SECTION 2: ABOUT THE SURVEY

2.1 Survey Area and Sampling

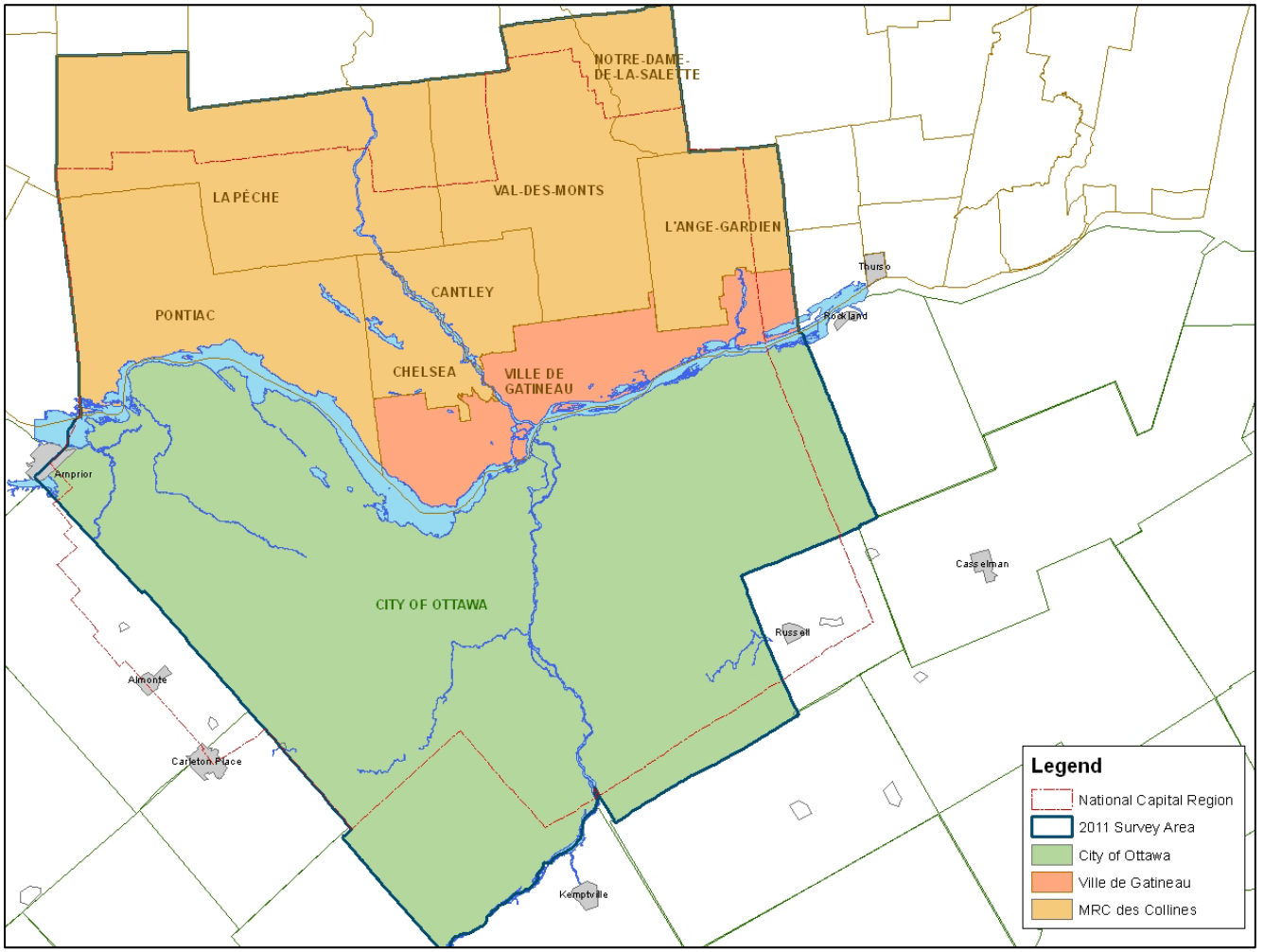
Households randomly selected to participate in the survey were from an area consisting of most of the National Capital Region; that is, from the City of Ottawa, Ville de Gatineau and the Municipalité régionale de comté des Collines de l'Outaouais. The municipalities of Almonte, Carleton Place and Russell (all in Ontario) which are within the National Capital Region were not included in the survey area. **Figure 2-1** depicts the surveyed areas. The 2011 survey area is the same as that of the 2005 survey.

