatched	**1,938**
Females, 0 to 19 years	270
Females, 20 to 24 years	101
Females, 25 to 34 years	165
Females, 35 to 44 years	135
Females, 45 years and over	190
Males, 0 to 19 years	283
Males, 20 to 24 years	110
Males, 25 to 34 years	206
Males, 35 to 44 years	184
Males, 45 years and over	282
Persons added to the Yukon Territory health care file	12

Source: Statistics Canada, 2006 Reverse Record Check.

Since the sample allocation depends on assumptions about the size of some populations such as the projected number of intercensal births and immigrants, the actual sample size for the provincial sample from the births, immigrants, and non-permanent residents frames is not known until after the final sample is selected. This is not the case for the territories' sample. Table 7.1.4 gives the final sample size for each province and territory. The 2006 RRC sample consisted of 67,813 persons in the provinces and 1,938 persons in the territories. In addition, 84,522 persons with a weight of one contributed to the territorial estimates.

Table 7.1.4 Sample size for Canada, provinces and territories

	Number of persons
Canada	**154,273**
All provinces	**67,813**
Newfoundland and Labrador	3,622
Prince Edward Island	3,802
Nova Scotia	5,443
New Brunswick	5,532
Quebec	9,093
Ontario	12,595
Manitoba	5,646
Saskatchewan	5,204
Alberta	6,360
British Columbia	10,516
All territories	**86,460**
Matched	84,522
Unmatched	1,938
Yukon Territory	**26,887**
Matched	26,307
Unmatched	580
Northwest Territories	**35,301**
Matched	34,448
Unmatched	853
Nunavut	**24,272**
Matched	23,767
Unmatched	505

Source: Statistics Canada, 2006 Reverse Record Check.

The sample design varies by frame according to the nature of the list that was used. In the 2001 Census frame, the sample design was a one-stage stratified design. The population was stratified by province of residence, sex, age, and marital status. People enumerated on Indian reserves in the 2001 Census were placed in separate strata. As mentioned, we used optimal allocation in each stratum. The sample was allocated to strata in order to obtain the largest possible number of 'missed' cases.

Sampling fractions were not the same in all strata. To make the sample design more efficient, higher sampling rates were applied in subgroups for which high undercoverage or a lower tracing rate was expected. For example, as in the 2001 RRC, single males aged 20 to 24 in 2006 had a greater probability of being selected, since it had been observed in previous RRCs that undercoverage was consistently higher in that stratum. As a result of increased interest in the aboriginal population, the size of the sample in the provincial strata for people on Indian reserves enumerated in the 2001 Census was double the 2001 sample size.

The missed frame is a conceptual frame since there is no list of all persons missed in the 2001 Census. The sample for this frame consists of all cases classified as 'missed' in the 2001 RRC. The sample is not stratified per se, though there is an implicit stratification since the 'missed' cases in 2001 were from different frames and strata.

For the births frame, copies of birth registrations for the intercensal period were obtained from vital statistics. The frame was then stratified by mother's province of residence. Provincial samples were selected systematically, after sorting by date of birth of the child.

The immigrants frame is constructed from immigration records obtained from Citizenship and Immigration Canada, and stratified by province. In 2006, unlike in 2001, there was no yearly stratification for the three provinces that receive the most immigrants (i.e., Quebec, Ontario and British Columbia). Provincial samples were selected systematically, after sorting by year of immigration.
The non-permanent residents frame (permit holders and refugee claimants) was constructed from records obtained from Citizenship and Immigration Canada. Records were sorted by province. Unlike in 2001, for Quebec, Ontario and British Columbia no strata containing refugee claimants or holders of study, minister's or work permits were created. Provincial samples were selected systematically, after sorting by type of permit and refugee status, to ensure each of these groups was adequately represented.

The methodology for the territories was changed in 2006. As with previous RRCs, the sampling frames of the three territories were created from their respective health care files. Some files from other sources were added for the Yukon Territory in order to improve basic coverage. The people listed in the sampling frames of each territory were then matched by name, sex and age with the 2006 Census response database using exact matching. A manual verification was also performed. Matched people were classified as enumerated, and given a weight of 1. People not classified as enumerated were then stratified by age and sex. After sorting by geography, a one-stage systematic sample was taken from each stratum.

The next step after selecting samples was to prepare the sample, which included checking the quality of information for different variables of interest (i.e., geographic or demographic). For example, we checked the accuracy of names and the validity of birth dates. Addresses were standardized to facilitate subsequent processing. To update the geographic information, especially for the census sample and the missed where the information was from 2001, we then matched to Canada Revenue Agency's 2000 to 2005 personal income tax files. We also used these files, along with vital statistics data, to verify whether any selected persons had died.

7.2 Processing and classification

7.2.1 Processing

The goal of processing is:

1. To determine whether each selected person (SP) was part of the census target population.
2. If so, to determine whether each SP was enumerated.
3. To provide further information for the non-response adjustment.

The results of processing were used to determine the classification assigned to an SP for estimation and tabulation (see Section 7.4 and Section 9).

Most of the work in processing involved searching the RRC version of the 2006 Census Response Database (RRC RDB) to determine whether the SP was enumerated at one of the addresses associated with him or her. The addresses were obtained from various sources including:

- the sampling frame for the selection address
- updates from tax records
- the computed-assisted telephone interview (CATI) and paper questionnaires(see Section 7.3)
- matches with the RRC Response Data Base (RDB) using birth date and sex of the SP and members of his or her household, or, the SP's name, postal code or telephone number.

The RRC RDB is an early version of the 2006 Census Response Database (RDB) that is available before the end of census processing. There are some minor differences between the RRC RDB and later versions of the census databases. In particular, the RRC RDB, which is a database of persons, contains all census records for persons with three exceptions. The first are imputed census records for imputations made during whole household imputation (WHI). The second group consists of census records with missing or invalid names, or incomplete or invalid birth dates. This group is also known as the 'incompletely enumerated.' The third group consists of all census records that were added late, after the start of RRC processing.

The first step after sample preparation was to process all SPs with the addresses available from the sampling frame and tax data to search the RRC RDB for each SP. There were two outcomes. When the SP was found, the classification of 'enumerated' was usually assigned and no further processing was required. An exception was SPs who were later identified as deceased before the census from vital statistics for deaths. When the SP was not found, the case was sent for collection. While collection was taking place, searching the RRC RDB continued. When data from the CATI interview was available, it could be determined whether or not each SP was part of the census target population. If so, the CATI data could enable further searching.

Searching was done both automatically and manually by clerical staff. Automated searching was done first as follows: for addresses obtained from a match with the RRC RDB, there was a corresponding census questionnaire. First, we calculated a measure of similarity between the census questionnaire and the RRC data. When this measure was above a specified threshold, it was automatically concluded that the SP was enumerated at that address. If so, neither this address nor the SP's other addresses needed to be processed by the clerical staff. Computer programs also determined when one address was a duplicate of another. These duplicate addresses also did not need to be processed.

To search manually, the clerical staff used a number of tools. There were often suggested census questionnaire or census collection units that matched the address. Staff could also search the RRC RDB using flexible parameters. Electronic telephone directories were also used. The results of the manual search were then automatically edited to minimize errors. A file containing the search results was then produced. It is the data from this file that was used to classify SPs.

7.2.2 Classification

Processing provides the information required to determine which SPs were:

(a) 'listed'
(b) 'mobile'
(c) included in the 'census target population'
(d) 'enumerated'
(e) 'missed.'

Some SPs belonged to three or four of these categories. Other SPs did not belong to any of these groups. This is explained in more detail in this section. The 'census target population' includes the groups of persons enumerated in Section 2.2. An SP is considered 'out of scope' if he/she is not part of the census target population. Each SP classified as out of scope is assigned a reason for the classification such as death, emigration, or representation by another sampling frame. In order to classify an SP as deceased, the death must have appeared in the vital statistics files as a registered death. SPs classified in the census target population are either 'enumerated'

or 'missed.' An SP is considered 'enumerated' if he/she was in the RRC RDB. The 'missed' classification was assigned to SPs in the census target population who were not enumerated.

The definitions of 'listed' and 'mobile' depend on whether or not the addresses and information from the CATI interview were required to determine the classification. In many cases, collection provided addresses that were not available from the other sources. In other cases, all of the addresses obtained during collection were also available from another source. An SP was 'listed' if he/she was classified without using data from the CATI interview. That is, even if collection data were obtained, the address/addresses collected during the interview was/were not required. An SP was considered 'mobile' if his or her usual place of residence, as defined in Section 2.4, was only available from the collection data. Further, by definition, SPs that are not in the census target population, and therefore classified as out of scope, are mobile.

Selected persons for whom one or more of characteristics (a) to (e) cannot be determined are considered non-respondents. There are two types of non-respondents:

- An SP is 'not identified' when it cannot be determined whether or not they are listed.
- An SP is 'not traced' when it cannot be determined whether or not they are included in the census target population.

Table 7.2 presents the distribution of the sample by classification and sampling frame. The classification is determined from specific combinations of characteristics (a) to (e). Data for the territories is divided into the matched stratum and the unmatched strata. Among the 67,813 SPs selected in the provinces, 56,789 were classified as 'enumerated,' 5,431 were classified as 'missed' and 2,901 were non-respondents. An adjustment for non-response is done during the estimation (see Section 7.4). Note that the definition of a non-respondent for classification, and therefore for estimation, is not the same as the usual definition of a non-respondent for whom data collection is attempted but not completed. This is because classification uses data from many sources of which one may be collection. To avoid confusion, Section 7.3 on collection refers to 'completed collection' rather than 'response.'

'Traced' SPs are SPs for whom it can be determined whether or not they are included in the census target population. For purposes of estimation and tabulation, traced SPs are our respondents. Since names, including those of household members, and addresses are available in the RRC RDB, and the tools for consulting the database are sufficiently powerful, it can be verified whether a SP is enumerated at an address even if the address is vague. This ensures that SPs are classified as traced only when it is known whether or not they are mobile and whether or not they are enumerated.

The usefulness of knowing whether a SP is enumerated is self-evident. Selected persons who are in the census target population who are not enumerated, and therefore classified as missed, are the basis for the estimate of undercoverage. We also wanted to classify the respondent SPs according to characteristics (a) to (c), in order to choose the most appropriate respondents to represent the non-respondents. The above definitions implied that:

- not identified SPs are also not traced
- not traced identified SPs are not listed
- enumerated not mobile SPs are listed
- enumerated mobile SPs are not listed.

We also determined the Census Day address (usual place of residence) of each SP in the census target population. This is the address where, according to census instructions, the SP should have been enumerated. If the SP was enumerated, the enumeration address is considered to be the Census Day address even if other information may have raised doubts about the proper interpretation of census instructions.

More information on classification can be found in Diallo (2008).

Table 7.2 Classification of selected persons, sampling frames for Canada

	Provincial strata										Territorial strata					
	2001 Census		Missed		Births		Immigrants		Non-permanent residents		Matched		Unmatched		Total	
Classification	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Total	**56,317**	**100.0**	**2,797**	**100.0**	**3,874**	**100.0**	**2,977**	**100.0**	**1,848**	**100.0**	**84,522**	**100.0**	**1,938**	**100.0**	**154,273**	**100.0**
Enumerated	**48,443**	**86.0**	**1,888**	**67.5**	**3,448**	**89.0**	**2,154**	**72.3**	**856**	**46.4**	**84,493**	**100.0**	**500**	**25.8**	**141,782**	**91.9**
Listed	48,039	85.3	1,875	67.0	3,439	88.8	2,147	72.1	853	46.2	84,493	100	487	25.1	141,333	91.6
Unlisted	404	0.7	13	0.5	9	0.2	7	0.2	3	0.2	0	0.0	13	0.7	449	0.3
Missed	**3,896**	**6.9**	**393**	**14.1**	**271**	**7.0**	**432**	**14.5**	**439**	**23.8**	**29**	**0.0**	**647**	**33.3**	**6,107**	**4.0**
Listed	497	0.9	35	1.3	51	1.3	35	1.2	32	1.7	29	0.0	66	3.4	745	0.5
Not mobile unlisted	2,245	4.0	240	8.6	155	4.0	286	9.6	247	13.4	0	0.0	342	17.6	3,515	2.3
Mobile unlisted	1,154	2.0	118	4.2	65	1.7	111	3.7	160	8.7	0	0.0	239	12.3	1,847	1.2
Out of scope	**2,212**	**3.9**	**196**	**7.0**	**49**	**1.3**	**123**	**4.1**	**112**	**6.1**	**0**	**0.0**	**576**	**29.7**	**6,384**	**4.1**
Listed	1,815	3.2	89	3.2	23	0.6	4	0.1	2	0.1	0	0.0	552	28.5	2,485	1.6
Unlisted	397	0.7	107	3.8	26	0.7	119	4.0	110	6.0	0	0.0	24	1.2	783	0.5
Non-response	**1,766**	**3.1**	**320**	**11.4**	**106**	**2.7**	**268**	**9.0**	**441**	**23.9**	**0**	**0.0**	**215**	**11.1**	**3,116**	**2.0**
Identified not traced	1,292	2.3	302	10.8	106	2.7	267	9.0	440	23.8	0	0.0	203	10.5	2,610	1.7
Not identified	474	0.8	18	0.6	0	0.0	1	0.0	1	0.1	0	0.0	12	0.6	506	0.3

Source: Statistics Canada, 2006 Reverse Record Check.

7.3 Data collection

7.3.1 Environment

Head office (HO) staff in Ottawa worked closely with staff in five Statistics Canada regional offices (ROs) to collect data during the survey phase of the RRC. These ROs were located in Halifax, Sherbrooke, Toronto, Winnipeg and Edmonton. The suggestions and recommendations made by the ROs as a result of conducting the 2001 RRC were incorporated into the design and operations of the 2006 survey. HO was responsible for providing a computer-assisted telephone interviewing (CATI) application that met the needs of the survey and was interviewer and respondent friendly.

Assignment of the sample to the ROs was based on HO's 'best guess' about where the selected person (SP) was residing during the collection period. Once a case was assigned to an RO, it was never transferred to another RO even if it was determined that the SP moved outside the RO collection area. RO coverage areas and survey counts are shown in Table 7.3.1.

Table 7.3.1 Geographic coverage for regional offices

Regional offices	Coverage	Number of cases
Total	**Canada**	**20,114**
Halifax	Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick	4,139
Sherbrooke	Quebec	2,174
Toronto	Ontario	3,454
Winnipeg	Manitoba, Saskatchewan, Alberta (part), Yukon Territory, Northwest Territories, Nunavut	5,562
Edmonton	Alberta (part), British Columbia	4,785

Source: Statistics Canada, 2006 Reverse Record Check.

A total of 20,114 cases were sent for collection. Section 7.1 describes the two sample designs used in the RRC for the provinces and for the territories. The number of cases requiring collection in the territorial sample was the sample of 1,938 taken from the unmatched strata. For the provincial sample, the number of cases requiring collection could not be determined until after all cases were sent for a first attempt at processing whereby the RRC census response database (RRC RDB) was searched. There were two outcomes to this search. When the SP was not found, it was sent for collection. There were a total of 8,453 such cases, referred to as the 'regular' sample. A sample of 11,231 SPs was selected from among the found SPs. These are referred to as the 'non-response adjustment (NRA)' sample. The collection results for the NRA sample were used to estimate a parameter of the RRC non-response adjustment model described in Section 7.4. RO staff was not made aware if a case was NRA or regular.

The 20,114 cases sent to the field represented 28.4% of the RRC sample. Most of the sample not sent for collection was related to SPs who were found on the RRC RDB during the first search. A classification of enumerated could therefore be assigned to these SPs and no further work was required. The remainder of the sample not sent for collection included 729 deceased SPs, SPs from the sample of 2006 births who were not available in time and 24 cases not sent for other reasons including frame overlap or insufficient information to determine exactly who the SP was.

There were three versions of the RRC Survey questionnaire; non-proxy, proxy, and deceased before Census Day. The content of the 2006 RRC Survey questionnaire focused on the collection of addresses, especially those where the SP lived on Census Day and in the month of May 2006.

Names and demographic data were collected for all Census Day household members. The three 2006 Census questions on Aboriginal identity were added to the RRC for the first time. Collection was proxy by design for everyone who was less than 18 years of age and SPs presumed deceased. Otherwise, proxy respondents were used when the SP was not available during the survey period or was difficult to reach.

When it was determined at the time of contact that an SP was deceased, it was important to ascertain with a proxy respondent if the SP had died before, on, or after Census Day. Different paper questionnaires and CATI flows were used depending on the date of death. In some cases, it was known that the SP was deceased prior to collection. If two sources such as taxation data and vital statistics indicated the SP was deceased, the case was not sent for collection. If one source indicated that the SP had died, the case was sent for collection with a flag indicating that he/she was 'presumed deceased.'

The main survey data collection method was CATI. The CATI application was developed using many of the standards set for all CATI questionnaires conducted at Statistics Canada. The application was constructed of various interrelated modules and was accessed through the generic interface for ROs. Interviewers were assigned to cases based on language and whether cases required interviewing or tracing.

The 2006 RRC Survey was a multiple collection mode survey. Paper questionnaires in both official languages were available for those SPs who were contacted but requested a paper questionnaire as opposed to giving information by telephone. Selected persons who the RO did not succeed to contact by telephone were mailed a paper questionnaire package prepared and sent to the best address as determined by the RO. Selected persons were asked to return their completed paper questionnaire to the RO. Finally, some responses were obtained by field interviewers using the paper questionnaires. Data capture from the paper questionnaires was done in the ROs using the CATI system. All of the coordination work necessary to operationalize a sequential multiple mode survey was done by the RO managers in partnership with HO. Unlike the 2001 RRC Survey, there was no follow-up survey.

Tracing was a key aspect of the 2006 RRC. Tracing refers to the work done to find telephone and address information for SPs or a proxy person. As part of sample preparation, cases were linked to tax data to provide updated contact data for the SP and their household members. In some cases, initial CATI data was outdated or incomplete and therefore tracing was required.

7.3.2 Operations

Data collection consisted primarily of interviewing and tracing. As data collection began, 82.8% of the cases sent for collection were placed in the queue for interviewing and the remaining 17.2% in the tracing queue. As required, cases were moved back and forth between interviewing and tracing. For SPs initially in the tracing queue, no telephone number had yet been found for the SP or any family member. As tracing leads were found, cases were moved to interviewing. When all tracing leads were exhausted for interviewing cases they were moved to tracing.

The CATI input data were loaded as sample preparation and the first search of the RRC RDB in processing was completed. Data collection began in the ROs on January 10, 2007. Active collection ended on July 15, 2007. In total there were 184 days where at least one RO was actively collecting data. A total of 16,984 questionnaires were completed in this time frame. Between July 16 and July 31, 2007, a passive collection took place wherein returned paper questionnaires or persons calling the RO to do the survey were handled. During this time, 112 questionnaires were completed. It should also be noted that even among the 16,984 questionnaires deemed complete by the ROs, some were later judged in HO to be either seriously incomplete or conducted with an incorrect SP.

Data collection was conducted using two shifts of interviewers working six or seven days per week. Interviewers were given the survey objectives and background along with a detailed training manual. Mock interviews were incorporated into the training sessions using the CATI application. A call scheduler assigned cases to interviewers in normal operations. On occasion, an interviewer could be assigned to manage specific cases. This may have been to take an in-coming call or to make a call to someone who preferred to speak in a non-official language. Calls were made overseas especially for SPs in the non-permanent resident (NPR) group who had left Canada. Quality management of the collection operation included monitoring of the interviewers, retraining and the discussion of specific data quality issues noted in HO relating to completed questionnaires. Regional office managers allocated resources to the survey while balancing the needs of other surveys taking place in their region. Sustained efforts to interview persons who initially refused to participate in the survey improved response rates.

Table 7.3.2 shows the distribution of cases sent to ROs from HO over time. Interviewing typically began in the RO as soon as new cases arrived. The adjusted total reflects cases that were dropped by the ROs as a result of an HO request. This was made because HO processing was able to resolve a regular sample case that had gone into the field. This may have been due to a SP being confirmed as being deceased from the 2005 and 2006 vital statistics (VS) files or the SP was found on the 2006 RRC RDB. Additionally, some cases were resent to an RO if a case that was originally completed and returned did not meet the data quality standard expected.

Table 7.3.2 Regional office workloads by date sent

	Regional offices					
2007 send date	**Halifax**	**Sherbrooke**	**Toronto**	**Winnipeg**	**Edmonton**	**Total**
Total	**4,151**	**2,183**	**3,485**	**5,620**	**4,816**	**20,255**
January 3	1,556	894	1,512	875	2,079	6,916
January 25	1,712	587	833	877	595	4,604
February 8	425	435	674	779	334	2,647
February 22	395	247	400	1,566	1,730	4,338
March 8	26	4	15	1,267	25	1,337
March 22	0	0	4	223	5	232
April 26	37	16	47	33	48	181
Cases dropped by head office[1]	12	9	31	58	31	141
Adjusted total	**4,139**	**2,174**	**3,454**	**5,562**	**4,785**	**20,114**

Note: 1. Collection no longer required because selected person had been located on the RRC RDB.
Source: Statistics Canada, 2006 Reverse Record Check.

Survey data were sent electronically to HO from the five ROs each night after interviewing came to a halt. Transmission reports and collected survey data were reviewed each morning by HO staff. Cases considered unsuitable for processing were reactivated and sent back to the RO for follow up.

Three detailed management reports were created at HO to document the progress of the survey. One report gave statistics on the cases currently in the RO (unopened cases, completed cases, and opened cases not yet completed). The second report presented very detailed statistics of a number of RO outcomes. This report was produced on a weekly basis. The third report included progress by variables such as case type, sampling frame and stratum. Case completion projections were made for the ROs to help them meet their collection targets.

Data collected in the field were analyzed at HO for completeness and accuracy. Cases were rejected if data were missing or ambiguous in key fields or if the data had mistakenly been acquired for someone other than the SP. Cases which were not rejected were compiled into batches for processing as described in Section 7.2.

The average duration of the CATI interview was 13 minutes. The actual time spent on each case however was larger given the number of contact attempts required and the amount of tracing that was involved.

