

Evaluation of the National Longitudinal Survey of Children and Youth Weighting Methodology

Agnes Waye

Statistics Canada, R. H. Coats Building, 100 Tunney's Pasture Driveway, Ottawa, ON
K1A 0T6

Abstract

The National Longitudinal Survey of Children and Youth is a longitudinal survey conducted by Statistics Canada. At cycle 7, there was a single non-response adjustment which was calculated using logistic regression models and response homogeneous groups (RHGs). Due to the difficulties encountered during collection and the low response rate at cycle 8 compared to previous cycles, several key changes were made to the weighting methodology with the goal of reducing the effect of higher non-response on the quality of the estimates. At cycle 8, two weight adjustments were made based on a sequential model to reflect the different aspects of non-response (refusals and non-contacts). The approach used to create RHGs was changed from a quantiles-based method at cycle 7 to cluster analysis at cycle 8, since the latter allows more flexibility. We also examined alternative weighting methods and compared them to the cycle 8 weighting methodology.

Key Words: weighting, nonresponse, cluster analysis, weight adjustments

1. Introduction

Nonresponse is a problem which occurs in almost all surveys. Nonresponse is a failure to collect required data from a sample member. There are two forms of nonresponse: unit nonresponse and item nonresponse. Unit nonresponse is the inability to collect any information at all from a sampled unit, while item nonresponse is a failure to collect a specific item of information from a sampled unit which has supplied other information (Lynn,1996). This paper focuses on unit nonresponse only, and the term nonresponse will refer to unit nonresponse.

One of the most important problems with nonresponse in surveys is the introduction of bias, which is frequently difficult to quantify. If the nonrespondents differ systematically in some way from the respondents, then the sample will produce biased population estimates. This is the main reason why survey statisticians are concerned about nonresponse.

There are two ways to handle the effects of nonresponse. One is to minimize the effects of nonresponse at the data collection stage. The other way is to make statistical adjustments at the analysis stage. Weighting adjustments are a popular method for handling unit nonresponse in sample surveys and can greatly reduce nonresponse bias. Adjustments for nonresponse are made by omitting the nonrespondents and adjusting the sampling weights of respondents to account for the nonrespondents. For example,

weighting class adjustments are made based on the assumption that sample members can be partitioned into weighting classes within which the response of nonrespondents, had they been attained, would be similar to those of respondents (Little and Rubin, 1987). The auxiliary variables used to define weighting classes must be available for both respondents and nonrespondents. Response homogeneous groups (RHGs) are often used to construct these weighting classes, where units with similar probabilities of response, or propensity scores, are grouped together in the same class.

This paper describes the changes in weighting methodology to the National Longitudinal Survey of Children and Youth (NLSCY) at cycle 8 and also explores two alternative weighting methods. The motivation for these changes arose from concerns about the low response rate at cycle 8. The most important change in the weighting methodology at cycle 8 was the use of a sequential model instead of a single nonresponse model. There were various other changes as well, and they will be described in more detail in the following sections.

In the remainder of the paper, we discuss the key changes made to the weighting methodology and explore two alternative weighting methods. In section 2, a brief overview of the NLSCY is given and key terms are defined. Section 3 describes the weighting steps at cycle 8. Section 4 explains the changes made to the weighting methodology. Section 5 describes the two alternative weighting methods explored and shows some results, and finally, section 6 gives a conclusion and discussion.

