



# CCPA Review

CANADIAN CENTRE FOR POLICY ALTERNATIVES - MANITOBA

**March 23, 2015**

## Low response in National Household Survey leads to bias

**T**his article is about data and non-response bias. Wait, wait, don't put the article down just yet, it's about politics too.

When the Government of Canada decided to cancel the long form census for 2011 and replace it with the voluntary National Household Survey (NHS) many expressed concern NHS data would be less reliable and vulnerable populations would be less likely to participate. These concerns have proven to be correct.

NHS data for small areas must be treated with caution but it is impossible to tell just how much non-response bias occurred and who really was missed. It may very well be that immigrant and Aboriginal persons or the poor are underrepresented in NHS data but we should not accept these assertions without evidence.

### **Correlation Analysis**

In order to examine data quality and suppression in the 2011 NHS I

**Correlation (Pearson's r value) for Selected 2006 Census Variables to 2011 National Household Survey Global Nonresponse Rate (GNR) by Census Tract for Winnipeg, Inner City and Non Inner City**

Selected 2006 Census Variables	Winnipeg	Inner City	Non Inner City
% reporting Non-official language Mother Tongue	0.169	0.426	-0.139
% of Recent immigrants (arrived 2001 to 2006)	0.234	0.208	0.112
% of Aboriginal identity population	0.503	0.462	0.322
Prevalence of low income after tax in 2005 %	0.260	0.089	0.113
Median 2005 after-tax income	-0.429	-0.576	-0.219
Average 2005 after-tax income	-0.377	-0.580	-0.184

Source: Statistics Canada 2006 Census table 94-578-XCB2006005, 2006 Census table 97-563-XCB2006043, 2006 Census table 97-563-XCB2006012

Inner City boundary based on Census Tracts

### **CCPA-MB**

205 - 765 Main St.  
Winnipeg, MB  
R2W 3N5

### **PHONE**

204.927.3200

### **EMAIL**

CCPAMB@  
POLICYALTERNATIVES.CA

### **WEBSITE**

WWW.  
POLICYALTERNATIVES.CA

### **BLOG**

WWW.POLICYFIX.CA

### **TWITTER**

@CCPAMB

RESEARCH - ANALYSIS - SOLUTIONS

In other words, the higher one's income, the more likely one is to respond to the voluntary National Household Survey; the lower one's income, by contrast, the less likely one is to respond.

looked at a few select variables from the 2006 Census at the Census Tract level for Winnipeg and ran a correlation between these variables and the Global Nonresponse Rate (GNR) from the 2011 National Household Survey.<sup>1</sup>

The correlations shown below are a proxy measure to see whether particular variables are associated with non-response. The data do suggest that certain segments of the population are less likely to have participated in the voluntary National Household Survey. The results provide both interesting and puzzling data.

There appears to be a slight relationship between non-official mother tongue (that is, neither English nor French), and the global nonresponse rate. City wide there is a positive relationship: as the percentage of population with a non-official language mother tongue increases, the global nonresponse rate increases. In the inner city the relationship is ( $r = 0.426$ ). For the non-inner city there is a slight negative relationship, that is, as the percentage reporting a non-official language mother tongue increases, the global nonresponse rate goes down.

For recent immigrants, there does appear to be a positive relationship; as the percentage of recent immigrants goes up, so too does the global nonresponse rate.

The "strongest" relationship among these selected variables is for the Aboriginal population. A correlation ( $r = 0.503$ ) suggests that as the percentage of the Aboriginal popula-

tion goes up, so too does the global nonresponse rate.

The data also suggest a negative relationship between income and response rate. As the median and average income goes up, global nonresponse rates go down, especially in the inner city. In other words, the higher one's income, the more likely one is to respond to the voluntary National Household Survey; the lower one's income, by contrast, the less likely one is to respond. This is consistent with what was feared by those who expressed concerns about the voluntary character of the NHS.

Of course, a correlation does not suggest one variable is causing the other so we have to be careful with words such as "strength of a relationship" but the data does suggest a relationship.

Calculation of the coefficient of determination (which is simply the  $r$  value squared) does provide some indication of how much that change in the dependent variable, in this case the global nonresponse rate (GNR), is explained by the independent variable.

For example, if we were to square the  $r$  value for the Aboriginal population, the coefficient of determination would be 0.253, suggesting

<sup>1</sup>Five (5) Census Tracts were not included in the correlation analysis because geographical boundaries changed between 2006 and 2011 and there is no GNR for them in 2011. These Census Tracts were: CT 0110.05; CT 0150.00; CT 0500.05; CT 0510.01, and; CT 0590.01.

<b>Coefficient of Determination ( r<sup>2</sup> ) for Selected 2006 Census Variables to 2011 National Household Survey Global Nonresponse Rate (GNR) by Census Tract for Winnipeg, Inner City and Non-Inner City</b>			
<b>Selected 2006 Census Variables</b>	<b>Winnipeg</b>	<b>Inner City</b>	<b>Non Inner City</b>
% reporting Non-official language Mother Tongue	0.028	0.181	0.019
% of Recent immigrants (arrived 2001 to 2006)	0.054	0.043	0.012
% of Aboriginal identity population	0.253	0.213	0.103
Prevalence of low income after tax in 2005 %	0.067	0.007	0.012
Median 2005 after-tax income	0.184	0.331	0.047
Average 2005 after-tax income	0.142	0.336	0.033
Source: Statistics Canada 2006 Census table 94-578-XCB2006005, 2006 Census table 97-563-XCB2006043, 2006 Census table 97-563-XCB2006012			
Inner City boundary based on Census Tracts			

that the percentage of Aboriginal population explains approximately 25 percent of the change in the global nonresponse rate.

What is really interesting, and somewhat puzzling, is that low income (poverty) appears to have little relationship to global non-response rates. The coefficient of determination for low income suggests poverty explains less than seven percent of changes in the global non-response rate, and less than one percent in the inner city. It may be because poverty is highly correlated with the overall inner city population but this is not fully clear and requires further investigation.

**Conclusion**

The voluntary nature of the NHS was controversial from the start, especially with respect to data quality concerns. Soon, data from the NHS at the neighbourhood level should be available. It will be interesting to see what the response rates look like, especially for inner city neighbourhoods. In the meantime there is a message we should all heed, data nerds and normal people alike—don't take data for granted. If we don't have solid and reliable data about what's going on, it will be even more difficult to bring about the kinds of changes that are necessary. For that reason, data is an important political issue.

*Darren Lezubski is with Ultra Insights Planning and Research Services. He can be reached at: Darren@ultrainsights.ca*