

## A NEW DESIGN FOR THE CANADIAN NATIONAL SEAT BELT SURVEY

*Aline Chouinard*

*Road Safety Directorate  
Transport Canada*

The National Seat Belt Survey is an observational survey designed in 1978 and conducted almost every year until 2001. Despite the overall national seat belt usage rate of 89.9%  $\pm 0.6$  in 2001 [1], an average of 40% of all vehicle occupants killed and 21% of those seriously injured are still unbelted at the time of a collision [2]. The National Seat Belt Survey was re-designed in 2001 to include rural communities.

The design comprises several stratification levels and multiple stages of sampling. The stratification levels are province/economic region and rural or urban strata. Strata definitions not only take into account the population size of communities but also their population density and proximity (or lack thereof) to a city.

The stages of sampling include the selection of the intersections, the day of the week and the time of day. The intersections are drawn from the Canadian Highway Information System, a spatial database. The estimation model for calculating the seat belt usage rate will be presented. Finally, sampling of intersections with low traffic volumes will be discussed.

### **1. INTRODUCTION**

The Canadian National Seat Belt Survey is an observational survey that was designed in 1978 and conducted almost every year until 2001. The purpose of the survey was to monitor the progress of Canada's National Occupant Restraint Program (NORP). To reduce fatalities and personal injuries among vehicle occupants involved in collisions, the NORP Task Force of the Canadian Council of Motor Transport Administrators (CCMTA), under the auspices of Road Safety Vision 2010, has set a goal of raising seat belt use in Canada to 95% by the year 2010 [3].

For some time, Transport Canada and our provincial and territorial counterparts have wanted to conduct a rural seat belt survey across Canada. Despite the overall national seat belt usage rate of 89.9%  $\pm 0.6$  in 2001 [1], an average of 40% of all vehicle occupants killed and 21% of those seriously injured are still unbelted at the time of a collision, and 63% of collisions resulting in a fatality occur on rural roads [2].

The rural seat belt survey and the National Seat Belt Survey need to be compatible with each other, so that comparisons can be made between them. As a result of new technological developments since 1978, and new data, the design of a rural seat belt survey has led to a re-design of the National Seat Belt Survey.

The results of the "pilot survey" for the new Canadian National Seat Belt Survey were presented at the Canadian Multidisciplinary Road Safety Conference [4]. The present paper presents the "final" design of the survey.

### **2. THE FRAME OF THE SURVEY**

The frame of the survey, in theory, would consist of all occupants of all vehicles on Canadian roads at any point in time throughout the year. The National Seat Belt Survey is an observational survey: observers are stationed at various controlled intersections across the country and they observe the seat belt use of the occupants of the light duty vehicles stopping at the intersection. The survey frame is: all occupants of private light duty vehicles (LDV), including automobiles, minivans, pick-up trucks and sport utility vehicles (SUVs), with Canadian licence plates travelling on a Canadian road during the sample week, during daylight hours. In short, the aim of the survey is to collect a snapshot of the seat belt usage rate of a representative sample of the traffic on Canadian roads.

### 3. THE STRATIFICATION OF THE SEAT BELT SURVEY

#### 3.1 Definition of rural and urban

The definition should correspond to an intuitive concept of “rural” as well as being supported by definitions in use among demographers. The definition of a rural road used by collision databases is a primary or secondary highway or a road where the speed limit exceeds 60km/h. Under such a definition, an arterial street in the middle of a large city could be classified as rural or urban, and the main street of a small village could be classified as urban or rural. Most people will find such a definition confusing. When police officers on duty fill out collision reports, they will check the box “rural” or “urban” according to their perception of whether the road segment where the accident occurred is rural or urban. Our definition should be in line with this intuitive concept of “rural”.

The Statistics Canada demographic definitions [5] use several criteria to classify census subdivisions into Census Metropolitan Areas (i.e. large cities), Census Agglomerations (i.e. small cities) or neither. The criteria used are the population size, the population density and the level of economic integration with the nearest city. All census subdivisions (CSD) together cover Canada entirely and do not overlap. Using the Statistics Canada demographic definitions will make it simpler to update the survey strata after each census.

The definition of “urban” opted for is: those CSDs that are either a Census Metropolitan Area (CMA) or a Census Agglomeration (CA) or are within one, and whose population is over 10,000. Other CSDs are rural.

Two special cases arise. CSDs that have a population larger than 10,000 but are not within a CA (or CMA) are considered rural, i.e. their population density is less than 400 inhabitants per square kilometre. CSDs that have a population less than 10,000 but are within a CA or a CMA are also considered rural, i.e. they are rural fringes.

For the purpose of the survey, the roads located within the boundaries of an urban stratum are urban roads regardless of their speed limit. Similarly, roads located within a rural stratum are rural roads regardless of their speed limit.

#### 3.2 Strata definitions

The first level of stratification is the jurisdiction (province or territory) and economic region. There are 76 distinct non-overlapping economic regions. The economic regions do not cross provincial/territorial boundaries. The second level uses the population size of CSDs as well as their demographic properties.