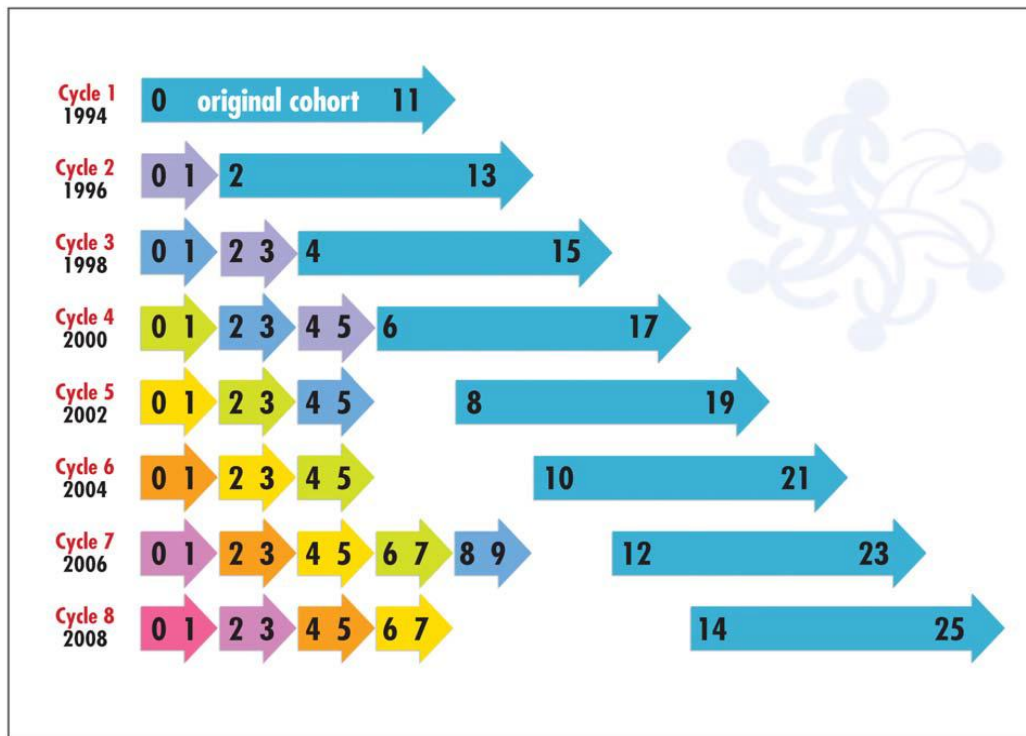
2. Overview of the NLSCY

The National Longitudinal Survey of Children and Youth (NLSCY) is a longitudinal survey, sponsored by Human Resources and Social Development Canada (HRSDC), which has been conducted by Statistics Canada from 1994 to 2008. There were 8 cycles in total, with cycle 8 being the last cycle of the NLSCY. The objective of the survey is to collect information on characteristics which impact the development and well-being of Canadian children and youth over time. The content of the NLSCY includes questions concerning the child's health, behaviour, learning, and social development. Every two years, data is collected through personal and telephone interviews of the person most knowledgeable (PMK) of the child, who is usually the biological mother. The survey samples households from the Labour Force Survey (LFS), which served as the frame for the NLSCY. The LFS is a monthly survey that provides estimates of employment and unemployment in Canada. It also collects information on various other characteristics of the working-age population, including age, educational attainment, and marital status.

Collection of cycle 1 data of the survey began in the fall of 1994 for children aged 0 to 11 living in one of the ten provinces in Canada. Units that entered the survey at cycle 1 will be referred to as the original cohort in this paper. Starting at cycle 2, Early Childhood Development (ECD) cohorts of 0 and 1 year olds were selected and followed until they were 4 to 5 years old, respectively. In the spring of 2009, the survey completed its data collection for cycle 8.

Figure 1 below describes the cohorts of the NLSCY as of cycle 8. The long arrow corresponds to the original cohort and the shorter arrows represent ECD cohorts. Each row is a cycle and the numbers indicate the ages of the children.

Age of children at each cycle, original cohort versus the early childhood development cohorts



Source: Statistics Canada, National Longitudinal Survey of Children and Youth.

