

Table 1 National Household Survey response rate by response mode, Canada, provinces and territories

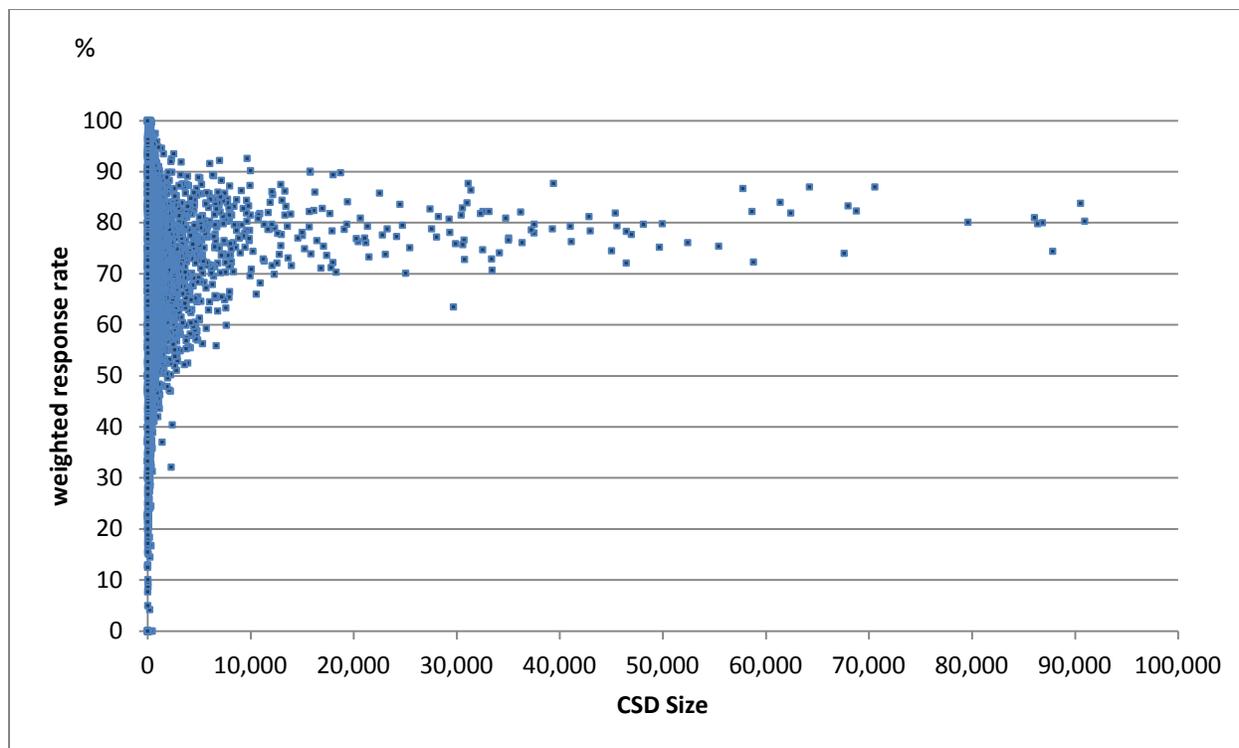
Provinces and territories	Unweighted				Weighted
	Internet	Printed version	Other	All modes	All modes
	%				%
Canada	43.5	11.7	13.4	68.6	77.2
Newfoundland and Labrador	27.6	18.5	17.2	63.3	72.5
Prince Edward Island	29.5	17.8	13.1	60.4	70.0
Nova Scotia	33.9	18.8	12.3	65.0	74.8
New Brunswick	38.1	13.4	12.4	63.9	74.2
Quebec	44.0	13.3	14.6	71.9	80.7
Ontario	45.4	10.5	11.7	67.6	76.3
Manitoba	36.9	14.1	18.1	69.1	76.3
Saskatchewan	31.3	15.6	16.9	63.8	73.1
Alberta	45.3	8.9	13.1	67.3	75.4
British Columbia	47.8	10.1	11.6	69.5	77.1
Yukon	19.1	7.7	38.1	64.9	72.7
Northwest Territories	3.8	0.4	79.7	83.9	83.8
Nunavut	--	0.7	75.6	76.3	76.3

Note: The response rates are based on the NHS's final sampling weights. The initial sampling weight of the dwellings that responded to the NHS before a specific date during the collection period is equal to the sampling fraction in their area. The dwellings that were in the non-response follow-up subsample and responded were assigned a larger weight to compensate for non-response. The weighted response rates are calculated as follows: the weighted number of sampled private dwellings that returned a questionnaire divided by the weighted number of sampled private dwellings classified as occupied.

