CEP 16-10

Response Bias in Voluntary Surveys: An Empirical Analysis of the Canadian Census

Kerry Nield
Carleton University

Ardyn T. Nordstrom
Carleton University

June 2016; revised 9 August 2016

CARLETON ECONOMIC PAPERS

Department of Economics

1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

Response Bias in Voluntary Surveys: An Empirical Analysis of the Canadian Census *

Kerry Nield† Ardyn Nordstrom‡

June 2016

Abstract

In 2011, the National Household Survey replaced the traditional Long Form Census in Canada. The questions in the National Household Survey were similar to the Long Form Census, but responding to this survey was no longer mandatory. This paper provides an empirical analysis of the information loss arising from the change to a voluntary response policy. Comparisons of the differences between the non-mandatory 2011 National Household Survey and the 1996, 2001, and 2006 mandatory Long Form Census are used to identify changes related to the response policy. Using two-sample Kolgomov-Smirnov tests, differences in income distributions are tested to find that high income earners are underrepresented in the voluntary survey. This finding is corroborated by comparisons of various inequality measures across these time periods. Differences in discrete variables are tested using differences in proportions and Pearsons chi-squared tests. Systematic misrepresentation of certain groups is found in the voluntary survey. The switch to a voluntary response policy in 2011 likely led to an over representation of women and married individuals.

JEL classification: C42, D31, J11.
Keywords: Census, Demographics, Nonresponse Bias.

*We are grateful to Marcel Voia and Lynda Khalaf for their support.
†Email: kerry.nield@carleton.ca
‡Email: ardyn.nordstrom@carleton.ca

1 Introduction

In 2010, the Canadian government cancelled the mandatory long form census that had previously been conducted every five years to collect information on the population of Canada. This change was made against the recommendations of many researchers, who argued that the shift to non-mandatory reporting would significantly skew results and yield less accurate analysis. However, the 2011 Long Form Census was replaced with the non-mandatory National Household Survey (NHS). The Short Form Census was still mandatory and completed in 2011. In November of this year, the mandatory long form census was reinstated. This analysis determines whether making responses obligatory is a necessary step to getting high quality information from social surveys. High quality data is data that accurately reflects the characteristics and makeup of the population. There is limited empirical research on this topic currently.

This research attempts to determine if there are any systematic differences in the distribution of individuals in the voluntary long form survey compared with individuals in the mandatory long form survey. These differences would exist since certain types of individuals may no longer answer the survey if it is not required, which would bias the results if these individuals are systematically different. We employ several different tests, including Pearson's chi-squared test for differences in categorical variables and the Kolmogorov-Smirnov two-sample test to look for significant changes in continuous distributions. This research compares the voluntary 2011 NHS with the 2006 Census. A comparison of the 2001 Census to the 2006 Census, as well as the 1996 Census to the 2001 Census act as a baseline. Significant differences from 2006 to 2011 with insignificant differences from 2001 to 2006 and 1996 to 2001 would indicate that the non-mandatory response policy change has had a significant impact on the quality of the data. Systematic differences between those who choose to respond to the voluntary survey as compared to the population would result in changes of the distribution in certain variables. The variables of interest that are analyzed have been selected from the previous literature as key targets.

We find this is indeed the case: the survey results reflect changes that are not representative of the real population. There exist demographic differences, as well as ideological differences between respondents in the mandatory and voluntary surveys. We find that females are overrepresented in the voluntary survey, as are married individuals. Young individuals are underrepresented in the NHS. High-income earners are also underrepresented. While demographic differences can be corrected for by using the mandatory 2011 census weightings, differences in behaviour responses is a larger concern. Behavioural differences could not be corrected for if there is a systematic underrepresentation of certain ideological opinions or minority groups. This response bias has large implications for the validity of policy decisions, as well as research completed using the National Household Survey.

2 Literature Review

Much of the literature in this area of research relies anecdotal or subjective evidence and lacks empirical support for most main conclusions. There are a few exceptions; Rindfuss, Choe, Tsuya, Bumpass, and Tamaki (2015) uses three large government surveys in Japan

and several types of regressions to identify areas where nonresponse rates leads to bias in the sample. In Canada, Green and Milligan (2010) use data from multiple data sources, including the National Household Survey, to identify differences in income distributions across optional or enforced response surveys in Canada. Wright (2015) uses data from surveys conducted by Brandeis University to compare the means for characteristics of survey respondents to the characteristics of the population overall, for which data was available. Within the literature more generally there is a consensus that survey nonresponse rates do not immediately cause bias in the sample but that specific features of the population, the sample, or the survey design will lead to bias being more evident. While many practitioners identify nonresponse bias as an issue in survey results (Walton-Roberts et al., 2014; Rindfuss et al., 2015; Groves, 2006; Veall, 2010), the lack of empirical evidence to support this provides an opportunity to test these assertions. Given the change in response policies in Canada's census over the last decade, and the little empirical work done to evaluate these changes, this is an opportune time to measure the effect nonresponse bias has in the National Household Survey, if any. This research attempts to fill some of this gap in the literature.

In an analysis of response rates over time, Atrostic, Bates, and Silberstein (2001) analyse the response rates between 1990-1999 in five American government surveys: the Current Population Survey, the Consumer Expenditure Diary, the Consumer Expenditure Quarterly Survey, and National Crime and Victimization Survey, and the Survey of Income and Program Participation. They find that response rates fell for most of the 1990s, and that refusal rates climb substantially during this period to contribute to the decline in response rates. These findings provide motivation for additional research into the effect that high nonresponse rates will have on research.

Further motivation can be found in Walton-Roberts et al. (2014), which presents the outcomes of a multidisciplinary meeting to discuss the implications of eliminating the long form census in Canada. They emphasize the need for good data to drive evidence for decision making, providing multiple examples within Canadian policy including the identification of locations for interventions to prevent infant death in high-risk Aboriginal communities. In addition, they describe the census' role in describing changes in the dynamics of Canadian household decisions and outcomes, such as the change in mortality among different income groups over time. They describe the role the census plays in determining the size of different groups, which acts as a key denominator in research across disciplines. However, no empirical evidence is provided to support these claims.