Figure 1: Various cohorts of the NLSCY for cycles 1 to 8

Note that the original cohort is purely longitudinal: it is not topped up to reflect changes that occur in the population over time as a result of immigration. Therefore, the original cohort does not have cross-sectional weights at cycle 8. The longitudinal population at cycle 8 of the NLSCY consists of children living in any province in Canada who met the following criteria:

- Original cohort: children aged 0 to 11 as of December 31, 1994
- Cycle 5 cohort: children aged 0 to 1 as of December 31, 2002
- Cycle 6 cohort: children aged 0 to 1 as of December 31, 2004
- Cycle 7 cohort: children aged 0 to 1 as of December 31, 2006

The cross-sectional population at cycle 8 of the NLSCY consists of children living in Canada who met the following criteria:

- New 0 to 1 year old children selected at cycle 8
- Returning 2 to 3 year old children from the cycle 7 ECD cohort of 0 to 1 year old children
- Returning 4 to 5 year old children from the cycle 6 ECD cohort of 0 to 1 year old children
- New top up of 2 to 5 year old children selected at cycle 8
- Returning 4 to 7 year old children from the top up sample of 2 to 5 year old children selected at cycle 7

- f) Returning 6 to 7 year old children from the top up sample of 2 to 5 year old children selected at cycle 6
- g) Returning 6 to 7 year old children from the ECD cohort of 0 to 1 year old children selected at cycle 5.

The NLSCY produced longitudinal and cross-sectional weights for the ECD cohorts and two sets of longitudinal weights for the original cohort at cycle 8. The two types of longitudinal weights were funnel and non funnel weights. Funnel weights were produced for units that responded at every cycle since entering the survey, and non funnel weights were produced for all units that were longitudinally in scope, regardless of whether or not they have responded at every cycle.

3. Changes made to the Weighting Methodology at Cycle 8

There were several changes made to the weighting methodology at cycle 8. We describe these changes briefly in the sections below. They will each be discussed in more detail in section 4, which describes the cycle 8 weighting steps.

3.1 Sequential Model

Previous work by Groves and Couper (1995) found that it is important to devote separate attention to the process of contact and refusal in surveys. In their research, they found a different set of predictors for their contact and refusal models. This shows the importance of model specification, since their results would have been quite different if they used a single nonresponse model instead of two distinct models.

Back at cycle 7, we used a nonresponse model at the weighting step and had one nonresponse adjustment. However, the response rates were low at cycle 8 of the NLSCY. There were more units that required tracing in cycle 8 compared to cycle 7. For example, the collection rate for the original cohort went from 80.5% in cycle 7 to 68.0% in cycle 8. In an effort to minimize the effects of the low collection and response rates, some changes were made to the weighting methodology. At cycle 8, instead of a single nonresponse model, a sequential model was used at the weighting stage. This sequential model is made up of a noncontact model and a refusal model to reflect these two distinct components of nonresponse in surveys. Due to the low collection and response rates at this cycle, the sequential model was a suitable choice for the weighting methodology. By using a sequential model, variables associated with noncontact and refusal could be specified separately. In this paper, the group referred to as refusals is a subset of those that were contacted and is made up mostly of units that refused to participate in the survey, but there were some special cases where we could not collect information for other technical reasons.

3.2 Cluster Analysis

Back at cycle 7, RHGs were formed using the equal quantile method (Girard, C. *et al.*, 2009). Estimated probabilities were taken from the logistic regression model and were sorted in ascending order. Groups of equal sizes were formed as the RHGs. However, one disadvantage of this equal quantile method is that units on each side of a given boundary could be very similar in terms of their score. Therefore, we would expect that by using cluster analysis, members belonging to a RHG would be more homogenous with respect to their scores and that the boundaries of the classes would fall in more natural places. It is for this reason that cluster analysis was used to create the RHGs at cycle 8.

Clusters were made up of units with similar probabilities of response. The algorithm for creating clusters will be explained in section 4.1.

3.3 Cooperation Variable

Cooperation scores were numbers (between 0 and 1) related to the level of cooperativeness of a given unit based on the amount of item response present at cycle 1. Only members of the original cohort were assigned these scores. The cooperation variable was a categorical variable based on the value of the cooperation score. At cycle 7 and cycle 8, the cooperation scores for individuals who have attained 18 years or more of age were derived based on item response rates of the now adults when they were 16 or 17 years old, instead of using the cooperation scores derived from cycle 1 data.