Of the self-administered response modes, the online mode was used most. At the national level, this mode had a response rate of 43.5%. Data collected with an online questionnaire are generally more complete and consistent.

The response rate varied by province and territory and by census subdivision (CSD) size. For large CSDs, the response rates varied appreciably in the same way as the provincial and territorial response rates: the weighted response rates were mostly between 70% and 80%. For the smallest CSDs, however, the rates varied much more. Figure 2 shows the weighted response rate in relation to the number of occupied private dwellings for CSDs with 100,000 or fewer occupied private dwellings. As the figure indicates, the response rates for CSDs with fewer than 20,000 occupied private dwellings were highly scattered, while the response rates for larger CSDs (between 20,000 and 100,000 occupied private dwellings) fell mostly between 70% and 90%.

Figure 2 Distribution of the NHS weighted response rate by census subdivision (CSD) size, CSDs with fewer than 100,000 dwellings occupied by the usual residents



Another measure of the survey's response rate is the response rate for each question on returned questionnaires. This rate varies substantially from one section of the questionnaire to another. The response rates for demographic characteristics range from 96.7% to 99.7%. The response rates for sociocultural, linguistic and mobility characteristics range from 95.1% to 99.2%. The response rates for education characteristics range from 89.4% to 95.8%. For the work, income and dwelling characteristics sections, however, the response rates range from 80.7% to 93.9%. For more details, please see the reference guides for the various release topics (see [Appendix 2](#)).

4. Data processing

4.1 Data Operations Centre

Statistics Canada's Data Operations Centre (DOC) was the central reception and storage point for electronic and printed questionnaires. Electronic questionnaires were transmitted directly to the DOC's servers, and printed questionnaires were scanned and stored as images. After the quality of the image was confirmed, the data were captured by optical mark recognition (OMR) and intelligent character recognition (ICR). If the image quality was inadequate, the data were captured manually by an operator.

Coding, the next stage of data processing, was also carried out in the Data Operations Centre. All write-in responses were submitted to an automated coding system that assigned each response a numeric code using Statistics Canada reference files, code sets and standard classifications. When the system was unable to assign a code to a particular response, the response was coded manually by an operator.

Coding was applied to the following variables: 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, place of birth of parents, major field of study, location of study, language of work, industry, occupation and place of work.

4.2 Data edit and non-response imputation

After data capture, initial edit and coding operations have been completed, the data are processed up to the final edit and imputation stage. The final edit detects invalid responses and inconsistencies. This edit is based on rules determined by Statistics Canada's subject-matter analysts. Unanswered questions are also identified. Imputation replaces these missing, invalid or inconsistent responses with plausible values. When carried out properly, imputation can improve data quality by replacing non-responses with plausible responses similar to the ones that the respondents would have given if they had answered the questions. It also has the advantage of producing a complete data set.

The nearest-neighbour method was used to impute NHS data. This method is widely used in the treatment of non-response. It replaces missing, invalid or inconsistent information about one respondent with values from another, 'similar' respondent. The rules for identifying the respondent most similar to the non-respondent may vary with the variables to be imputed. Donor imputation methods have good properties and generally will not alter the distribution of the data, a drawback of many other imputation techniques. Following nearest-neighbour imputation, the data are checked for consistency.

4.3 Weighting

The final responses are weighted so that the data from the sample accurately represent the NHS's target population. The weighting process involves calculating sampling weights, adjusting the weights for the survey's total non-response, and calibrating the weights against census totals.

First, an initial sampling weight of about 3 is assigned to each sampled household. The initial weight of 3 is the inverse of the probability of being selected in the NHS sample. As noted in [Section 3.2](#), about 3 of 10 households were selected in the sample, which yields an initial weight of just over 3 (10/3). Then the sampling weights are adjusted to reflect the selection of the subsample. As mentioned in [Section 3.4](#), the subsample was selected from the set of households that had not responded to the NHS by mid-July 2011. It is important to note that at the end of these two weighting steps, some households have a weight of 1 because in some regions, all households are selected in the NHS sample.