7.3.3 Tracing

Tracing was undertaken by both HO and the ROs and was critical to the success of the RRC. Of the 3,456 cases that started in tracing, successful leads that yielded interviews were found for 66% of them. Among the 16,658 cases that started in the interviewing queue and required tracing, the trace rate was higher at 88%. Numerous valuable leads were also found for these cases. Overall, 11,339 of the 16,944 completed cases, 67%, required some tracing effort.

To increase response rates, RO managers contacted provincial government agencies and departments to obtain addresses and telephone numbers for cases where contact had not been established. Once collection began, HO was engaged in providing tracing leads using several large administrative files containing names and addresses but not necessarily telephone numbers. These files included motor vehicle registration, taxation, GST rebate on new homes and the Canada Post change of address. Additional information specific to SPs on the immigrant and the NPR frames was obtained from Citizenship and Immigrant Canada in paper format. Vital statistics files for 2005 and 2006 were also searched.

Interviewers used a variety of tracing tools, on-line electronic directories being the most popular. However, the most effective tracing leads came from the CATI application itself. Information loaded into the application included addresses from the RRC RDB (which is from the 2006 Census itself) and older taxation files. In cases where the RO received an address lead from HO, an on-line site such as Canada 411 was used to find a telephone number. If this was not successful, then a paper questionnaire package could be sent to the address. In comparing the HO and RO tracing addresses found independently, it was concluded that the larger HO files offered more tracing information that was unique and useful. There was overlap between the two efforts in that the same or very similar addresses were often obtained.

The response rate achieved was high, 84.2% of the 20,114 cases sent for collection. This accomplishment was due to the extensive tracing carried out by HO and the ROs. Another factor was the persistence of RO staff in calling an SP when they had the correct telephone number but no one was answering. The median number of contact attempts made for completed cases was seven. For cases that were never completed, the median number of contact attempts was 28. These numbers vary by the province and territory of selection, sampling frame and demographic variables. A case may have a high number of contact attempts though it may not ever have required any tracing.

Table 7.3.3 shows the number of contact attempts for completed and not completed cases by the sampling frame. Close to 222,400 calls were made for completed cases and about 104,300 calls were made for cases for which no completed survey was ever obtained.

Table 7.3.3 Median number of contact attempts, completion status, sampling frames for Canada

Sampling frames	Completed cases	Cases not completed
Canada	**7**	**28**
2001 Census	7	30
Missed	8	35
Births	5	22
Immigrants	9	23
Non-permanent residents	11	15
Health care files	7	25

Source: Statistics Canada, 2006 Reverse Record Check.

It was expected that the work involved in initially contacting an SP from the NRA sample would be easier than for an SP in the regular sample because the initial CATI contact data included the SP's most recent address from the RRC RDB. However, there were many NRA sample cases that took more contact attempts to talk with the SP compared to persons in the regular sample.

7.3.4 Collection statistics

Many statistics were monitored throughout the data collection period. An analysis was done after collection was complete.

Table 7.3.4.1 shows provincial and territorial completion rates by type of case as either regular or NRA. The table shows that completion rates are higher for the NRA cases. This is expected because the initial CATI data included the more recent address specified in the 2006 Census. These SPs would have, with a probability close to 1, been classified enumerated. The distribution of the SPs in the regular sample is different. Compared to the entire RRC sample, persons in the regular sample come from strata with a lower expected probability of being classified enumerated and a higher expected probability of being classified 'missed' or 'out of scope.' Evidence from past RRCs indicates that such persons are more difficult to contact.

Table 7.3.4.1 Completion counts and rates, type of sample for Canada, provinces and territories

Provinces and territories	Regular sample			Non-response adjustment sample			Total		
	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)
Canada	**11,231**	**8,491**	**75.6**	**8,883**	**8,453**	**95.2**	**20,114**	**16,944**	**84.2**
Newfoundland and Labrador	396	323	81.6	349	345	98.9	745	668	89.7
Prince Edward Island	500	418	83.6	372	357	96.0	872	775	88.9
Nova Scotia	815	652	80.0	585	557	95.2	1,400	1,209	86.4
New Brunswick	722	566	78.4	591	554	93.7	1,313	1,120	85.3
Quebec	1,125	904	80.4	1,031	1,002	97.2	2,156	1,906	88.4
Ontario	2,025	1,469	72.5	1,271	1,212	95.4	3,296	2,681	81.3
Manitoba	770	606	78.7	716	686	95.8	1,486	1,292	86.9
Saskatchewan	692	565	81.6	556	526	94.6	1,248	1,091	87.4
Alberta	998	718	71.9	977	905	92.6	1,975	1,623	82.2
British Columbia	2,162	1,452	67.2	1,243	1,166	93.8	3,405	2,618	76.9
Yukon Territory	353	268	75.9	508	492	96.9	861	760	88.3
Northwest Territories	453	364	80.4	428	410	95.8	881	774	87.9
Nunavut	220	186	84.5	256	241	94.1	476	427	89.7

Source: Statistics Canada, 2006 Reverse Record Check.

Table 7.3.4.2 gives completion statistics by frame and case type. The low response rate for the SPs in the NPR frame is partially because many NPRs appear to have left Canada prior to the end of their permit expiry date. Also, in many cases, the permit expiry date came before the start of survey operations. It was frequently very difficult to locate these SP or a suitable proxy. This was especially true for NPRs with a permit to study in Canada where the completion rate was just 62.0%.

Table 7.3.4.2 Completion counts and rates, sampling frames, type of sample for Canada

Sampling frames	Regular sample			Non-response adjustment sample			Total		
	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)
All	**11,231**	**8,491**	**75.6**	**8,883**	**8,453**	**95.2**	**20,114**	**16,944**	**84.2**
2001 Census	7,184	5,654	78.7	5,496	5,289	96.2	12,680	10,943	86.3
Missed	780	583	74.7	676	621	91.9	1,456	1,204	82.7
Births	366	284	77.6	436	424	97.2	802	708	88.3
Immigrants	831	564	67.9	642	607	94.5	1,473	1,171	79.5
NPR	1,044	588	56.3	441	369	83.7	1,485	957	64.4
Health care files[1]	1,026	818	79.7	1,192	1,143	95.9	2,218	1,961	88.4

Note: 1. From the unmatched strata. By definition, no collection is required for samples in the matched strata.
Source: Statistics Canada, 2006 Reverse Record Check.

Table 7.3.4.3 gives completion statistics by stratum and type of case for the sample selected from the demographic strata. As discussed in Section 7.1, demographic strata were used for the 2001 Census frame and the unmatched frames in the territories.

Table 7.3.4.3 Completion counts and rates, strata, type of sample for 2001 Census and territories unmatched sampling frames, Canada

	Regular sample			Non-response adjustment sample			Total		
Strata	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)	Cases sent	Cases completed	Completion rate (%)
All provinces, 2001 Census frame[1]	**7,184**	**5,654**	**78.7**	**5,496**	**5,289**	**96.2**	**12,680**	**10,943**	**86.3**
Females, 0 to 14 years	407	318	78.1	344	331	96.2	751	649	86.4
Females, 15 to 24 years	949	796	83.9	617	591	95.8	1,566	1,387	88.6
Females, 25 to 34 years, married	159	120	75.5	207	201	97.1	366	321	87.7
Females, 25 to 44 years, not married	509	368	72.3	491	468	95.3	1,000	836	83.6
Females, 35 years and over, married	306	271	88.6	227	226	99.6	533	497	93.2
Females, 45 years and over, not married	504	445	88.3	230	224	97.4	734	669	91.1
Males, 0 to 14 years	455	351	77.1	452	434	96.0	907	785	86.5
Males, 15 to 24 years	1,309	1,039	79.4	883	850	96.3	2,192	1,889	86.2
Males, 25 to 34 years, married	163	120	73.6	290	285	98.3	453	405	89.4
Males, 25 to 44 years, not married	1,070	729	68.1	892	839	94.1	1,962	1,568	79.9
Males, 35 years and over, married	473	395	83.5	274	271	98.9	747	666	89.2
Males, 45 years and over, not married	419	326	77.8	315	307	97.5	734	633	86.2
On reserve	461	376	81.6	274	262	95.6	735	638	86.8
All territories, unmatched frames[2]	**1,026**	**818**	**79.7**	**1,192**	**1,143**	**95.9**	**2,218**	**1,961**	**88.4**
Females, 0 to 19 years	117	99	84.6	133	129	97.0	250	228	91.2
Females, 20 to 24 years	56	47	83.9	60	56	93.3	116	103	88.8
Females, 25 to 34 years	83	69	83.1	110	103	93.6	193	172	89.1
Females, 35 to 44 years	65	47	72.3	104	101	97.1	169	148	87.6
Females, 45 years and over	94	79	84.0	90	89	98.9	184	168	91.3
Males, 0 to 19 years	129	113	87.6	165	163	98.8	294	276	93.9
Males, 20 to 24 years	66	58	87.9	58	55	94.8	124	113	91.1
Males, 25 to 34 years	126	95	75.4	137	128	93.4	263	223	84.8
Males, 35 to 44 years	116	80	69.0	143	132	92.3	259	212	81.9
Males, 45 years and over	167	127	76.0	171	167	97.7	338	294	87.0
Unknown	7	4	57.1	21	20	95.2	28	24	85.7

Notes: 1. Age five years ago at the time of the 2001 Census. In Quebec persons in common-law relationships are included in the married strata.
2. Age at the time of the 2006 Census.

Source: Statistics Canada, 2006 Reverse Record Check.

Another statistic of interest is the degree to which questionnaires were completed by proxy. Collection was proxy by design for everyone who was less than 18 years of age and SPs presumed deceased. Otherwise, proxy was used when the SP was not available during the survey period or was difficult to reach. Overall, 6,363 cases representing 37.6% of the completed sample were done by interviewing a suitable proxy.

Table 7.3.4.4 gives, for Canada and the provinces and territories, the number of cases sent for collection, the number of these that required tracing, and the percentage of cases sent for collection that required tracing. The tracing rate was highest among the provinces for Ontario and British Columbia and for Yukon Territory and Nunavut.

Table 7.3.4.4 Cases requiring tracing for Canada, provinces and territories

Provinces and territories	Number of cases sent	Number of cases that required tracing	Percentage of cases that required tracing (%)
Canada	**20,114**	**11,339**	**56.4**
Newfoundland and Labrador	745	283	38.0
Prince Edward Island	872	425	48.7
Nova Scotia	1,400	774	55.3
New Brunswick	1,313	650	49.5
Quebec	2,156	1,076	49.9
Ontario	3,296	2,055	62.3
Manitoba	1,486	796	53.6
Saskatchewan	1,248	659	52.8
Alberta	1,975	1,110	56.2
British Columbia	3,405	2,135	62.7
Yukon Territory	861	605	70.3
Northwest Territories	881	479	54.4
Nunavut	476	292	61.3

Source: Statistics Canada, 2006 Reverse Record Check.

There were three modes of collection, CATI, self-enumeration using the paper questionnaire, and personal interview also using the paper questionnaire. Of the 16,944 completed questionnaires. 94.1% were done by CATI, 4.5% were done by self-enumeration, and 1.4% by personal interview. These data show the importance of the multiple mode approach. Without the use of self-enumeration and in-person interviewing, the national completion rate could have been less than 80%. The collection mode varied by province and territory. This may reflect different operational methods in the ROs or differences in the characteristics of the persons requesting a questionnaire, or different demographic distributions. Self-enumeration was particularly important in Ontario where 6.9% of the completed cases were done by self-enumeration, and in British Columbia where 10.6% of the completed cases were done by self-enumeration.

7.4 Estimation

The final weights of the selected persons (SP) began with their initial (or design) weights. The initial weight of an SP from the missed frame is the final weight assigned to him or her during the previous Reverse Record Check (RRC) when the SP was classified as missed. For the other sampling frames, the initial weights are generally equal to the inverse of the probability of selection. The exception is the non-permanent residents frame where the initial weight is higher

to account for the small number of non-permanent residents who were not in the sampling frame when the sample was selected. Final non-permanent resident counts were only available after the sample was selected. Initial weights were adjusted to add to these counts.

In order to reduce bias, the initial weights of the respondents had to be adjusted to account for non-response. The weight of the non-respondents was redistributed among the respondents. Where possible, this was done by ensuring that the weight of non-respondents with certain characteristics was redistributed only to respondents with the same characteristics. The following characteristics (or 'metadata') were used: sampling stratum; indication that the SP filled out a tax return for the year preceding the census year thus providing us with an indication that the SP is in the target population; and whether or not the SP was listed, mobile, or part of the target population.

The weight adjustments were done with the aid of the StatMx module of Statistics Canada's Generalized Estimation System (GES). In order to accomplish the redistribution of the weight of the non-respondents, the RRC was viewed as a sample in three phases where each phase corresponds to the 'selection' of a nested sample as follows. Selection of the SPs from the sampling frames was the first phase, then selection of the identified SPs from the all of the SPs and, last, selection of the traced SPs from the identified SPs. When a respondent with the same characteristics as a non-respondent could not be identified in a stratum, the stratum was grouped with another stratum deemed similar.

After adjusting for non-response, the estimated number of enumerated persons in the territories has traditionally been lower than the comparable census count. This is likely due to undercoverage of the census target population in the health care files. To address this bias, the weight of SPs selected in a territory was adjusted so that the estimated number of enumerated persons equalled the comparable census count for that territory.

Table 7.4 presents the weighted distribution of the sample by classification and sampling frame. Refer to Section 7.2 for the definition of the classification. Note that only SPs found in the RRC RDB are classified as enumerated. The RRC RDB differs from the final census database in that it does not include imputations made during whole household imputation (WHI), enumerations with an invalid or missing name or an incomplete or invalid birth date, or enumerations added after the start of the RRC data processing phase. People from the target population who are not in the RRC RDB are classified as missed. Census population undercoverage is estimated by the number (weighted) of missed persons less the number of persons excluded from the RRC RDB. This is the 'X' for the database extraction factor referred to in Section 9.

Last, in order to calculate the variance of the estimates, the RRC sampling frame was viewed as a stratified design with selection probabilities proportional to size. The size measures were constructed so as to reproduce the final weights.

You can obtain more information on the 2006 RRC estimation methods from Théberge (2008).

Table 7.4 Weighted classification of selected persons, sampling frames for Canada

	Provincial strata										Territorial strata					
	2001 Census		Missed		Births		Immigrants		Non-permanent residents		Matched		Unmatched		Total	
Classification	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Total	**29,334,767**	**100.0**	**1,795,025**	**100.0**	**1,672,476**	**100.0**	**1,109,472**	**100.0**	**368,539**	**100.0**	**89,178**	**100.0**	**24,179**	**100.0**	**34,393,636**	**100.0**
Enumerated	**25,885,054**	**88.2**	**1,226,462**	**68.3**	**1,499,660**	**89.6**	**808,722**	**72.9**	**185,205**	**50.2**	**89,147**	**100.0**	**6,563**	**27.1**	**29,700,814**	**86.3**
Listed	25,590,698	**87.2**	1,213,415	67.6	1,489,137	89.0	802,454	72.3	184,406	50.0	89,147	100.0	6,298	26.0	29,375,555	85.4
Not listed	294,356	1.0	13,047	0.7	10,523	0.6	6,268	0.6	799	0.2	0	0.0	265	1.1	325,259	0.9
Missed	**2,001,536**	**6.8**	**350,237**	**19.5**	**136,557**	**8.2**	**212,114**	**19.1**	**135,602**	**36.8**	**31**	**0.0**	**10,259**	**42.3**	**2,846,337**	**8.3**
Listed	242,894	0.8	23,746	1.3	18,712	1.1	11,413	1.0	7,253	2.0	31	0.0	833	3.4	304,883	0.9
Not mobile not listed	1,031,898	3.5	179,303	10.0	71,336	4.3	122,710	11.1	56,009	15.2	0	0.0	4,483	18.5	1,465,740	4.3
Mobile not listed	726,744	2.5	147,188	8.2	46,509	2.8	77,991	7.0	72,340	19.6	0	0.0	4,943	20.4	1,075,714	3.1
Out of scope	**1,448,175**	**5.0**	**218,325**	**12.2**	**36,258**	**2.2**	**88,636**	**8.0**	**47,733**	**12.9**	**0**	**0.0**	**7,358**	**30.4**	**1,846,486**	**5.3**
Listed	1,104,035	3.8	57,172	3.2	9,209	0.6	1,371	0.1	487	0.1	0	0.0	6,871	28.4	1,179,146	3.4
Not listed	344,140	1.2	161,153	9.0	27,049	1.6	87,265	7.9	47,246	12.8	0	0.0	487	2.0	667,340	1.9

Source: Statistics Canada, 2006 Reverse Record Check.

8. Census Overcoverage Study

8.1 Introduction

Following the 2001 Census of Population, the level of overcoverage due to duplication of individuals was measured by three studies, each one covering a part of the overcoverage: the Automated Match Study (AMS), the Collective Dwelling Study (CDS) and the Reverse Record Check (RRC). The introduction of names to the 2006 Census Response Database provides an opportunity to use name matching to measure overcoverage and therefore estimate overcoverage with a single study, the Census Overcoverage Study (COS). The COS is based on a series of automated exact and probabilistic matching operations and manual work. These matching operations also involve the use of various administrative data files. Therefore, the 2006 RRC measures just undercoverage and the CDS is no longer conducted as collective dwellings are covered by the COS.

8.2 Methodology

The methodology for estimating 2006 overcoverage was based on matching persons while the Automated Match Study (AMS)[5] was based on matching households of persons. The 2006 Census Overcoverage Study (COS) took advantage of the fact that the 2006 Census Response Database (RDB) contained respondent names. For the first time, the names were captured and were available for computer processing. It was anticipated that the inclusion of names in measuring overcoverage would maximize the proportion of total overcoverage covered by automated matching methods. Since the RRC no longer measures overcoverage, the new methodology reduces coverage study costs associated with the collection of additional addresses by the RRC for overcoverage measurement. The COS also produces a more precise estimate without geographic restrictions such as those applied to the 2001 AMS. Persons who were living in collective dwellings and hence completed a Form 3A or 3B were in scope for the COS.

In principle, the RDB could have been matched to itself to detect duplicate enumerations. However, on a practical level, and for methodological considerations, the COS was conducted in two steps as outlined below. It should be noted that the RRC version of the 2006 Census Response Database (RRC RDB) was not the same as the database that the COS used, since some records excluded for the RRC did not need to be excluded for the COS.

8.2.1 Step 1: Exact matching with administrative data

The first step was based on exact matching procedures, and involved matching the RDB with a set of administrative data files representing a large portion of the census target population. It was expected that this process would directly identify cases of overcoverage. In particular, RDB records assigned to the same administrative record through 'many-to-one' matches were declared to be cases of overcoverage without further review, since they pointed to the same individual from the administrative data files.

8.2.1.1 Administrative data files

Since there is no single administrative data file covering the entire Canadian Census target population, it was necessary to combine several files, each one covering a different segment of the population, in order to carry out the COS. The aim was to maximize the coverage of the Canadian Census target population while avoiding duplication of individuals among the administrative data files.

5. For a detailed description of the AMS methodology, see the 2001 Technical Report on Coverage Studies.

The following administrative data files were used:

- 2005 income tax records, supplemented with additional records for taxation years 2000 to 2004.
- Birth files for Canadian citizens born between 1985 and 2003.
- Immigration files for immigrants born outside Canada between 1985 and 2003, to cover children of immigrants not present in the birth files of Canadian citizens born between 1985 and 2003.
- Immigration files for immigrants who arrived in Canada between 2004 and May 16, 2006 (Census Day), given that they would not be on the income tax file for 2005.
- Non-permanent residents files.
- Health care files from the Yukon Territory, the Northwest Territories and Nunavut.

The income tax files for 2000 to 2004 were included with the 2005 income tax file in order to improve the coverage of the Canadian census target population. The personal income tax records accounted for approximately 80% of all administrative records used at the first step. The Health care files for the three territories were used to represent all persons living in the territories, whereas the other administrative data files, as listed above, were used to represent persons living in the provinces. As a variety of administrative data files were used, every effort was made to remove duplicates, so the first step exact match would be effective.

8.2.1.2 Using names

Without the presence of names in the 2006 RDB, the new methodology for measuring person duplication would not have been developed. The names used in the RDB for matching purposes were taken from Step B of the census questionnaire, which contains a list of all reported members of the household. Family name(s) and given name(s) were included in the same 80-character field. Respondents were asked to list their family name(s) first and then their given name(s). In order to use this field for matching, it was necessary to standardize the names and separate the 80 characters into a family name and a given name.

However, despite the instructions, not all respondents wrote their family names and given names in the correct order. Since this could lead to problems when matching with administrative data files, a strategy to separate names into family name and given name was developed to address this issue.

Family name(s) are separate from given name(s) in the adminstrative data files. This made it possible to compute the probability that a particular name is either used as a given name or a family name based on frequencies in the Canadian population. The name frequencies were broken down by sex and year of birth. This acknowledged that the use of a name as a given or family name may vary between males and females and over time. The name frequencies were then used to parse each name into the part most likely to be the given name(s) and the part most likely to be the family name(s). The same strategy was applied to the names from the RDB and to the names from the administrative sources. It was important that the name was parsed in the same manner on both files to ensure that the exact match was effective.

8.2.1.3 Exact match

Since the goal of the exact match was to identify each individual rather than to find a number of suggested matches for each record in the RDB, it was necessary to take a very conservative approach and only consider overcovered cases where a high degree of certainty was achieved. The variables used for this process were name, sex and date of birth.

Overcoverage was identified when two or more RDB records matched to the same administrative record. For evaluation purposes, a sample of these overcoverage matches was manually verified, as well as a sample of the one-to-one cases. An adjustment to the estimate of overcoverage,

based on the results of the verification sample, was done to account for false matches whereby two or three records had the same administrative record but did not represent the same individual.

A record in the RDB may have been a match for more than one administrative record, and vice-versa, thus creating a many-to-many match. For example, this can occur when two individuals have the same name and date of birth. When two RDB records matched to two administrative records, it was assumed that this grouping contained two valid one-to-one matches. However, a sample of the two-to-two matches was taken to verify this assumption. Following this review, the two-to-two now considered overcoverage were weighted up and added to the total estimate of overcoverage coming from the first step. All other combinations of many-to-many matches were manually verified and either classified as overcoverage or not. In this way, all of the many-to-many matches were resolved at the first step of the COS.