For the initial nonrespondents, which were units that did not respond at cycle 1, the cooperation variable was imputed. At cycle 7, the cooperation variable was assigned according to a uniform random distribution for these initial nonrespondents. These variable values were imputed for these initial nonrespondents since we cannot assume that they would all be uncooperative, i.e. have 0 as the cooperation score. However, at cycle 8, frame information was used to assign cooperation score variables instead. A logistic regression model was created to model the probability of response at cycle 1. Then, RHGs were created using the estimated probabilities to determine the score variables. By using a logistic regression model, we no longer assume that the level of cooperativeness of an initial non respondent is random. In fact, most of the initial nonrespondents ended up having the lowest value for the cooperation score variable, indicating that they were not likely to cooperate. This result is not surprising since those that did not respond at cycle 1 should share similar characteristics with those who were not very cooperative with survey response.

4. Weighting Methodology at Cycle 8

The final survey weights at cycle 8 were calculated based on the following formula:

$$wt_{final} = wt_{design} * adj_{noncontact} * adj_{refusal} * adj_{poststratification}$$

The final weight was the product of the design weight, noncontact adjustment, refusal adjustment, and post stratification adjustment. As seen from the above, the starting point of the final survey weight was the NLSCY design weight, which were calculated based on the LFS subweight. The LFS subweight was a weight which had been adjusted for LFS household nonresponse. The subsections below describe each of the weighting steps in detail.

4.1 Noncontact Model, RHG formation, and Noncontact Adjustments

Units that were eligible for weighting were merged with the frame file (LFS data) to get variables for modelling. There were 60 frame variables in total, and there were no paradata variables used in the model. Some examples of the frame variables were: full time or part time status of current job, marital status, sex, or reason for leaving his or her last job. Next, the units were split into those that were contacted, and those that were not. Chi-squared tests were used to determine which frame variables were significantly associated with the contact status at the 0.05 level. The selected variables entered a stepwise logistic regression model with the contact status as the outcome variable. (For example, the cross-sectional refusal model ended up with 17 variables.) Some examples of the variables that were significant in the model for the cross-sectional population were

education level and dwelling status of the PMK. Cooperation variables were used in addition to the frame variables for the logistic regression models of the original cohort. This was one of the most significant variables in the noncontact model.

The resulting estimated contact probabilities, or propensity scores, were used to create RHGs by using the PROC FASTCLUS procedure in SAS. This procedure creates the RHGs by performing cluster analysis on the basis of distances computed from the score. PROC FASTCLUS uses the “*k*-means” algorithm that consists of selecting a set of points called cluster seeds as a first guess of the means of the cluster. Each observation is assigned to the nearest seed to form temporary clusters. The seeds are then replaced by the means of the temporary clusters, and the process is repeated until no further changes occur in the clusters or the maximum number of iterations has been reached.

Our goal was to find the smallest number of classes that were homogenous with respect to the estimated probabilities of contact obtained from the logistic regression model. We were careful not to create too many clusters because this resulted in smaller cluster sizes. Smaller cluster sizes were often associated with higher adjustments since often the contact rate was low within each cluster, and we did not want to create instability in the weights by having overly large adjustments.

In order to select the number of RHGs, we created R^2 graphs where we plotted the number of classes against the value of R^2 , or the coefficient of determination. The coefficient of determination is a single measure used to assess the homogeneity of classes (Alavi and Beaumont, 2003). A value close to 1 indicates that the classes are homogeneous. The value of R^2 increases as the number of classes increases. Figure 2 below shows an example of a graph used to determine the number of classes. This graph was used for the noncontact model of the original cohort. Note that in PROC FASTCLUS, the requested number of classes is always larger than the actual number of classes created. For example, only 4 classes were created even when 5 classes were requested. This situation results from the constraint that each class must contain a pre-specified minimum number of contacted units, and classes that do not satisfy this constraint are deleted in PROC FASTCLUS.