For the purposes of tabulating the survey results in the Québec parts of the survey area, the Ville de Gatineau and the MRC des Collines de l'Outaouais are referred to collectively in this report as the "Outaouais."

The Consultant developed a survey sampling procedure that ensured the sample was distributed uniformly in proportion to the sample frame across the entire survey area, with some oversampling from sampling districts with smaller populations in order to keep the sampling errors within acceptable limits. The survey sample was composed of listed telephone numbers as well as randomly generated telephone numbers (to allow for contact with households with unlisted numbers). The sample plan aimed to achieve survey completions from 25,300 households, representative of the region.

The geography of the survey area was stratified into 42 different sampling districts. It should be noted that for reporting purposes, these sampling districts are amalgamated into 26 TRANS Districts. These sampling districts were based on a review of Census Tract boundaries (on the Québec side), traffic zone boundaries (on the Ottawa side), populations and household counts, Transitways, and the location of mid- and high-rise apartments. The sampling frame included 31 survey areas in Ontario, and 11 in Québec. A map of the 26 TRANS districts is provided in **Figure 5-1** in Section 5.

Figure 2-1: Map of the Survey Area



2.2 Survey Instrument

The same survey instrument used in past O-D surveys was used for the 2011 iteration to allow for maximum comparability with previous results in the National Capital Region. However, a few changes to the survey instrument were made, including the addition of new questions on household income, nature of occupation (if worker), arrival time, number of car passengers (if driving), and the relationship with the driver (if passenger).

The survey instrument is designed to collect information through a logical progression of events divided in three categories: household data, person data and trip data. **Table 2-1** outlines the primary data points of the survey.[4] The items listed in bold were newly collected in 2011.

Table 2-1: TRANS 2011 O-D Survey – Data Collected

Household	Person	Trip
Household No. Location Type of dwelling Household size Number of vehicles **Household income**	Person No. Age and gender Driver's license? Transit pass? (type, if any) Occupation status (worker, student, retiree, etc.) **Part-time student? (if worker, retiree, etc.)** **Part-time worker (if student, retiree, etc.)** **Occupation type (management, health etc.)** Usual place of work or school and parking arrangement Telecommuted the day before? (if no work trip)	Trip No. Origin and destination Departure and **arrival** time Purpose Mode(s) used (car – driver or passenger, transit, bicycle, walk, etc.) If transit: mode of access to stop/station, line(s) used, transfer point(s) **If car-driver: number of persons in the car** **If car-passenger: relationship with the driver**

Note: Bolded text indicates newly collected information in 2011.

The new items allow an improved understanding of the basis for modelling travel. In particular, household income has been found to be a key variable in trip generation and mode choice. Part-time student and worker status and occupation type similarly are important in determining trip generation and mode choice. Auto occupancy (number of persons in the vehicle) was used to corroborate and expand upon other data sources, notably vehicle classification and occupancy counts. Finally, the relationship between the vehicle passenger and the driver (i.e., whether they were members of the same household or not) is important in modelling ridesharing as a mode.

[4] Some administrative and control variables also are included in the final database.

It may be noted that the survey was intended to capture personal travel for all purposes, including work-related business trips. However, the survey excludes trips made by householders who made trips as a commercial driver (e.g., taxi driver) and those for work on the road (e.g., plumbing service calls).

See **Appendix A** for the full English survey instrument, which outlines the structure of the survey and the flow of information collected. Note that the survey was also conducted in the respondent's language of choice, English or French. All survey materials, including the survey instrument, were developed in both languages.

2.3 Survey Process

The 2011 O-D survey was conducted between September 20, 2011 and December 10, 2011, by way of computer-aided telephone interviews. A total of 26,342 surveys was completed among the NCR households.[5]

During the data collection period, surveys (i.e., interviews) were conducted from Tuesday to Saturday. This allowed interviewers to capture the survey participants' travel on the previous day; that is, on a working weekday. The majority of surveys were completed during weeknights between 16:30 and 21:15 and on Saturdays between 10:00 and 17:00. However, in order to maximize productivity, surveys were also completed on weekdays to reach residents requiring an appointment or who could not be contacted in the evening.