Groves (2006) identifies the decline in response rates of voluntary surveys in the United States and notes that within academic circles response rates of 70-85% are deemed to be the threshold necessary to avoid bias. Groves argues that these thresholds are not necessary since bias does not always come from high nonresponse rates. Using discussion and mathematical proofs, he describes how bias in a variable's mean comes from the probability that an individual will respond and how that probability is correlated with the variable of interest. Extending this logic, Groves shows that a higher response rate could actually lead to higher mean biases in cases where the variable is highly correlated with the individual's probability of responding to the survey.

One of the few empirical analyses of voluntary response policies and nonresponse bias by Rindfuss et al. (2015) used data from three Japanese surveys (the 2000 and 2009 National Survey on Family and Economic Conditions, along with the 2009 National Survey on Family

and Economic Conditions). Using ANOVA, OLS, and probit in their analysis they find that biases are most prevalent when considering the distribution of certain variables but that analyses of the relationships between variables are less affected by high nonresponse rates. This paper highlights that over time response rates have been declining in voluntary surveys, although they show that these attitudes differ significantly across different countries. Other characteristics that may cause systematically different response behaviours between different groups of people include the interviewer's traits, hours of work, and a neighbourhood's social characteristics. In this survey in particular, the researchers found that response rates differed by gender, age groups, marital status, living situation, and education level (and that some of these group specific differences varied over time).

These results are corroborated by a literature review by Groves and Peytcheva (2008), which consisted of a survey of 59 studies that involved nonresponse issues they identify the factors of surveys most commonly associated with nonresponse bias. Specifically, the largest biases between the population mean and the mean of those who respond to the surveys were found in cases when only certain individuals has a history with the body conducting the survey (such as individuals who had been a patient of a hospital conducting a survey compared to someone who had never been treated there), or if the population is more general (such as an entire country rather than a university campus where the population may be more homogeneous). Biases are also more likely to arise with questions that address respondent attitudes than with questions focusing on behaviour.

Empirical work by Wright (2015) employed survey data from surveys sent by Brandeis University between 2006-2010, to determine the effect that nonresponses have on the distribution of characteristics in the sample. The university had demographic details on the entire population, so Wright is able to compare means of the response group to the means of the population overall. Although the population is fairly homogeneous the author finds statistically significant difference in the average gender, age, and religiosity distributions of the population compared to the survey respondents, which suggests that there are systematic differences between those who responded and those who did not. Another empirical analysis by Green and Milligan (2010) used the Survey of Labour and Income Dynamics (SLID) and its precursor; tax data from the Canadian Revenue Agency; and then compared these data sources' income distributions to the Long Form Census in comparable time periods in the late 1990s. By doing this, the authors find that voluntary surveys have a narrower income distribution so groups at the top of the distribution have underestimated incomes, while those at the bottom of the distribution end up with overestimated incomes. In addition to analyzing the effect of nonresponse bias on income this paper looks at the broader impacts of a voluntary census. Specifically, they emphasize that the census is used to determine the weights for all datasets in Statistics Canada so if these weights are biased it could spread bias to all Statistics Canada surveys. This research is particularly interesting for our current project since it focuses on the National Household Survey and provides a foundation for further research on the implications of the switch to a voluntary long form census.

(Veall, 2010) also focuses on the implications of Canada's transition to the voluntary long form census, however this article does not employ any kind of empirical work. Veall describes some of the key issues with the National Household Survey, such as the higher cost, and that inaccurate population estimates impact the weights used in other Statistics Canada surveys. This analysis notes that the National Household Survey may be better able

to avoid false responses since liars may have the incentive not to respond instead.

3 Data

For the purpose of this research, the public use microdata files for the Canadian National Household Survey from 2011, and the Canadian Census of the Population from 1996, 2001, and 2006 censuses are compared. The 1996, 2001, and 2006 censuses are conducted to collect social, economic, and demographic information from a representative sample of the Canadian population and contain data on 792,448, 801,055 and 844,476 individuals, respectively. The NHS from 2011 contains nearly identical questions and has a similar format to the previous censuses; however, responding to this survey was optional. The NHS dataset contains records for 887,012 individuals. The response rates for the Canadian Census in 1996, 2001, and 2006 were 99.2%, 98.4%, and 96.5%, compared with 77.2% for the 2011 NHS (Statistics Canada, 1996, 2001, 2006, 2011). The population of interest for these surveys includes individuals living in Canada who are not living in an institution at the time of the survey, as well as Canadians temporarily living abroad as diplomats or for other government service. The population excludes individuals in the Armed Forces or living on aboriginal reserves.

Since this research focuses on specific variables, observations with missing information on household income, education level, immigration status, age, or home ownership status are excluded from the sample. Instances of missing data are rare and result in less than 1% of the sample being excluded from each survey. In analyses of variables such as employment, household chores, education, and income additional observations have to be excluded since these questions are not addressed to individuals 15 or younger. This accounts for 15-20% of each dataset being excluded when considering variables that are not relevant to youth. The descriptive statistics for this adult subset are shown in Table 2.

The unweighted summary statistics[1] in Table 1 and Table 2 show that there are noticeable differences in the means of certain variables. For example, in 2011, 59% of the survey's adult respondents were married, compared with only 49% in 2006, 50% in 2001, and 52% in 1996. Similarly, in 2001 and 2006 the proportion of the sample that was single was 33% and 35%, respectively, however in 2011 only 30% of the sample was single. There was only a slight difference in the proportion of the samples that were male or female: 50.8% of 2006 sample, but 50.9% of the 2011 sample is female. Average nominal household income has been steadily rising within these samples: the average individual's real income is $28,695, $30,585, $35,067, and $36,707 in 1996, 2001, 2006, and 2011, respectively[2]. Education characteristics over time have also been changing, however these changes have been less consistent. In 2001, 15% of the sample reported having a university degree, however in 2006 and 2011, 22% and 26% of individuals reported having a university degree, respectively. Similarly, the proportion of people who had not completed high school fell from 33% to 24% between 2001 and 2006, and then fell again to 20% in 2011.

[1] Unweighted statistics are used since this analysis is considering the differences between individuals who have responded to the survey to determine if they are not accurately representing the population. If these are fundamentally different, this indicates the weights may not be appropriate to attempt to represent the population.

[2] These values refer to nominal incomes adjusted to 2006 dollars.