Note that in the first step, for technical reasons, RDB records for the provinces were matched to provincial administrative records, and RDB records for the territories were matched to the records in the territorial administrative Health Care Files. Hence, cases of overcoverage between the provinces and the territories were missed at Step 1, but they were included in Step 2.

The exact match rate in Step 1 was 66.5%, which means 66.5% of RDB records were involved in a match with an administrative record. Among all the RDB records, we note that:

- 64.68% of RDB records were part of a one-to-one match
- 1.76% of RDB records were involved in a many-to-one match (case of overcoverage)
- just 0.05% of matches were part of a many-to-many relationship
- 33.52% have not been matched.

A total of 260,708 persons involved in multiple enumerations were identified in Step 1. Estimating the number of persons involved in multiple enumerations was done by assigning a weight to each enumeration. Two-to-one matches identified in Step 1, for example, represent one person who was enumerated twice. In order to estimate the number of persons involved in duplicate enumerations, each RDB record was given a weight of ½. The premise was that the usual residence of the person is equally likely to be that of the first enumeration as that of the second enumeration. Cases of overcoverage whereby the enumerations are in more than one province (interprovincial overcoverage) were of particular interest since each province is assigned an equal portion of the total weight of 1[6]. Table 8.2.1.3 presents the total overcoverage in Step 1 for intraprovincial, intraterritorial, interprovincial and interterritorial pairs.

6. Some of these weights were adjusted for false matches, as mentioned earlier in this section.

Table 8.2.1.3 Step 1 intraprovincial, intraterritorial, interprovincial and interterritorial overcoverage for provinces and territories

Provinces and territories	Intraprovincial and intraterritorial overcoverage		Interprovincial and interterritorial overcoverage		Step 1 total
	Estimated number	Percentage of total overcoverage (%)	Estimated number	Percentage of total overcoverage (%)	Estimated number
Total	**246,982**	**94.7**	**13,726**	**5.3**	**260,708**
Newfoundland and Labrador	4,400	89.7	506	10.3	4,906
Prince Edward Island	1,101	87.8	153	12.2	1,254
Nova Scotia	6,642	88.7	846	11.3	7,488
New Brunswick	5,741	90.0	639	10.0	6,381
Quebec	58,076	97.0	1,774	3.0	59,851
Ontario	91,868	96.3	3,530	3.7	95,398
Manitoba	7,902	93.8	525	6.2	8,427
Saskatchewan	7,225	91.4	679	8.6	7,904
Alberta	24,469	90.0	2,724	10.0	27,193
British Columbia	38,651	94.3	2,339	5.7	40,990
Yukon Territory	297	99.3	2	0.7	299
Northwest Territories	424	98.7	6	1.3	429
Nunavut	185	98.1	4	1.9	189

Source: Statistics Canada, 2006 Census Overcoverage Study.

A total of 246,982 persons were overcovered within the same province or territory, and 13,726 between provinces or territories, for a total of 260,708 persons overcovered in Step 1. Only 5.3% of total overcoverage in Step 1 was interprovincial or territorial. The highest rates of interprovincial or interterritorial overcoverage were in the Atlantic provinces and Alberta. In percentage terms, interprovincial/interterritorial overcoverage was much smaller in the territories since Step 1 only applied to overcoverage between territories.

At this stage in the process, the RDB was split into two parts. Part A consisted of all RDB records that were matched to at least one administrative record, whether overcovered or not. Part B consisted of all RDB records that were not matched to an administrative record, as well as all territorial records. The latter was done to take into account provincial-territorial matches that were missed in Step 1. A probabilistic match was then done between Part B and the entire RDB to identify cases of overcoverage that were not identified in Step 1.

8.2.2 Step 2: Probabilistic match with the RDB

Step 2 of the COS is a probabilistic record linkage between RDB records that were not matched with an administrative record (Part B), about 10.2 million records, and the complete RDB (Part A + Part B) consisting of about 30.6 million records. Statistics Canada's Generalized Record Linkage System (GRLS) was used for this step.

8.2.2.1 Using the Generalized Record Linkage System

First, the rules governing the probabilistic match were established. Within the framework of GRLS, variables such as first name, last name, sex, date of birth, and some variables related to geography (listed in the next paragraph), were considered during the record linkage. The output

from GRLS results in pairs of individuals with an associated weight that indicates the strength of the match. The higher the matching weight is, the more likely the pair is a good match, thus resulting in overcoverage.

The Generalized Record Linkage System allows for variations in the spelling of names and variations in the agreement on date of birth. Geography was also considered in the linkage via the PRCDCU field (combination of province code, census division and collection unit), postal code and city (when postal code is missing). All the variables involved in the probabilistic record linkage were subject to different rules in a preliminary step called the selection criteria and rules applied for the purpose of the actual record linkage. Frequency weights for all variables, except for sex (because male and female are approximately in the same proportion in the population), were also used within GRLS. Frequency weights allow for matches on more common values to be weighted less heavily than matches on less common values.

The standard Fellegi-Sunter (1969) approach is implemented in the GRLS. An upper threshold, S2, is established, above which matches were accepted as overcoverage without verification. The threshold S2 was set conservatively so as to minimize the probability of finding false matches of overcoverage above S2. A lower threshold, S1, below, which matches are rejected without further review (i.e., no overcoverage), was also determined to minimize cases of overcoverage below threshold S1.

8.2.2.2 Manual verification

Due to time and resource constraints, it was impossible to verify all cases in the middle zone i.e., pairs whose matching weight was between S1 and S2 (1.1 million pairs). Instead, a sample of these matches was selected.

The sampling method used was systematic sampling with selection probabilities P_i proportional to size measure θ_i (1- θ_i). Pairs were ordered by province or territory, sex and date of birth. θ_i is the matching weight standardized on the interval [0,1]. The matching weight itself is from GRLS. θ_i is correlated with the probability of being a true match (i.e., a case of overcoverage). Pairs with θ_i close to 0 or 1 had the lowest probability of being selected for manual verification. The total sample size was 19,802 pairs.

The standardized GRLS matching weight, θ_i, was determined as follows:

$$\theta_i = (X_i - C)/(D - C)$$

Where: X_i is the GRLS matching weight for each pair i
C = S1-1
D= S2+1

By definition, we did not want S1 and S2 to be the boundary points of the interval [0,1]. This is why C is equal to S1-1 and D is equal to S2+1.

The sample was selected using the SAS statistical software PROC SURVEY SELECT procedure. The first-order inclusion probabilities were calculated in SAS. However, due to time and resource constraints, second-order inclusion probabilities which were needed to calculate the variance estimates, were not determined. As a result, the variance estimate is only an approximation and overestimates the true variance. Estimation is discussed in more detail in the next section.

The selection probabilities for a sample design with probability proportional to size using $\theta_i(1-\theta_i)$ as the size measure were calculated as follows:

$P_{ki} = \theta_{ki}(1-\theta_{ki}) \Big/ \sum_{i \in k} \theta_{ki}(1-\theta_{ki})$; k represents the stratum, i represents the pair

With the methodology outlined in this section, a pair whose weight was in the middle of the interval S1 to S2 had a greater chance of being verified. This is because these were the cases that we were more uncertain about. When the matching weight was close to S1, it was more likely not to be a case of overcoverage. In contrast, when the matching weight was close to S2, it was more likely to be a case of overcoverage. Therefore, there was no need to select a large sample near the end points of the interval to obtain good estimates.

A team of clerks examined information from the RDB to determine whether or not there was overcoverage. When necessary, they referred to census questionnaire images to verify RDB data to determine whether or not there was overcoverage. Quality control samples were selected as part of the manual verification process, to assess the quality of the coding. When the clerks were unsure about a case, it was referred to experts.

Table 8.2.2.2.1 provides estimates of total Step 2 overcoverage, overcoverage above S2 and overcoverage between S1 and S2 by province and territory. We note that most of the overcoverage comes from between S1 and S2. The total estimate is 235,946, of which 180,523 comes from between the thresholds and 55,423 comes from pairs above S2. The last are pairs with a matching weight sufficiently high as to be declared overcoverage without manual verification. Note that the coefficients of variation (CVs) are all under 10% (except for the Yukon Territory, with 11.54%). There is, of course, no variance associated with the overcoverage found above S2.

Table 8.2.2.2.1 Step 2 overcoverage for Canada, provinces and territories

Provinces and territories	Above the GRLS upper threshold	Between the GRLS upper and lower thresholds		Total Step 2		
	Number	Estimated number	Standard error	Estimated number	Standard error	Coefficient of variation (%)
Canada	**55,423**	**180,523**	**3,001**	**235,946**	**3,001**	**1.27**
Newfoundland and Labrador	545	2,618	253	3,163	253	8.00
Prince Edward Island	154	809	74	963	74	7.68
Nova Scotia	917	4,296	292	5,213	292	5.60
New Brunswick	918	2,763	183	3,681	183	4.97
Quebec	11,583	51,630	1,605	63,213	1,605	2.54
Ontario	21,677	63,770	1,931	85,447	1,931	2.26
Manitoba	1,085	6,688	442	7,773	442	5.69
Saskatchewan	944	5,695	356	6,639	356	5.36
Alberta	3,948	15,961	743	19,909	743	3.73
British Columbia	13,555	25,564	1,280	39,119	1,280	3.27
Yukon Territory	16	192	24	208	24	11.54
Northwest Territories	63	329	34	392	34	8.67
Nunavut	19	208	20	227	20	8.81

Source: Statistics Canada, 2006 Census Overcoverage Study.

Table 8.2.2.2.2 provides the interprovincial, interterritorial, intraprovincial and intraterritorial overcoverage. At the national level, 3.7% of the total overcoverage comes from the inter provincial-territorial overcoverage. In percentage terms, the inter/intra overcoverage is higher in the territories. This is expected because interprovincial/territorial overcoverage was not measured in Step 1. For the provinces, the highest proportions of inter/intra are, as in Step 1, in the Atlantic provinces and in Alberta.

Table 8.2.2.2.2 Step 2 intraprovincial, intraterritorial, interprovincial and interterritorial overcoverage for provinces and territories

Provinces and territories	Intraprovincial and intraterritorial overcoverage		Interprovincial and interterritorial overcoverage		Step 2 total
	Estimated number	Percentage of total overcoverage (%)	Estimated number	Percentage of total overcoverage (%)	Estimated number
Total	**227,281**	**96.3**	**8,665**	**3.7**	**235,946**
Newfoundland and Labrador	2,878	91.0	285	9.0	3,163
Prince Edward Island	878	91.1	85	8.9	963
Nova Scotia	4,677	89.7	536	10.3	5,213
New Brunswick	3,358	91.2	323	8.8	3,681
Quebec	62,282	98.5	931	1.5	63,213
Ontario	83,340	97.5	2,107	2.5	85,447
Manitoba	7,371	94.8	402	5.2	7,773
Saskatchewan	6,169	92.9	470	7.1	6,639
Alberta	18,097	90.9	1,811	9.1	19,908
British Columbia	37,757	96.5	1,362	3.5	39,119
Yukon Territory	111	53.5	96	46.5	208
Northwest Territories	254	64.9	138	35.1	392
Nunavut	109	47.9	118	52.1	227

Source: Statistics Canada, 2006 Census Overcoverage Study.

8.3 Estimation of overcoverage

In 2006, overcoverage was measured primarily by the Census Overcoverage Study (COS). The total overcoverage estimate comprises individuals overcovered in Step 1, and those deemed overcovered during the probabilistic matching in Step 2. Individuals deemed overcovered in Step 2 whose matching weight was above the upper threshold S2, had a weight of 1. The weight of overcoverage cases identified from the sample between the lower threshold S1 and the upper threshold S2 was determined by the sample design.

To evaluate the COS, the Automated Match Study (AMS) was repeated in 2006. The COS estimates were compared to those of the AMS. The comparison revealed a bias in the COS estimates whereby some pairs identified in the AMS were not found in the COS frames. Since the AMS provided an estimate of overcoverage not included in the COS, the last step in estimating overcoverage was to account for this bias by using the AMS estimates to adjust the COS estimates. This step is discussed at the end of the section. More information on evaluation of the COS is in Section 10.2.

The variance of the estimate of total overcoverage comes primarily from the sample between the thresholds S1 and S2 in Step 2. Another portion is from samples of two-to-two cases in Step 1, and a very small portion is obtained from the samples used to adjust for false matches in Step 1. As with the 2001 AMS, overcoverage observed between two provinces or territories is divided equally between the provinces or territories in question. The same principle applies to other domains of estimation.(Two individuals, for example, who do not belong to the same age group).

Statistics Canada's StatMx software was used to calculate point estimates, as well as the Step 2 sample variance between S1 and S2. As explained in Section 8.2.2.2, the sample was selected with probability proportional to $\theta_i(1-\theta_i)$, where θ_i represents the standardized GRLS matching weight defined on [0,1]. Since StatMx cannot produce variances for a PPSWOR (probability proportional to size without replacement) design, a variance estimate for a PPSWR (probability proportional to size with replacement) design was used. Consequently, the variance is overestimated. As explained in Section 8.2.2.2, this approximation comes from not deriving the second order inclusion probabilities.

Table 8.3.1 presents the total overcoverage estimates for Step 1 and Step 2.

Table 8.3.1 COS overcoverage for Canada, provinces and territories

	Step 1	Step 2		Total		
	Estimated number	Estimated number	Standard error	Estimated number	Standard error	Coefficient of variation (%)
Canada	**260,708**	**235,946**	**3,001**	**496,653**	**3,001**	**0.60**
Newfoundland and Labrador	4,906	3,163	253	8,069	253	3.14
Prince Edward Island	1,254	963	74	2,218	74	3.34
Nova Scotia	7,488	5,213	292	12,700	292	2.30
New Brunswick	6,381	3,681	183	10,062	183	1.82
Quebec	59,851	63,213	1,605	123,064	1,605	1.30
Ontario	95,398	85,447	1,931	180,844	1,931	1.07
Manitoba	8,427	7,773	442	16,200	442	2.73
Saskatchewan	7,904	6,639	356	14,543	356	2.45
Alberta	27,193	19,909	743	47,101	743	1.58
British Columbia	40,990	39,119	1,280	80,109	1,280	1.60
Yukon Territory	299	208	24	507	24	4.73
Northwest Territories	429	392	34	821	34	4.14
Nunavut	189	227	20	415	20	4.82

Source: Statistics Canada, 2006 Census Overcoverage Study.

Table 8.3.2 provides the total overcoverage estimate based on intraprovincial and interprovincial or territorial overcoverage. Some 4.5% of overcoverage is interprovincial or territorial. In Step 1, 53% of overcoverage is inter-provincial or territorial, while the figure is 3.7% for Step 2. Ontario and Quebec have the least interprovincial overcoverage, while the territories, Atlantic provinces and Alberta have the most.

Table 8.3.2 Intraprovincial, intraterritorial, interprovincial and interterritorial COS overcoverage for provinces and territories

Provinces and territories	Intraprovincial and intraterritorial overcoverage		Interprovincial and interterritorial overcoverage		Total
	Estimated number	Percentage of total overcoverage (%)	Estimated number	Percentage of total overcoverage (%)	Estimated number
Total	**474,262**	**95.5**	**22,391**	**4.5**	**496,653**
Newfoundland and Labrador	7,278	90.2	791	9.8	8,069
Prince Edward Island	1,979	89.2	239	10.8	2,218
Nova Scotia	11,319	89.1	1,381	10.9	12,700
New Brunswick	9,099	90.4	963	9.6	10,062
Quebec	120,358	97.8	2,706	2.2	123,064
Ontario	175,208	96.9	5,637	3.1	180,844
Manitoba	15,273	94.3	927	5.7	16,200
Saskatchewan	13,395	92.1	1,149	7.9	14,543
Alberta	42,566	90.4	4,535	9.6	47,101
British Columbia	76,408	95.4	3,701	4.6	80,109
Yukon Territory	408	80.6	98	19.4	507
Northwest Territories	678	82.6	143	17.4	821
Nunavut	294	70.7	122	29.3	415

Source: Statistics Canada, 2006 Census Overcoverage Study.

As described above, comparison of the COS estimates and the AMS estimates revealed a bias in the COS estimates. Consequently, an adjustment was made to the COS estimates using the AMS estimate of the undercoverage not covered by the COS. The adjusted estimates are the final estimates of total population overcoverage that appear in Section 1. Table 8.3.3 presents the overcoverage estimates before and after the AMS adjustment. The biggest increases are in Nunavut (6.11%) and Alberta (5.77%), while the smallest are in Quebec (2.56%) and the Yukon Territory (2.69%).

Table 8.3.3 Overcoverage estimates, before and after adjustment, for Canada, provinces and territories

Provinces and territories	Before adjustment	Adjustment		After adjustment	Percentage increase (%)
		Due to GRLS	Under S1 threshold		
Canada	**496,654**	**16,724**	**2,337**	**515,715**	**3.70**
Newfoundland and Labrador	8,069	228	12	8,309	2.89
Prince Edward Island	2,217	54	10	2,282	2.84
Nova Scotia	12,700	400	53	13,153	3.44
New Brunswick	10,062	399	24	10,485	4.04
Quebec	123,064	3,044	188	126,296	2.56
Ontario	180,845	6,450	999	188,294	3.96
Manitoba	16,200	409	174	16,783	3.47
Saskatchewan	14,543	544	52	15,140	3.94
Alberta	47,101	2,513	370	49,984	5.77
British Columbia	80,109	2,614	448	83,171	3.68
Yukon Territory	507	11	3	521	2.69
Northwest Territories	821	31	2	854	3.86
Nunavut	415	26	1	442	6.11

Source: Statistics Canada, 2006 Census Overcoverage Study.

8.4 Types of overcoverage

In 2006, the possible types of overcoverage were examined for the first time. The most frequent types are described below.

Some 20% of COS overcoverage is from the 'consecutive/quasi consecutive identifier' category. This refers to overcoverage from two identical households with exactly the same address or in very close geographic proximity (and therefore have a similar household identifier). Two households were considered identical if they contained the same people with the same demographic characteristics.

Another 20.5% of COS overcoverage is from the ' identical households: not consecutive/quasi consecutive' category. These are identical households that are not geographically close. A further 16.9% of COS overcoverage is from the 'child(ren) of parents living in separate households' category.

We then find 12.0% of COS overcoverage in the 'student/young adult who has recently left home' category, and 11.1% in the 'non-identical households: one household is included in another' category (whereby the members of one household can all be found in the other larger household).

9. Estimation

Estimation for the DCS, RRC, and the COS are covered in Section 6.2, Section 7.4, and Section 8.3 respectively. This section describes how the results of the census coverage studies are combined to produce estimates of census population undercoverage (U), population

overcoverage (O), and population net undercoverage (N) for a variety of domains. The impact of sampling error on the quality of the estimates is also produced by calculating an estimated standard error for each estimate. The results of the Reverse Record Check (RRC) and census data are used to construct estimates of undercoverage while the results of the Census Overcoverage Survey (COS) provide estimates of overcoverage. Net undercoverage is the difference between undercoverage and overcoverage. This section details the calculation of these estimates and their estimated standard errors.

Let:

C = published census count of the number of persons in the census target population

$\hat{U}$ = estimate of undercoverage
= estimate of the number of persons not included in C that should have been

$\hat{O}$ = estimate of overcoverage
= estimate of the number of persons included in C who should not have been

$\hat{N}$ = estimate of net undercoverage
= estimate of the number of persons not included in C who should have been net of the number of persons included in C who should not have been
= $\hat{U} - \hat{O}$

$\hat{R}_U$ = estimate of undercoverage rate
= $100 * \frac{\hat{U}}{C + \hat{N}}$

$\hat{R}_O$ = estimate of overcoverage rate
= $100 * \frac{\hat{O}}{C + \hat{N}}$

$\hat{R}_N$ = estimate of net undercoverage rate
= $100 * \frac{\hat{U} - \hat{O}}{C + \hat{N}}$

The estimate of overcoverage from the COS is constructed by summing the weights for each person found to be involved in more than one enumeration. If, for example, a case of overcoverage involving two enumerations is found in Step 1 of the COS, then each enumeration receives a weight of ½. Assuming that overcoverage other than from multiple enumerations, such as enumeration of fictitious persons, is negligible, then:

If $\hat{O}^{COS}$ = estimate of overcoverage from the COS, then $\hat{O} = \hat{O}^{COS}$.

$\hat{U}$ is constructed from the results of the RRC and census data as follows.

Let:

$\hat{M}$ = estimate of the number of persons in the RRC target population that have not been enumerated
= sum of the final weights of persons classified as 'missed'

X = number of persons included in C that cannot be identified in the RRC as enumerated

Then $\hat{U} = \hat{M} - X$.

X , for database extractions, can be determined from the final census database. Persons in the RRC sample (SPs) who are in scope for the census, but for whom the RRC cannot determine whether or not they have been enumerated at their Census Day address, are classified as Missed. There are a number of reasons why an SP could not be identified as enumerated:

1. The SP's Census Day address points to a dwelling that contains imputed enumerations. This is the case, for example, for non-response dwellings for which the data of another household was used as a result of whole-household imputation (WHI).

2. Some enumerations on the census database were deemed too incomplete to be used by the RRC to identify an SP as enumerated. Incomplete enumerations in this context usually involve invalid data in the date of birth or the name field such as a name of '?' or 'Mr.' or 'Unknown' or 'Person 1.' Any SP pointing to such an enumeration was classified as missed. These are the 'RRC incomplete enumerations.'

3. There were some enumerations added to the census database after the data were extracted to create the RRC database. These late enumerations were not available to the RRC so the RRC could not identify any enumeration at these dwellings.

At the national level, X is about half of $\hat{M}$. This is a notable increase from 2001 when X was only about ⅓ of $\hat{M}$. The increase is largely due to an increase in both the number of non-response dwellings and the number of misclassified dwellings that resulted in doubling the number of persons imputed during the WHI step of census processing. The following table gives the components of the estimated population coverage error for Canada.