Prior to being called, sampled households were sent a brochure describing the study. The brochure was designed to notify the household of the forthcoming telephone call, and to inform on the nature of the survey and the confidentiality of the interview. A bilingual brochure was sent to households in Ontario and a French brochure (with information on how to receive a bilingual version) was sent out to households in Québec. Sending out advanced notification via letters, brochures and emails is a common and valuable practice that serves to enhance the participation rate and provides validation of the study.

As noted, survey participants were asked questions about all the trips made during the previous 24-hour weekday period, by them and by each person in their household (5 years of age and older). This includes trips made by foot, bicycle, rollerblade and motorized vehicle. Interviews were conducted with a representative of the household, 16 years of age and older, who could best provide information on the trips of all fellow household members.

Data were collected using SAQE, a system developed by the Ministère des Transports du Québec. The system also includes data quality and coding tools.

2.4 Data Quality and Processing

Many quality control mechanisms were in place during and after the data collection period.

[5] This represents the total number of surveyed households. However, a small number of these surveys was excluded from analysis due to insufficient data collected during the telephone interview. The summaries described in this report are based on the usable surveys (which constituted the vast majority of the surveys).

Supervisors regularly monitored interviews and checked completion statistics to ensure the quality of collected data. Completions were also monitored to ensure that a sufficient number of surveys were completed in each sampling district.

Multivariate validation tests integrated into SAQE allowed surveyors to complete basic data validation activities while conducting the survey. A substantial proportion of geo-coding was also completed during the survey process using the look-up tables integrated into SAQE, which contained most generators, intersections and address ranges.[6]

Once the survey was complete, SAQE conducted another series of validation checks for logic and completeness of information. A list of errors was generated for further investigation by coders. A list of generators and streets created by surveyors was also created. Coders checked if the new generator or street could be located in the existing look-up tables, and if not, validated the creation of a new generator.

Throughout the data collection, TRANS agency staff monitored the data collected and validated the coders' work. During and after the data collection, they geo-coded newly created generators and streets as well as households that had changed addresses by applying appropriate XY coordinates. They also performed checks to identify any logic errors that may have been missed by SAQE's integrated testing systems. Coders worked to correct any errors identified at this stage. During this phase of the data processing, approximately 1% of the household surveys were rejected due to too many vague or ambiguous origins or destinations, too many refused responses, or other problems identified with the survey data. MTQ also conducted a systematic geographic refinement of cases coded on street addresses, using assessment roles or street files to estimate the best possible XY coordinates.

Once the initial geocoding and cleaning phase was concluded, the entire dataset was extracted from SAQE to an MS Access database for final data processing. Survey response data were reviewed to ensure that all data fields contained values where expected. Outstanding geocoding issues were resolved to the extent possible. Treatments were applied to certain kinds of geocoded locations, such as offsets for locations coded to intersection and imputations of locations coded to street only without civic address. Once the geocoding was finalized, the data were also subjected to a number of standard trip logic tests (such as testing the distance between origin and destination for reasonableness with respect to the mode of travel used) and other validation tests.

The household, person, and trip records were screened against data acceptance criteria, and households with insufficient or poor information were rejected. Household surveys were rejected if household or demographic data critical to the analysis or data expansion were missing or could not be determined (e.g., household location, age and gender of persons within the household). Trips missing essential information such as departure time, purpose, mode, origin, or destination were flagged as 'poor trip records.' The household survey was rejected if the majority of persons five years of age or older (those for whom travel was captured) either had 50% or more poor trip records or had unknown mobility (i.e., the respondent refused to provide trip information for the given householder, or did not have knowledge of the trips made). At this stage, approximately 2.7% of the surveys collected were rejected. For accepted households, persons within the household with too many poor trips had their full 24-hour trip chains suppressed, and were treated as if they

[6] Surveyors were also able to manually enter a generator name, address or intersection if it was not found in the look-up tables.

had unknown mobility. For persons with a minority of poor trip records, the unknown trip information were imputed, where possible, using algorithms tailored to randomly assign values with the same distributions as those observed for cases with similar households, people, trip purposes, points of origin, modes of travel and/or other characteristics.