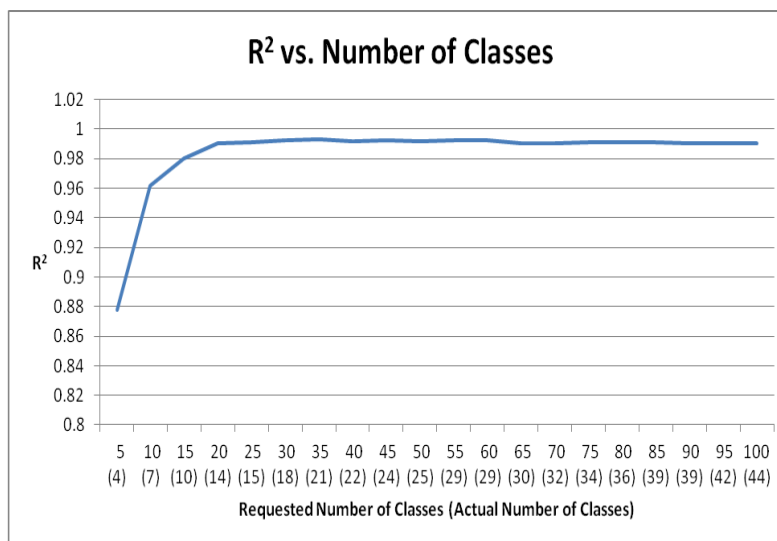


Figure 2: Graph to determine the number of classes to choose for cluster analysis

After the creation of the RHGs, noncontact adjustments were made. Within each RHG, the noncontact adjustment was the inverse of the design weighted contact rate. The contact adjusted weight was equal to the design weight multiplied by the noncontact adjustment.

4.2 Refusal Model, RHG formation, and Refusal Adjustments

After creating the noncontact model, RHGs, and noncontact adjustments, the refusal model was created. The steps for the refusal model were similar to those of the noncontact model. First, only units that were contacted were used to create a refusal model. Again, chi-squared tests were conducted to see which of the 60 frame variables were significantly associated with the response status at the 0.05 level. The selected variables were chosen to create a stepwise logistic regression model with the response status as the outcome variable. (For example, the cross-sectional refusal model ended up with 14 variables.) The estimated probabilities of response were used to create the RHGs through cluster analysis as before. Within each RHG, the refusal adjustment was the inverse of the weighted response rate, where the weight was the contact adjusted weight calculated at the end of section 4.1. The refusal adjusted weight was equal to the contact adjusted weight multiplied by the refusal adjustment.

Some examples of the variables in the refusal model for the cross-sectional population were: total weekly family earnings and spouse education level. These two variables were not in the noncontact model, and in fact, there were no significant income variables at all in the noncontact model. This shows that sequential modelling allows the components of noncontact and refusal to be modelled separately, which is the benefit over using a single nonresponse model.

Table 1 shows the range of the adjustments for the noncontact and refusal models, and as mentioned earlier, it was not desirable to have overly large adjustments. The ranges of the adjustments below show that most adjustments were fairly low and below 3, with the exception of the refusal adjustment of the cycle 5 funnel weights. This was due to the lower funnel response rate for the cycle 5 cohort units. Table 2 below shows the number of RHGs created for each set of weights for the noncontact and refusal models. The number of RHGs created was mostly under 10, with the exception of a few cases. For example, the cycle 5 cohort had 13 noncontact RHGs because this was the lowest number of RHGs for which the R^2 value was satisfactory. In the table, ‘OC’ is an abbreviation of original cohort.

Table 1: Range of Adjustments for Weights

<i>Weights</i>	<i>Cycle</i>	<i>Range</i>
OC Longitudinal/Funnel noncontact	1	1.0049 - 2.6393
OC Longitudinal Nonresponse	1	1.3667 – 2.3377
OC Funnel Nonresponse	1	1.6457– 3.9748
ECD Longitudinal/Funnel Noncontact	5	1.0137 – 1.3405
	6	1.0159 – 2.3416
	7	1.0036 – 1.9423
ECD Cross-Sectional Noncontact	-	1.0303 – 2.0783
ECD Longitudinal Refusal	5	1.0137 – 1.3405
	6	1.0159 – 2.3416

	7	1.0036 – 1.9423
ECD Funnel Refusal	5	1.4018 – 3.3237
	6	1.1379 – 2.3500
	7	1.0895 – 1.9946
ECD Cross-Sectional Refusal	-	1.0900 – 2.0631