Table 9.1 Components of estimate of population coverage error for Canada

Components	Number of persons
Estimate of M	2,846,337
Total X	**1,461,965**
X for imputed persons	933,176
X for late enumerations	105,833
X for RRC incomplete enumerations	422,956
Estimate of N	**868, 657**
Estimate of U	1,384,372
Estimate of O	515,715
C	31,612,897
C + Estimate of N	32,481,554

Source: Statistics Canada, 2006 Census Coverage Studies.

The estimated standard errors are defined as follows:

Let:

$\hat{T}$ = estimate of the number of persons in the census target population derived from the census count and the estimate of net population undercoverage

= $C + \hat{N}$

$v(\hat{M})$ = estimated variance of $\hat{M}$ as determined by the design of the RRC

$v(\hat{O})$ = estimated variance of $\hat{O}$ as determined by the design of the COS

Then,

$$se(\hat{U}) = \sqrt{v(\hat{M})}\ ,\ se\left(\hat{R}_u\right) = \sqrt{\left(\frac{\hat{U}^2 + \hat{T}^2 + 2\hat{U}\hat{T}}{\hat{T}^4}\right)v\left(\hat{M}\right) + \frac{\hat{U}^2}{\hat{T}^4}v\left(\hat{O}\right)}$$

$$se\left(\hat{O}\right) = \sqrt{v(\hat{O})}\ ,\ se\left(\hat{R}_O\right) = \sqrt{\left(\frac{\hat{O}^2}{\hat{T}^4}\right)v(\hat{M}) + \left(\frac{\hat{O}^2 + \hat{T}^2 + 2\hat{O}\hat{T}}{\hat{T}^4}\right)v(\hat{O})}$$

$$se\left(\hat{N}\right) = \sqrt{v(\hat{M}) + v(\hat{O})}\ ,\ se\left(\hat{R}_N\right) = \sqrt{\left(\frac{(\hat{U} - \hat{O})^2 + \hat{T}^2 + 2(\hat{U} - \hat{O})\hat{T}^2}{\hat{T}^4}\right)\left[v(\hat{M}) + v(\hat{O})\right]}$$

10. Evaluation of Coverage Studies

10.1 Reverse Record Check

10.1.1 Introduction

The results of the largest coverage study, the Reverse Record Check (RRC), can be evaluated by comparing its estimates with data on the same characteristics from other sources such as the 2006 Census database. Comparisons with RRC estimates serve to evaluate RRC estimates and to quantify conceptual and measurement differences.

Despite some conceptual differences between the RRC and the 2006 Census, the RRC estimate of persons enumerated in the 2006 Census can be compared with the census count. However, to render the two numbers comparable, certain adjustments were made to the census counts before comparing them.

RRC intercensal components of growth estimates can be compared with estimates from other sources. The RRC estimate of persons who died between the 2001 Census and the 2006 Census can be compared with the count from vital statistics files. Estimates of counts of net interprovincial migration from Canada Customs and Revenue Agency data can be compared with RRC estimates. It is not possible, however, to construct strict comparisons for this characteristic, since reasonable adjustments for conceptual differences cannot be derived. Last, RRC estimates of population growth components can be compared with similar estimates from administrative data.

10.1.2 Comparisons with census counts

Since the RRC single-stage stratified sampling design results in unbiased estimators, differences between RRC estimates and census counts are due to sampling error on the part of the RRC estimates, conceptual differences between the two sources, and/or systematic biases in the two sources that result in an underestimate or overestimate of the characteristic under study.

10.1.2.1 Enumerated

Provincial and national comparisons are presented in Table 10.1.2.1 with the standard error of the RRC estimate and the *t*-value for testing the hypothesis that there is no difference between the RRC estimate and the comparable census count. The following adjustments were made to published census counts to account for conceptual differences between the two sources:

- Imputations added in the whole household imputation stage of the census based on the results of the Dwelling Classification Study are not included. This is because while they are included in census counts, they are not part of the RRC estimate of enumerated persons.

- 2006 Census overcoverage is subtracted. This is because the census counts contain overcoverage whereas the RRC estimate is based on the number of unique persons enumerated rather than the number of enumerations.

- The census count of persons living outside Canada five years ago (based on Form 2B data) excluding immigrants from the intercensal period and non-permanent residents, is also subtracted. This is because the RRC estimates do not include these persons.

- Last, 2001 Census overcoverage is added. This is because there is overcoverage in the RRC estimates via the initial weights in the 2001 Census sampling frame. These weights were not adjusted for this overcoverage.

Table 10.1.2.1 Comparison of RRC estimates of the number of enumerated persons and comparable census counts for Canada, provinces and territories

	Enumerated persons				
	RRC				
Provinces and territories	Estimated number	Standard error	Comparable census count	Difference	t-value[1]
Canada	**29,700,814**	**60,913**	**29,702 519**	**-1,705**	**-0.03**
Newfoundland and Labrador	477,780	4,998	480,933	-3,153	-0.63
Prince Edward Island	129,689	2,851	129,539	150	0.05
Nova Scotia	855,894	8,146	864,467	-8,573	-1.05
New Brunswick	696,566	7,025	698,562	-1,996	-0.28
Quebec	7,166,868	30,743	7,117,212	49,656	1.62
Ontario	11,486,571	49,339	11,492,043	-5,472	-0.11
Manitoba	1,083,564	9,030	1,084,518	-954	-0.11
Saskatchewan	909,187	8,066	917,063	-7,876	-0.98
Alberta	3,032,038	21,619	3,054,382	-22,344	-1.03
British Columbia	3,766,948	23,393	3,768,090	-1,142	-0.05
Yukon Territory	28,601	0	28,601	0	…
Northwest Territories	38,781	0	38,781	0	…
Nunavut	28,328	0	28,328	0	…

...not applicable
Note: 1. A t-value greater than 1.96 or less than -1.96 indicates that the difference is significant at the 95% level.
Source: Statistics Canada, 2006 Census Reverse Record Check.

Nationally, the RRC estimate of the number of persons enumerated in the 2006 Census is slightly lower, -0.03%, than the comparable 2006 Census count. In 2001, the RRC overestimated the comparable census count by 0.07% while in 1996 the RRC underestimated the census by 0.08%. Provincially, the greatest difference occurs for Quebec (*t*-value of 1.62) where the RRC estimate of the number of enumerated persons exceeds the comparable census count by 49,656. In the majority of provinces the difference is negative, though relatively small in most cases. None of the observed differences are statistically significant.

The most significant differences must be investigated, since they may be due to a bias in RRC classification (including, for example, province of residence on Census Day). However, other factors also play an important role. Apart from sampling errors, the difference may be explained by biases in the adjustments applied to the published census count to obtain a conceptually comparable figure (e.g., returning Canadians). The RRC non-response bias may also affect this difference, since the non-response adjustment is designed to obtain the best result for estimating missed persons rather than enumerated persons. Though there are few significant differences, the fact that most of them are negative may indicate a slight bias.

10.1.3 Comparison with population estimates

10.1.3.1 Deceased persons

Table 10.1.3.1 compares the estimated number of persons deceased during most of the intercensal period (i.e. May 15, 2001 to December 31, 2005) by RRC province of classification with counts from vital statistics files of deaths (VS). Deaths in 2006 are excluded, since vital statistics for 2006 were not yet available at the time of the analysis. At the national level, the RRC estimate is higher than the VS count by 9,134 (0.9%). The highest relative difference is observed in Manitoba (-6,635/45,687, or 14.5%). In absolute value terms, the differences vary from 0.8% to 14.5%. In t-value terms, the highest t-value occurs for (2.05) where the RRC estimate is higher than the VS count, and in Manitoba (-1.87) where the RRC estimate is lower than the VS count. All other estimates are well below the 95% confidence levels. Despite the slightly higher difference in British Columbia, these results indicate no need for further investigation.

Table 10.1.3.1 Comparison of RRC estimates of the number of deceased persons and vital statistics count for the provinces

Provinces	Persons deceased May 15, 2001 to December 31, 2005				
	RRC				
	Estimated number	Standard error	Vital statistics count	Difference	t-value[1]
Total	**1,049,974**	**31,499**	**1,040,840**	**9,134**	**0.29**
Newfoundland and Labrador	19,157	1,967	19,790	-633	-0.32
Prince Edward Island	4,897	471	5,335	-438	-0.93
Nova Scotia	37,845	3,239	37,542	303	0.09
New Brunswick	29,586	2,499	28,780	806	0.32
Quebec	240,062	15,892	253,962	-13,900	-0.87
Ontario	388,878	23,282	386,267	2,611	0.11
Manitoba	39,052	3,539	45,687	-6,635	-1.87
Saskatchewan	39,027	3,313	40,857	-1,830	-0.55
Alberta	95,953	8,239	86,324	9,629	1.17
British Columbia	155,518	9,380	136,296	19,222	2.05

Note: 1. A t-value greater than 1.96 or less than -1.96 indicates that the difference is significant at the 95% level.
Source: Statistics Canada, 2006 Census Reverse Record Check.

10.1.3.2 Interprovincial migration

Table 10.1.3.2 compares RRC estimates of net interprovincial migration for the intercensal period with corresponding figures from Canada Customs and Revenue Agency (CCRA) files. In general, in-migration and out-migration statistics are not comparable since the RRC only takes into account migration flows that occurred between the sampling frame reference date e.g., May 15, 2001 for the census frame, and Census Day 2006, while estimates based on CCRA data take annual migration into account. Accordingly, only net migration estimates are presented.

The difference is significant for Alberta (t-value of 2.38) where the RRC estimates a much higher positive net migration than estimates derived from CCRA data. While both sources estimate a strong positive net migration, the net amount differs depending on the source. It is recognized that there has been considerable migration to Alberta, and that it may be hard to distinguish between permanent and temporary migration. Some who migrate to Alberta to work have settled there permanently. Others have gone there to work, but have retained their residence in their province of origin and return there on a more or less frequent basis. Census respondents do not always correctly identify the location where they should be enumerated. Furthermore, the census question used to identify mobility refers to the place where the person lived one year ago and five years ago but does not specify the concept of usual residence. It is therefore possible that the respondent will provide a temporary place of residence, leading to a misinterpretation of his or her mobility. The combination of these two factors may affect the accuracy of census estimates.

For all provinces except British Columbia, both estimates show net migration in the same direction. The Reverse Record Check estimate for British Columbia indicates a low negative net migration of 2,316, while the demographic data indicate a positive net migration of 12,887. The t-values for this comparison do not suggest a need for further investigation.

Table 10.1.3.2 Comparison of RRC estimates of net interprovincial migration and count from Canada Customs and Revenue Agency data for provinces

	Net interprovincial migration					
	RRC[1]					
Provinces	Sample size	Estimated number	Standard error	CCRA count	Difference	t-value[2]
Newfoundland and Labrador	408	-17,566	4,425	-14,530	-3,036	-0.69
Prince Edward Island	381	-707	2,724	-434	-273	-0.10
Nova Scotia	710	-16,011	7,626	-6,987	-9,024	-1.18
New Brunswick	552	-10,187	6,326	-7,491	-2,696	-0.43
Quebec	397	-20,524	14,679	-20,789	265	0.02
Ontario	1,293	-29,777	23,213	-24,884	-4,893	-0.21
Manitoba	430	-25,718	8,169	-24,655	-1,063	-0.13
Saskatchewan	587	-41,151	6,883	-36,009	-5,142	-0.75
Alberta	1,512	163,956	17,241	122,892	41,064	2.38
British Columbia	1,008	-2,316	16,693	12,887	-15,203	-0.91

Notes: 1. The RRC excludes persons living in a province on May 15, 2006 who had lived in one of three territories five years earlier on May 15, 2001.
2. A t-value either greater than 1.96 or less than -1.96 indicates that the difference is significant at the 95% level.

Source: Statistics Canada, 2006 Reverse Record Check.

10.1.4 Components of population growth

An extensive comparison of RRC estimates of the components of intercensal population growth with census counts and population estimates derived from administrative data was done by the Demography Division (This topic is also discussed in Section 10.3.). The RRC estimates are a by-product of the RRC and therefore not necessarily precise. Table 10.1.4 compares the two estimates of population growth by component. Note that estimates of returning Canadians, and persons living on Indian reserves or settlements who were incompletely enumerated in 2001 and enumerated in 2006 were added to the RRC estimates to make them comparable to the administrative data counts.

The administrative data counts are a combination of a number of estimates of population growth component: births, deaths, immigration, internal migration, emigration, net number of non-permanent residents, growth of non-enumerated Indian reserves. These counts are subject to varying amounts of measurement error depending on the source. This is particularly so for the net number of non-permanent residents. It is also important to note that the RRC is not designed to produce estimates of the components of growth. The components estimates are a by-product of the RRC. Therefore, differences between the RRC estimates and the administrative data counts are to be expected.

Nationally, RRC estimates differ by 5.1% from the administrative data estimates. The largest differences occur for British Columbia (-77,192) and Ontario (-49,371). As a percentage of the administrative data estimates, these differences amount to 32.0% and 6.0% respectively.

Table 10.1.4 Comparison of RRC estimates of population growth and estimates from administrative data for the provinces

	Population growth May 15, 2001 to May 15, 2006		
	RRC	Administrative data	
Provinces	Estimated number	Estimated number	Difference
Total	**1,546,667**	**1,629,624**	**-82,957**
Newfoundland and Labrador	-14,969	-12,213	-2,756
Prince Edward Island	1,285	1,393	-108
Nova Scotia	-3,082	3,197	-6,279
New Brunswick	-5,533	629	-6,162
Quebec	282,784	253,150	29,634
Ontario	773,672	823,043	-49,371
Manitoba	31,185	26,816	4,369
Saskatchewan	-17,131	-13,529	-3,602
Alberta	336,990	308,480	28,510
British Columbia	161,466	238,658	-77,192

Source: Statistics Canada, 2006 Reverse Record Check.

10.2 Census Overcoverage Study

Due to methodology changes in the way overcoverage was measured in 2006, we need a tool to evaluate whether the changes would cause a significant variation in overcoverage rates.

In general, the 2006 overcoverage can be divided into two parts: the part that would have been covered in 2001 by the Automated Match Study (AMS), and the part that would have been

covered by the Reverse Record Check (RRC). For simplicity, and since it covers only a small portion of the total overcoverage (less than 1%), the Collective Dwelling Study (CDS) is not considered in this evaluation.

10.2.1 Comparison of 2001 and 2006 Automated Match Study (AMS)

The AMS was conducted again in 2006, to compare the 2001 overcoverage estimates with those of 2006 and to ensure differences observed between 2001 and 2006 were not due to methodology changes. The 'Monster Match' and 'Mini-Monster Match' computer programs were executed, to find similar pairs of households in the 2006 AMS sample frame. This portion of the overcoverage can be measured using the 2001 methodology. We should point out that the geographic variables used to identify a dwelling in 2001, i.e., province, provincial electoral district (PED) and enumeration area (EA), were replaced by a combination of province, census division (CD) and collection unit (CU). Since we wished to duplicate the 2001 methodology as closely as possible, CDs were converted into PEDs. However, we could not do so with as much accuracy for the 2006 collection units. Nonetheless, the concept of collection unit is similar to the concept of enumeration area.

Table 10.2.1 compares the 2001 AMS estimates with those of 2006.

Table 10.2.1 AMS estimated overcoverage for Canada, provinces and territories, 1996, 2001 and 2006 censuses

	1996			2001				2006			
	Estimated number	Standard error	Coefficient of variation (%)	Estimated number	Standard error	Coefficient of variation (%)	Percentage increase (%)	Estimated number	Standard error	Coefficient of variation (%)	Percentage increase (%)
Canada	**93,688**	**3,505**	**3.74**	**146,412**	**3,430**	**2.34**	**56.28**	**292,594**	**4,578**	**1.56**	**99.84**
Newfoundland and Labrador	1,366	113	8.27	1,657	118	7.12	21.32	4,710	129	2.74	184.20
Prince Edward Island	445	25	5.62	439	31	7.06	-1.32	1,293	64	4.95	194.55
Nova Scotia	2,098	174	8.29	2,875	201	6.99	37.04	6,696	254	3.79	132.90
New Brunswick	1,609	157	9.76	2,608	230	8.82	62.08	5,807	261	4.49	122.67
Quebec	22,893	1,777	7.76	35,061	1,635	4.66	53.15	68,373	2,008	2.94	95.01
Ontario	37,387	2,822	7.55	53,378	2,490	4.66	42.77	108,488	3,572	3.29	103.24
Manitoba	3,445	301	8.74	4,128	303	7.34	19.84	8,873	420	4.73	114.92
Saskatchewan	2,034	199	9.78	3,629	272	7.49	78.43	7,601	347	4.56	109.43
Alberta	5,226	473	9.05	13,263	834	6.29	153.79	26,574	1,327	5.00	100.36
British Columbia	16,697	858	5.14	28,710	1,386	4.83	71.95	53,338	1,392	2.61	85.78
Yukon Territory	81	10	12.35	173	26	14.99	114.20	230	16	6.75	32.81
Northwest Territories	407	50	12.29	325	25	7.70	-20.26	446	30	6.79	37.40
Nunavut	…	…	…	165	27	16.40	…	164	19	11.66	-0.10

… not applicable
Sources: Statistics Canada, Automated Match Study for 1996, 2001 and 2006.

We noted a steady increase in overcoverage between 1996 and 2006. Nationally, there was a 56% increase between 1996 and 2001 and a 99.8% increase between 2001 and 2006. The largest increases were observed in the Atlantic provinces, while the smallest ones occurred in the territories.

Since the AMS methodology did not change between 1996 and 2006, we may conclude that the increases are necessarily due to an actual increase in overcoverage in the Census. We noted that coefficients of variation (CVs) decreased between 1996 and 2006. Nationally, the CV decreased from 3.74% in 1996 to 1.56% in 2006.

10.2.2 Comparison of 2006 AMS and 2006 COS

One evaluation tool involves comparing COS and AMS overcoverage estimates. The procedure consists of matching all COS pairs that contain overcoverage with the AMS survey frame (this is referred to as the AMS domain of the COS). For matching pairs, we use COS weighting to estimate the AMS domain of the COS.

Here are the comparative results of overcoverage estimates of the 2006 AMS and of the AMS domain of the COS.

Table 10.2.2 AMS estimated overcoverage and COS estimated AMS domain overcoverage for Canada, provinces and territories

	AMS estimate		COS estimate of AMS domain			
Provinces and territories	Estimated number	Standard error	Estimated number	Standard error	Difference[1]	Percentage difference (%)[2]
Canada	**292,594**	**4,578**	**268,372**	**2,999**	**-24,222**	**-8.28**
Newfoundland and Labrador	4,710	129	4,178	253	-532	-11.30
Prince Edward Island	1,293	64	1,192	74	-101	-7.82
Nova Scotia	6,696	254	6,001	292	-695	-10.38
New Brunswick	5,807	261	5,096	183	-711	-12.24
Quebec	68,373	2,008	63,490	1,603	-4,883	-7.14
Ontario	108,488	3,572	100,017	1,930	-8,471	-7.81
Manitoba	8,873	420	7,497	442	-1,376	-15.51
Saskatchewan	7,601	347	6,890	356	-711	-9.35
Alberta	26,574	1,327	23,874	742	-2,700	-10.16
British Columbia	53,338	1,392	49,496	1,280	-3,842	-7.20
Yukon Territory	230	16	189	24	-41	-17.93
Northwest Territories	446	30	342	34	-104	-23.25
Nunavut	164	19	110	20	-55	-33.33

Notes: 1. AMS estimate minus COS estimate.
2. AMS estimate minus COS estimate as a percentage of AMS estimate.
Sources: Statistics Canada, 2006 Automated Match Study, 2006 Census Overcoverage Study.

Nationally, we noted a difference of 8.28%. The largest differences were in the territories, especially in Nunavut where the AMS domain of the COS was 33.33% lower than the 2006 AMS. The difference was also consistently negative.

The final part of the evaluation entailed verifying whether cases identified as overcoverage by the AMS were also identified as overcoverage by the COS. This evaluation identified a bias in the COS estimates. Estimates from the AMS were used to adjust COS estimates for this bias, as outlined in Table 8.3.3, Section 8.3.

10.2.3 Reliability

Table 10.2.3 shows the reliability of the 2001 and 2006 overcoverage estimates in terms of estimated coefficient of variation. In 2006 all coefficients of variation (CVs) were below 5%, and we noted a significant reduction in CVs. This is because more than half the estimate was based on overcoverage cases with a weight of one. The reduction was also partly due to the fact that in 2001, standard errors deriving from the RRC were high compared to those of the AMS, the two main contributors to the overcoverage estimate.

Table 10.2.3 Estimated coefficient of variation for estimated overcoverage for Canada, provinces and territories, 2001 and 2006 censuses

	2001 Census	2006 Census
Provinces and territories	**Estimated coefficient of variation (%)**	**Estimated coefficient of variation (%)**
Canada	**4.30**	**0.60**
Newfoundland and Labrador	16.26	3.14
Prince Edward Island	18.88	3.34
Nova Scotia	16.64	2.30
New Brunswick	21.62	1.82
Quebec	9.12	1.30
Ontario	8.30	1.07
Manitoba	17.44	2.73
Saskatchewan	18.48	2.45
Alberta	12.76	1.58
British Columbia	8.49	1.60
Yukon Territory	20.77	4.73
Northwest Territories	15.49	4.14
Nunavut	16.36	4.82

Source: Statistics Canada, 2006 Census Overcoverage Study.

10.3 Population estimates

10.3.1 Error of closure

Statistics Canada's Population Estimates Program (PEP) determines provincial and territorial population counts on Census Day by adding census population counts and estimates of census population net undercoverage[7]. The PEP then extends these adjusted census counts to July 1, whereupon they become the base population for postcensal population estimates. For more

7. The PEP also adds estimates of populations living on incompletely enumerated Indian reserves (IEIR) to these figures.

information on population estimates, see Estimates of Total Population, Canada, Provinces and Territories.

When determining the adjusted census counts, the PEP evaluates the quality of postcensal estimates for the five-year period preceding the census by comparing postcensal estimates for Census Day with the adjusted census counts. The difference between the two is referred to as the error of closure. A detailed review of this error is the main quality evaluation of the postcensal estimates.