The finalized dataset for this origin-destination survey includes person records for a small proportion of householders whose mobility is unknown (whether due to initial refusal to provide any information or later excluded due to any missing trip information at all after data processing). Only complete 24-hour trip chains with all essential trip data known or imputed for all trips are used in the analysis presented in this report and the transportation modelling. The existence of persons with unknown mobility was compensated by using trip correction factors in the data expansion, as explained in the following section.

2.5 Data Expansion

Data weighting is used to adjust survey respondents' contributions to the overall survey results by a multiplication factor so as to compensate for both planned for and unexpected disproportionate results. The sampling plan oversampled certain geographic districts with smaller populations in order to obtain better data for areas with smaller populations. In addition to this planned-for disproportion by district, the unweighted survey data may also prove less than representative due to non-response bias, which occurs when certain types of household and/or individual are less likely to respond to the survey.

In the case of this survey, the intent of the data weighting is to adjust for correctible sources of disproportion or bias in the survey sample so that it represents (as best as possible) the actual travel patterns of all residents in the region, while expanding the household and person level data so that it represents (as best as possible) the actual number of residents and number of trips of the entire population. That is, the expanded survey data should provide survey estimates of the number of trips between areas, number of transit trips, etc., that closely align with actual counts within the region.

The following household and population controls were adjusted for in the data weighting and expansion, using 2011 Census data:

- Household counts by district (the study area stratified into the 42 sampling districts which were then amalgamated into the 26 TRANS districts for reporting (see **Figure 5-1**));
- Dwelling types within each district, grouped into single-detached, other ground-oriented dwellings (semi-detached and row/townhouses), and apartments/condominiums;
- Household sizes within each district (one-person, two-person, three to four persons, five or more persons);
- Population counts by district; and
- Age (nine groups) and gender distributions of the population within each district.

An iterative proportional fitting (IPF) method was employed to balance household weights and person weights for the multiple factors described above. In this method, incremental adjustments to the household weights are made in succession for each of the household controls, as well as a

composite adjustment to each household weight to account for the disproportionate distribution by age/gender amongst the members of each household. To limit the effect of extreme data weights, upper and lower limits were placed on weighting factors (with limits set relative to the base weight for household counts for each sampling district). Using the IPF method, any number of controls can be introduced. Each successive adjustment to balance a given control may slightly or significantly unbalance the correction previously introduced for a different control (which would normally be a disadvantage for multiplicative weighting). However, iteratively cycling through each control results in convergence to a solution where all household and population controls have expected distributions (to within reasonable tolerance). In this manner, all persons within each household typically carry the same base weight as the household, although it may be noted that small calibrations were made separately to the records at the household level and at the person level to ensure that the weighted counts of total households and of total persons matched Census figures.

Of note, this data weighting was undertaken for each of 42 sampling districts, whereas Section 5 of this report presents data for 26 TRANS districts, which are aggregations of the sampling districts (although occasionally sampling districts are sub-divided by TRANS district boundaries).

Following the data weighting and expansion, trip correction factors were introduced to correct for the following:

- A correction factor to account for disproportionate distribution (if any) of survey responses by weekday (Monday through Friday) in each district; and
- A correction factor to account for persons in each age/gender group within each district who have unknown mobility (e.g., respondent refused to provide trip data, or too many trips with poor or refused information).

The aforementioned correction factors were applied to the trip expansion factors only.

The expanded survey responses for household, person, and trip characteristics were compared to Census and other benchmark data (e.g., transit ridership in terms of both revenue trips and boardings) in order to validate the data expansion, with positive overall results. The expanded data were found to match very closely to the various controls adjusted for in the weighting, and also match closely other benchmark figures (such as number of employed persons in the planning area).[7]

2.6 Survey Completions

A total of 26,342 randomly selected households participated in the survey and 25,374 surveys were retained for expansion and analysis. **Table 2-2** below summarizes the number of households per district, per region and for the survey area as well as the number of households with valid telephone numbers sampled, interviewed households and validated households, meaning those retained for analysis.