Table 2: Number of RHGs Created

<i>Weights</i>	<i>Cycle</i>	<i>RHGs created</i>
OC Longitudinal/Funnel Noncontact	1	12
OC Longitudinal Nonresponse	1	6
OC Funnel Nonresponse	1	6
ECD Longitudinal/Funnel Noncontact	5	13
	6	11
	7	11
ECD Cross-Sectional Noncontact	-	6
ECD Longitudinal Refusal	5	9
	6	9
	7	8
ECD Funnel Refusal	5	9
	6	9
	7	9
ECD Cross-Sectional Refusal	-	7

4.3 Poststratification

After the adjustments were made, the weights were post stratified. For the cross-sectional sample, the weights were poststratified to known population counts within cells formed by the child's age, sex, and province as of the reference year of the cycle. The population counts were taken from the 2001 Canadian census projections. Similarly, for the longitudinal sample, the same thing was done except the initial year of the cohort was used instead. For example, the initial year of the original cohort would be 1994.

Table 3 below shows the post stratification adjustments for all the sets of weights produced. As seen below, all the means of the adjustments are quite low and are all less than 1.3.

Table 3: Poststratification Adjustments for Weights

<i>Weights</i>	<i>Cycle</i>	<i>Mean</i>	<i>CV</i>	<i>Min</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Max</i>
OC Long	1	1.2265	19.9703	0.6638	1.0499	1.2065	1.3854	2.1748
OC Funnel	1	1.2560	27.0356	0.6091	1.0335	1.2021	1.4044	3.0200
ECD Long	5	1.2513	15.8141	0.9155	1.0998	1.2425	1.3543	2.1907
	6	1.2010	11.9713	0.9051	1.1320	1.1773	1.2534	1.7409
	7	1.2538	17.9099	0.9464	1.1179	1.2111	1.2622	2.0862
ECD Funnel	5	1.2419	16.0214	0.8897	1.0836	1.1931	1.4093	1.9545
	6	1.2115	14.1001	0.9058	1.1182	1.1558	1.3025	1.9836
	7	1.2606	19.2343	0.9640	1.0865	1.2074	1.3055	2.1439
ECD Cross	5,6,7,8	1.2945	15.3898	0.8966	1.1784	1.2569	1.3894	2.3197

4.4 Comparison with Cycle 7 Weighting Methodology

Table 4 below shows some examples of estimates obtained using cycle 7 weights and cycle 8 weights. The information below came from data related to the PMK from the longitudinal master files. Even with all the changes made to the weighting methodology, the cycle 8 estimates did not seem to be drastically different from those of cycle 7 at the national level. As seen in the ‘Change in Distribution’ column, the differences in percentages between the two cycles are very small. The numbers in Table 4 do not add up to 100% for each variable since refusals, not stated, and other groups were not included.

Table 4: Comparison of Examples of Weighted Estimates (Cycle 7 Method and Cycle 8 Method)

<i>Variable</i>	<i>Level</i>	<i>Cycle 7 %</i>	<i>Cycle 8 %</i>	<i>Change in Distribution (Cycle 8% - Cycle 7%)</i>
Are you currently attending a school, college or	Yes	5.13%	4.12%	-1.01%
	No	92.06%	91.47%	-0.59%
Do you use a public library?	Daily	0.7%	0.57%	-0.13%
	Weekly	6.68%	6.34%	-0.34%
	Monthly	16.86%	13.45%	-3.41%
	Several times a year	16.38%	13.57%	-2.81%
	Once or twice a year	22.88%	21.13%	-1.74%

5. Alternative Weighting Methods

In this section, we explore two alternative weighting methods to the cycle 8 weighting method and discuss the results. In the next few subsections, we will refer to the weighting method described in section 4 as the cycle 8 weighting method, since this was the actual weighting method used. These two alternative weighting methods were applied to the cross-sectional weights only. Subsections 5.1 and 5.2 will describe the alternative methods, and 5.3 will discuss and compare the results. The actual cycle 8 weights were produced based on the cycle 8 method described in section 4. Due to time constraints, we did not have time to explore these two alternative weighting methods until after the cycle 8 weights were produced.