The urban strata are defined as follows:

U1: CSDs within a CMA and having a population over 500,000

U2: CSDs within a CMA and having a population between 100,000 and 499,999

U3: CSDs within a CMA and having a population between 50,000 and 99,999

U4: CSDs within a CMA and having a population between 10,000 and 49,999

U5: CSDs within a CA and having a population over 50,000

U6: CSDs within a CA and having a population between 10,000 and 50,000

The rural strata are defined as follows:

R1: CSDs not within a CMA /CA but having a population over 10,000 (i.e. density less than 400 per square kilometre despite their large population size)

R2: CSDs within a CMA but having a population less than 10,000 (rural fringes of CMA)

R3: CSDs within a CA but having a population less than 10,000 (rural fringes of CA)

R4: CSDs **not** within a CMA/CA and having a population between 5,000 and 9,999

R5: CSDs **not** within a CMA/CA and having a population between 2,500 and 4,999

R6: CSDs **not** within a CMA/CA and having a population between 1,000 and 2,499

R7: CSDs having a population less than 1,000.

With these definitions, each CSD in Canada will belong to one and only one stratum. The allocation of the CSDs among the strata will need to be reviewed after each census. Canada’s population is distributed among the strata defined above as shown in Table 1. Stratum R7 is excluded from the survey frame as it encompasses only 3.88% of the population and the locations involved are often difficult to reach.

The assumption underlying stratification is that the seat belt usage behaviour may vary between strata while being uniform within a stratum. Communities of the same size are assigned to a stratum on the basis of their demographic properties.

**Table 1. Distribution of the population by strata**

Strata	Population	Percent
U1	7,618,210	25.39
U2	5,641,573	18.80
U3	2,491,428	8.30
U4	2,816,435	9.39
U5	1,538,569	5.13
U6	2,242,470	7.47
R1	954,470	3.18
R2	695,257	2.32
R3	704,129	2.35
R4	1,581,936	5.27
R5	1,273,302	4.24
R6	1,285,282	4.28
R7	1,164,633	3.88
<b>All strata</b>	<b>30,007,094</b>	<b>100%</b>

**3.3 Communities with low population density**

Intersections located in communities with low population density tend to have very low traffic volume. Therefore, communities having a population density of less than 3 inhabitants per square kilometres were excluded from the CSD frame. This exclusion comprises 2.5% of the Canadian population. The net survey frame includes 93.6% of the Canadian population.

**4. SAMPLING OF THE SURVEY SITES**

**4.1 Database of intersections**

In the first stage of sampling, the survey sites are selected. The biggest risk involved in seat belt surveys is the low traffic volume that may be encountered at some intersections, a problem recognized by Shapiro et al. [6].

The problem is to balance the number of sites and the cost-effectiveness of the data collection process while obtaining a large number of observations per stratum. The intersections must also be selected according to a random process in order not to introduce selection bias. Moreover, the intersections selected must represent the distribution of traffic among the categories of roads and not the distribution of the roads themselves.

Transport Canada’s Canadian Highway Information System (CHIS) [7] is a GIS database that contains all road segments in the country; the segments are sorted into categories, from highways to ice roads. Table 3 shows the

distribution of the road segments by category. Category five (i.e. residential streets/ country roads) is the largest category of road segments; however these road segments generally have low traffic volumes. To ensure sufficient traffic volume and the cost-effectiveness of the data collection, only road segments of the first four categories were selected (expressways and arterial roads). The intersections of all these road segments were calculated and the CSD to which they belong was determined, and controlled intersections were separated from the others.

**Table 3. Number of road segments by category**

Category	Frequency	Percent
1	35,094	1.81%
2	90,608	4.67%
3	65,547	3.38%
4	122,621	6.32%
5	1,625,081	83.81%
Other	53,316	2.68%
<b>All</b>	<b>1,992,267</b>	<b>100.00%</b>

A database that includes 23,056 distinct intersections has thus been created for the purpose of the survey. Further work in validating the suitability of these intersections is ongoing. Table 4 shows the number of intersections by stratum. Interestingly, the distribution of the intersections among the strata mimics the distribution of the population among the strata.

**Table 4. Number of intersections per strata**

Strata	Frequency	Percentage
U1	3,774	16.37
U2	3,052	13.24
U3	1,517	6.58
U4	2,124	9.21
U5	1,184	5.14
U6	1,591	6.90
R1	1,371	5.95
R2	687	2.98
R3	620	2.69
R4	1,941	8.42
R5	1,352	5.86
R6	1,633	7.08
R7	2,210	9.59
<b>Total</b>	<b>23,056</b>	<b>100.00</b>

## 4.2 First sampling stage

In the first stage of sampling, one survey site would be selected with equal probability from each stratum (as determined by the economic region and the population size stratum) in two replicates. Strata with larger populations should have more sites; for each 500,000 population in a stratum, two additional sites (one per replicate) will be drawn. In other words, all strata with a population size under 500,000 will have two survey sites (one per replicate); those with more than 500,000 but less than 1,000,000 will have four sites (two per replicate) and so on. A systematic sampling scheme for the list of intersections, ordered by longitude and latitude, will be used for the larger strata so that the entire geographical area of the stratum may be covered in the sample; a simple random sample could result in all sites being in the same area of the stratum.

