Innovative Northern Design Improvements in the Canadian Labour Force Survey

Scott Meyer and Edward J. Chen
Statistics Canada, 18th floor, R.H. Coats Building, Tunney's Pasture,
Ottawa, Ontario, Canada K1A 0T6

Abstract:

The Canadian Labour Force Survey (LFS) first started sampling in the Yukon in 1991 with some earlier attempts made on a trial basis. Over time, the LFS sample design has expanded to include the Northwest Territories and Nunavut in conjunction with other household surveys conducted by Statistics Canada. The low population and vast land areas in the territories require alternate approaches to accommodate some of the operational difficulties inherent to remote communities. During the past two decades, there have been a number of sample design updates and coverage changes in the North. This paper describes the latest design updates and improvements, focusing on the most recent adjustments which began to be phased in starting in January, 2011.

Key Words: survey design improvements, low population, data quality.

1. Introduction

Canada's vast northern territories include the Yukon, Northwest Territories (NWT) and Nunavut with a growing combined population of 33,000 households in 2006. Whitehorse, Yellowknife and Iqaluit are the respective capital cities. Statistics Canada started to conduct the Labour Force Survey (see Reference) in Yukon in 1991, with some earlier attempts made on a trial basis. Over the years, there have been a number of minor design and coverage changes to the Northern frame. This paper describes the latest design update which is being phased in starting in January, 2011.

Much of the former design was based on information from the 1996 Census of Population or earlier. Many changes, including mine closures and population growth, have occurred since then. Some data quality indicators and sample sizes started to be problematic by 2010. Under the constraints of using current systems and keeping sample size and costs constant, we sought to: 1. Introduce one-stage sample designs in Whitehorse, Yellowknife and some other communities; 2. Update the sample allocation and sample rates to reflect the 2006 Census of Population estimates; 3. Alleviate excessive respondent burden and distribute the samples more equitably; 4. Re-stratify outlying communities in the Yukon; 5. Introduce Address Register (AR) based listing and improve the labelling convention in the North.

This design update was the product of extensive consultation with the major stakeholders including the representatives from the territorial governments, analysts from the Labour Statistics Division, collection specialists from Collection Planning and Management Division (CPMD) and from household surveys that are interested in using the LFS frame in the North. In addition to the LFS, these surveys include the Canadian Community Health Survey (CCHS), and Survey of Household Spending (SHS).

2. Rotation Pattern

In the North, since inception in 1991 and continuing through the latest design update, we have adopted a 24-month rotation pattern where selected dwellings spend one month in survey, two months out of survey, repeated eight times.

Considering the first quarter of 2011, the set of households selected in January would birth (i.e., have their first survey interview) in January 2011 and be re-interviewed in the first month of each quarter thereafter (March 2011, July 2011, etc.) for a total of 8 quarters. The February set of households, distinct from January, would similarly be re-interviewed the second month of each quarter, and March, the third month.

The quarters over a two year period are assigned rotation numbers 1-8. Rotation 1-4 is applied to the four quarters in even years, while 5-8 is applied to odd years. For example, the first quarter of 2011 has rotation 5 as the birth rotation, with rotation 4 households having their 2nd interview, rotation 3 households having their 3rd interview, and so on.

3. Design Update and Improvements

The 2011 Canadian Labour Force Survey frame design updates and improvements in the North are described below**.**

3.1 Expanded Use of One-Stage Design

A one-stage design was introduced in Iqaluit and some large communities in the 1998 design improvement while the sample design for Whitehorse and Yellowknife was a two-stage approach with clusters (PSUs) selected at the first stage and then dwellings selected within clusters. We sought to expand the use of one-stage design in more communities in this design update. In the Whitehorse Census Subdivision (CSD), there were 93 clusters (average size of about 90 households) on the frame and 24 were active (i.e., had survey sample) for any given quarter. In Yellowknife, there were 26 clusters on the frame with 8 active each quarter.

Over time, the household counts in these clusters changed and the prescribed sample rates began to be problematic. Clusters had been set-up approximately equal in size based on Census 1991 data, but this was no longer true by 2010. In Yellowknife, clusters ranged from 25 to 750 dwellings. In Whitehorse, large developments (Arkell and Copper Ridge – about 2000 homes or 25% of the city) remained out of sample for many years and the production maps (based on Census 1991 enumeration areas) were not clear for these areas. Sample sizes would fluctuate substantially depending on which clusters were active and growth outside of the selected clusters was not reflected in the sample.

In the updated design, we decided to create single stage strata in Whitehorse and Yellowknife, collapsing the former clusters together. We use the term single cluster strata for these newly delineated parcels of land. After listing, we would draw a systematic sample in each single cluster stratum. Resulting samples are essentially systematic samples dispersed throughout the entire community.

This strategy is very robust in terms of coverage as any new growth belongs in exactly one of the single cluster strata and can be added to the frame via well established list

maintenance procedures. This approach also eliminates sample size vagaries due to PSU rotation.