5.1 Method A

In the first alternative method, which we will call method A, we used a single nonresponse model instead of a sequential model. First, the probability of contact was calculated by creating a stepwise logistic regression model with the contact status as the outcome variable. Then, a nonresponse model was created by using the probability of contact as a variable along with the frame variables in a stepwise logistic regression model, where the outcome variable was the response status. Nonresponse adjustments were made as before by creating RHGs using cluster analysis, where the adjustment was the inverse of the design weighted response rate. There were 7 nonresponse RHGs created. The range of the nonresponse adjustments were from 1.14 to 2.65. Finally, the weights were post stratified as described in section 4.

This approach is a mix of the cycle 7 method, which used a single model, and the cycle 8 method, where a sequential model was used. This method enabled us to use both

probabilities without having a sequential model, and we were also able to incorporate contact and response probabilities without having a large number of classes.

5.2 Method B

The second alternative method, method B, incorporated contact and refusal probabilities by first creating contact classes, then creating refusal classes within each of the contact classes. First, a noncontact model was created, and noncontact classes were created using cluster analysis. Then, refusal classes were created within each of the contact classes using cluster analysis once again. Noncontact adjustments were applied in the same manner as the cycle 8 method, but refusal adjustments were applied within each contact class. For the 5 noncontact classes created, 4 of the classes in turn had 3 refusal classes created, and 1 of them had 2 refusal classes. Note that the number of classes was low because the refusal classes were created within a contact class, so the number of units in the class was not very high to begin with. We did not want to end up with overly small refusal classes, so we were careful not to create too many within each noncontact class. The range of the refusal adjustments were from 1.14 to 2.65. Again, the weights were poststratified as described in section 4.

For the cycle 8 method, there was a single nonresponse model created for all those that were contacted. For method B, there is more flexibility because different refusal models were specified for those that were contacted, depending on the score they received from the contact model. The resulting variables in each of these response models were different. This method did not assume that all contacted units had the same characteristics associated with the response status. By allowing the refusal models to be flexible, this may provide some protection against model misspecification.

5.3 Results and Discussion

This section shows some results from the two alternative methods. In Table 5 and Table 6 below, the mean poststratification adjustment for method A and B were 1.29 and 1.28, respectively, which were similar to the range of the poststratification adjustment for the cycle 8 method. Similarly, the minimum and maximum values below were in line with those of the cycle 8 method.

Table 5: Method A: Poststratification Adjustments for Cross-Sectional Weights

<i>Mean</i>	<i>CV</i>	<i>Min</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Max</i>
1.2948	15.3745	0.8980	1.1693	1.2619	1.3714	2.3405

Table 6: Method B: Poststratification Adjustments for Cross-Sectional Weights

<i>Mean</i>	<i>CV</i>	<i>Min</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Max</i>
1.2806	15.3613	0.8901	1.1540	1.2610	1.3806	2.1800

Table 7 shows the estimates using the two alternative methods compared with the cycle 8 weighting method. The variances were estimated using bootstrap weights. As seen in Table 6, there were no major differences between the two alternative methods and the cycle 8 method (refer to the ECD Cross row in Table 3). In general, the differences in standard error were small except for cases where the percentages were low (for example, for child's parent status: biological father and step mother).

There were two possible reasons for the similar results when comparing the alternative methods and the cycle 8 method. First, attempts were made to choose RHGs resulting in low adjustment factors for the alternative methods and for the cycle 8 method. Secondly, alternative method and cycle 8 weights were all post stratified to the same population counts (the projected counts based on the 2001 Canadian Census of population). Therefore, these two factors may have helped produce results that were similar to the cycle 8 method ones.