Table 1: Descriptive Statistics - Overall Population

Variable	**1996**	**2001**	**2006**	**2011**
Female	0.509	0.509	0.508	0.509
	(0.500)	(0.500)	(0.500)	(0.500)
Immigrant	0.224	0.206	0.190	0.179
	(0.417)	(0.404)	(0.393)	(0.384)
Minority	0.198	0.162	0.134	0.140
	(0.399)	(0.368)	(0.341)	(0.347)
Aboriginal	0.043	0.038	0.031	0.028
	(0.202)	(0.191)	(0.173)	(0.164)
Lives in Census Metropolitan Area	0.691	0.670	0.620	0.604
	(0.462)	(0.470)	(0.485)	(0.489)
Provinces of Residence				
British Columbia	0.133	0.130	0.131	0.129
	(0.339)	(0.337)	(0.337)	(0.336)
Alberta	0.107	0.104	0.100	0.094
	(0.310)	(0.305)	(0.300)	(0.291)
Saskatchewan	0.029	0.031	0.032	0.034
	(0.169)	(0.172)	(0.176)	(0.182)
Ontario	0.385	0.385	0.381	0.373
	(0.487)	(0.487)	(0.486)	(0.484)
Manitoba	0.036	0.036	0.037	0.039
	(0.186)	(0.187)	(0.189)	(0.193)
Quebec	0.241	0.238	0.240	0.247
	(0.427)	(0.426)	(0.427)	(0.431)
New Brunswick	0.021	0.023	0.024	0.026
	(0.144)	(0.150)	(0.154)	(0.158)
Nova Scotia	0.026	0.029	0.030	0.032
	(0.160)	(0.168)	(0.172)	(0.175)
Prince Edward Island	0.004	0.004	0.004	0.005
	(0.062)	(0.065)	(0.067)	(0.068)
Newfoundland and Labrador	0.014	0.016	0.017	0.019
	(0.119)	(0.126)	(0.130)	(0.137)
Observations	792,448	801,055	844,476	887,012

Note: Standard errors in brackets.

Table 2: Descriptive Statistics - Adult Population

Variable	**1996**	**2001**	**2006**	**2011**
Owns Their Place of Residence	0.699	0.718	0.734	0.749
	(0.459)	(0.450)	(0.442)	(0.434)
Real Income	28,149	30,585	35,067	36,707
	(28,695)	(30,190)	(55,697)	(51,393)
Dwelling Value	157,031	153,598	282,752	339,073
	(68,612)	(62,458)	(218,810)	(250,253)
Marital Status				
Married	0.524	0.503	0.484	0.589
	(0.499)	(0.500)	(0.500)	(0.492)
Single	0.321	0.338	0.349	0.279
	(0.467)	(0.473)	(0.477)	(0.449)
Widowed	0.052	0.051	0.057	0.049
	(0.222)	(0.219)	(0.231)	(0.216)
Divorced or Separated	0.103	0.108	0.111	0.083
	(0.304)	(0.311)	(0.314)	(0.277)
Highest Level of Education				
Did not Complete High School	0.364	0.328	0.237	0.194
	(0.481)	(0.470)	(0.425)	(0.395)
High School Graduate	0.232	0.231	0.256	0.252
	(0.422)	(0.422)	(0.436)	(0.434)
College Diploma	0.247	0.259	0.281	0.288
	(0.431)	(0.438)	(0.450)	(0.453)
University Degree	0.135	0.154	0.223	0.261
	(0.341)	(0.361)	(0.416)	(0.439)
Labour Force Status				
Employed	0.594	0.622	0.624	0.615
	(0.491)	(0.485)	(0.484)	(0.487)
Unemployed	0.067	0.050	0.044	0.052
	(0.249)	(0.217)	(0.205)	(0.221)
Not in the Labour Force	0.339	0.328	0.332	0.334
	(0.473)	(0.469)	(0.471)	(0.472)
Adult Observations	620,982	635,509	685,513	726,120

Notes: Income and dwelling value have been adjusted to 2006 dollars. Standard errors in brackets.

These are only sample means, which are not sufficient in determining whether the samples are significantly different. The fact that the means are exhibiting differences between 2006 and 2011 that are not similar to the differences between 2001 and 2006, and the differences between 1996 and 2001, suggest that changes in the sample distribution may be a result of the new response policy rather than changes in the population. The next section of this proposal discusses how these differences are evaluated more rigorously.

4 Methodology

By the nature of our research question, we have a model-free hypothesis. To study the variation caused by voluntary responses to a survey as compared to non-voluntary survey responses, this research aims to check if there are changes in distributions. The challenge is to isolate changes in the distribution that are due to response bias and not due to real changes in the population. If the 2011 NHS does have response bias, this would mean that the quality of the data is diminished. High quality data reflects the characteristics and makeup of the population. The mandatory censuses results are high quality data since the mandatory response policy ensures that virtually all individuals are represented. This information is to be used to analyze whether or not the 2011 NHS is also representative.

This research compares the distributions of key variables in the 2006 Census to the 2011 NHS. In order to ensure that these differences are related to the survey response and not real changes in the population, we compare these results with the 2001 Census to the 2006 Census, as well as the 1996 Census to the 2001 Census. Significant differences from 2006 to 2011 with minor differences from 2001 to 2006 and 1996 to 2001 would indicate that the non-mandatory response policy change impacted the quality of the data. Discrepancies between these surveys is an indication that the change in response policy has resulted in a lower quality dataset.

The variables of interest for this research are variables that previous research has identified as potentially related to response rates. These indicators are likely to influence whether or not an individual will respond to a non-voluntary survey. The variables that we looked at are: gender, age, immigrant status, minority status, residence in a metropolitan area, province of residence, home ownership, education, employment status, marital status and income.

4.1 Continuous Variables

This research employs several techniques to empirically analyze the differences in the distributions. For continuous variables, the Kolmogorov-Smirnov two-sample test is used. The resulting statistic answers the hypothesis that the distributions are statistically similar. The two sample statistic is computed for three cases: comparing the 2011 NHS to the 2006 Census, and comparing the 2006 Census to the 2001 Census, and the 2001 Census to the 1996 Census. A finding that there is a statistically significant difference in distribution from 2006 to 2011, but not in the other cases would indicate that there is a systematic difference in the response rate to the mandatory census as compared to the non-mandatory NHS. This analysis is completed for income, as it is the only continuous variable that is available in the

public use microdata files. The income values are normalized to 2006 level dollars across all years.