Next, since a number of households in the subsample were still non-respondent at the end of collection operations, the sampling weight is adjusted for the survey's residual non-response. This is done by transferring the weights of non-respondent households to the nearest-neighbour respondent households. The latter are identified in a manner similar to the imputation process described in [Section 4.2](#), using known variables for respondent and non-respondent households, including census variables and a few variables resulting from matches to administrative databases.

Lastly, the weights are calibrated against census totals at the level of geographic calibration areas. Those areas contain an average of about 2,300 dwellings or 5,600 people in the NHS target population. They are formed by grouping dissemination areas so that they are contiguous, have enough respondent households to make calibration easy to perform, and do not straddle census division boundaries or,

Any significant change in survey method or content can affect the comparability of the data over time, and that applies to the NHS as well. It is impossible to determine with certainty whether, and to what extent, differences in a variable are attributable to an actual change or to non-response bias. Consequently, at every stage of processing, verification and dissemination, considerable effort was made to produce data that are as precise in their level of detail, and to ensure that the NHS's published estimates are of good quality in keeping with Statistics Canada standards.

Caution must be exercised when NHS estimates are compared with estimates produced from the 2006 Census long form, especially when the analysis involves small geographies. Users are asked to use the NHS's main quality indicator, the global non-response rate (see [Section 6.3](#)), in assessing the quality of the NHS estimates and determining the extent to which the estimates can be compared with the estimates from the 2006 Census long form. Users are also asked to read any quality notes that may be included in dissemination products.

Discrepancy between 2011 Census counts and 2011 NHS estimates

The final weights are selected so as to reduce or eliminate differences between the 2011 Census population counts and the NHS estimates. However, some discrepancies may persist because the weighting constraints sometimes have to be discarded. In addition, since the final weight adjustment is based on calibrated areas, some of which are made up of several small municipalities, there may be discrepancies between the NHS estimates and the census counts for small municipalities. The discrepancy between the population counts and the sample estimates is the difference between the NHS estimate and the 2011 Census count divided by the 2011 Census count.

Whether there is a discrepancy or not is an indication of the quality of the NHS estimates. For a given census subdivision (CSD) or any other geographic area, users are invited to compare the 2011 Census count with the NHS estimate for the same target population to get an idea of the quality of the NHS estimates. The larger the discrepancy is, the greater the risk of having poor-quality NHS estimates.

For CSDs with a population of 25,000 or more, the census count and the NHS estimate are practically identical. That is not always the case for smaller CSDs.

Comparisons of the 2011 Census population counts and the NHS population estimates at the CSD level for the same target population are presented in three figures in [Appendix 3](#). Comparisons are provided for CSDs with a population between 5,000 and 25,000, CSDs with a population between 1,000 and 5,000, and CSDs with a population between 40 and 1,000. Each figure shows the ratio of the NHS population estimate to the 2011 Census population count. If the ratio is equal or close to 1, the NHS population estimate is equal to the 2011 Census population count. If the ratio is greater than 1, the NHS estimate is greater than the 2011 Census count, and if the ratio is less than 1, the NHS estimate is less than the census count. The farther the ratio is from 1, the greater the risk of having poor-quality NHS estimates.

An analysis of the three figures shows that for small CSDs, there can be large discrepancies between the 2011 Census population count and NHS population estimate. As explained in [Section 4.3](#), those discrepancies are due to weighting, and as in any survey, they may be larger for small geographic areas. A similar analysis comparing the NHS estimates and the 2011 Census counts for common questions would also provide an idea of the quality of the NHS estimates.

5.5 Indicators of non-response bias

As noted in [Section 3.1](#), the higher a survey's non-response is, the greater the risk of non-response bias. During collection, the purpose of non-response follow-up, especially the subsample follow-up, was to maximize the survey's response rate and control potential non-response bias due to the survey's voluntary nature.

To assess the quality of the NHS estimates, in addition to the usual procedures (see [Section 5.3](#)), indicators of non-response bias were calculated and analyzed.