Table 10.3.1 provides errors of closure for 2006 and 2001 by province and territory. Note that a positive error means the postcensal estimate has overestimated the population. For Canada in 2006 the error of closure was +105,352, an error rate of +0.32%. This is double the 2001 rate. For eight provinces, the 2006 error rates were between -0.5% and +0.5%. The rates were higher for the Yukon Territory, British Columbia, the Northwest Territories, Alberta and Nunavut. Of these five regions, only British Columbia had a positive error. Compared to 2001, the 2006 rates were higher for the Yukon Territory, British Columbia and Alberta, and lower for the Atlantic provinces and Saskatchewan. Overall, as in 2001, the majority of provinces had small errors of closure. However, unlike 2001, the largest errors in 2006 ocurred in two of the most populous provinces.

Table 10.3.1 Error of closure of population estimates for Canada, provinces and territories, 2001 and 2006

	2001		2006	
Provinces and territories	**Number**	**Rate (%)**	**Number**	**Rate (%)**
Canada	**49,948**	**0.16**	**105,352**	**0.32**
Newfoundland and Labrador	11,381	2.18	-1,137	-0.22
Prince Edward Island	1,483	1.09	74	0.05
Nova Scotia	9,005	0.97	-2,778	-0.30
New Brunswick	4,587	0.61	3,404	0.46
Quebec	-222	0.00	26,088	0.34
Ontario	11,288	0.10	50,173	0.40
Manitoba	-1,035	-0.09	-5,396	-0.46
Saskatchewan	16,017	1.60	-3,061	-0.31
Alberta	1,604	0.05	-47,801	-1.40
British Columbia	-4,347	-0.11	88,057	2.08
Yukon Territory	- 360	-1.20	- 979	-3.04
Northwest Territories	497	1.22	- 875	-2.03
Nunavut	50	0.18	-417	-1.36

Source: Statistics Canada, Demography Division.

10.3.2 Error in two terms

The size of the error of closure depends, on the one hand, on error in the postcensal population estimates and, on the other hand, on error in the estimate of census net population net undercoverage. In order to evaluate error in the postcensal estimates, it is useful to express the error of closure in two terms:

$$E = P_{PEP} - P_{AC} = \left(P_{RRC} - P_{AC}\right) - \left(P_{RRC} - P_{PEP}\right)$$

where

P_{RRC}: Population estimated directly from the Reverse Record Check (RRC)

= RRC estimate of the number of persons enumerated + census population net undercoverage + persons on Incompletely Enumerated Indian Reserves (IEIR)

P_{AC}: Adjusted census population

= census enumerations + census population net undercoverage + persons on on IEIR

P_{PEP}: Postcensal estimate of population on Census Day

The first term compares the Census Day population estimated directly by the RRC with the adjusted census population. This difference, which is the difference between the RRC estimate of the number of enumerated persons and the number of census enumerations, should be due mainly to RRC sampling error[8].

The second term compares the population estimated directly by the RRC with the postcensal estimate of population. This difference is a comparison of the RRC estimate of population growth with the PEP estimate. This term helps to determine whether the PEP estimates of population growth have important errors that may have contributed to the error of closure.

Table 10.3.2 presents the two error of closure terms for 2006 and their associated standard errors. First, we note that the error of closure is significant for Alberta, British Columbia, the Yukon Territory and the Northwest Territories.

8. The calculation of the differences in the number of enumerations requires certain adjustments to make the census and RRC numbers comparable. In particular, returning Canadians among those enumerated in the census but not among those enumerated by the RRC must be considered.

Table 10.3.2 Error of closure in two terms for Canada, provinces and territories, 2006

Provinces and territories	$(P_{RRC} - P_{AC}) - (P_{RRC} - P_{PEP})$			$P_{RRC} - P_{AC}$			$P_{RRC} - P_{PEP}$		
	Estimated number	Standard error	t-value[1]	Estimated number	Standard error	t-value[1]	Estimated number	Standard error	t-value[1]
Canada	**105,352**	**58,259**	**1.81**	**-1,705**	**60,913**	**-0.03**	**-107,057**	**47,426**	**-2.26**
Provinces	**107,623**	**58,254**	**1.85**	**-1,705**	**60,913**	**-0.03**	**-109,328**	**47,421**	**-2.31**
Newfoundland and Labrador	-1,137	2,907	-0.39	-3,153	4,998	-0.63	-2,016	4,941	-0.41
Prince Edward Island	74	941	0.08	150	2,851	0.05	76	2,785	0.03
Nova Scotia	-2,778	5,287	-0.53	-8,573	8,146	-1.05	-5,795	8,248	-0.70
New Brunswick	3,404	3,934	0.87	-1,996	7,025	-0.28	-5,400	6,913	-0.78
Quebec	26,088	25,638	1.02	49,656	30,743	1.62	23,568	24,794	0.95
Ontario	50,173	45,499	1.10	-5,472	49,339	-0.11	-55,645	43,102	-1.29
Manitoba	-5,396	7,510	-0.72	-954	9,030	-0.11	4,442	9,477	0.47
Saskatchewan	-3,061	5,426	-0.56	-7,876	8,066	-0.98	-4,816	7,878	-0.61
Alberta	-47,801	17,623	-2.71	-22,344	21,619	-1.03	25,457	20,747	1.23
British Columbia	88,057	20,020	4.40	-1,142	23,393	-0.05	-89,199	23,111	-3.86
Territories	**-2,271**	**..**	**..**	**..**	**..**	**..**	**2,271**	**..**	**..**
Yukon Territory	-979	420	-2.33	0	0	**.**	979	420	2.33
Northwest Territories	-875	432	-2.02	0	0	**.**	875	432	2.02
Nunavut	-417	447	-0.93	0	0	**.**	417	447	0.93

. not available for this reference period
.. not available for any reference period
Note: 1. A t-value greater than 1.96 or less than -1.96 indicates that the difference is significant at the 95% level.
Source: Statistics Canada, Demography Division.

The error of closure is clearly dominated by the difference between the estimates of growth and the PEP estimates. This is true for British Columbia and Ontario as well, although the difference is not significant for Ontario. For Alberta on the other hand, significant error of closure is equally due to both error types, with neither being significant. For most of the other provinces, the two types of error are smaller. In fact, sampling error in the RRC estimates as expressed by differences in the number of enumerated does not contribute significantly to the error of closure for any of the provinces, while error in the postcensal estimates does contribute significantly to the error of closure for Canada and British Columbia.

Analysis of the two error terms is not appropriate for the territories since the last step in RRC estimation, for the territories only, is to calibrate to census counts. In addition, the RRC cannot directly estimate population growth for 2001 to 2006 because only a sampling frame for the 2006 Census Day is used. In general, migration patterns for the territorial population, especially for short-term moves which are more difficult to estimate, account for the higher rate of errors of closure in the provinces.

10.3.3 Growth

The error term may be broken down further. Table 10.3.3 presents the differences between RRC and PEP estimates of population growth by growth component for the two provinces, Alberta and British Columbia, where the error of closure is significant, and for the provinces combined.

Table 10.3.3 Difference between RRC and PEP estimates of 2001 to 2006 population growth components for Alberta, British Columbia and provinces

Population growth components	Alberta	British Columbia	Provinces
Total	**25,457**	**-89,199**	**-109,328**
Natural increase	-9,380	-17,806	-10,351
Net interprovincial migration	43,791	-13,034	0
Net international migration	-6,872	-55,635	-93,028
Other	-2,081	-2,724	-5,949

Source: Statistics Canada, Demography Division.

For Canada excluding the territories, the difference between the RRC and PEP estimates of population growth is almost entirely due to the difference between the RRC and PEP estimates of net international migration. The PEP estimate includes immigration, a very reliable component, net non-permanent residents which may also be considered reliable, and emigration. The RRC-PEP difference is therefore somewhat due to the emigration difference. For Canada excluding the territories, the difference between the RRC and PEP estimates of emigration are also significant. That is, the PEP estimate of the number of persons exiting the country is significantly lower than the RRC estimate. This gap comes largely from British Columbia and Ontario.

For British Columbia, as for Canada, the difference in net international migration comprises the largest portion of the difference between RRC and PEP growth estimates. PEP emigration estimates for this province are significantly lower than those of the RRC. The difference in natural growth from a high RRC estimate of deaths also contributes to the overall gap. Last, the difference in net interprovincial migration is a factor, though the difference is not significant.

For Alberta, the difference between RRC and PEP growth is not significant. This is due to a difference that is positive and significant for net interprovincial migration, but negative for the other components. The RRC estimate of net interprovincial migration is significantly higher than

that of the PEP. Without this difference the overall gap between RRC and PEP would have been smaller, as would the error of closure. A slightly low PEP estimate of net interprovincial migration for Alberta would thus have contributed to the negative error of closure for this province.

10.3.4 Conclusion

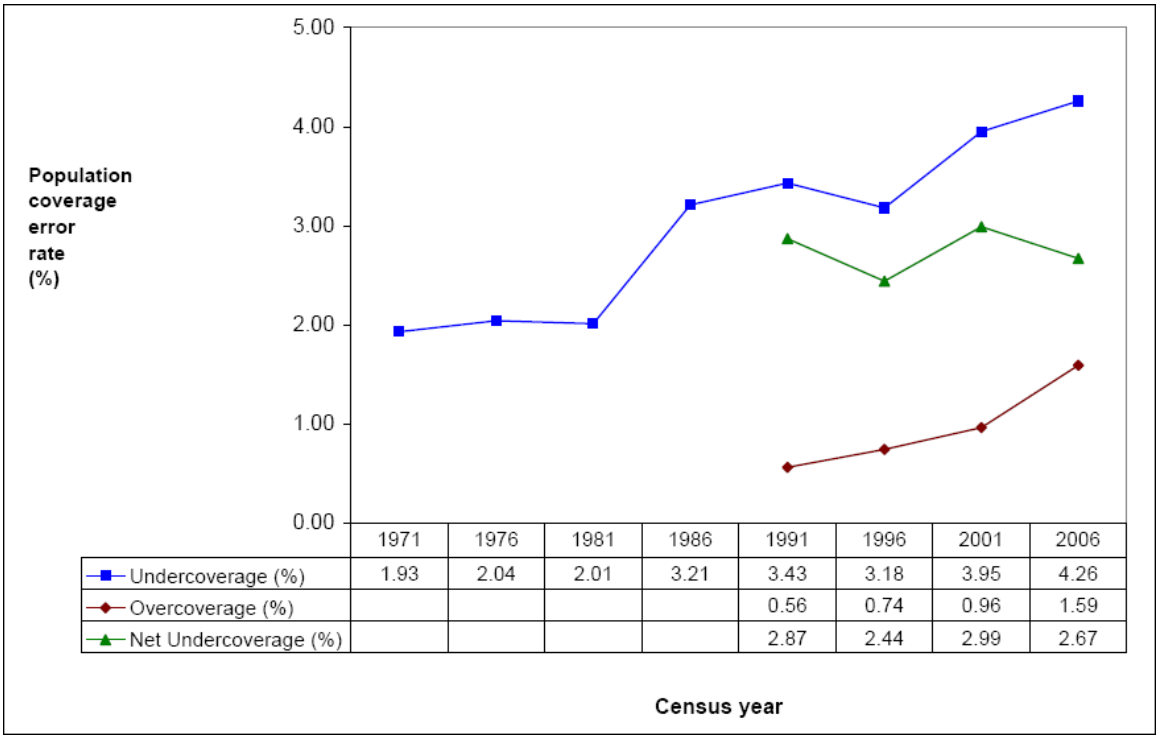
In conclusion and allowing for RRC sampling error, the PEP estimates are consistent with census data adjusted for population net undercoverage. Only two provinces and territories have significant errors of closure. Review of the error of closure reveals that an emigration estimate that may have been too low would have significantly contributed to the positive error for British Columbia. To this may be added a slightly high net interprovincial migration, probably due to an underestimate of people leaving for Alberta. A slightly low estimate of net interprovincial migration would have contributed to the negative error of closure for Alberta. This province received a considerable number of migrants in the year preceding the census, many of whom were temporary migrants. As was the case with errors for the territories, the greater difficulty in estimating this migration is almost certainly responsible for the larger error.

11. Historical estimates of population coverage error

11.1 Estimates

This section presents historical estimates of population coverage error. Chart 11.1 presents the estimated population undercoverage rate $\hat{R}_U$ for the 1971 Census to the 2006 Census, and the estimated population overcoverage rate $\hat{R}_O$ and the estimated population net undercoverage rate $\hat{R}_N$ for the 1991 Census to the 2006 Census. The series for overcoverage and net undercoverage begin in 1991 because the overcoverage was first estimated for the 1991 Census following an experimental study done for the 1986 Census.

Chart 11.1 Estimated rates of population coverage error for Canada, 1971 Census to 2006 Census



	1971	1976	1981	1986	1991	1996	2001	2006
Undercoverage (%)	1.93	2.04	2.01	3.21	3.43	3.18	3.95	4.26
Overcoverage (%)					0.56	0.74	0.96	1.59
Net Undercoverage (%)					2.87	2.44	2.99	2.67

Note: Blank cells indicate data not available.
Sources: Statistics Canada, 1971 to 2006 Census Coverage Studies.

Population coverage error is a growing data quality concern; undercoverage has doubled since 1981 and overcoverage has doubled since 1996. Changes in net undercoverage from one census to the next reflect changes in undercoverage and/or overcoverage which, in turn reflect changes in population demographics, changes in the living arrangements of Canadians, changes in census methodology, and changes in the methodology of the coverage studies. The last is discussed in Section 11.2.

As seen in Chart 11.1, there is an increasing trend in both the population undercoverage rate and the overcoverage rate. First measured at 1.93% for the 1971 Census, the rates of undercoverage were similar for 1976 and 1981 at 2.04% and 2.01% respectively. Undercoverage jumped to 3.21% for the 1986 Census, increased to 3.43% for the 1991 Census, and then decreased by about the same amount to 3.18% for the 1996 Census. The rate of undercoverage increased notably to 3.95% for 2001 and then increased again to 4.26% for the 2006 Census.

The overcoverage rate increased from 0.74% for the 1996 Census to 0.96% for the 2001 Census. The increase from 1991 to 1996 is due to a change in the methodology of the coverage studies. The increase in overcoverage from 0.96% for 2001 to 1.59% for the 2006 Census is the largest increase in the series. From 2001 to 2006, the most significant increases were for the province of Newfoundland and Labrador and for the Northwest Territories (i.e., 1.00% and 0.98% respectively). We also noted very significant increases for the Yukon Territory and Nunavut (i.e., 0.76% and 0.85% respectively).

Although net undercoverage diminished slightly from 2001 to 2006, both undercoverage and overcoverage increased. Coverage error reflects error on the part of the respondent such as when the rules on whom to include are applied incorrectly, and on the part of census operations

such as when new dwellings recently under construction are erroneously excluded. As for many surveys, reduced respondent participation continues to be an issue for the census. This is evidenced by both increased non-response and increased undercoverage.

It should also be noted that the methodology of the 2006 Census included a number of changes and, therefore, there is the potential for changes in undercoverage and overcoverage. Even though there are high quality standards governing all census operations, these changes may have resulted in changes in population coverage error. The 2006 Census of Population moved from a decentralized, manual operation to a more centralized and automated one:

- Questionnaires were mailed by Canada Post Corporation in a majority of urban areas.
- The Address Register, which is updated by listing operations, provided the mailing addresses.
- Follow-up became centralized.

Further, in some regions, it was difficult to recruit enough staff.

Looking back at undercoverage since the 1981 Census, the increase in undercoverage observed in the 1986 Census led to introducing the Address Register (AR) for the 1991 Census. The AR provided a separate list of those urban dwellings which should have been enumerated. For the 1996 Census, the introduction of enumeration by an enumerator (EN) rather than self-enumeration in some large city inner-city enumeration areas (EAs) reduced undercoverage. Also, moving Census Day from early June to mid-May helped to control undercoverage because people were more likely to be at home and less likely to be moving.

Table 11.1 and Table 11.2 present estimates of undercoverage. Note that 1971 is not included in Table 11.2 because estimates were produced for different age groups for those above age 24.

Table 11.1 Estimated population undercoverage rates and standard errors for Canada, provinces and territories, 1971 Census to 2006 Census[1]

Provinces and territories	1971		1976		1981		1986		1991		1996		2001		2006	
	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error
	%															
Canada	**1.93**	**0.09**	**2.04**	**0.10**	**2.01**	**0.09**	**3.21**	**0.13**	**3.43**	**0.12**	**3.18**	**0.09**	**3.95**	**0.13**	**4.26**	**0.17**
Newfoundland and Labrador	2.25	0.72	1.10	0.39	1.74	0.45	1.92	0.33	2.47	0.30	2.45	0.29	2.43	0.32	2.62	0.54
Prince Edward Island	1.23	1.13	0.38	0.25	1.17	0.54	2.14	0.80	1.67	0.23	1.76	0.28	1.89	0.53	3.04	0.52
Nova Scotia	1.33	0.45	0.86	0.34	1.05	0.34	2.15	0.34	2.25	0.36	2.70	0.27	3.44	0.41	4.02	0.54
New Brunswick	1.65	0.56	2.16	0.37	1.81	0.30	2.71	0.33	3.71	0.42	2.49	0.28	3.57	0.42	3.56	0.43
Quebec	2.10	0.19	2.95	0.25	1.91	0.21	2.91	0.31	3.18	0.20	2.46	0.18	2.93	0.26	2.46	0.32
Ontario	1.68	0.12	1.52	0.17	1.94	0.14	3.43	0.19	4.23	0.28	3.40	0.18	4.56	0.25	5.18	0.34
Manitoba	1.13	0.38	1.07	0.33	0.98	0.35	2.94	0.40	2.31	0.36	2.55	0.29	3.49	0.43	4.32	0.57
Saskatchewan	1.00	0.37	1.33	0.34	0.99	0.37	2.38	0.37	2.15	0.32	3.30	0.32	3.18	0.37	3.81	0.50
Alberta	2.55	0.44	1.49	0.26	2.54	0.36	3.00	0.32	2.51	0.27	2.99	0.24	3.18	0.33	4.74	0.49
British Columbia	2.89	0.39	3.13	0.31	3.16	0.33	4.48	0.36	3.42	0.24	4.58	0.24	5.30	0.34	4.83	0.41
Yukon Territory	..	..	..	..	..	..	..	..	4.12	0.58	3.92	0.51	5.59	1.16	7.23	0.64
Northwest Territories	..	..	..	..	..	..	..	..	5.73	0.57	4.28	0.67	9.10	0.80	5.74	0.57
Nunavut	…	…	…	…	…	…	…	…	…	…	6.54	0.63	5.07	1.39	5.55	0.60

.. not available for this reference period
... not applicable

Note: 1. Excludes incompletely enumerated Indian reserves. Includes non-permanent residents and territories in 1991, 1996, 2001 and 2006. Includes revisions to 1986 original publication. Excludes estimates of persons missed in dwellings incorrectly classified as unoccupied in 1971 and 1976.

Sources: Statistics Canada, 1971 to 2006 Census Coverage Studies.

Table 11.2 Estimated population undercoverage rates and standard errors, sex and age group for Canada, 1976 Census to 2006 Census[1]

Sex and age group	1976		1981		1986		1991		1996		2001		2006	
	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error
	%													
Both sexes	**2.04**	**0.10**	**2.01**	**0.09**	**3.21**	**0.13**	**3.43**	**0.12**	**3.18**	**0.09**	**3.95**	**0.13**	**4.26**	**0.17**
0 to 4 years	2.31	0.28	1.21	0.22	2.14	0.49	3.55	0.49	2.89	0.36	4.42	0.71	4.07	0.65
5 to 14 years	1.20	0.16	1.23	0.21	2.08	0.26	2.49	0.27	1.45	0.14	2.90	0.38	3.10	0.46
15 to 17 years[2]	1.99	0.38	2.96	0.52	3.58	0.60	3.75	0.42	3.48	0.42	4.36	0.53	1.56	0.60
18 to 19 years	..	..	..	..	..	..	..	..	..	..	..	..	8.86	1.58
20 to 24 years	5.31	0.38	5.51	0.29	8.66	0.46	8.18	0.52	8.00	0.34	9.85	0.62	10.50	0.74
25 to 34 years	2.85	0.28	2.31	0.28	4.51	0.35	5.65	0.35	5.81	0.29	8.07	0.36	9.43	0.56
35 to 44 years	1.54	0.26	2.20	0.26	2.32	0.31	2.84	0.29	2.78	0.24	4.04	0.33	5.36	0.50
45 to 54 years	1.22	0.33	0.81	0.23	1.58	0.29	1.61	0.27	1.90	0.21	1.79	0.29	2.64	0.43
55 to 64 years	0.92	0.20	0.91	0.29	2.06	0.31	1.69	0.28	2.23	0.34	1.22	0.37	0.95	0.53
65 years and over	1.20	0.25	0.71	0.30	1.76	0.31	1.51	0.28	1.52	0.26	1.29	0.34	0.21	0.40
Males	**2.46**	**0.17**	**2.37**	**0.13**	**3.75**	**0.16**	**3.95**	**0.16**	**3.89**	**0.14**	**4.90**	**0.19**	**5.51**	**0.26**
0 to 4 years	2.53	0.46	1.32	0.33	2.22	0.67	2.79	0.58	2.56	0.47	3.36	0.89	4.24	0.95
5 to 14 years	1.14	0.21	1.27	0.29	1.98	0.32	2.32	0.34	1.46	0.24	2.38	0.49	3.04	0.64
15 to 17 years[2]	1.93	0.48	3.12	0.68	4.09	0.74	3.55	0.60	3.68	0.43	5.49	0.80	1.88	0.88
18 to 19 years	..	..	..	..	..	..	..	..	..	..	..	..	10.06	2.45
20 to 24 years	5.99	0.52	6.03	0.48	10.36	0.57	8.98	0.81	9.48	0.50	11.68	0.92	12.21	1.12
25 to 34 years	3.64	0.46	2.70	0.44	5.43	0.45	7.28	0.56	7.74	0.42	10.67	0.55	11.42	0.86
35 to 44 years	2.33	0.48	3.42	0.40	3.29	0.51	3.65	0.41	3.94	0.39	5.71	0.51	7.77	0.79
45 to 54 years	1.63	0.41	1.21	0.38	1.95	0.52	2.05	0.45	2.12	0.27	2.50	0.44	4.14	0.69
55 to 64 years	1.28	0.34	0.91	0.40	1.88	0.47	2.04	0.44	2.50	0.54	1.35	0.54	2.13	0.77
65 years and over	1.90	0.44	0.69	0.47	1.57	0.50	1.41	0.50	1.64	0.45	1.50	0.53	-0.05	0.56
Females	**1.61**	**0.10**	**1.65**	**0.12**	**2.68**	**0.17**	**2.93**	**0.17**	**2.49**	**0.12**	**3.02**	**0.18**	**3.04**	**0.23**
0 to 4 years	2.07	0.36	1.10	0.33	2.06	0.62	4.35	0.71	3.24	0.55	5.50	1.14	3.88	0.92
5 to 14 years	1.26	0.27	1.19	0.31	2.20	0.33	2.65	0.39	1.45	0.22	3.44	0.58	3.17	0.66
15 to 17 years[2]	2.05	0.51	2.80	0.73	3.05	0.76	3.96	0.54	3.28	0.55	3.13	0.69	1.23	0.83
18 to 19 years	..	..	..	..	..	..	..	..	..	..	..	..	7.58	1.96
20 to 24 years	4.62	0.48	4.98	0.43	6.89	0.72	7.36	0.71	6.45	0.48	7.91	0.84	8.70	0.98
25 to 34 years	2.03	0.38	1.92	0.32	3.59	0.45	3.98	0.37	3.84	0.40	5.41	0.46	7.43	0.73
35 to 44 years	0.72	0.24	0.93	0.31	1.33	0.32	2.01	0.35	1.62	0.28	2.35	0.43	2.90	0.61
45 to 54 years	0.81	0.38	0.41	0.26	1.20	0.35	1.16	0.34	1.68	0.33	1.09	0.37	1.13	0.51
55 to 64 years	0.58	0.25	0.92	0.34	2.23	0.50	1.35	0.33	1.97	0.40	1.09	0.52	-0.22	0.73
65 years and over	0.64	0.38	0.71	0.42	1.89	0.44	1.58	0.36	1.43	0.32	1.13	0.45	0.40	0.56

.. not available for this reference period

Notes: 1. Excludes incompletely enumerated Indian reserves. Includes non-permanent residents and territories in 1991, 1996, 2001 and 2006. Includes revisions to 1986 original publication. Excludes estimates of persons missed in dwellings incorrectly classified as unoccupied in 1976.