The updated design has 24 single cluster strata in Whitehorse CSD and 16 single cluster strata in Yellowknife. These strata are allocated equally into 8 rotation groups as described in Section 2. Each quarter there is birth sample in 1/8 of the clusters (three in Whitehorse and two in Yellowknife) and follow-up surveys continuing in the other 7/8 of the clusters (21 in Whitehorse and 14 in Yellowknife).

The same approach was taken in Whitehorse Unorganized, the area within the Whitehorse Census Agglomeration (CA) boundary, but outside the Whitehorse CSD. Formerly, two clusters were selected from six on the frame. After the updates, we have two single cluster strata that together cover the entire area of Whitehorse Unorganized.

The updated design in Whitehorse and Yellowknife is similar to what already existed for Iqaluit – an exhaustive set (7 for Iqaluit) of continuously surveyed single cluster strata that cover the entire community. Other communities had been using a one-stage design for several years: Hay River (NWT), Inuvik (NWT), Norman Wells (NWT), Rankin Inlet (Nunavut), Cambridge Bay (Nunavut), and Kugluktuk (Nunavut). In the updated design, Dawson (Yukon) and Watson Lake (Yukon) were promoted to the same "always in survey" status.

3.2 New Allocation and Sample Parameters

Another symptom of the former design was that sample size yields, based on fixed sampling rates that had not been adjusted for many years, were over the intended targets and continuing to trend upward. This was especially true in Nunavut and was driving up collection costs.

The updated design used Census 2006 household counts and a recent housing survey in Nunavut to re-set the sampling rates in each stratum. At the same time, some of the extremely high sampling fractions at the stratum level that existed in the former design (e.g., 1 in 5 in Kugluktuk, Nunavut; 1 in 5 in the stratum of Haines Junction and Mayo, Yukon) were reduced.

Sampling fractions remain high as data requirements are demanding, and operational reasons play a role (minimum yields to make it "worth the trip" to a remote village). Still, an effort was made to distribute the response burden more evenly within each territory. The smallest basic weight (inverse of the stratum sampling fraction) in the updated design is 8 in the stratum of Coral Harbour and Repulse Bay, Nunavut. The largest basic weight is 22 in Hay River, NWT.

Many strata had slight or no change to the sample rate – only what was needed to bring the overall sample size back to the intended target. Comparability to the former design has been largely preserved.

3.3 New Stratification of the Outlying Communities in the Yukon

While some communities are continuously surveyed under one-stage designs described in Section 3.1, other small towns, villages and hamlets are grouped into strata where two-

stage sampling is used – communities randomly selected with probability proportional to size (RPPS) at the first stage and a systematic sample of dwellings within those communities selected at the second stage.

As characteristics of some communities have changed drastically since the last design in the mid 1990s, re-stratification of outlying communities in the Yukon was urgently needed.

In the former Yukon design, Faro was grouped with Watson Lake and Dawson and there were two communities selected from this stratum. Faro had shrunk considerably in population after a mine closure in the late 1990s and no longer fit with the other two towns that continued to grow steadily.

In the updated design, Dawson and Watson Lake were promoted to always in survey one-stage strata and Faro was grouped with other smaller outlying villages in a stratum with one community selected at the first stage.

Pelly Crossing was not covered in the former design and has been added to frame. The Yukon age-sex population control totals for 2011 were updated accordingly, with a contribution from Pelly Crossing.

3.4 New Address Register Listing and Clusters on National Geographic Database System

Coinciding with the design update, we have also improved cartographical and mapping capability for the North, with LFS cluster boundaries part of the National Geographic Database (NGD) throughout Canada including all provinces and territories.

The Address Register (AR) is a database of Canadian residential addresses that is regularly updated using various administrative sources. The AR is primarily used as a frame for the mail out of Census packages. Since 2004, the AR also facilitates initial listing of Canadian Labour Force Survey clusters in the provinces by providing a pre-filled list of dwellings that fall within the cluster boundary. Further, the AR pre-fills are sequenced by block and block face to be consistent with the cluster maps.

As part of the geography updates, we were able to introduce the use of AR listing for the first time in the North - reducing cluster listing costs. Further, batches of cluster maps reflecting the latest road network and block structure can now be automatically generated by the Generalized Mapping System (GMS) whenever required.

3.5 New Labelling Conventions

The design update provided an opportunity to change the labelling convention which should lead to easier management of the frame and better interpretation of the sample design.

Specifically, single cluster strata now have “000” embedded in the cluster ID to indicate there is no cluster level stage of selection. The cluster ID codes within two-stage design strata have also been revamped.

4. Transition Strategies for the 2011 Design Update

The new theoretical weights (i.e., the initial stratum level weights before any non-response or calibration adjustments are applied) consistent with the updated North design were implemented in January, 2011. However, there will be a period of time when the observed sample is a mixture of dwellings selected with the old parameters and IDs, that continue their follow-up interviews, and dwellings selected with the new parameters and IDs. The dwellings selected with the old parameters (i.e., the old sample rates and corresponding theoretical weights) and IDs are called the *carry forward* sample. A few different approaches for dealing with carry forward sample are described below.