4.2 Inequality

This research also analyzes inequality measures on income. The measures are completed for the 1996, 2001, and 2006 Census, as well as the 2011 NHS. The income inequality measures that are analyzed are the Gini coefficient, the Atkinson index, and the Generalized entropy index. These indices are chosen because of their varying sensitivities to different parts of the distribution. The Gini coefficient is sensitive to changes in the middle of the distribution. The Atkinson index displays increasing sensitivity towards the lower tail of the distribution with increasing values of ϵ. The Generalized Entropy index is more sensitive to changes in the upper distribution with increasing values of α. Using these measures, we are able to identify relative changes in the tails of the distribution. A difference in inequality from the 1996, 2001 and 2006 Census to the 2011 NHS would provide further evidence that there has been a difference in the distribution of individuals answering the voluntary survey. The calculated inequalities are also compared with the inequalities available from CANSIM. As above, the income values are normalized for comparison. using CPI indexes, all variables are transformed to be real variables in 2006 dollars.

4.3 Discrete Variables

The public use microdata files do not allow access to any other continuous variables. Accordingly, other methods must be employed to measure differences in discrete variables. By measuring the difference in proportion of binomial and multinomial categorical variables across survey years, we can determine if there is a significant relationship between different surveys and the categorical variables of interest. The test statistic associated with these differences in proportions evaluates the null hypothesis that there is no difference in the proportion of individuals in each category from one census period to the next. However, the proportion of a population in a given demographic group will likely change across time. As such, the difference in difference of proportions is used to evaluate if the sample deivates from the trend. This statistic is calculated for two cases. In the first case, the null hypothesis tested is that there is no difference in trend across three time periods, while in the second case the null hypothesis tests that the trend has not changed in four time periods. These higher order differences in proportions allow us to determine whether the new sample deviates from trend. Any evidence to refute the null in this case would suggest that one or more of the census years is leading to a deviation from previous population trends. By combining these with first and second order differences in proportions provide we have insight into which census year is of interest and, consequently, whether the response policy of NHS may be responsible for changes seen in 2011. Specifically, we test for differences in gender, immigrant status, marital status, education levels, employment status, homeownership status, and province representation.

As a robustness check we have employed Pearson's chi-squared tests. This test predicts the probability that the difference between two survey years could have occurred by chance. We evaluate the null hypothesis that the values assumed by the categorical variables of

interest are independent of the survey response year. If the test's results reject the null hypothesis, it would indicate that there is a dependency between the categorical variables and survey response policy. These results are then compared to tests conducted for previous time periods to determine if any differences found from 2006 to 2011 are consistent with differences found from 2001 to 2006 and differences from 1996 to 2001. If these differences in differences are found to be significantly different, this suggests that the 2011 census deviates from the previous trend in the variable being examined and that the voluntary census captures a sample that is not representative of the population. This additional step is required because the difference from one year to another is not interesting by itself; it must be compared to previous trends to determine if the difference from 2006 to 2011 is economically or statistically significant.

The combined results from all of the above statistical tests provide insight on the quality of the 2011 NHS. Comparisons of previous high quality datasets (the 1996, 2001, and 2006 censuses) provide evidence that discrepancies in distributions in 2011 are not reflective of real changes in the population, but reflect changes in the census response policy. Evidence that even a single variable is subject to systematic response bias indicates that the quality of the voluntary survey is diminished.

5 Results

5.1 Continuous Variables

There are differences in the reported values in the public use microdata file between the different surveys. Income is top coded in the public use files for these surveys. This is a common practice, done for privacy related reasons. The 1996 and 2001 Census report income truncated to a maximum value of $200,000, in nominal values. The 2011 NHS and 2006 Census reports income to a maximum value of $1,124,045 and $1,285,586 respectively, again in nominal terms. In order to make a comparison of these distributions, we use the CPI index to adjust the income to real values in 2006 dollars.

The results of the Kolmogorov-Smirnov two-sample test reject the null hypothesis of equal distributions in all cases. This illustrates that the income distributions are statistically different from each other in all cases. Table 3 shows the results of this analysis. Distance represents the greatest distance found between the two distributions. The P-value shows the probability that the two distributions are the same. These results show that the 2001 distribution is statistically different from and dominates the 1996 distribution. We also find that the 2006 distribution is statistically different from the 2001 distribution. However, there is evidence that 2006 dominates 2001, and vice versa. This provides evidence for higher order stochastic dominance. These tests are calculated using a subset of real incomes for which all years report values. We also compare the distributions for income in the range of $0 to less than $223,108 for all years in real 2006 dollars. This is done since the last observation ($223,108) in the 2001 Census has 2,430 observations. These are individuals that earn greater than $200,000 dollars in 2001 dollars ($223,108 in 2006 dollars.) Therefore, including this last value is not representative of the distribution function. The case comparing the 2011 distribution to the 2006 distribution also finds evidence for high order stochastic dominance.

Table 3: Kolgomov Smirnoff Test 2-sample test statistic

Smaller Group	Distance	P-value	Data Range
2006	0.0396	0.00	full
2011	-0.0307	0.00	full
COMBINED 2006-2011	0.0396	0.00	full
2006	0.0401	0.00	subset
2011	-0.0310	0.00	subset
COMBINED 2006-2011	0.0401	0.00	subset
2001	0.0308	0.00	subset
2006	-0.0173	0.00	subset
COMBINED 2001-2006	0.0308	0.00	subset
1996	0.0406	0.00	subset
2001	-0.0004	0.90	subset
COMBINED 1996-2001	0.0406	0.00	subset

Note: These calculations are performed using adjusted real income values in 2006 dollars. "Full" refers to calculations completed for the real 2006 income range of ($0, $1,022,797). "Subset" refers to calculations completed using the real 2006 income range of ($0, $223,108).

The two-sample Kolmogorov-Smirnoff test rejects the null hypothesis of equal distributions and provides evidence of higher stochastic dominance. For a robustness check, the 2011 and 2006 distributions were also compared using an expanded range of income values. Using a range of $0 to less than $1,022,797, the null hypothesis is again rejected, and again evidence of higher level stochastic dominance is found.

From this analysis, we conclude that the distributions of income have changed for every Census and the NHS. This is illustrated in Figure 1, as the CDF (cumulative distribution function) of each year does not consistently directly overlap with other years. In particular interest, looking at the lower end of the distribution, we see that there is an irregularity in the CDF of 2011 as compared to the other years' distributions. This can be noted in Figure 2. The CDF of 2011 follows a different path than previous years. The 2006 CDF appears to dominate the 2011 CDF for incomes below approximately $7,500. This is an indication that in 2011, individuals who made lower incomes were a larger percentage of the surveyed population than was the case in 2006.