2. Data for all years except 2006 is for persons aged 15-19.

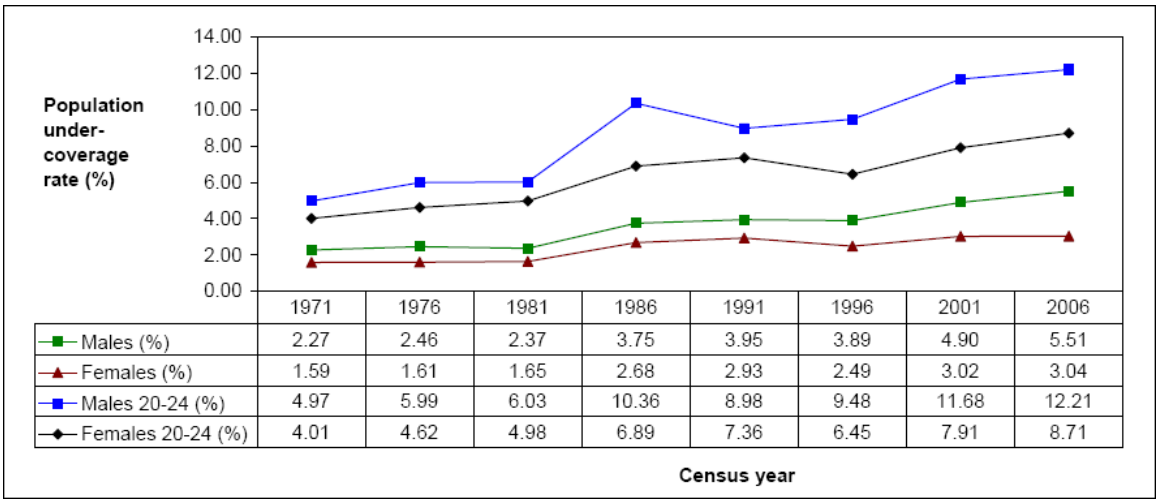
Sources: Statistics Canada, 1976 to 2006 Census Coverage Studies.

These tables reveal that:

Undercoverage is usually higher in British Columbia. Among the provinces, British Columbia had the highest rate of undercoverage in every census from 1971 to 2001 except for 1991 and 2006. Ontario had the highest rate in 1991 at 4.23% and in 2006 at 5.18%. Undercoverage rates for Quebec, the Atlantic provinces and the Prairie provinces tend to be lower than the national rate.

Undercoverage is higher for young adults and higher for males. There are two persistent demographic trends. First, undercoverage for males is higher than undercoverage for females. Second, undercoverage is highest for young adults regardless of sex. As seen in Chart 11.2, undercoverage for males is higher than undercoverage for females for every census year since 1971, increasing from 2.27% to 5.51% for males and from 1.59% to 3.04% for females. Chart 11.2 also shows that undercoverage for young males aged 20 to 24 is higher than undercoverage for all males. This is also the case for females, but the young adult female rates are lower than the young adult male rates. The 2006 Census marked the highest rates of undercoverage for both young adult males and young adult females, 12.21% for males 20 to 24 and 8.71% for females 20 to 24. Higher undercoverage for young adults is due in part to less stable living arrangements. Young adults are more likely than older adults or children to change their living arrangements because they are, for example, moving away from home to attend a post-secondary institution or moving in with friends or spouses.

Chart 11.2 Estimated rates of population undercoverage, sex and age group for Canada, 1971 Census to 2006 Census



	1971	1976	1981	1986	1991	1996	2001	2006
Males (%)	2.27	2.46	2.37	3.75	3.95	3.89	4.90	5.51
Females (%)	1.59	1.61	1.65	2.68	2.93	2.49	3.02	3.04
Males 20-24 (%)	4.97	5.99	6.03	10.36	8.98	9.48	11.68	12.21
Females 20-24 (%)	4.01	4.62	4.98	6.89	7.36	6.45	7.91	8.71

Sources: Statistics Canada, 1971 – 2006 Census Coverage Studies.

Table 11.3 and Table 11.4 present estimates of overcoverage.

Table 11.3 Estimated population overcoverage rate and standard errors for Canada, provinces and territories, 1991 Census to 2006 Census[1]

Provinces and territories	1991		1996		2001		2006	
	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error
	%							
Canada	**0.56**	**0.04**	**0.74**	**0.04**	**0.96**	**0.05**	**1.59**	**0.01**
Newfoundland and Labrador	0.48	0.09	0.77	0.12	0.63	0.10	1.63	0.05
Prince Edward Island	0.74	0.15	0.91	0.14	0.92	0.18	1.66	0.06
Nova Scotia	0.36	0.09	0.47	0.07	0.81	0.14	1.40	0.03
New Brunswick	0.46	0.09	0.60	0.10	0.89	0.19	1.41	0.03
Quebec	0.51	0.07	0.85	0.08	1.03	0.10	1.66	0.02
Ontario	0.59	0.07	0.67	0.07	0.88	0.09	1.49	0.02
Manitoba	0.45	0.11	0.88	0.15	0.80	0.15	1.42	0.04
Saskatchewan	0.35	0.08	0.55	0.11	1.06	0.20	1.53	0.04
Alberta	0.51	0.09	0.59	0.10	0.89	0.13	1.47	0.02
British Columbia	0.68	0.10	0.89	0.09	1.26	0.12	1.96	0.03
Yukon Territory	0.29	0.07	0.70	0.17	0.86	0.16	1.62	0.08
Northwest Territories	0.29	0.07	1.32	0.22	1.00	0.11	1.98	0.08
Nunavut	…	…	0.99	0.22	0.59	0.10	1.44	0.07

… not applicable
Note: 1. Excludes incompletely enumerated Indian reserves. Includes non-permanent residents.
Sources: Statistics Canada, 1991 to 2006 Census Coverage Studies.

Table 11.4 Estimated population overcoverage rate and standard errors, sex and age group for Canada, 1996 Census to 2006 Census[1, 2]

Sex and age groups	1996		2001		2006	
	Estimated rate	Standard error	Estimated rate	Standard error	Estimated rate	Standard error
	%					
Both sexes	**0.74**	**0.04**	**0.96**	**0.05**	**1.59**	**0.01**
0 to 4 years	0.61	0.10	0.96	0.18	1.35	0.07
5 to 14 years	0.96	0.09	1.52	0.15	2.24	0.07
15 to 17 years[3]	1.24	0.15	1.85	0.26	2.33	0.14
18 to 19 years	..	..	..	..	2.65	0.17
20 to 24 years	2.44	0.28	2.66	0.32	2.88	0.11
25 to 34 years	0.66	0.08	0.92	0.09	1.43	0.06
35 to 44 years	0.38	0.06	0.49	0.06	1.05	0.05
45 to 54 years	0.48	0.11	0.39	0.04	1.13	0.05
55 to 64 years	0.52	0.11	0.38	0.05	1.24	0.06
65 years and over	0.36	0.07	0.77	0.21	1.60	0.06
Males	**0.70**	**0.04**	**0.92**	**0.06**	**1.62**	**0.02**
0 to 4 years	0.52	0.09	0.69	0.07	1.35	0.09
5 to 14 years	0.99	0.15	1.59	0.21	2.25	0.10
15 to 17 years[3]	1.12	0.24	1.45	0.31	2.37	0.20
18 to 19 years	..	..	..	..	2.28	0.21
20 to 24 years	2.34	0.34	2.44	0.45	2.75	0.15
25 to 34 years	0.65	0.11	1.03	0.14	1.51	0.08
35 to 44 years	0.38	0.06	0.46	0.06	1.10	0.06
45 to 54 years	0.35	0.07	0.34	0.03	1.16	0.07
55 to 64 years	0.37	0.12	0.33	0.04	1.30	0.09
65 years and over	0.33	0.02	0.74	0.21	1.69	0.10
Females	**0.77**	**0.06**	**1.00**	**0.08**	**1.56**	**0.01**
0 to 4 years	0.69	0.18	1.25	0.36	1.35	0.10
5 to 14 years	0.92	0.14	1.44	0.21	2.23	0.10
15 to 17 years[3]	1.36	0.29	2.27	0.43	2.28	0.19
18 to 19 years	..	..	..	..	3.04	0.28
20 to 24 years	2.55	0.46	2.89	0.46	3.01	0.17
25 to 34 years	0.66	0.11	0.81	0.12	1.35	0.08
35 to 44 years	0.37	0.10	0.53	0.11	0.99	0.06
45 to 54 years	0.61	0.20	0.43	0.07	1.11	0.06
55 to 64 years	0.66	0.19	0.42	0.09	1.18	0.07
65 years and over	0.38	0.11	0.80	0.33	1.53	0.08

.. not available for this reference period

Notes: 1. Estimates by sex and age groups are not available for the 1991 Census.
2. Excludes incompletely enumerated Indian reserves.
3. Data for all years except 2006 is for persons aged 15 to 19.

Sources: Statistics Canada, 1996 to 2006 Census Coverage Studies.

These tables reveal that:

Overcoverage is consistently higher for British Columbia than for the other provinces. Among the provinces, the rate of population overcoverage is highest for British Columbia. This has also been the case for the past three censuses.

Overcoverage is more common for school-aged children and young adults. There is a trend of higher overcoverage for children aged 5 to 17 and for young adults aged 18 to 24. For school-aged children, it is largely due to children whose parents do not live in the same household who are often enumerated with each parent. Overcoverage for young adults likely reflects the same less stable living arrangements that can also lead to undercoverage.

Note that in Table 11.4 the age group 15 to 17 contains data for those 15 to 19 for the 1996 Census and the 2001 Census. The 2006 estimates revealed that persons aged 18 and 19 behaved more like young adults than like children in their propensity for undercoverage and overcoverage.

Provincially and nationally, differences in the total overcoverage estimate for 2001 and 2006 were all statistically significant at the 95% level, except for Saskatchewan. In comparing 2006 provincial and territorial estimates of overcoverage, there are some small methodological differences to be aware of. We should remember that overcoverage for Saskatchewan was high in 2001, due to the important impact of a Reverse Record Check (RRC) observation with a higher weight. We should also note that for Nunavut in 2001, only the Automated Match Study (AMS) and the Collective Dwelling Study (CDS) contributed to the overcoverage. We therefore expected to see an increase in the number of overcounted persons in Nunavut. In addition, the AMS covered four distinct regions in 2001: Eastern Canada (Newfoundland and Labrador, Prince Edward Island, Nova Scotia and New Brunswick); Quebec; Ontario; and Western and Northern Canada (Manitoba, Saskatchewan, Alberta, British Columbia, the Yukon Territory, the Northwest Territories and Nunavut). Accordingly, no overcoverage was measured between Nunavut (or any other territory) and the Eastern provinces, Quebec or Ontario. This may also explain the significant increases in Newfoundland and Labrador, the Northwest Territories and the Yukon Territory.

11.2 Coverage studies design changes

Differences in the design of the coverage studies over time mean that the rates in Table 11.1, Table 11.2, Table 11.3, and Table 11.4 are not strictly comparable. A list of methodological changes since the 1976 coverage studies is given below. It is remarkable that the fundamentals of the Reverse Record Check (RRC) approach for measuring undercoverage has not changed in any substantive manner since it was first carried out for the 1966 Census of Population. A sample is taken from a frame representing the census target population that is independent of the census. Census records are then checked ('Reverse Record Check') to determine if the sampled persons were indeed enumerated. There have been more changes in the measurement of overcoverage. Multiple studies were carried out for 1991, 1996 and 2001. In 1996, the RRC was expanded to include the measurement of overcoverage. For 2006 there was a new study to measure all overcoverage that exploited exact and probabilistic matching involving names, age and sex.

2006:

Both the RRC and the Census Overcoverage Study (COS) made optimal use of the name field added to the 2006 Census Response Database (RDB) in their matching and searching operations. Further:

(a) The measurement of overcoverage was restricted to the COS. The methodology of the RRC was subsequently changed so that not all cases were sent for field collection. New for the

2006 RRC, a processing step was carried out prior to collection in order to determine whether or not collection was required. A search of the RRC version of the 2006 Census response database (RRC RDB) for the persons selected in the sample using data from the sampling frame and the various update sources such as tax data was done. If the search resulted in locating the sampled person on the RRC RDB, collection was not required. The exception was a sample of those that had been found in order to collect data required for the non-response adjustment.

(b) The three studies used for the 2001 coverage studies to measure overcoverage were replaced by the 2006 COS. The study used a methodology that was different from any previous overcoverage study. Essentially, the COS exploited the use of name matching to identify overcoverage.

As for 1996 and 2001, the 2006 RRC did not estimate the number of persons missed in incompletely enumerated Indian reserves and Indian settlements. You can find more information on this topic in Section 12.2.

2001:

(a) The institutional component of the Collective Dwelling Study (CDS) was dropped and overcoverage estimates in this population were produced by the RRC.

(b) The Dwelling Classification Study (DCS) replaced the Vacancy Check (VC) which was used in previous censuses to re-examine dwellings classified as unoccupied by the enumerator. The DCS is an extension of the VC in order to estimate the number of persons living in non-response dwellings.

1996:

(a) The 1996 RRC did not estimate the persons missed on incompletely enumerated Indian reserves.

(b) The Temporary Residents Study was cancelled because of concerns about the quality of the data, and because it was recognized that the RRC would measure most of this type of undercoverage with sufficient quality.

(c) Compared to 1991, a more comprehensive measure of overcoverage was produced due to integrating the Private Dwelling Study into the RRC so that each sampled person could be identified as having been enumerated more than once. This approach resulted in an increase of addresses to be processed where overcoverage could have occurred. Second, the Automated Match Study (AMS) was substantially expanded from the 1991 approach of measuring overcoverage within an enumeration area (EA) to measuring overcoverage within a large region (Atlantic, Quebec, Ontario, rest of Canada).

1991:

(a) Non-permanent residents were included in the target population for the first time.

(b) Following experimental studies in 1986, the measurement of population overcoverage commenced with the 1991 coverage studies. The results of three studies were combined to form a comprehensive estimate: the Private Dwelling Study (PDS), the Collective Dwelling Study (CDS), and the Automated Match Study (AMS).

1986: The rates in Table 11.1 for the 1986 Census differ from the results published in the *User's Guide to the Quality of 1986 Census Data: Coverage* as they include revisions made after the 1986 publication when incompletely enumerated Indian reserves were included as missed. In the

original 1986 publication, they were included as 'enumerated' since published provincial census counts included an estimate of persons missed on such reserves.

1976: Census counts did not include estimates from the vacancy check (VC) of persons missed in dwellings incorrectly classified as unoccupied. The 1976 population undercoverage rate would have been 1.78% had it included the results of the 1976 VC. There was no VC in the 1971 Census.

12. Special topics

12.1 Persons not enumerated

This section introduces the concept of census population collection undercoverage. Although this new measure does not specify the error in census data from coverage error, it provides a picture of how well the census was able to enumerate its target population. Also, collection undercoverage removes the issue of bias between the true number of persons living in occupied non-response dwellings, the estimate of this number resulting from imputation, and the RRC estimate.

It is instructive to expand the concept of undercoverage to include persons not enumerated for any reason. Undercoverage is defined to be the number of persons not included in the census count. As discussed in Section 3.3, the census count C is composed of two elements: $C = E + I$ where E = the number of enumerations, and I = the number of imputed persons. Undercoverage, therefore, is a subset of all persons who were not listed on a census form but should have been. It does not include those who were not enumerated either because no completed census form was returned for the dwelling (non-response dwelling) or the dwelling was not subject to non-response follow-up because they were erroneously classified as unoccupied (misclassified occupied dwelling).

Also from Section 3.3, an estimate of the true number of persons in the census target population T is given by

$$\hat{T} = C + \hat{N} = C + \hat{U} - \hat{O}$$

Combining these two equations:

$$\hat{T} = C + \hat{N} = C + \hat{U} - \hat{O} = E + \left(I + \hat{U}\right) - \hat{O}$$

This formulation of $\hat{T}$ has three components:

1. E = the number of persons who were listed on a census form[9];
2. $\hat{O}$ = an estimate of the number of excess enumerations[10]; and
3. $\left(I + \hat{U}\right)$ = an estimate of the number of persons who were not listed on a census form who should have been.

9. It is possible that some of the persons listed on the form may not appear in the final census database. So, in the strictest sense, 'persons listed on the form' is used in this section to represent persons in the final census database.

10. Most cases of overcoverage involve duplicate enumerations where the same person appears twice on the database. In a small number of cases, however, the same person appears more than twice. $\hat{O}$ estimates the number of excess enumerations rather than the number of persons involved in multiple enumerations.

The last component, $\left(I + \hat{U}\right)$, estimates the number of persons missed in the census for any reason. Let us define census population collection undercoverage, denoted by L, as persons not enumerated for any reason. Then, the estimate of census population collection undercoverage is $\hat{L} = \left(I + \hat{U}\right)$ and the corresponding estimate of the rate of census population collection undercoverage rate is:

$$\hat{R}_L = 100 * \frac{\hat{L}}{\hat{T}} = 100 * \left(\frac{I + \hat{U}}{C + \hat{N}}\right)$$

Net census collection undercoverage can be defined by subtracting overcoverage $\hat{O}$ from $\hat{L}$.

Then:

$$\hat{T} = C + \hat{N} = E + \left(I + \hat{U}\right) - \hat{O} = E + \hat{L} - \hat{O}$$

Although net collection undercoverage cannot be applied to census counts to adjust for coverage error, $\hat{L}$ and $\hat{R}_L$ provide a broader picture of how well the census was able to enumerate its target population. Also, collection undercoverage removes the issue of bias between the true number of persons living in occupied non-response dwellings, the estimate of this number resulting from imputation, and the RRC estimate.

Table 12.1 gives the estimated 2006 Census population collection undercoverage $\hat{L}$ and $\hat{R}_L$, and population undercoverage $\hat{U}$ and $\hat{R}_U$ (also in Table 1.3) and their estimated standard errors for provinces and territories and for national age and sex groups. At the national level, the 2006 Census achieved an enumeration for 92.87% of its target population ($100 - \hat{R}_L$). This compares to 95.74% of the target population included in the 2006 Census count of 31,612,897 persons ($100 - \hat{R}_U$). The difference between these two rates is simply the inclusion of the imputations in $\hat{R}_L$. The 2001 Census achieved an enumeration for 94.17% of its target population compared to 97.57% of the target population included in the 2001 Census count of 30,007,094 persons. The decrease in these two rates from 2001 to 2006 is from an increase in undercoverage and an increase in imputations.