[7] Note, however, that the absence of unlisted telephone numbers from the sampling universe constitutes a known bias risk. This includes households with no land lines or which use only cell phones.

Table 2-2: Household Interview Completions per Sampling District, 2011

Sampling District	Occupied Dwellings (2011 Census)	Valid Household Telephones Sampled (1)	Interviewed Households (2)	Valid Survey Response Rate (3)	Validated Household Surveys (4)	Sampling Rate (5)	Estimated Sampling Error (6)
101 Hull Centre	10,938	2,363	678	29%	619	5.7%	± 3.8%
102 Plateau	18,878	2,399	730	30%	742	3.9%	± 3.5%
103 Aylmer	10,259	1,614	512	32%	500	4.9%	± 4.3%
104 Nord de Hull-Aylmer	5,996	1,168	391	33%	374	6.2%	± 4.9%
105 Hull, Nord de St-Raymond	14,079	2,202	751	34%	699	5.0%	± 3.6%
106 Gatineau (ouest)	18,009	2,917	915	31%	899	5.0%	± 3.2%
107 Gatineau (centre)	13,526	2,234	714	32%	681	5.0%	± 3.7%
108 Gatineau (est)	11,518	1,644	516	31%	508	4.4%	± 4.3%
109 Nord de Gatineau, Val-des-Monts, Notre-Dame-de-la-Sallette, L'Ange-Gardien	11,171	1,580	508	32%	523	4.7%	± 4.2%
110 Buckingham / Masson-Angers	9,841	1,496	486	32%	464	4.7%	± 4.4%
111 Chelsea / Cantley	5,991	1,150	372	32%	359	6.0%	± 5.0%
RDD, Outaouais prefixes (7)	n/a (7)	1,577	96	6%	n/d (7)	n/d (7)	n/a (7)
201 Ottawa Central	6,251	1,273	417	33%	372	6.0%	± 4.9%
202 Ottawa Inner Central	13,764	2,692	821	30%	742	5.4%	± 3.5%
203 Ottawa Inner West	12,125	1,816	613	34%	586	4.8%	± 3.9%
204 Ottawa Inner East	13,094	1,813	662	37%	623	4.8%	± 3.8%
205 Ottawa Inner South	6,451	1,109	408	37%	393	6.1%	± 4.8%
206 Ottawa East (n)	12,560	1,879	620	33%	613	4.9%	± 3.9%
207 Ottawa East (s)	12,678	2,377	652	27%	649	5.1%	± 3.7%
208 Beacon Hill	14,029	2,374	784	33%	785	5.6%	± 3.4%
209 Alta Vista (n)	12,034	1,743	600	34%	575	4.8%	± 4.0%
210 Alta Vista (s)	16,807	2,568	869	34%	844	5.0%	± 3.3%
211 Alta Hunt C (e)	13,592	1,935	601	31%	608	4.5%	± 3.9%
212 Alta Hunt C (w)	12,292	1,854	606	33%	599	4.9%	± 3.9%
213 Merivale (s)	19,244	2,927	966	33%	944	4.9%	± 3.1%
214 Merivale (n)	13,748	1,996	670	34%	656	4.8%	± 3.7%
215 Ottawa West (e)	12,120	1,830	603	33%	605	5.0%	± 3.9%
216 Ottawa West (w)	11,953	1,835	577	31%	588	4.9%	± 3.9%
217 Bayshore/Cedarview (s)	18,563	2,849	912	32%	919	5.0%	± 3.2%
218 Bayshore/Cedarview (n)	13,666	1,949	664	34%	641	4.7%	± 3.8%
219 Orleans (w)	12,516	1,999	648	32%	637	5.1%	± 3.8%
220 Orleans (e)	18,897	2,847	880	31%	866	4.6%	± 3.3%
221 Orleans (s)	11,538	1,717	548	32%	541	4.7%	± 4.1%
222 Rural East	4,093	1,153	397	34%	386	9.4%	± 4.7%
223 Rural SouthEast	9,316	1,427	467	33%	453	4.9%	± 4.5%
224 SouthGloucest/Leitirm	6,236	1,153	361	31%	352	5.6%	± 5.1%
225 South Nepean (ne)	12,467	1,916	609	32%	595	4.8%	± 3.9%
226 South Nepean (rest)	13,796	1,965	592	30%	599	4.3%	± 3.9%
227 Rural SouthWest	9,188	1,534	550	36%	536	5.8%	± 4.1%
228 Kanata (s)	16,167	2,560	798	31%	776	4.8%	± 3.4%
229 Kanata (n)	12,592	1,875	562	30%	563	4.5%	± 4.0%
230 Stittsville	9,253	1,561	472	30%	482	5.2%	± 4.3%
231 Rural West	8,751	1,433	478	33%	478	5.5%	± 4.4%
RDD, Ottawa prefixes (7)	n/a (7)	3,712	267	7%	n/d (7)	n/d (7)	n/a (7)
Outaouais Total	**130,206**	**22,344**	**6,669**	**30%**	**6,368**	**4.9%**	**± 1.2%**
Ottawa Total	**379,781**	**63,671**	**19,674**	**31%**	**19,006**	**5.0%**	**± 0.7%**
Survey Area Grand Total	**509,987**	**86,015**	**26,343**	**31%**	**25,374**	**5.0%**	**± 0.6%**