The indicators were calculated using a data file matching the 2006 and 2011 censuses. By means of a complex matching method using surnames, addresses and birthdates, 73% of 2011 Census respondents were linked to their 2006 records. As a result, we have 2006 Census data (including data from the long form) for a large portion of the NHS sample, whether the household responded or not.

These data made it possible (1) to compare NHS respondents and non-respondents for various characteristics measured in 2006, and (2) to calculate and analyze bias indicators and assess the quality of the NHS estimates. However, these analyses have some limitations, due to the nature of the matching file. It was impossible to match the entire NHS sample to the 2006 Census, and indicators could only be calculated for large geographic areas such as the provinces and territories, census divisions and census metropolitan areas.

It is important to keep in mind that these bias indicators are based on data from the previous census and not bias estimates calculated directly with 2011 NHS data. The indicators were used to assess the potential risk of bias for each geographic area. Analysis of these indicators and additional quality assessment analyses (see [Section 5.3](#)) provided assurance that the published NHS estimates meet Statistics Canada's quality standards. Notes are provided for variables and geographic areas for which some limitations on the quality of the NHS estimates must be taken into account.

6. Data dissemination for NHS standard products

6.1 Data suppression

The data that Statistics Canada disseminates are subject to various automated and manual processes to determine whether they should be suppressed. These processes are carried out to maintain **confidentiality** and **data quality**.

6.2 Suppression for confidentiality reasons

Confidentiality refers to the assurance that Statistics Canada will not disclose any information that could be used to identify respondents. Various confidentiality rules are applied to all data that are released or published to prevent the publication or disclosure of any information deemed confidential. If necessary, data are suppressed to prevent direct or residual disclosure of identifiable data. Consequently, geographic areas whose population is below a certain threshold are not published. For details on the confidentiality suppression thresholds, please see the Data Quality and Confidentiality Standards and Guidelines for the NHS.

6.3 Suppression due to estimate quality

Following the review of data quality, the dissemination of data whose quality is not considered satisfactory can be restricted if necessary. Quality indicators are produced for all standard place of residency geographies for which data are released.

The global non-response rate is an important measure of the quality of NHS estimates. It combines household and item non-response. This measure is used for the 2011 Census, just as it was in 2006 for dissemination of the Census, including the long form. In the specific case of the NHS, the global non-response rate is weighted to take account of the initial sample and the subsample used in non-response follow-up. It is calculated and presented for each geographic area.

As noted in [Section 3.1](#), there is non-response bias when a survey's non-respondents are different from its respondents. The higher the non-response is, the greater the risk of non-response bias. For the NHS, a number of measures were taken to mitigate the potential effects of non-response bias. Despite those efforts, the risk of non-response bias remains.

The global non-response rate is also used as a main dissemination criterion associated with the quality of the NHS estimates. For example, the NHS estimates for any geographic area with a global non-response rate greater than or equal to 50% are not published in the standard products. The estimates for such areas have such a high level of error that they should not be released under most circumstances.

The 50% threshold is based on studies of the global non-response rate in relation to the indicators of non-response bias (see [Section 5.5](#)). The studies showed that with a global non-response rate of 50% or more, the bias was so large that the estimates were not of sufficiently high quality.

At the Canada level, the NHS's global non-response rate is 26.1%. Item non-response made a much smaller contribution to the global non-response rate than household non-response. Table 2 shows the NHS's global non-response rate for Canada and for each province and territory.

Table 2 Global non-response rate of the 2011 National Household Survey, Canada, provinces and territories

Provinces and territories	Global non-response rate (%)
Canada	26.1
Newfoundland and Labrador	31.4
Prince Edward Island	33.4
Nova Scotia	28.2
New Brunswick	28.6
Quebec	22.4
Ontario	27.1
Manitoba	26.2
Saskatchewan	29.3
Alberta	27.4
British Columbia	26.1
Yukon	29.9
Northwest Territories	16.1
Nunavut	25.2

6.4 Coverage of published NHS data

Canada has a total of 147 census metropolitan areas (CMAs) and census agglomerations (CAs). For all of these areas, the global non-response rate is less than 50%, and published NHS data are available in standard products. In addition, NHS standard products are available for all 293 census divisions (CDs) and all 308 federal electoral districts (FEDs).