Table 12.1 Estimated population collection undercoverage, population undercoverage and standard errors for various characteristics, 2001 and 2006 censuses

Characteristics	Population collection undercoverage				Population undercoverage			
	Estimated number	Standard error	Estimated rate (%)	Standard error (%)	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	**2,317,548**	**53,831**	**7.13**	**0.18**	**1,384,372**	**53,831**	**4.26**	**0.17**
Provinces and territories								
Newfoundland and Labrador	24,602	2,698	4.82	0.55	13,355	2,698	2.62	0.54
Prince Edward Island	7,043	697	5.11	0.53	4,185	697	3.04	0.52
Nova Scotia	65,177	4,875	6.95	0.56	37,711	4,875	4.02	0.54
New Brunswick	41,392	3,099	5.55	0.44	26,543	3,099	3.56	0.43
Quebec	439,516	24,014	5.78	0.33	187,047	24,014	2.46	0.32
Ontario	949,518	41,310	7.52	0.35	654,118	41,310	5.18	0.34
Manitoba	76,380	6,453	6.46	0.58	51,113	6,453	4.32	0.57
Saskatchewan	61,630	4,791	6.22	0.51	37,734	4,791	3.81	0.50
Alberta	260,063	16,072	7.65	0.51	161,337	16,072	4.74	0.49
British Columbia	383,648	16,539	9.06	0.43	204,722	16,539	4.83	0.41
Yukon Territory	3,058	193	9.51	0.66	2,325	193	7.23	0.64
Northwest Territories	3,429	233	7.96	0.58	2,475	233	5.74	0.57
Nunavut	2,091	174	6.80	0.61	1,706	174	5.55	0.60
Sex and age group								
Both sexes, all ages	**2,317,548**	**53,831**	**7.13**	**0.18**	**1,384,372**	**53,831**	**4.26**	**0.17**
0 to 4 years	117,038	10,902	6.74	0.67	70,670	10,902	4.07	0.65
5 to 14 years	219,442	17,434	5.59	0.47	121,698	17,434	3.10	0.46
15 to 17 years	53,053	7,751	4.07	0.62	20,368	7,751	1.56	0.60
18 to 19 years	100,574	12,797	11.40	1.62	78,170	12,797	8.86	1.58
20 to 24 years	301,515	15,156	13.39	0.76	236,589	15,156	10.50	0.74
25 to 34 years	540,183	22,405	12.41	0.58	410,458	22,405	9.43	0.56
35 to 44 years	410,186	23,689	8.15	0.51	269,695	23,689	5.36	0.50
45 to 54 years	278,103	21,014	5.50	0.44	133,175	21,014	2.64	0.43
55 to 64 years	149,382	19,220	4.08	0.55	34,708	19,220	0.95	0.53
65 years and over	148,071	16,903	3.46	0.41	8,840	16,903	0.21	0.40
Males, all ages	**1,345,021**	**40,339**	**8.35**	**0.27**	**887,171**	**40,339**	**5.51**	**0.26**
0 to 4 years	61,504	8,084	6.91	0.97	37,760	8,084	4.24	0.95
5 to 14 years	110,739	12,427	5.51	0.65	61,012	12,427	3.04	0.64
15 to 17 years	29,323	5,770	4.37	0.90	12,583	5,770	1.88	0.88
18 to 19 years	57,385	10,186	12.55	2.51	45,985	10,186	10.06	2.45
20 to 24 years	173,793	11,572	15.01	1.15	141,357	11,572	12.21	1.12
25 to 34 years	313,973	16,880	14.41	0.89	248,935	16,880	11.42	0.86
35 to 44 years	268,892	18,535	10.59	0.81	197,141	18,535	7.77	0.79
45 to 54 years	176,622	16,732	7.00	0.71	104,490	16,732	4.14	0.69
55 to 64 years	94,345	13,672	5.18	0.79	38,835	13,672	2.13	0.77
65 years and over	58,445	10,338	3.15	0.57	-927	10,338	-0.05	0.56
Females, all ages	**972,526**	**36,300**	**5.94**	**0.23**	**497,200**	**36,300**	**3.04**	**0.23**
0 to 4 years	55,535	7,533	6.55	0.95	32,911	7,533	3.88	0.92
5 to 14 years	108,703	12,246	5.68	0.68	60,686	12,246	3.17	0.66
15 to 17 years	23,731	5,177	3.75	0.85	7,786	5,177	1.23	0.83
18 to 19 years	43,189	7,750	10.17	2.01	32,185	7,750	7.58	1.96
20 to 24 years	127,722	9,837	11.67	1.00	95,232	9,837	8.70	0.98
25 to 34 years	226,209	14,845	10.40	0.75	161,522	14,845	7.43	0.73
35 to 44 years	141,294	14,846	5.66	0.63	72,554	14,846	2.90	0.61
45 to 54 years	101,481	12,728	4.01	0.52	28,685	12,728	1.13	0.51
55 to 64 years	55,037	13,514	2.99	0.76	-4,127	13,514	-0.22	0.73
65 years and over	89,626	13,384	3.70	0.57	9,767	13,384	0.40	0.56

Sources: Statistics Canada, 2006 Census, 2006 Reverse Record Check and 2006 Census Overcoverage Study.

12.2 Aboriginal peoples

12.2.1 Introduction

Users should also be aware of the extent to which Indian reserves and Indian settlements participated in the 2006 Census. In some cases enumeration was not permitted or was interrupted before it could be completed. In other cases the quality of the enumeration was considered inadequate. These geographic areas, a total of 22, are called incompletely enumerated Indian reserves and Indian settlements. Data for 2006 are therefore not available for the incompletely enumerated reserves and settlements, and are not included in tabulations. Similar problems have occurred in previous censuses. In the 2001 Census there were 30 Indian reserves and Indian settlements that were declared incompletely enumerated. Among these, 14 became participating reserves in the 2006 Census.

This section presents estimates of 2006 Census net population undercoverage for **participating reserves**. For **incompletely enumerated Indian reserves and Indian settlements,** model-based estimates are presented. Since no reliable source exists to verify the assumptions used in the models, these estimates must be used with caution.

12.2.2 Participating reserves

The following table gives estimates of 2006 Census net undercoverage for all persons living on participating reserves including those without Aboriginal identity for Canada, for the eastern region: Newfoundland and Labrador, Prince Edward Island, Nova Scotia, New Brunswick, Quebec and Ontario, and for the western and northern region: Manitoba, Saskatchewan, Alberta, British Columbia, the Yukon Territory and the Northwest Territories: Census population net undercoverage 'on reserve,' for Canada.

Limitations of the coverage studies do not permit the production of estimates by Aboriginal identity. The rate of census net undercoverage indicates what proportion of the entire population that should have been enumerated is, on a net basis, not included in 2006 Census tabulations. Users are advised to consult the standard error of an estimate to determine its suitability for use.

The estimate of net undercoverage is the estimate of population undercoverage less the estimate of population overcoverage. One limitation of the estimate of overcoverage is that for a particular geography such as participating reserves, the estimate includes persons who appear on questionnaires for two dwellings where at least one of the dwellings is on reserve. The other dwelling may be on the same reserve, on a different reserve, or not on a reserve. Since the COS does not determine at which dwelling an individual should have been listed at, the assumption is made that it is equally likely that the individual should have been listed at the first dwelling as at the second dwelling. Therefore, in order to produce estimates of overcoverage, half of the weight for the person is assigned to each dwelling. This concept is important for small domains such as the 'on reserve' population. About half of the overcoverage cases involving a dwelling on reserve also involved a dwelling off reserve.

12.2.3 Incompletely enumerated Indian reserves and settlements

Neither the 2006 Census nor the Reverse Record Check is in a position to produce an estimate of the population living in the 22 incompletely enumerated Indian reserves and settlements. In order to produce official estimates of population, a model-based methodology was used to prepare estimates of population for these geographical areas. **The resulting estimates should be used with caution as they are based entirely on a model whose assumptions cannot be verified. The validity of these estimates depends on the extent to which the model assumptions capture the true underlying situation.**

The national model results can be found at: Model estimates for incompletely enumerated Indian reserves and settlements for Canada.

In the 2001 Census, 30 reserves, with approximately 34,500 persons, were classified as 'incompletely enumerated.' Among the 22 reserves and settlements considered as incompletely enumerated in the 2006 Census, six were considered to have had complete enumerations in the 2001 Census while the other 16 were 'incompletely enumerated' or 'refusal.' The population for the 22 incompletely enumerated reserves and settlement was estimated at approximately 40,000, an increase from 2001.

The estimation model is as follows. A two step model was developed to estimate the population. The first step uses a simple linear regression to predict the Census count in 2006. The linear regression was constructed using all Indian reserves that were completely enumerated in both the 2001 and the 2006 Census. The model assumes a linear growth from 2001 to 2006 for all provinces with separate estimates, for the intercept and the regression parameters for each province. The model was evaluated for the basic regression assumptions of independence of errors, homogeneity of variances and normality of errors.

For each incompletely enumerated reserve, the input variable for the regression model was either the actual census count in 2001 or the best predicted census count from the 2001 model. The output of the model was the estimated census count in 2006.

The second step is done to produce consistency with the results of the census coverage studies. An adjustment was made to the estimated 'census' count to account for net undercoverage of all subjected census counts. Net undercoverage for the incompletely enumerated reserves was estimated by calculating the net undercoverage rate for all completely enumerated reserves in each province and then applying that rate to the estimated 'census' count of all the incompletely enumerated Indian reserves in the province. The estimated 'census' count and the 'estimated net missed persons' in each reserve were then summed to create an 'estimated' population for the incompletely enumerated Indian reserves.

For provincial estimates please refer to: Incompletely enumerated Indian reserves and Indian settlements.

Appendix A. Reverse Record Check Survey questionnaires

Non-proxy questionnaire

Proxy questionnaire

Deceased before Census Day questionnaire

Non-proxy

2006 Reverse Record Check Census Quality Survey

CONFIDENTIAL once completed. Collected under the Authority of the Statistics Act (Revised Statutes of Canada, 1985, chapter S19)

FOR REGIONAL OFFICE USE ONLY

SPIN: __ __ __ __ __ __ __ __

INT#: __ __ __ __ __ __ __

Given name(s) | Family name

1. Did you change your name since May 15, 2001, or is it different than the name recorded on the line above?

O_2 No
O_1 Yes ➔ *Print new name(s) below*

Given name(s)	Family name

2. Was your usual home on Census Day, MAY 16, 2006 in Canada or outside Canada? For a definition of "usual home", please see the Questionnaire Instructions on the back page of the covering letter.

O_1 In Canada
O_2 Outside Canada ➔ *Go to Question 19 at the bottom of Page 3*

3. The HOUSEHOLD ROSTER is used to list all persons living at your usual home on CENSUS DAY, MAY 16, 2006, even if they were temporarily away. If your usual home on Census Day was in a collective dwelling (hotel, motel, camp, rooming house, etc.) or an institution then a Household Roster is not required ➔ *Go to Question 4 on Page 3. However,* if you were living in a Hutterite colony, then complete the Household Roster.

WHOM TO INCLUDE IN THE HOUSEHOLD ROSTER:

- **EVERYONE WHO USUALLY LIVED** at your address, including newborn babies and room-mates,
- **STUDENTS** who return to live with their parents during the year should be included at their parent's address even if they live elsewhere while attending school or working at a summer job,
- **CHILDREN IN JOINT CUSTODY** who live here most of the time. Children who spend equal time with each parent should be included if they were staying with you on May 16, 2006,
- **SPOUSES OR COMMON-LAW PARTNERS WHO LIVED ELSEWHERE** while working or studying, but who returned home periodically,
- **LANDED IMMIGRANTS** who usually lived with you or **REFUGEES**, that is, persons claiming refugee status and family members living with you,
- **PERSONS FROM ANOTHER COUNTRY WITH A WORK OR STUDY PERMIT** and family members living here with them,
- **PERSONS** who usually live here but are now **IN AN INSTITUTION** (such as a home for the aged, a hospital or a prison), **IF THEY HAVE BEEN THERE LESS THAN SIX MONTHS,**
- **PERSONS** staying here on May 16, 2006, **WHO HAD NO USUAL HOME ELSEWHERE.**

DO NOT INCLUDE THESE PERSONS IN THE HOUSEHOLD ROSTER:

- Persons who had their **usual home at another address in Canada** and who were staying with you temporarily (for example, persons visiting or persons who have their secondary residence here, at this address),
- Residents of another country who were **visiting** Canada (for example, persons on a business trip or on vacation),
- Government representatives of another country or members of the Armed Forces of another country and their family members,
- Persons living in a collective dwelling (other than Hutterite) or an institution such as a hotel, senior's residence, hospital, etc.

Question 3 continued

List below in the Household Roster all persons who lived at your usual address, on May 16, 2006, even if they were temporarily away, then indicate the "Date of Birth, Sex, Marital Status and Relationship to you" for each person. **Put your own information on the first line.**

Household Roster Codes

Sex	Marital Status	Relationship to you
F – Female M – Male	1 – Never legally married (single) 2 – Legally married (and not separated) 3 – Separated, but still legally married 4 – Divorced 5 – Widowed	1 – Spouse or common-law partner 2 – Son or daughter * 3 – Father or mother or step father/mother 4 – Brother or sister or step brother/sister 5 – Other person related 6 – Other person not related

** Stepchildren, adopted children and children of a common-law partner are to be considered sons and daughters.*

HOUSEHOLD ROSTER – CENSUS DAY, MAY 16, 2006

Roster #	Given Name(s)	Family Name	Date of Birth DD-MM-YYYY	Sex	Marital Status	Relationship to you
01						
02						
03						
04						
05						
06						
07						
08						
09						
10						
11						
12						
13						

Information that you wish to add about the Household Roster

YOUR CURRENT AND RECENT ADDRESSES

4. Is the address where you received this questionnaire considered to be your current usual home?

○1 Yes ➔ *Go to Question 6*
○2 No

5. Is your current usual home in Canada?

○1 Yes
○2 No ➔ *Go to Question 9*

6. What is the address or exact physical location of your usual home? If you have a mailing address based on a Post Office box, general delivery, rural route or business, please record as specifically as possible a home address based on civic style street name or 911 numbering.

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

7. When did you move to this address?

○1 **On or before May 14, 2001** ➔ *Go to Question 10*
○2 **Between May 15, 2001 and May 15, 2006**
○3 **On or after May 16, 2006**

8. Were you living in Canada before you moved to your new address?

○1 Yes
○2 No ➔ *Specify country below*

➔ *Go to Question 10*

9. What was the address or exact physical location of your last usual home in Canada before moving to your current address?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

CENSUS DAY 2006

10. What was the address or exact physical location of your usual home on CENSUS DAY, Monday, MAY 16, 2006?

○1 Not in Canada ➔ *Go to Question 19*
○2 Same address as in Question 6
○3 Same address as in Question 9
○4 In Canada – Different address than in Question 6 or 9
➔ *Specify address below*

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

11. If this address is rural (no civic style address), what are the names of two crossroads or streets close to this address?

12. What are the names and characteristics of an adult, if any, living with you at that address on Census Day, Monday, May 16, 2006? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)
D D / M M / Y Y Y Y

○ No other adult

13. Was your usual home on Census Day any one of the following:

- **a lodging or rooming house,**
- **a hotel, motel or tourist home,**
- **a nursing home, hospital or staff residence,**
- **communal quarters of a military camp,**
- **a group home, mission, work camp, prison or**
- **another dwelling of this type?**

○1 Yes
○2 No ➔ *Go to Question 15*

14. What is the name of this dwelling or residence?

➔ *Go to Question 17*

15. Was your usual home on Census Day:

○1 **Owned by you or a member of your household** (even if it was still being paid for)
○2 **Rented** (even if no cash rent was paid)
○3 **Band housing**

16. Was your usual home on Census Day a(n):

○1 **Single-detached house**
○2 **Semi-detached house**
○3 **Row house**
○4 **Apartment or flat in a detached duplex**
○5 **Apartment in a building that has five or more storeys**
○6 **Apartment in a building that has fewer than five storeys**
○7 **Other single attached house**
○8 **Mobile home**
○9 **Other movable dwelling**

17. Were you actually living at that address recorded in Question 10 at the time of the Census, that is, on May 16, 2006?

○1 Yes ➔ *Go to Question 30*
○2 No

18. On May 16, 2006 were you outside Canada?

○1 Yes
○2 No ➔ *Go to Question 30*

19. When did you leave Canada?

If exact date is not known, give best estimate
Day | Month | Year
D D / M M / Y Y Y Y

20. What is the main reason you were outside Canada on Census Day, May 16, 2006?

- O_1 **Vacation**
- O_2 **Studies**
- O_3 **Posted with either the Canadian Armed Forces or Foreign Affairs and International Trade**
- O_4 **Work or looking for work**
- O_5 **Retirement in another country**
- O_6 **Accompanying or joining a spouse or family member**
  ➔ *Continue with Question 21*
- O_7 **Other reason** – *Specify below*

Only answer this question if you answered "**Accompanying or joining a spouse or family member**" in the previous question.

21. What is the main reason this family member was outside Canada on May 16, 2006?

- O_1 **Vacation**
- O_2 **Studies**
- O_3 **Posted with either the Canadian Armed Forces or Foreign Affairs and International Trade**
- O_4 **Work or looking for work**
- O_5 **Retirement in another country**
- O_6 **Other reason** – *Specify below*

22. In what country were you living or staying on May 16, 2006?

23. When did you return to live in Canada?

If exact date is not known, give best estimate

Day / Month / Year ➔ *Go to Question 25*

D D / M M / Y Y Y Y

○ I have not returned to live in Canada

24. Do you intend to return to live in Canada?

O_1 Yes ➔ *Specify date (If exact date is not known, give best estimate)*

Day / Month / Year

D D / M M / Y Y Y Y

O_2 No ➔ *Specify reason, such as: Emigrated, foreign student that left Canada after completing studies, foreign worker that left Canada after completing work, etc.*

O_3 Do not know if I will return to live in Canada.

25. What was the last address in Canada where you lived before residing outside of Canada?

○ Same address as in Question 9 OR specify below

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

26. What are the names and characteristics of an adult, if any, living with you at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)

D D / M M / Y Y Y Y

○ No other adult

27. On May 16, 2006, did you have a residence in Canada either occupied by one or more members of your family or available for your immediate occupancy?

O_1 Yes
O_2 No ➔ *Go to Question 30*

28. What is the address or exact physical location of this other residence?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

29. What are the names and characteristics of an adult, if any, living at that address on Census Day, Monday, May 16, 2006? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)

D D / M M / Y Y Y Y

○ No other adult

OTHER CANADIAN ADDRESSES

30. During the month of May, 2006, did you live or stay elsewhere *in Canada*?

O_1 Yes
O_2 No ➔ *Go to Question 33*

31. What is the address or exact physical location of this other dwelling?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

32. What are the names and characteristics of an adult, if any, living at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)
☐ | ☐☐ / ☐☐ / ☐☐☐☐
D D M M Y Y Y Y

○ No other adult

33. Is there a secondary residence *in Canada*, such as a cottage or condominium, that you or another household member owns or rents and where you stay on occasion?

○$_1$ Yes – *Specify below*

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number
☐☐☐ ☐☐☐ - ☐☐☐☐

○$_2$ No ➔ *Go to Question 35*

34. What are the names and characteristics of an adult, if any, living at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)
☐ | ☐☐ / ☐☐ / ☐☐☐☐
D D M M Y Y Y Y

○ No other adult

35. Is there any other dwelling *in Canada* where someone, such as a parent, relative or friend, may have included your name on a 2006 Census questionnaire?

○$_1$ Yes
○$_2$ No ➔ *Go to Question 38*

36. What is the address or exact physical location of the dwelling?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number
☐☐☐ ☐☐☐ - ☐☐☐☐

37. What are the names and characteristics of an adult, if any, living at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to you

Sex: M or F | Date of Birth (Day / Month / Year)
☐ | ☐☐ / ☐☐ / ☐☐☐☐
D D M M Y Y Y Y

○ No other adult

CENSUS DAY 2001

38. Were you living in Canada five years before Census Day, that is, on May 15, 2001?

○$_1$ Yes ➔ *Go to Question 40*
○$_2$ No

39. In what country did you live on May 15, 2001?

☐☐☐ ➔ *Go to Question 41*

40. What was the address or exact physical location where you lived five years ago, that is on May 15, 2001?

○ Same address as in Question 6 OR
○ Same address as in Question 9 OR

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number
☐☐☐ ☐☐☐ - ☐☐☐☐

QUESTIONS ABOUT YOU

41. *If you were aged 16 or under on May 16, 2006, go to Question 44.*

42. Were you living with a common-law partner on May 16, 2006?

○$_1$ Yes
○$_2$ No ➔ *Go to Question 44*

43. What are the names and characteristics of your common-law partner?

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Sex: M or F | Date of Birth (Day / Month / Year)
☐ | ☐☐ / ☐☐ / ☐☐☐☐
D D M M Y Y Y Y

44. What is the language that you first learned at home in childhood and still understand?

If you no longer understand the first language learned, indicate the second language learned.

○ 1 English
○ 2 French
○ 3 Other – *Specify below*

45. Are you an Aboriginal person, that is, North American Indian, Métis or Inuit (Eskimo)?

○ 4 No
○ 1 Yes, North American Indian
○ 2 Yes, Métis
○ 3 Yes, Inuit (Eskimo)

46. Are you a member of an Indian Band / First Nation?

○ 2 No
○ 1 Yes, member of an Indian Band / First Nation – *Specify Indian Band / First Nation (for example, Musqueam)*

47. Are you a Treaty Indian or a Registered Indian as defined by the *Indian Act* of Canada?

○ 2 No
○ 1 Yes, Treaty Indian or Registered Indian

48. Of what country were you a citizen on May 16, 2006?

Note that "Canada by naturalization" refers to the process by which an immigrant is granted citizenship to Canada, under the Citizenship Act.

Indicate more than one citizenship, if applicable.

○ 1 Canada, by birth
○ 2 Canada, by naturalization
○ 3 Other country – *Specify below*

49. Are you now, or have you ever been a landed immigrant?

Note that a "landed immigrant" is a person who has been granted the right to live in Canada permanently by immigration authorities.

○ 2 No ➔ *Go to end of questionnaire*
○ 1 Yes

50. In what year did you first become a landed immigrant?

○ 1 2006
○ 2 Other – *Specify below*

If exact year is not known, indicate best estimate.

Y Y Y Y ➔ *Go to end of questionnaire*

51. When in 2006 did you become a landed immigrant?

○ 1 Before May 16, 2006 (Census Day)
○ 2 On or after May 16, 2006

YOU HAVE NOW COMPLETED YOUR QUESTIONNAIRE

Please return your completed questionnaire in the pre-paid envelope provided.

Thank you

THE LAW PROTECTS WHAT YOU TELL US

The confidentiality of your completed questionnaire is protected by law. All Statistics Canada employees have taken an oath of secrecy. The personal information you provided cannot be given to anyone outside Statistics Canada – not the police, not another government department, not another person. This is your right!

COMMENTS?

Proxy

2006 Reverse Record Check Census Quality Survey

CONFIDENTIAL once completed. Collected under the Authority of the Statistics Act (Revised Statutes of Canada, 1985, chapter S19)

FOR REGIONAL OFFICE USE ONLY

SPIN: __ __ __ __ __ __ __ __

INT#: __ __ __ __ __ __ __

Given name(s) | Family name

1. Did … change his or her name since May 15, 2001, or is it different than the name recorded on the line above?

○2 No
○1 Yes ➔ *Print new name(s) below*

Given name(s)	Family name

1A. Is … deceased?