As noted above, high values were top coded in 2006 and 2011. The top coded value in 2011 was $1,022,797 (2006 dollars), and 201 individuals are coded to have earned this income (or greater). In 2006, the top coded value was $1,285,586 (2006 dollars). There were 202 individuals who were coded to have earned this income (or greater.) In fact, if we look at number of individuals coded to have earned greater than $500,000 (2006 dollars), there are 1,817 respondents in 2006, and only 1,597 respondents in 2011. This is a decrease of 12.1% in this top income group. This provides strong evidence that agrees with Green and Milligan (2010). We find that the upper tail of the distribution has fallen out, and thus the lower tail of distribution is over represented.

Figure 1: CDF of Income

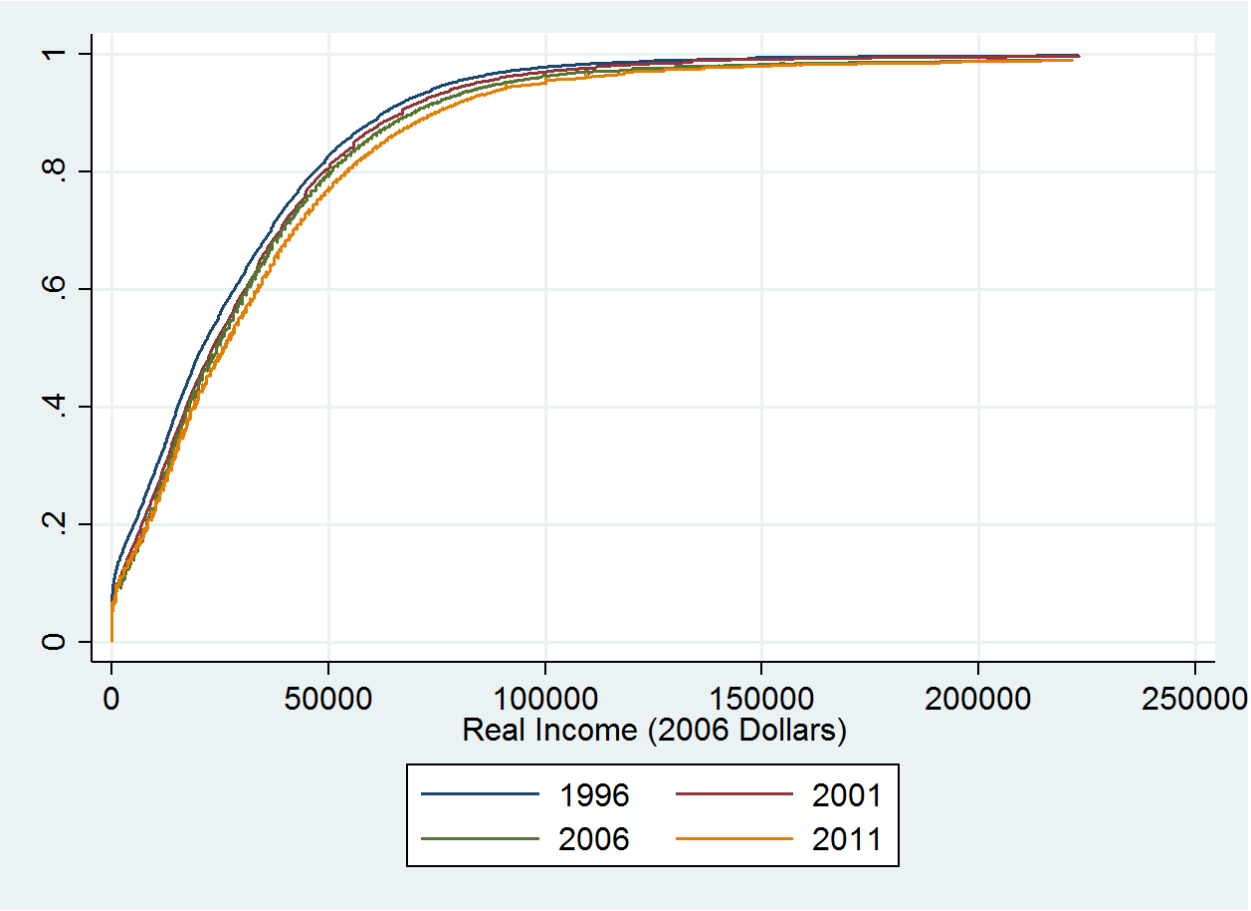

Figure 2: CDF of Income

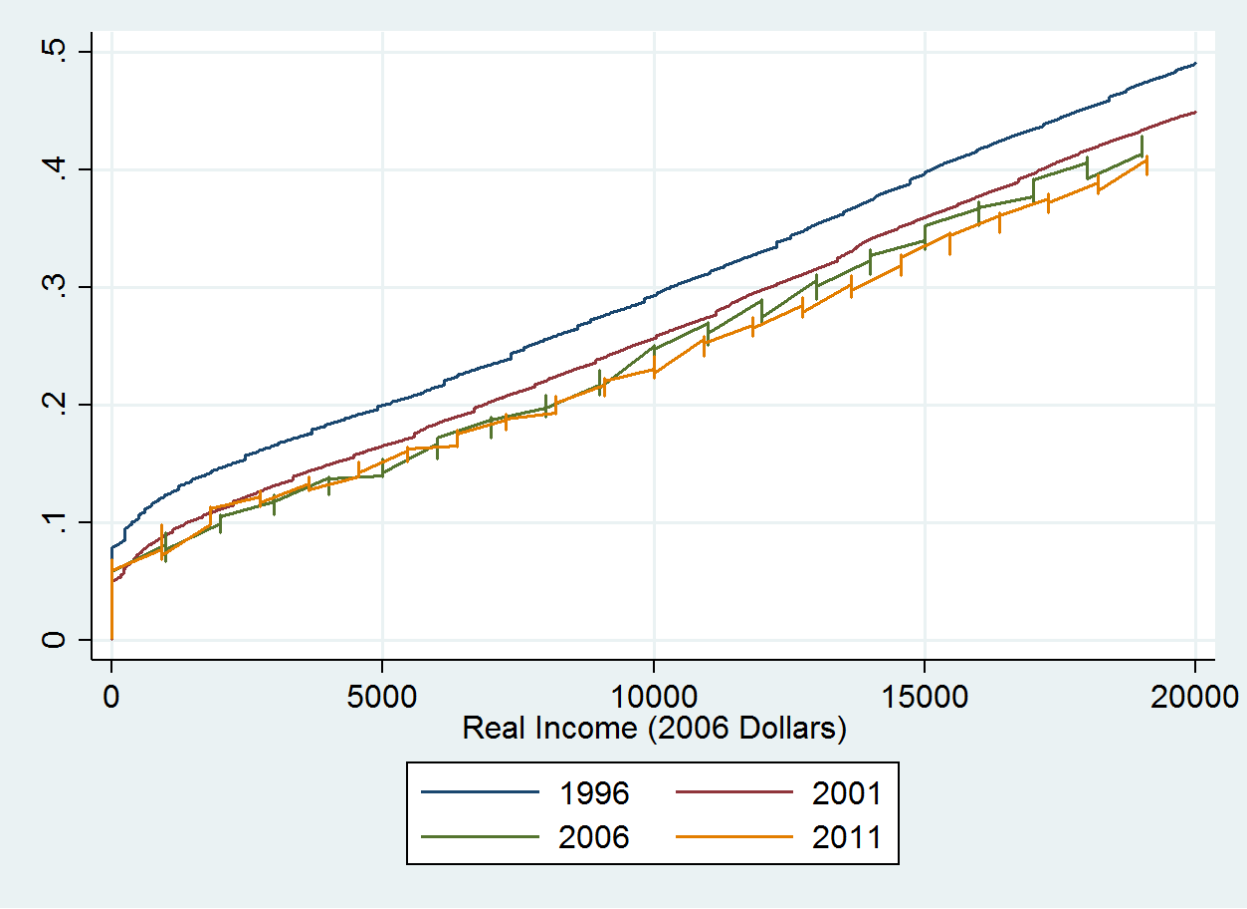

Table 4: Income Inequality Indices

	GE(-1)	GE(0)	GE(1)	GE(2)	A(0.5)	A(1)	A(2)	Gini
1996	5.20781	0.53394	0.35414	0.39003	0.18813	0.41371	0.9124	0.45693
2001	4.98961	0.51488	0.34822	0.38553	0.18400	0.40243	0.90892	0.45327
2006	345.85096	0.59859	0.34473	0.39300	0.18143	0.45042	0.99856	0.45033
2011	361.55527	0.60735	0.35213	0.40014	0.18587	0.45521	0.99862	0.45458

Note: The Inequality Indices in this table have been calculated using adjusted real income values in 2006 dollars. They are calculated for the range ($0, $223,108).

5.2 Inequality

Comparing the inequality measures from 2011 to 2006 with this information in mind, further evidence to support the shift in income representation in the voluntary survey is found. As mentioned above, the generalized entropy index increases the focus on the upper tail as α increases. This indicates that lower values of α emphasis the lower tail of the distribution. Table 4 shows the results of these inequality measures calculated for each year using real income values from 2006 truncated to values less than $223,108. These are values that all four years have data. As can be noted in Table 4, the generalized entropy index finds that for all values of α, inequality has increased from 2006 to 2011. As α decreases, the difference in inequality measures between 2011 and 2006 for the generalized entropy indexes increases. This means that as the focus shifts to the lower end of the distribution, larger differences in inequality are perceived. This is due to the lower portion of the income distribution being over represented, as shown above.

In Table 5, the shift in representation is also noted. In this table, the indexes are calculated for real 2006 incomes from 0 to less than $1,022,797. This is the range (not including top coded values) for which data in 2006 and 2011 is available. Inequality is a measure of the spread of income distribution. As the top portion of the income responding to the survey has decreased, this would imply less inequality. The GE(2) and A(0.5), as they focus on the upper end of the distribution, find evidence that inequality has decreased. On the reverse, the inequality indices that focus on the lower tail find increased inequality due to the relatively larger portion of the population that is very poor. Focusing on the lower distribution of income, the GE(-1), GE(0), and A(2) index show that inequality has increased. These findings agree with the previous evidence: high income earners do not respond to the voluntary survey, and thus low income earners are over represented.