With a global non-response rate threshold of 50% for the release of NHS data, estimates are published for a majority of census subdivisions (CSDs), or municipalities. Of the 4,567 CSDs with an estimated population of more than 40 (for confidentiality reasons, those with a population of less than 40 are not published), NHS estimates are available in standard products for 3,439 (75.3%). Table 4 shows the distribution of published CSDs by province and territory. The proportion ranges from 100% for the Northwest Territories to 57.4% for Saskatchewan. Table 3 also shows the proportion of each province's and territory's population covered by published CSDs. It ranges from 100% for the Northwest Territories to 79.4% for Prince Edward Island. Overall, data are available for 96.6% of the Canadian population targeted by the NHS.

Table 3 Published data by census subdivision, 2011 National Household Survey, Canada, provinces and territories

Provinces and territories	Published census subdivisions		
	number	%	target population %
Canada	3,439	75.3	96.6
Newfoundland and Labrador	241	69.5	83.9
Prince Edward Island	77	70.0	79.4
Nova Scotia	76	85.4	96.4
New Brunswick	191	71.5	88.7
Quebec	979	84.3	97.8
Ontario	429	81.4	98.5
Manitoba	190	70.6	92.1
Saskatchewan	456	57.4	81.7
Alberta	293	75.1	96.7
British Columbia	437	82.6	97.2
Yukon	15	62.5	84.4
Northwest Territories	34	100.0	100.0
Nunavut	21	84.0	87.0

Note: CSDs not published for confidentiality reasons are excluded from this table. They have an estimated population less than 40.

The NHS is the largest voluntary survey ever conducted by Statistics Canada. During data collection, Statistics Canada used a wide variety of tools to encourage as many people as possible to complete the NHS. As a result, the final response rate was 68.6%, similar to the rates for Statistics Canada's other voluntary surveys.

In some small areas, the response rate was not high enough to produce a valid statistical picture. For those cases, users are encouraged to use data for a higher geography. For most areas, however, the responses received made it possible to produce good-quality estimates that will meet the needs of many users.

7. Appendices

Appendix 1 – List of questions in the 2011 NHS

Questions about the individual

- Q.1 Name
- Q.2 Sex
- Q.3 Date of birth and age
- Q.4 Marital status
- Q.5 Common-law
- Q.6 Relationship to Person 1
- Q.7 Difficulties with activities of daily living
- Q.8 Reduction of activities due to a physical or mental condition or health problems:
 - (a) at home
 - (b) at work or at school
 - (c) in other areas, for example, transportation or leisure
- Q.9 Place of birth
- Q.10 Citizenship
- Q.11 Landed immigrant status
- Q.12 Year of immigration
- Q.13 Knowledge of English or French
- Q.14 Knowledge of languages other than English or French
- Q.15 Language(s) spoken at home
 - (a) most often
 - (b) on a regular basis, but not as often as the main language reported in part (a)
- Q.16 First language learned at home in childhood and still understood
- Q.17 Ethnic and cultural origins
- Q.18 Aboriginal identity
- Q.19 Population group
- Q.20 Registered or Treaty Indian status
- Q.21 Membership in a First Nation or Indian band
- Q.22 Religion
- Q.23 Place of residence 1 year ago
- Q.24 Place of residence 5 years ago
- Q.25 Place of birth of parents
 - (a) father
 - (b) mother
- Q.27 Secondary (high) school diploma or equivalent
- Q.28 Registered Apprenticeship or other trades certificate or diploma
- Q.29 College, CEGEP, or other non-university certificate or diploma
- Q.30 University certificate, diploma or degree
- Q.31 Major field of study
- Q.32 Province, territory or country in which the certificate, diploma or degree was completed

- Q.33 School attendance
- Q.34 Hours worked for pay or in self-employment
- Q.35 Lay-off or absence from work
- Q.36 Arrangements to start a new job
- Q.37 Recent search for paid work
- Q.38 Availability for work
- Q.39 Date of last job
- Q.40 Name of employer
- Q.41 Kind of business, industry or service
- Q.42 Work or occupation
- Q.43 Main activities at work
- Q.44 Class of worker
- Q.45 Legal status of business (for self-employed workers)
- Q.46 Place of work
- Q.47 Method of travel to work
- Q.48 Length of commute
- Q.49 Language of work
 - (a) most often
 - (b) on a regular basis, but less often than main language reported in part (a)
- Q.50 Number of weeks worked in 2010
- Q.51 Full-time or part-time work
- Q.52 Amount paid for child care
- Q.53 Amount of support payments
- Q.54 Option to permit use of income tax return files
- Q.55 Income in 2010 (sources and amounts)

