NOVA SCOTIA
Finance
Community Counts

Census of Population
Cautions on Data Use

Prepared by:

Dennis W. Pilkey,
Director, Community Counts Division
Department of Finance

Disclaimer:
These notes are a combination of facts and considered speculation on the effects of Census coverage and the use of Census data. The opinions are those of the author and do not necessarily reflect the opinions of the Nova Scotia Government or of Statistics Canada. Questions and comments can be referred to the author at pilkeydw@gov.ns.ca

June 2005

NOVA SCOTIA
Finance
Community Counts

Census of Population
Cautions on Data Use

Purpose

The purpose of this paper is to understand some of the limitations of Census coverage and their implications when using Census data for analysis.

Introduction

Many people use custom and standard tabulations of data from Statistics Canada's Census of Population, which is carried out every five years. These rich demographic data about Canada's population are a major statistical resource for many policy planning and business decisions. These data are the most reliable available and one of the few sources of statistical information for small areas. The Census short form is intended to cover the entire population (100%), and data are therefore not subject to the usual sampling errors associated with surveys. For example, many surveys are often presented as being accurate within five percent, 19 times out of 20. Although the Census of Population does not have these sampling errors, there are limitations in the data that need to be considered. The Census long form is given to one in five persons (20%). This very large sample size has the effect of minimizing sampling errors, but does not eliminate them.

Census Coverage

Unadjusted Census numbers

100% population – every household gets a Census form to complete, which includes questions about basic demographic data (age, gender, marital status, relationships of persons in the household). Data are collected and reported based on where the person was on Census Day. There are exception rules for persons such as students, tourists and persons on travel or short term visas. These cases have the potential to result in Census coverage errors because of incomplete reporting of information related to the situation. For Nova Scotia, **908,007** people were reported in the 2001 Census. Census demographic tabulations use this number.

20% sample – one in five households receives the long form to complete, which includes the above information as well as a number of other questions related to such things as education, employment, income, ethnicity, and languages. The results of the 20% sample are used to estimate the corresponding numbers for these additional questions for the total population. For Nova Scotia, the 20% sample provides estimates for the **897,570** people that are represented by the sample. The difference between the 100% data (908,007) and the 20% data (897,750) is because people living in institutions such as homes for acute care (seniors) are not included in the long form information.

Adjusted Census numbers

The Census of Population attempts to get information from every person in Canada, without duplication, and counting people according to their legal residence and not necessarily where they are on Census Day. When collecting and reporting on the over 30,000,000 people in Canada, it is extremely difficult to make this ideal a reality. For example, students are asked to be recorded at their permanent address, but many fill out census forms for their school year residence. If parents at the permanent residence also list the student on their census form, the student is counted twice. Another potential error in this situation is that the parents do not list the student on their census form, resulting in a higher population reported for the university town than it should be, and a lower population reported for the home area of the student.

Statistics Canada carries out a series of studies that estimate the number of people that were missed, those that were double counted, those that were counted in the wrong place (provincial and territorial level only) and those that should not have been counted at all, e.g. out-of-country visitors. There are usually more people missed than incorrectly counted and as a result there is a net undercoverage. In 2001, the net undercoverage for Nova Scotia was 24,521 or 2.7% of the population. The adjusted population for Nova Scotia as of July 1, 2001, was estimated to be **932,389.** Appendix A shows the undercoverage, overcoverage and net undercoverage for Nova Scotia for the 1991, 1996 and 2001 Census of Population.

Appendix B shows the relative population differences by age group and gender for the short (100% data) and long (20% data) forms for the 2001 Census of Population.

Observations

The following are observations and implications on use of tabulations from the above:

- Data from the 100% population compared to demographic estimates:
  - Total population is understated by 24,521.
  - The 20-34 age group is most understated, especially for males.
  - The 20-34 age group is most likely relatively overstated for university and economic centres, and understated for rural areas. Nova Scotia's high percentage of university students, especially those from out-of-province, accentuates these distortions.
  - Net undercoverage adjustments are at the provincial/territorial level and do not affect the distribution of the unadjusted population numbers. As a result, 39,643[1] coverage errors remain in the unadjusted numbers.

- Data from the 20% sample compared to demographic estimates:
  - Total population is underreported by 34.958[2].
  - Underreporting is disproportionately higher for seniors, because of relatively large numbers in institutional care.

[1] Sum of undercoverage (32,082) and overcoverage (7,561)
[2] Sum of net undercoverage (24,521) and those not covered in the 20% long form sample (10,437).

- Undercoverage is higher for most age groups than in the 100% sample, possibly because of increased reporting burden.
- The 20-34 group is even more underrepresented than for the 100% sample. The gaps of reporting are relatively higher right up to age 59.
- For persons 75 years of age and over, 12.9% of females and 7.8% of males are not included in the 20% sample.
- Net undercoverage adjustments are at the provincial/territorial level and do not affect the distribution of the unadjusted numbers. As a result, 50,080[3] coverage errors and additional persons not covered by the 20% sample remain in the unadjusted numbers.