Comparing the Gini coefficient from 2006 to 2011 indicate that inequality has decreased. The provides further evidence that the upper portion of the income distribution has fallen out of the sample. These findings agree with Green and Milligan (2010), and can be evidenced by Hou, Picot, et al. (2014). This report shows that the inequality measures found using taxation data displays greater inequality than the inequality measures found using voluntary surveys.

Table 5: Income Inequality Indices: 2006 and 2011

	GE(-1)	GE(0)	GE(1)	GE(2)	A(0.5)	A(1)	A(2)	Gini
2006	383.73822	0.68133	0.49945	1.01312	0.22619	0.49406	0.99870	0.50058
2011	394.59514	0.67491	0.47104	0.81991	0.22152	0.49080	0.99873	0.49619

Note: The Inequality Indices in this table have been calculated using adjusted real income values in 2006 dollars. They are calculated for the range ($0, $1,022,797).

5.3 Discrete Variables

In order to compare the differences in distributions of categorical variables in the NHS it is insufficient to compare the 2011 NHS to the 2006 census values, as variables often change over time. Table 6 shows why this is necessary; apart from gender, each categorical or binary variable represents a significantly different proportion of each census sample when compared to the previous census. Therefore, we compare these differences to previous trends using higher-order differences in proportions.

The proportion of the population that identified as a minority falls by 1.8% from 1996 to 2001; falls by 1.6% from 2001 to 2006; then fell by 1.1% from 2006 to 2011. The proportions of married individuals falls by 2.1% from 1996 to 2001; falls by 1.9% from 2001 to 2006; then rises by 10.4% from 2006 to 2011. The proportion of females in the population changes slightly from year to year as well: it remained constant from 1996 to 2001, but fell by 0.1% from 2001 to 2006 then rose by 0.1% from 2006 to 2011. While these differences are interesting on their own, they do not provide conclusive evidence that the 2011 survey's unique response policy was the potential of these differences. Further differences in differences must be considered to make these conclusions.

5.3.1 Higher-Order Differences

Using point estimates, these first differences could be compared to one another. By looking at the difference between 2011 and 2006, and comparing this to the change between 2006 and 2001, this analysis finds that there are significantly higher proportions of women responding to the NHS, leading us to reject the null hypothesis that there is no difference in the trend across these three time periods. We know this is the case because there was an increase from 2006 to 2011 but that when compared to previous trend this is a deviation from previous population trends. These differences are summarized in Table 7.