Questions about the dwelling

- E.1 Who pays the rent or mortgage, taxes, electricity, etc., for this dwelling?
- E.2 Is this dwelling owned by you or rented?
- E.3 Is this dwelling part of a condominium development?
- E.4 How many rooms and bedrooms are there in this dwelling?
- E.5 When was this dwelling originally built?
- E.6 Is this dwelling in need of any repairs?
- E.7 Is this dwelling located on an agricultural operation?
- E.8 What are the yearly payments for
 - (a) electricity?
 - (b) oil, gas, coal, wood or other fuels?
 - (c) water and other municipal services?
- E.9 What is the monthly rent?
- E.10 What are the owner's costs?

Appendix 2 – List of reference guides for NHS domains of interest

1. Aboriginal Peoples Reference Guide, National Household Survey
2. Ethnic Origin Reference Guide, National Household Survey
3. Languages Reference Guide, National Household Survey
4. Place of Birth, Generation Status, Citizenship and Immigration Reference Guide, National Household Survey
5. Religion Reference Guide, National Household Survey
6. Visible Minority and Population Group Reference Guide, National Household Survey
7. Education Reference Guide, National Household Survey
8. Labour Reference Guide, National Household Survey
9. Mobility and Migration Reference Guide, National Household Survey
10. Journey to Work Reference Guide, National Household Survey
11. Housing Reference Guide, National Household Survey
12. Income Reference Guide, National Household Survey

Appendix 3 – Comparison of 2011 Census population count and NHS population estimate by size of census subdivisions

Figure 3.1 Distribution of the ratio of the NHS population estimate to the 2011 Census population count, census subdivisions (CSDs) with a population of 5,000 to 24,999