For both of the above data sets, the numbers do not include any possible errors of persons counted in the wrong place within the province. This is most likely to happen for students and others who are not actually living at home at the time of the Census. In addition, any numbers for missing persons are added at the provincial level and people counted twice, or in the wrong province or territory, are removed from the provincial totals only. The resulting adjusted population estimates are then pro-rated to the Census Division (county) level based on the original counts. The most likely effect of this is to understate rural numbers for these people and overstate economic and university centres.

While overall national, provincial and territorial numbers are based on a *de jure* (legal address) principle, the sub-provincial data is a combination of *de jure* and *de facto* (address reported on day of Census) principles.

Concluding Comments

In spite of the above observations, the Census of Population remains the best source of small area data that is available. Probably the biggest weaknesses that affect Nova Scotia are due to over-reporting of post-secondary students, and non-reporting for seniors and others in institutions. For health system planning, especially for any initiatives that consider general societal indicators, this represents a significant weakness. Plans are underway to try to get better coverage of this population in the future.

The actual population counts and basic numbers may not be 100% accurate as many people expect. While the exact numbers may not be completely accurate, the various derived indicators such as percentages and percent share (i.e. percent of people with a degree) are broadly comparable across the various jurisdictions and are considered a valid source for analytical purposes.

[3] Sum of total people missed or miscounted by the 100% sample (39,643) and total people missed in the 20% sample (10.437).

NOVA SCOTIA
Finance
Community Counts

Appendix A

2001 Census of Population
Comparison of Coverage Error

	1991		1996		2001	
	#	%	#	%	#	%
Unadjusted Census (100 % Sample)	899,942		909,282		908,007	
Undercoverage	26,072	2.9	26,977	3.0	32,082	3.5
Overcoverage	3,283	0.4	4,328	0.5	7,561	0.8
Net Undercoverage	22,789	2.53	22,649	2.49	24,521	2.70
Adjusted Census Estimate (as of July 1, 2001)	922,731		931,931		932,389	
Unadjusted Census (20 % Sample)	890,950		899,970		897,570	
As percent of Adjusted		96.6		96.6		96.27
Total missed or counted incorrectly **	29,355		31,305		39,643	
20% sample deficit	8,992		9,312		10,437	
Total potential errors in 20% data (provincial level only*)	38,347		40,617		50,080	
Percent of Adjusted Census estimate		4.3		4.5		5.6

* These do not include possible errors of persons counted in the wrong place within the province. This is most likely to happen for students and others who are not actually living at home at the time of the Census. In addition, any numbers for missing persons are added at the provincial level and people counted twice or in the wrong province or territory are removed from the provincial totals only. The resulting adjusted population estimates are then prorated to the CD level (county) based on the original counts. The most likely effect of this is to understate rural numbers for these people and overstate economic and university centres. All lower levels of census data retain the original coverage errors.

While overall national, provincial and territorial numbers are based on a de jure (legal address) principle, the sub-provincial data is based on a combination of de jure and de facto (address reported on day of census) principles.

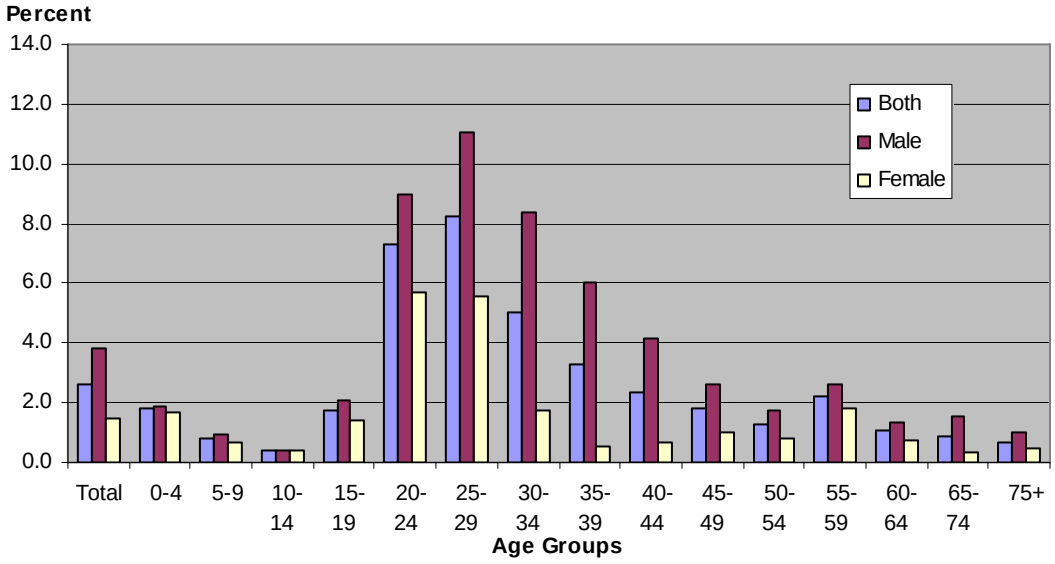
** sum of overcoverage and undercoverage.

NOVA SCOTIA
Finance
Community Counts

Appendix B

2001 Census of Population
Comparison of Adjusted and Unadjusted Census Counts

Short Form Data (100%)
Nova Scotia Population Missed in 2001 Census



Long Form Data (20%)
Nova Scotia Population Missed in 2001 Census