By looking at the third column in Table 7, which describes the ultimate difference from the overall trend from 1996 through to 2011, it is clear that this overall deviation is statistically significant[3] and amounts to an overrepresentation of females in the NHS of 0.3%. This is equivalent to approximately 2,661 more women in the sample than would be if the pre-existing trend extended to 2011, which is not an unreasonable assumption since the proportion of males to females is not intuitively expected to change.

There are similar patterns in the proportion of married individuals. Using this same

[3]This difference is significant at the 1% level, leading us to reject the null hypothesis that there is no difference in gender proportion trends between 2001 and 2011.

Table 6: First Differences in Proportions Between Subsequent Years

Variable	2001 - 1996	2006 - 2001	2011 - 2006
Overall Population			
Female	0.000	-0.001	0.001
Immigrant	-0.018***	-0.016***	-0.011***
Minority	-0.036***	-0.028***	0.006***
Aboriginal	-0.005***	-0.007***	-0.003***
Lives in Census Metropolitan Area	-0.021***	-0.050***	-0.016***
Province of Residence			
British Columbia	-0.002***	0.000	-0.001***
Alberta	-0.003***	-0.004***	-0.006***
Saskatchewan	0.001***	0.001***	0.002***
Manitoba	0.000*	0.001***	0.001***
Ontario	0.000	-0.004***	-0.008***
Quebec	-0.003***	0.002***	0.007***
New Brunswick	0.002***	0.001***	0.001***
Nova Scotia	0.003***	0.001***	0.001***
Prince Edward Island	0.000***	0.000*	0.000*
Newfoundland and Labrador	0.002***	0.001***	0.002***
Adult Population			
Marital Status			
Owns Their Place of Residence	0.018***	0.016***	0.015***
Married	-0.021***	-0.019***	0.104***
Single	0.017***	0.011***	-0.070***
Widow	-0.001***	0.006***	-0.007***
Divorced or Separated	0.005***	0.002***	-0.027***
Highest Level of Education			
Did Not Complete High School	-0.036***	-0.092***	-0.043***
High School Graduate	0.000	0.024***	-0.003***
College Diploma	0.012***	0.022***	0.007***
University Degree	0.019***	0.069***	0.038***
Labour Force Status			
Employed	0.028***	0.002***	-0.010***
Unemployed	-0.017***	-0.006***	0.008***
Not in the Labour Force	-0.011***	0.004***	0.002***

*** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level

Table 7: Differences in Differences in Proportions

Variable	(2006-2001) - (2001-1996)	(2011-2006) - (2006-2001)	(2011-2006) - (2006-2001) - (2006-2001) - (2001-1996)
Overall Population			
Female	-0.0012	0.0017**	0.003***
Immigrant	0.0026***	0.0047***	0.002 ***
Minority	0.0083***	0.0337***	0.025***
Aboriginal	-0.0018***	0.0033***	0.005***
Lives Census Metropolitan Area	-0.0290***	0.0344***	0.063***
Province of Residence			
British Columbia	0.0022***	-0.0013***	-0.003***
Alberta	-0.0009**	-0.0020***	-0.001**
Saskatchewan	-8.87E-05	0.0012***	0.001***
Manitoba	0.0005*	0.0004	0.000
Ontario	-0.0042***	-0.0036***	0.001
Quebec	0.0050***	0.0041***	-0.001
New Brunswick	-0.0007***	9.44E-05	0.001***
Nova Scotia	-0.0013***	-0.0002	0.001***
Prince Edward Island	-0.0002*	1.26E-05	0.000**
Newfoundland and Labrador	-0.0006***	0.0010***	0.002***
Adult Variables			
Owns Their Place of Residence	-0.0024***	-0.0006	0.002***
Marital Status			
Married	0.0017*	0.1236***	0.122***
Single	-0.0060***	-0.0804***	-0.074***
Widow	0.0073***	-0.0134***	-0.021***
Divorced or Separated	-0.0030***	-0.0298***	-0.027***
Highest Level of Education			
Did Not Complete High School	-0.0557***	0.0486***	0.104***
High School Graduate	0.0245***	-0.0276***	-0.052***
College Diploma	0.0098***	-0.0151***	-0.025***
University Degree	0.0497***	-0.0308***	-0.081***
Labour Force Status			
Employed	-0.0261***	-0.0118***	0.014***
Unemployed	0.0112***	0.0136***	0.002***
Not in the Labour Force	0.0149***	-0.0017**	-0.017***

*** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level

method, there are 12.2% more married individuals in the 2011 NHS than there should have been according to previous population trends[4]. Alternatively, looking at the deviations from the trend in the proportions of individuals who own their residences, these higher order differences show that there was not a significant change in the proportion of individuals who own a home. Previous trends would have suggested that the deviation in the change in homeownership should have lower than it actually was, suggesting that homeowners are also overrepresented in the NHS. The fact that such a wide variety of variables are exhibiting this large, sudden, deviation from 15-year trends suggests that the change in response policy in 2011 led to these changes in proportions, rather than some social phenomenon.

Table 7 shows that there are many characteristics which have non-constant fluctuation between all time periods. This makes it impossible to conclude what impact the census response policy has on the representation of groups in the NHS sample. For example, the proportion of immigrants in the population falls by 1.83% from 1996 to 2001, by 1.57% from 2001 to 2006, and by 1.10% from 2006 to 2011. However, the higher order differences in Table 7 show that these differences are highly non-constant, so it is not really possible to conclude whether there are over or under representations of immigrants in the NHS with this methodology. This holds true for many characteristics, including employment statuses and education levels. This is not entirely surprising since so many policies and exogenous trends influence these variables other than the census response policy. However, the fact that meaningful deviations are found in fundamentally important factors like gender supports the conclusion that the NHS led to systematic differences between the NHS sample and the Canadian population.

As a robustness check, these results are compared to results from Pearson's chi-squared tests, which are used to test whether differences between categories in different survey years were likely just the result of chance. The results of the Pearson's chi-squared tests, briefly described in Table 8 showed that there were significant differences in the proportion of married individuals between all back-to-back years. However when the difference in these differences was compared so the change from 2006 to 2011 was compared to the change from 2001 to 2006, there were deviations from the trend across 2001 to 2011 that were significantly different[5] from the difference between 1996 and 2006 overall. Similarly, the same deviation from this trend was found for the proportion of the population that owns their residence[6]. This supports the conclusion that individuals who respond to the 2011 NHS are significantly different from the population they are sampled from since the marriage and home ownership characteristics deviate from the trends that the population had previously been following. Intuition suggest that these differences are due to the survey's response policy rather than just socio-cultural movements differences since major breaks from 15-year trends in such a wide variety of variables would be extremely unlikely. Moreover, these kinds of differences from the trend were not seen until 2011, which lends further support to this conclusion.