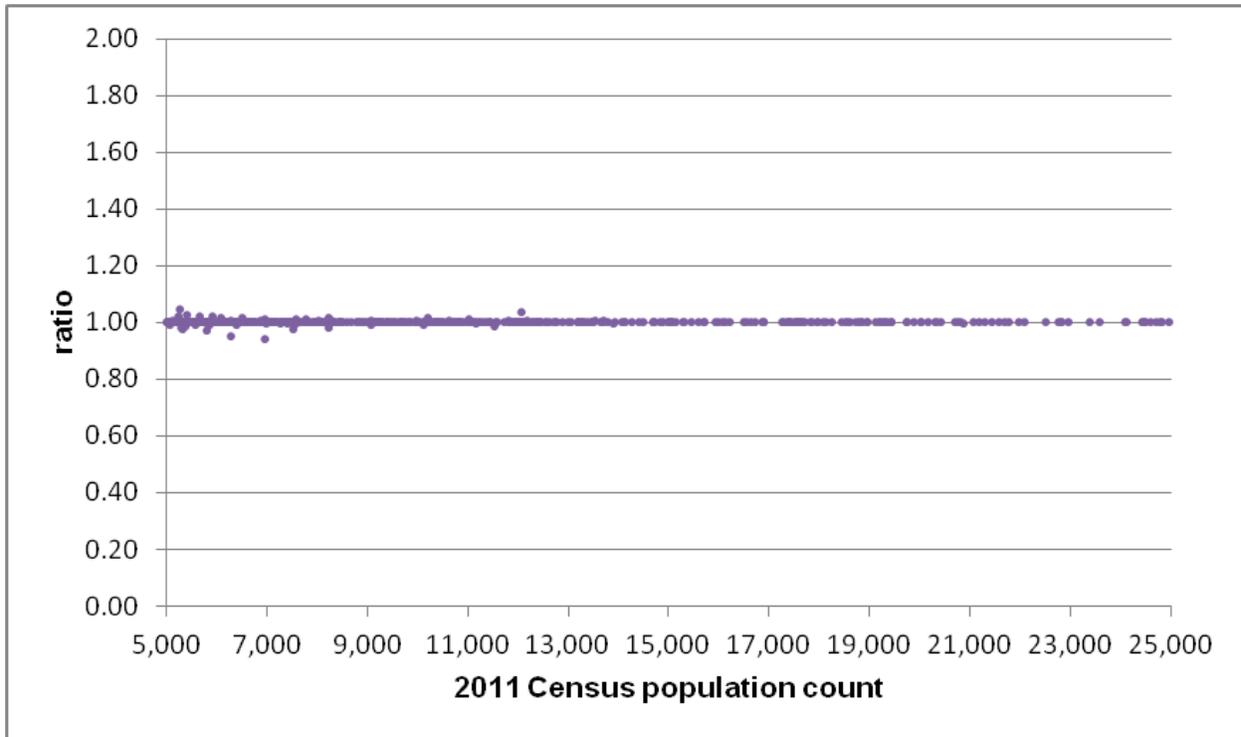


Figure 3.2 Distribution of the ratio of the NHS population estimate to the 2011 Census population count, census subdivisions (CSDs) with a population of 1,000 to 4,999

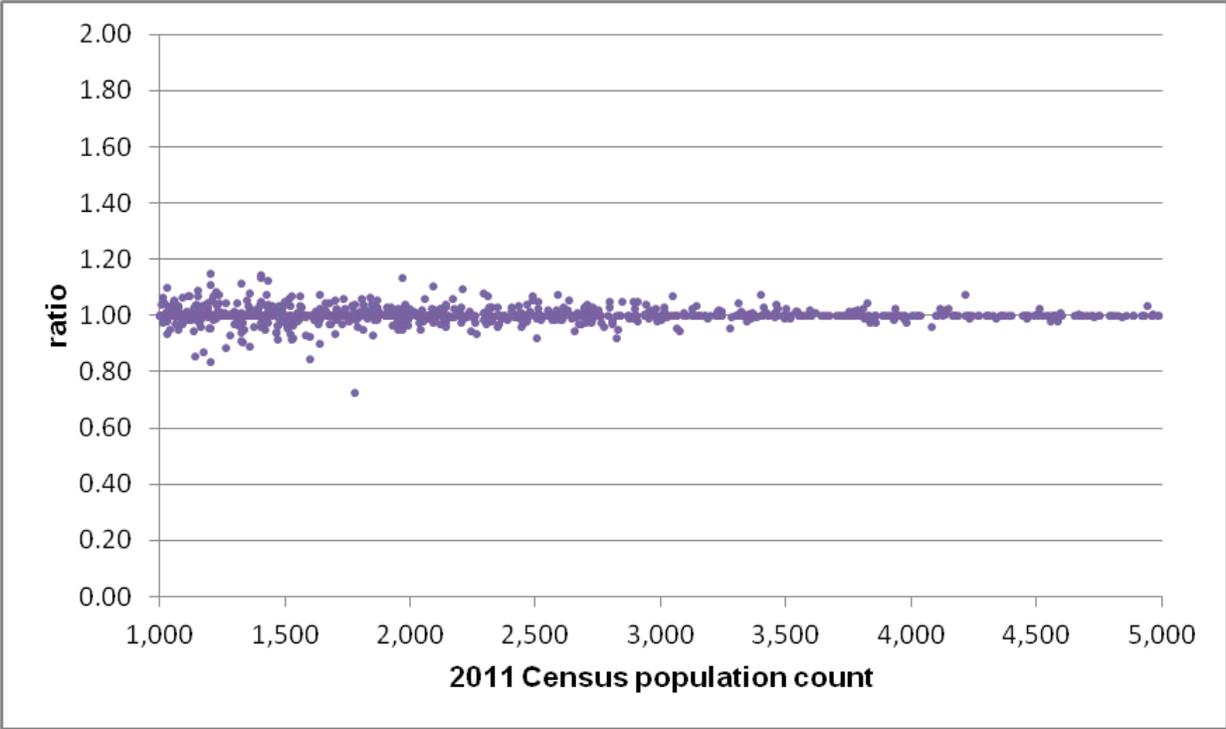


Figure 3.3 Distribution of the ratio of the NHS population estimate to the 2011 Census population count, census subdivisions (CSDs) with a population of 40 to 999