4.1 Weight Adjustment

During the transition period, the carry forward sample in certain clusters or PSUs will require compensatory weights to adjust for the change to the theoretical weights. The theoretical weights for the carry forward sample should reflect the probability of selection *at the time of selection* and the weight adjustment recovers the proper value.

The number of clusters requiring a weight adjustment was largest in the first quarter of 2011 and reduces each subsequent quarter as the carry forward samples are replaced on rotation by birth samples drawn with the new parameters.

4.2 Random Reduction

In other clusters, the old and new parameters have greater differences. Although weight adjustments are logically correct, retaining 100% of the carry forward sample has cost and capacity drawbacks. For these clusters, we retained an appropriate proportion of the originally sampled units. This approach was used for two Yellowknife clusters, rotations 1 and 2, with identical cluster boundaries before and after the updates.

Essentially, within a stratum (new single cluster stratum), the sample was drawn with the old parameters, but we are able to modify that sample to resemble a sample drawn with the new parameters. No compensatory weight adjustments are required.

Through random reduction, we can take advantage of previously contacted dwellings while controlling the sample size.

4.3 Off-Rotation Births

The third option is not to carry forward the sample – cease follow up interviews and select a replacement birth sample off rotation. This occurs in Whitehorse and much of Yellowknife where the new delineation is very different. There was an exceptionally large LFS birth sample for first quarter of 2011 to introduce the updated design in Whitehorse and Yellowknife.

To allow a smooth transition to the new design, two major tasks were completed:

- Establish the new single cluster stratum boundaries on the NGD, allowing the production and printing of maps. This was accomplished through a joint effort of Methodology and Geography Divisions.
- List the communities of Whitehorse and Yellowknife according to the new cluster delineations. This challenging task was carried out successfully in the fall of 2010 under the direction of CPMD.

Further, there was an effort to avoid recently surveyed dwellings in the samples from the updated lists.

5. Summary Results

5.1 Coverage

The territorial design has been modified several times over the years since its inception in 1991. The coverage rates vary over time depending on the design and population shifts in the covered and excluded communities. The coverage after the design updates is now 92% in the Yukon Territories with 11 communities included in the frame, 96% in NWT (with 24 communities) and 93% in Nunavut (with 19 communities).

5.2 Yukon

The Yukon has been in the active LFS sample since 1991, with even older attempts on a trial basis. Since 1991, there have been a number of minor improvements in some areas. In the latest updates, a request was first received from the Yukon territorial government representative regarding the stratification of the communities outside of Whitehorse. Specifically, the grouping of the town of Faro with Dawson and Watson Lake was considered problematic as Faro had undergone a substantial decline in population after a mine closure.

5.3 Northwest Territories

Sampling in the Northwest Territories started with Survey of Household Spending (SHS) in 1992, to match the available sample in the Yukon. Initially, only Yellowknife was included. Since 1992, there have been a number of design updates and minor improvements in some areas. In the latest design updates, Yellowknife was re-delineated and now has a one-stage design and the sample parameters were updated throughout. There was no change to the stratification of the outlying communities.

5.4 Nunavut

Nunavut is a relatively new territory. Although certain communities had samples when they were considered part of the Northwest Territories, sampling in Nunavut, considered as a distinct territory, started in 1998 with the SHS. Since 1998, there have been a number of design changes and minor improvements. Much population growth has occurred in recent years. The 2006 Census showed remarkable gains in Iqaluit especially. Adjustments to the allocation were required.

In the latest design updates, the sample parameters were updated throughout the territory. There were neither changes to the stratum groupings nor any change to the delineation within Iqaluit.

6. Concluding Remarks

This design improvement incorporated information from the 2006 Census of Population to update the stratification, allocation and sample rates to better reflect the growth and changes in the northern territories. We introduced timely innovations to efficiently manage the sample frame in the North. At the same time, the new sample frame facilitates better coordination among the major surveys in Statistics Canada; the list of surveys includes the Labour Force Survey, Canadian Community Health Survey, the Survey of Household Spending and many post-censual surveys.

This update was long overdue because the former design was largely based on the 1996 Census information or earlier. The new design will alleviate excessive sample burden imposed on some communities by distributing the sample more equitably in the territories. Moreover, we eliminated cluster rotation in Whitehorse and Yellowknife to improve sample efficiency and coverage. The re-stratification in the Yukon was necessary to better reflect the community changes that have occurred over time. The introduction of AR listing and the new labelling convention will enable the LFS design to be more efficient and consistent with the design in the South. Undoubtedly, the changes will ease the management of the frame and the production of survey estimates in the North.

Survey data quality in the North will be improved over time as the new sample is phased in over a two year period. This design update was a timely response to clients' concerns on data quality. Slippage rate, a data quality indicator should decrease over time and remain low and more stable, especially in the communities where we introduced a one-stage design.

REFERENCE

Methodology of the Canadian Labour Force Survey (2008). Statistics Canada. Catalogue No. 71-526X.