[4]This difference is significant at the 1% level

[5]At the 5% significance level

[6]At the 5% significance level

Table 8: Pearson's Chi-Squared Difference in Proportions Tests

Variable	Difference in Significance (At 5% Significance Level)				
	1996 vs. 2001	**2001 vs. 2006**	**2006 vs. 2011**	**(2006 - 2001) vs. (2001 - 1996)**	**(2011 - 2006) vs. (2006 - 2001)**
Married	Different	Different	Different	Not Different	Different
Owns Place of Residence	Different	Different	Different	Not Different	Different
University Degree	Different	Different	Different	Different	Different

5.3.2 Age Distribution

This analysis considers changes in the distribution of age groups in each of the census years. Table 9 shows that until 2011 there were no significant changes in the proportion of the population that was within the ages of 20-24. In 2011 there were significantly fewer people within this age group than the 2006 adjusted ages would have predicted[7]. The fact that this deviates from a previous trend of no change may signal that this age group is more likely to not respond to the census, leading to their being underrepresented[8]. This is consistent with previous research by Rindfuss et al. (2015). Similarly, the proportion of middle-aged individuals was significantly different from previous trends, leading individuals ages 50-54 to be overrepresented (based on previous trends) and individuals ages 55-59 to be underrepresented (based on previous trends)[9].

6 Conclusion

This analysis focuses on changes in trend associated with unconditional proportions. By evaluating the income distribution of each sample year with the Kolmogorov-Smirnov two-sample test, we found that the income distributions have changed across all the time periods. Additional tests using a truncated income distribution suggest that high income individuals respond less to the voluntary survey. This conclusion was robust and that the upper tail of income distribution is not represented in the 2011 NHS while the lower tail is overrepresented.

Difference in proportions tests to compare the proportions of married, female, immigrant, and homeowners in a the 2011 NHS. By comparing these proportions to previous years, then comparing these differences to previous trends in by using differences in differences this analysis finds that the proportions of married individuals, females, and homeowners in the 2011 NHS are statistically and intrinsically different from previous trends.

These findings suggest that there is a change in the distribution of individuals who

[7]The number of individuals in each age group are adjusted to reflect the size each group as it would have been in the most recent year of the two years being compared. For example, 1996 age groups were adjusted up by 5 years to be compared to 2001 age groupings.

[8]This deviation is significant at the 1% significance level

[9]These deviations are significant at the 1% significance level

Table 9: Age Group Differences Between Subsequent Years

Age Group	2006 vs. 2011	2001 vs. 2006	1996 vs. 2001
20-24	Decrease***	No change	No change
25-29	Increase***	Increase*	Decrease***
30-34	Increase***	Increase***	Increase***
35-39	Increase***	Increase***	Decrease***
40-44	Increase***	Increase***	Decrease***
45-49	Increase**	No change	Decrease***
50-54	Increase***	No change	No change
55-59	Decrease***	No change	No change
60-64	Decrease***	Decrease***	No change
65-69	Decrease***	Decrease***	No change
70-74	Decrease***	Decrease***	Decrease***
75-79	Decrease***	Decrease***	Decrease***

*** significant at 1% level; ** significant at 5% level; *significant at 10% level

Note: These ages are adjusted to represent the ages at the time of the most recent year in each two-year comparison.

responded to the previous censuses. By looking at previous changes in trends, this analysis finds that these changes in distribution are actually result of the change in response policy rather than the change in the population itself. This is particularly concerning because the census has such far-reaching implications. If groups of individuals with certain traits or behavioural characteristics are over or underrepresented, as we find in this study, then the voluntary census will lead to poor policy decisions.

In the future, this analysis would benefit from the use of the confidential census files. The public use microdata files censor the upper tails of the income distribution and change continuous variables like age into discrete variables. Access to the confidential files would allow this study to more accurately determine differences in distributions.

References

Atrostic, B. K., Bates, N., & Silberstein, A. (2001). Nonresponse in US government household surveys: consistent measures, recent trends, and new insights. *Journal of Official Statistics*, *17*(2), 209.

Green, D. A., & Milligan, K. (2010). The importance of the long form census to canada. *Canadian Public Policy*, *36*(3), 383–388.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 646–675.

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias a meta-analysis. *Public opinion quarterly*, *72*(2), 167–189.

Hou, F., Picot, G., et al. (2014). *Immigration, low income and income inequality in canada: What? s new in the 2000s?* (Tech. Rep.). Statistics Canada, Analytical Studies Branch.

Rindfuss, R. R., Choe, M. K., Tsuya, N. O., Bumpass, L. L., & Tamaki, E. (2015). Do low survey response rates bias results? Evidence from Japan. *Demographic Research*, *32*, 797.

Statistics Canada. (1996). *Surveys and statistical programs - census of population: Detailed information for 1996.* `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=2999`. ((Accessed on 02/14/2016))

Statistics Canada. (2001). *Surveys and statistical programs - census of population: Detailed information for 2001.* `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=30217`. ((Accessed on 02/14/2016))

Statistics Canada. (2006). *Surveys and statistical programs - census of population: Detailed information for 2006.* `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=30219`. ((Accessed on 02/14/2016))

Statistics Canada. (2011). *National household survey: Final response rates.* `https://www12.statcan.gc.ca/nhs-enm/2011/ref/about-apropos/nhs-enm_r012.cfm?Lang=E`. ((Accessed on 02/14/2016))

Veall, M. R. (2010). 2B or not 2B? What should have happened with the Canadian long form Census? What should happen now? *Canadian Public Policy*, *36*(3), 395–399.

Walton-Roberts, M., Beaujot, R., Hiebert, D., McDaniel, S., Rose, D., & Wright, R. (2014). Why do we still need a census? views from the age of "truthiness" and the "death of evidence". *The Canadian Geographer/Le Géographe canadien*, *58*(1), 34–47.

Wright, G. (2015). An empirical examination of the relationship between nonresponse rate and nonresponse bias. *Statistical Journal of the IAOS*, *31*(2), 305–315.