○1 No, … is not deceased ➔ *Go to Question 2*
○2 Yes, …died before Census Day, May 16, 2006 ➔ *Go to Question 1 of the pink questionnaire*
○3 Yes, …died on or after Census Day, May 16, 2006 ➔ *Go to Question 3*

2. Was …'s usual home on Census Day, MAY 16, 2006 in Canada or outside Canada? For a definition of "usual home", please see the Questionnaire Instructions on the back page of the covering letter.

○1 In Canada
○2 Outside Canada ➔ *Go to Question 19 at the bottom of Page 3*

3. The HOUSEHOLD ROSTER is used to list all persons living at ….'s usual home on CENSUS DAY, MAY 16, 2006, even if they were temporarily away. If …..'s usual home on Census Day was in a collective dwelling (hotel, motel, camp, rooming house, etc.) or an institution then a Household Roster is not required ➔ *Go to Question 4 on Page 3. However,* if …. was living in a Hutterite colony, then complete the Household Roster.

WHOM TO INCLUDE IN THE HOUSEHOLD ROSTER:

- **EVERYONE WHO USUALLY LIVED** at …..'s address, including newborn babies and room-mates,
- **STUDENTS** who return to live with their parents during the year should be included at their parent's address, even if they live elsewhere while attending school or working at a summer job,
- **CHILDREN IN JOINT CUSTODY** who live here most of the time. Children who spend equal time with each parent should be included if they were staying with …. on May 16, 2006,
- **SPOUSES OR COMMON-LAW PARTNERS WHO LIVED ELSEWHERE** while working or studying, but who returned home periodically,
- **LANDED IMMIGRANTS** who usually lived with … or **REFUGEES**, that is, persons claiming refugee status and family members living with them,
- **PERSONS FROM ANOTHER COUNTRY WITH A WORK OR STUDY PERMIT** and family members living here with them,
- **PERSONS** who usually lived with … but are now **IN AN INSTITUTION** (such as a home for the aged, a hospital or a prison), **IF THEY HAVE BEEN THERE LESS THAN SIX MONTHS,**
- **PERSONS** staying here on May 16, 2006, **WHO HAD NO USUAL HOME ELSEWHERE.**

DO NOT INCLUDE THESE PERSONS IN THE HOUSEHOLD ROSTER:

- Persons who had their **usual home at another address in Canada** and who were staying with … temporarily (for example, persons visiting or persons who have their secondary residence, at this address),
- Residents of another country who were **visiting** Canada (for example, persons on a business trip or on vacation),
- Government representatives of another country or members of the Armed Forces of another country and their family members,
- Persons living in a collective dwelling (other than Hutterite) or an institution such as a hotel, senior's residence, hospital, etc.

Question 3 continued

List below in the Household Roster all persons who lived at …'s usual address, on May 16, 2006, even if they were temporarily away, then indicate the "Date of Birth, Sex, Marital Status and Relationship to …" for each person. **Put …'s information on the first line.**

Household Roster Codes

Sex	Marital Status	Relationship to …
F – Female M – Male	1 – Never legally married (single) 2 – Legally married (and not separated) 3 – Separated, but still legally married 4 – Divorced 5 – Widowed	1 – Spouse or common-law partner 2 – Son or daughter * 3 – Father or mother or step father/mother 4 – Brother or sister or step brother/sister 5 – Other person related to … 6 – Other person not related to … 7 – Unknown relationship to …

* *Stepchildren, adopted children and children of a common-law partner are to be considered sons and daughters.*

HOUSEHOLD ROSTER – CENSUS DAY, MAY 16, 2006

Roster #	Given Name(s)	Family Name	Date of Birth DD-MM-YYYY	Sex	Marital Status	Relationship to …
01						
02						
03						
04						
05						
06						
07						
08						
09						
10						
11						
12						
13						

Information that you wish to add about the Household Roster

…'s CURRENT AND RECENT ADDRESSES

4a. If … is deceased, ➔ *Go to Question 10*

4. Is the address where this survey was received considered to be …'s current usual home?

○1 Yes ➔ *Go to Question 6*
○2 No

5. Is …'s current usual home in Canada?

○1 Yes
○2 No ➔ *Go to Question 9*

6. What is the address or exact physical location of …'s usual home? If … has a mailing address based on a Post Office box, general delivery, rural route or business, please record as specifically as possible a home address based on civic style street name or 911 numbering.

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

7. When did … move to this address?

○1 **On or before May 14, 2001** ➔ *Go to Question 10*
○2 **Between May 15, 2001 and May 15, 2006**
○3 **On or after May 16, 2006**

8. Was … living in Canada before he or she moved to his or her new address?

○1 Yes
○2 No ➔ *Specify country below*

➔ *Go to Question 10*

9. What was the address or exact physical location of …'s last usual home in Canada before moving to his or her current address?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

CENSUS DAY 2006

10. What was the address or exact physical location of …'s usual home on CENSUS DAY, Monday, MAY 16, 2006?

○1 Not in Canada ➔ *Go to Question 19*
○2 Same address as in Question 6
○3 Same address as in Question 9
○4 In Canada – Different address than in Question 6 or 9 ➔ *Specify address below*

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

11. If this address is rural (no civic style address), what are the names of two crossroads or streets close to this address?

12. What are the names and characteristics of an adult, if any, who lived with … at that address on Census Day, Monday, May 16, 2006? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* [ ][ ] *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) D D / M M / Y Y Y Y

○ No other adult

13. Was …'s usual home on Census Day any one of the following:

- **a lodging or rooming house,**
- **a hotel, motel or tourist home,**
- **a nursing home, hospital or staff residence,**
- **communal quarters of a military camp,**
- **a group home, mission, work camp, prison or**
- **another dwelling of this type?**

○1 Yes
○9 No or Do not know ➔ *Go to Question 15*

14. What is the name of this dwelling or residence?

➔ *Go to Question 17*

15. Was …'s usual home on Census Day:

○1 **Owned by … or a member of …'s household** (even if it was still being paid for)
○2 **Rented** (even if no cash rent was paid)
○3 **Band housing**

16. Was …'s usual home on Census Day a(n):

○1 **Single-detached house**
○2 **Semi-detached house**
○3 **Row house**
○4 **Apartment or flat in a detached duplex**
○5 **Apartment in a building that has five or more storeys**
○6 **Apartment in a building that has fewer than five storeys**
○7 **Other single attached house**
○8 **Mobile home**
○9 **Other movable dwelling**

17. Did … live at that address recorded in Question 10 at the time of the Census, that is, on May 16, 2006?

○1 Yes ➔ *Go to Question 30*
○2 No

18. On May 16, 2006 was … outside Canada?

○1 Yes
○2 No or Do not know ➔ *Go to Question 30*

19. When did … leave Canada?

If exact date is not known, give best estimate

Day | Month | Year
D D / M M / Y Y Y Y

20. What is the main reason … was outside Canada on Census Day, May 16, 2006?

- ○ 1 Vacation
- ○ 2 Studies
- ○ 3 Posted with either the Canadian Armed Forces or Foreign Affairs and International Trade
- ○ 4 Work or looking for work
- ○ 5 Retirement in another country
- ○ 6 Accompanying or joining a spouse or family member
  ➔ *Continue with Question 21*
- ○ 7 Other reason – *Specify below*

Only answer this question if you answered "Accompanying or joining a spouse or family member" in the previous question.

21. What is the main reason this family member was outside Canada on May 16, 2006?

- ○ 1 Vacation
- ○ 2 Studies
- ○ 3 Posted with either the Canadian Armed Forces or Foreign Affairs and International Trade
- ○ 4 Work or looking for work
- ○ 5 Retirement in another country
- ○ 6 Other reason – *Specify below*

22. In what country was … living or staying on May 16, 2006?

23. Has ... returned to live in Canada?

○ 1 Yes ➔ *Specify date (If exact date is not known, give best estimate)*

Day / Month / Year (D D / M M / Y Y Y Y) ➔ *Go to Question 25*

○ 2 No
○ 3 Do not know

➔ *If … is deceased, Go to Question 25*

24. Does … intend to return to live in Canada?

○ 1 Yes ➔ *Specify date (If exact date is not known, give best estimate)*

Day / Month / Year (D D / M M / Y Y Y Y)

○ 2 No ➔ *Specify reason, such as: Emigrated, foreign student that left Canada after completing studies, foreign worker that left Canada after completing work, etc.*

○ 3 Do not know if … intends to return to live in Canada.

25. What was the last address in Canada where … lived before residing outside of Canada?

○ Same address as in Question 9 OR specify below

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

26. What are the names and characteristics of an adult, if any, who lived with … at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) (D D / M M / Y Y Y Y)

○ No other adult

27. On May 16, 2006, did … have a residence in Canada either occupied by one or more members of their family or available for their immediate occupancy?

○ 1 Yes
○ 2 No ➔ *Go to Question 30*
○ 3 Do not know ➔ *Go to Question 30*

28. What is the address or exact physical location of this other residence?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

29. What are the names and characteristics of an adult, if any, who lived with … at that address on Census Day, Monday, May 16, 2006? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) (D D / M M / Y Y Y Y)

○ No other adult

OTHER CANADIAN ADDRESSES

30. During the month of May, 2006, did … live or stay elsewhere *in Canada*?

○ 1 Yes
○ 2 No or Do not know ➔ *Go to Question 33*

31. What is the address or exact physical location of this other dwelling?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

32. What are the names and characteristics of an adult, if any, who lived with … at that address during the month of May, 2006? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) D D / M M / Y Y Y Y

○ No other adult

33. Is there a secondary residence *in Canada*, such as a cottage or condominium, that … or another household member owns or rents and … stays/stayed on occasion?

○$_1$ Yes – *Specify below*

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

○$_2$ No or Do not know ➔ *Go to Question 35*

34. What are the names and characteristics of an adult, if any, living with … at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) D D / M M / Y Y Y Y

○ No other adult

35. Is there any other dwelling *in Canada* where someone, such as a parent, relative or friend, may have included …'s name on a 2006 Census questionnaire?

○$_1$ Yes
○$_2$ No ➔ *Go to Question 38*

36. What is the address or exact physical location of the dwelling?

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

37. What are the names and characteristics of an adult, if any, living at that address? (not required if a collective dwelling or institution)

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day / Month / Year) D D / M M / Y Y Y Y

○ No other adult

CENSUS DAY 2001

38. Was … living in Canada five years before Census Day, that is, on May 15, 2001?

○$_1$ Yes ➔ *Go to Question 40*
○$_2$ No or Do not know

39. In what country did … live on May 15, 2001?

☐ ➔ *Go to Question 41*

40. What was the address or exact physical location where … lived five years ago, that is on May 15, 2001?

○ Same address as in Question 6 OR
○ Same address as in Question 9 OR

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

○ Do not know

QUESTIONS ABOUT …

41. If the age of … was 16 or under on May 16, 2006, go to Question 44.

42. Was … living with a common-law partner on May 16, 2006?

○$_1$ Yes
○$_2$ No ➔ *Go to Question 44*

43. What are the names and characteristics of …'s common law partner?

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Sex: M or F | Date of Birth (Day / Month / Year) D D / M M / Y Y Y Y

44. What is/was the language that … first learned at home in childhood and still understands?

If … no longer understands the first language learned, indicate the second language learned.

○ 1 English
○ 2 French
○ 3 Other – *Specify below*

45. Is/was … an Aboriginal person, that is, North American Indian, Métis or Inuit (Eskimo)?

○ 4 No
○ 1 Yes, North American Indian
○ 2 Yes, Métis
○ 3 Yes, Inuit (Eskimo)

46. Is/was … a member of an Indian Band / First Nation?

○ 2 No
○ 1 Yes, member of an Indian Band / First Nation – *Specify Indian Band / First Nation (for example, Musqueam)*

47. Is/was … a Treaty Indian or a Registered Indian as defined by the *Indian Act* of Canada?

○ 2 No
○ 1 Yes, Treaty Indian or Registered Indian

48. Of what country was … a citizen on May 16, 2006?

Note that "Canada by naturalization" refers to the process by which an immigrant is granted citizenship to Canada, under the Citizenship Act.

Indicate more than one citizenship, if applicable.

○ 1 Canada, by birth
○ 2 Canada, by naturalization
○ 3 Other country – *Specify below*

49. Is … now, or has … ever been a landed immigrant?

Note that a "landed immigrant" is a person who has been granted the right to live in Canada permanently by immigration authorities.

○ 2 No or Do not know ➔ *Go to Question 52*
○ 1 Yes

50. In what year did … first become a landed immigrant in Canada?

○ 1 2006
○ 2 Other – *Specify below*

If exact year is not known, indicate best estimate.

Y Y Y Y ➔ *Go to Question 52*

51. When in 2006 did … become a landed immigrant?

○ 1 Before May 16, 2006 (Census Day)
○ 2 On or after May 16, 2006

52. What are the names and characteristics of the person who completed this survey?

○ *Listed in Question 3, enter Roster Number* ☐☐ *OR*

Given name(s)

Family name

Relationship to …

Sex: M or F

Date of Birth (Day / Month / Year)

D D / M M / Y Y Y Y

YOU HAVE NOW COMPLETED YOUR QUESTIONNAIRE

Please return your completed questionnaire in the pre-paid envelope provided.

Thank you

THE LAW PROTECTS WHAT YOU TELL US

The confidentiality of your completed questionnaire is protected by law. All Statistics Canada employees have taken an oath of secrecy. The personal information you provided cannot be given to anyone outside Statistics Canada – not the police, not another government department, not another person. This is your right!

COMMENTS?

Deceased

2006 Reverse Record Check Census Quality Survey

CONFIDENTIAL once completed. Collected under the Authority of the Statistics Act (Revised Statutes of Canada, 1985, chapter S19)

FOR REGIONAL OFFICE USE ONLY

SPIN: __ __ __ __ __ __ __ __

INT#: __ __ __ __ __ __ __ __

Given name(s)

Family name

Questions relating to a person deceased before CENSUS DAY, May 16, 2006

1. Did … change his or her name since May 15, 2001, or was it different than the name recorded on the line above?

O_2 No
O_1 Yes ➔ *Print new name(s) below*

Given name(s)

Family name

2. What was …'s date of birth?

If exact date is not known, indicate best estimate.

Day / Month / Year

D D / M M / Y Y Y Y

3. On what date did … die?

If exact date is not known, indicate best estimate.

Day / Month / Year

D D / M M / Y Y Y Y

4. What was …'s sex?

O_1 Female
O_2 Male ➔ *Go to Question 6*

5. What was …'s maiden name?

O Do not know

6. In what province, territory or country did … die?

O_1 Newfoundland and Labrador
O_2 Prince Edward Island
O_3 Nova Scotia
O_4 New Brunswick
O_5 Quebec
O_6 Ontario
O_7 Manitoba
O_8 Saskatchewan
O_9 Alberta
O_{10} British Columbia
O_{11} Yukon Territory
O_{12} Northwest Territories
O_{13} Nunavut
O_{14} United States
O_{15} Outside North America – *Specify country*

O Do not know

7. What was the address or exact physical location of …'s last known usual home in Canada? If … had a mailing address based on a Post Office box, general delivery, rural route or business please record as specifically as possible …'s home address based on civic style street name or 911 numbering.

If … had no usual home, enter the place where … last stayed.

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

O Do not know

8. What are the names and characteristics of an adult, if any, who was living with … at that address? (not required if a collective dwelling or institution)

Given name(s)

Family name

Relationship to …

Sex: M or F | Date of Birth (Day, Month, Year)

D D / M M / Y Y Y Y

O No other adult
O Do not know

9. Is there another residence *in Canada* such as a cottage, condominium or any type of second home that … or another household member owned or rented and where … stayed on occasion?

O_1 Yes – *Specify*

Number and street name or lot and concession number

Apartment Number | City, municipality, town, village, Indian reserve

Name of residence, if any

Province/Territory | Postal Code

Tel. Area Code | Telephone Number

O_2 No ➔ *Go to Question 11*
O_3 Do not know ➔ *Go to Question 11*

OVER ➔

10. What are the names and characteristics of an adult, if any, living at that address? (not required if a collective dwelling or institution)

○ 1 The same person as in Question 8 OR

Given name(s)

Family name

Relationship to …

Sex: M or F

Date of Birth (Day / Month / Year)

D D / M M / Y Y Y Y

○ 2 No other adult
○ 3 Do not know

11. What are the names and characteristics of the person who completed this questionnaire?

○ The same person as in Question 8 OR
○ The same person as in Question 10 OR

Given name(s)

Family name

Relationship to …

Sex: M or F

Date of Birth (Day / Month / Year)

D D / M M / Y Y Y Y

YOU HAVE NOW COMPLETED THE QUESTIONNAIRE

Please return your completed questionnaire in the pre-paid envelope provided.

Thank you

THE LAW PROTECTS WHAT YOU TELL US

The confidentiality of this completed questionnaire is protected by law. All Statistics Canada employees have taken an oath of secrecy. The personal information you provided cannot be given to anyone outside Statistics Canada – not the police, not another government department, not another person. This is your right!

COMMENTS ?

Appendix B. Acronyms

The following acronyms are used in this report.

AMS	Automated Match Study
CA	census agglomeration
CANCEIS	Canadian Census Edit and Imputation System
CATI	computer-assisted telephone interviewing
CCRA	Canada Customs and Revenue Agency
CDS	Collective Dwelling Survey
Census RDB	Census Response Database
CHL	Census Help Line
CMA	census metropolitan area
COS	Census Overcoverage Study
CSD	census subdivision
CU	collection unit
CV	coefficient of variation
CLD	crew leader district
DCS	Dwelling Classification Survey
DPC	data processing centre
EA	enumeration area
EN	enumerator
GRLS	Generalized Record Linkage System
HO	head office (of Statistics Canada)
IEIR	incompletely enumerated Indian reserves
L/L	list/leave
MO	mail-out
NPR	non-permanent resident
NRFU	non-response follow-up
PDS	Private Dwelling Study
PED	provincial electoral district
PEP	Population Estimation Program
RO	regional office (of Statistics Canada)
RRC	Reverse Record Check
RRC RDB	RRC version of the census RDB
SP	selected person
StatMx	Statistical Macro Extensions
VC	vacancy check
VR	Visitation Record
VS	vital statistics
WHI	whole household imputation

Bibliography

Armstrong, M. 2008. *2006 Reverse Record Check Survey – Data Collection Report*. Internal report of the Social Survey Methods Division. Ottawa. Statistics Canada.

Diallo, M. S. 2008. *Contre-vérification des dossiers Recensement 2006 : Classification 2006 et fichier maître des personnes choisies 2006*. Internal report. Ottawa. Statistics Canada.

Dick, P. 2007. Results of the 2006 Dwelling Classification Study (DCS). Meeting of the Federal-Provincial-Territorial Committee on Demography. Statistics Canada, Ottawa, Ontario, September 20-21. 33 slides.

Felligi, I.P. and Sunter, A.B. (1969). *A Theory of Record Linkage*. Journal of the American Statistical Association, Vol. 64, p. 1,183 – 1,210.

Generalized Systems Methods, 2007. *Methodology Software Library, StatMx, Calibration Estimation Function, Version 2.0, Description of the Function Parameters*. Internal report. Ottawa. Statistics Canada.

Morel, J. and Farr, H. 2008a. *2006 Census Overcoverage Study, Evaluation Report*. Internal report of the Social Survey Methods Division. Ottawa. Statistics Canada.

---. 2008b. *2006 Census Overcoverage Study, Methodology Report*. Internal report of the Social Survey Methods Division. Ottawa. Statistics Canada.

Théberge, A. 2008. *Non-response Adjustment and Variance Estimation for the 2006 Reverse Record Check*. Internal report of the Social Survey Methods Division. Ottawa. Statistics Canada.

Statistics Canada. No date*. Incompletely enumerated Indian reserves and Indian settlements.* 2006 Census. Notes Series. http://www12.statcan.gc.ca/census-recensement/2006/ref/notes/aboriginal-autochtones-eng.cfm (accessed November 27, 2009)

---. No date b. *Estimates of Total Population, Canada, Provinces and Territories*. http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3601&lang=en&db=imdb&adm=8&dis=2#4 (accessed November 27, 2009)

---. 2009. *2006 Census Technical Report: Sampling and Weighting*. Statistics Canada Catalogue no. 92-568-X. Ottawa, Ontario. 2006 Census Technical Reports Series. August. http://www12.statcan.ca/census-recensement/2006/ref/rp-guides/rp/sw-ep/sw-ep_index-eng.cfm (accessed November 27, 2009)

---. 2009b. “2006 Census population net undercoverage 'on reserve,' for Canada, Region” (table). *2006 Census Technical Report: Aboriginal Peoples*. Statistics Canada Catalogue no. 92-569-X. Ottawa, Ontario. 2006 Census Technical Reports Series. March. http://www12.statcan.ca/census-recensement/2006/ref/rp-guides/rp/ap-pa/table/table_6-eng.cfm (accessed November 27, 2009)

---. 2009c. “Model estimates for incompletely enumerated Indian reserves and settlements for Canada” (table). *2006 Census Technical Report: Aboriginal Peoples*. Statistics Canada Catalogue no. 92-569-X. Ottawa, Ontario. 2006 Census Technical Reports Series. March. http://www12.statcan.ca/census-recensement/2006/ref/rp-guides/rp/ap-pa/table/table_7-eng.cfm (accessed November 27, 2009)

---. 2004. *Coverage*. Statistics Canada Catalogue no. 92-394-XIE. Ottawa, Ontario. 2001 Census Technical Reports Series. November 25. http://www12.statcan.ca/english/census01/Products/Reference/tech_rep/coverage/index.cfm (accessed November 9, 2009)

---. 1999. *Coverage*. Statistics Canada Catalogue no. 92-394-XIE. Ottawa, Ontario, December 14. Technical Reports: 1996 Census of Population Series. http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=92-370-X&CHROPG=1&lang=eng (accessed November 9, 2009)

---. 1994. *Coverage*. Statistics Canada Catalogue no. 92-341-XIE. Ottawa, Ontario, March. 1991 Census Technical Reports Series.

---. 1990. *Coverage*. Statistics Canada Catalogue no. 99-135E. Ottawa, Ontario, March. User's Guide to the Quality of 1986 Census Data Series.