Table 7: Examples of Weighted Estimates (Method A and Method B Compared to Cycle 8 Method)

<i>Variable</i>		<i>Percent (A)</i>	<i>Percent (B)</i>	<i>Percent (Cycle 8)</i>	<i>Relative Difference of Standard Error (A)</i>	<i>Relative Difference of Standard Error (B)</i>
What type of school is this child currently in?	Public school	31.98%	32.03%	31.88%	0.45%	-0.74%
	Catholic school	7.38%	7.42%	7.42%	-0.55%	0.63%
	Private school	2.93%	2.92%	3.00%	-3.00%	-3.97%
Child's parent status	Both biological parents	83.49%	83.55%	83.58%	1.04%	-1.21%
	Biological mother and step father	1.84%	1.90%	1.83%	1.19%	5.97%

Out of the three methods (the two alternative methods and the cycle 8 method), method A was the simplest method to implement because it only required the construction of one set of refusal RHGs. Method B was the most difficult of the three methods because it required the creation of refusal RHGs within noncontact RHGs. The risk of creating high weight adjustments was one of the major challenges of method B. For method B, we had to avoid creating a high number RHGs with only a few units each, since this would result in large weight adjustments, as mentioned in section 4.1. Also, the resulting the number of RHGs was quite high for method B, where 5 noncontact RHGs, and 14 refusal RHGs were created in total. Compared to the cycle 8 method, where 6 noncontact RHGs and 7 refusal RHGs were produced, method B required a lot more time in creating and choosing the optimum RHG groupings.

In terms of flexibility of the models, method A provided the least flexibility because it only created a single refusal model for all units, regardless of the contact status. Method B was the most flexible because it allowed for the creation of different refusal models depending on the contact status and contact probability that was associated with the unit. The level of flexibility of the cycle 8 method was in between those of the alternate methods. It allowed for the contact probability to be modelled, but only created a single refusal model for all those that were contacted.

When comparing these three methods, cycle 8 would be the preferred method due to its ease of implementation compared to method B and the fact that it captured noncontact

and refusal separately. Although method A was the easiest to implement, it only accounted for noncontact as a variable in the refusal model. For method B, multiple refusal and noncontact models were created, but it was very time consuming to implement and we had to be careful to avoid creating too many RHGs. The three methods showed very little differences, so we think it is acceptable that the cycle 8 method was used as opposed to one of the alternative methods.

6. Conclusion

The weighting methodology for cycle 8 of the NLSCY was altered in order to minimize any additional bias that may have been introduced as a result of the lower collection and response rates. The decrease in the response rate indicated a need for a more sophisticated model that captured contact and refusal separately, and therefore a sequential model was used for cycle 8 weighting. There were no major differences between the cycle 7 and cycle 8 estimates even though numerous changes made to the weighting methodology. This lack of major differences shows that the new methodology is reasonable. The two alternative weighting methods produced results that were quite similar to the cycle 8 weighting method. As shown in section 5, the results from the alternative methods were very similar to those of cycle 8. However, from the two alternative methods and the cycle 8 method, the cycle 8 method is still the preferred method due to its ease of implementation and flexibility for modelling.

Acknowledgements

I would like to thank Mike Tam and Dany Faucher for all their help with this paper.

References

- Alavi, Asma and Beaumont, Jean-François. (2003) A Nonresponse Adjustment Strategy using Modelling in the Canadian Labour Force Survey. *Proceedings of the Joint Statistical Meeting Conference*, Statistics Canada.
- Girard, C., Ferland, M., Franklin, S. and M. Tremblay. (2009) Rethinking the NLSCY weighting methodology: the lessons learned and the innovations introduced. HSMD working paper. HSMD-2009-002E.
- Groves, Robert M., and Couper, Mick P. (1995) Theoretical Motivation for Post-Survey Nonresponse Adjustment in Household Surveys. *Journal of Official Statistics*, 11, 93-106.
- Haziza, David and Beaumont, Jean-François. (2007) On the Construction of Imputation Classes in Surveys. *International Statistics Review*, 75, 25-43.
- Iannacchione, V. G., Milne, J.G., and Folsom, R.E. (1991) Response Probability Weight Adjustments Using Logistic Regression. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 637-642.
- Little, R.J.A. and Rubin, D.B. (1987) *Statistical Analysis with Missing Data*. J. Wiley & Sons, New York.

Lynn, Peter. (1996) Weighting for Non-response. *Survey and Statistical Computing*, 205-213.