## 5. SECOND SAMPLING STAGE

### 5.1 Day of the week / Time of day / Direction of traffic

For each survey site, two time periods of observation are selected, one for each replicate. The observation period consists of a day of the week, a time of day and a direction of traffic to be observed. The day of the week is selected without replacement according to the distribution of the kilometres driven per day of the week as found in the Canadian Vehicle Survey (CVS) [8]. This approach ensures that more data are collected on busier days. The day is selected without replacement in order to avoid scheduling conflicts in the survey process. The time of day is selected with replacement from the available daylight hours according to the distribution of trip starting times taken from the CVS data (normalized over daylight hours). This approach ensures that more observations are taken at peak traffic times, thereby somewhat alleviating the issue of low traffic volumes at some intersections. The direction of traffic is selected without replacement with equal probability of selection.

## 6. ESTIMATION METHOD

The estimation method applies the following weights to each record to obtain the seat belt usage rate by stratum.

- a) The observation period weight is the ratio of the total traffic through the intersection counted during the observation period to the number of cars observed, times the inverse probability of selection for the time of day, times the inverse probability of selection for the day of the week.
- b) The stratum weight is the ratio of the total number of intersections in the stratum to the number of intersections selected in the sample.

The province/region seat belt usage rate is then calculated as the weighted sum of the strata seat belt usage rates. The stratum weight is calculated as the ratio of the population of the stratum to the total population of the province/region.

The overall national seat belt usage rate is calculated as the weighted sum of the provincial seat belt usage rates. The province weight is the ratio of the number of vehicles registered (in the categories included in the survey) in each province to the total number of vehicles registered in Canada.

## 7. DISCUSSION

### 7.1 Combining of strata

When stratifying by region rather than by province only, strata had to be combined in order to conduct the survey within budget by reducing the number of sites. Strata U3,U4 and R2 were combined, U5,U6 and R3 were combined, R1 and R4 were combined and R5 and R6 were combined. In most cases, these strata had similar seat belt usage rates [4].

### 7.2 Ignoring roads of category five

Ignoring roads of category five may bias the survey results because seat belt use could be lower on these low volume roads. However, if these roads (category five) were included in the intersections frame, they would have a very high probability of being drawn for the sample, which is undesirable. In theory, as suggested by Shapiro [6], the best approach would be to stratify for the category of road intersections; however, creating a new stratum for intersections of roads of category five would result in doubling the number of sites, which also would double the cost of the survey.

## 8. CONCLUSION

The new stratification of the census subdivisions is based not only on the population size but also on the population density and the level of economic interaction with the nearest urban core.

A database of intersections that meet the criteria of being controlled by a traffic sign and of having sufficient traffic volume was created for the purpose of the Canadian National Seat Belt Surveys. Roads with low traffic volumes are screened out from the intersections frame for cost-effectiveness reasons.

The new design of the Canadian National Seat Belt Survey comprises several stratification levels and multiple stages of sampling. The stratification levels are province/economic region and rural or urban. In the first stage of sampling, the intersections are selected. The second stage of sampling is the selection of the time period of observation and the direction of traffic.

## References

- [1] Transport Canada, “Results of Transport Canada’s July 2001 Survey of Seat Belt Use in Canada”,  
<http://www.tc.gc.ca/roadsafety/tp2436/rs200107/en/menu.htm>
- [2] Transport Canada, “Canadian Motor Vehicle Traffic Collision Statistics 2001”, TP3322, February 2003,  
<http://www.tc.gc.ca/roadsafety/tp/tp3322/2001/en/menu.htm>
- [3] Transport Canada, “Road Safety Vision 2010”, TP13347E, June 2002,  
[http://www.tc.gc.ca/roadsafety/tp/tp13736/pdf/CRS\\_Target.pdf](http://www.tc.gc.ca/roadsafety/tp/tp13736/pdf/CRS_Target.pdf)
- [4] Chouinard A., “A New Design For The Canadian National Seat Belt Survey”, Canadian Multidisciplinary Road Safety Conference XIV, Ottawa, June 2004.
- [5] Statistics Canada, “Geosuite, 2001 Census”, Catalogue No. 92F0085XCB
- [6] Shapiro, et al., “Sample design issues for surveys involving the observation of drivers”, 2001 Proceedings of the American Statistical Association, Section on Survey Research Methods,  
[http://www.amstat.org/meetings/jsm/2001/index.cfm?fuseaction=abstract\\_details&abstractid=301652](http://www.amstat.org/meetings/jsm/2001/index.cfm?fuseaction=abstract_details&abstractid=301652)
- [7] DMTI Spatial, “CanMap Streetfiles V5.0, User Manual”, 2001 DMTI Spatial Inc.
- [8] Statistics Canada, “Canadian Vehicle Survey”, Catalogue No. 53F0004XIE.