Notes:

(1) Valid household phone numbers excluding those with phone numbers determined to be incorrect (i.e., not-in-service, non-residential) or outside the survey area. In total 106,825 phone numbers were dialled, but, of these, 20,810 were invalid. Counts reflect the initial geocoding to sampling district based on telephone listing address.
(2) Valid survey response rate = total survey interviews completed as % of valid household phone numbers.
(3) Interviewed Households. Represents the number of households interviewed from each original sampling district. Figures in this column reflect initial geocoding to sampling district based on telephone listing address, which may have later have been updated. Households from the Random Digit Dialled (RDD) strata were not geocoded to district, but were later assigned the appropriate district based on the address provided by the respondent.
(4) Validated household surveys = surveys with sufficient good-quality information retained in the final survey dataset. In total, 969 surveys were rejected at various stages during geocoding, data cleaning, and validation. The figures in this column reflect the final geocoded locations of households.
(5) Sampling rate = validated household survey completions in final data set as % of total occupied dwellings.
(6) Sampling error = the margin of error of the survey results for households within each district due to random sampling, at a 95% confidence level (19 times out of 20), and for response proportions of 50% (maximum sampling error). This calculation has not been adjusted to account for data expansion.
(7) RDD = random digit dialled phone numbers. The survey included a random selection of RDD phone numbers for land line and cell phone prefixes in the survey area; however, at client request, only limited dialling of landline RDD numbers was undertaken late in survey administration. After survey completion, interviews with RDD phone numbers were geocoded to the appropriate sampling district (increasing the number of valid surveys in certain districts).

The table also presents the margin of error associated with random sampling,[8] which is, overall, approximately ±0.6% at a 95% confidence level (theoretically, for a given survey question, the true value for the population would be somewhere within the margin of error of the survey results, 19 times out of 20). Data weighting may slightly increase the sampling error beyond this. As with any survey, other possible sources of error may include: under-coverage error (e.g., exclusion of certain population groups such as institutional populations and households without telephones), non-response bias (corrected for in part by data weighting for key demographic variables), measurement error (controlled for by using a well developed survey instrument and in-depth interviewer training); processing errors during data coding and editing (controlled for through extensive data validation); and/or weighting errors (Census statistics used as weighting controls may be subject to over-/under-counts or other error; the weighting required aggregation of certain dwelling type, household size and age categories in order to prevent extreme weights; and the weighting processes, while sophisticated, may not perfectly synchronize the survey results with the Census counts).

In total, the final survey data set includes information on 25,374 households and 62,897 householders. Of the 59,663 householders five years of age or older and eligible to provide information on their travel for the day surveyed, 56,804 had known mobility (i.e., were confirmed as not travelling or provided good trip information for the majority of their trips). These individuals furnished information on 153,248 valid trips. After data weighting and expansion was applied, this robust set of origin-destination survey data serves as a snapshot of the travel patterns for households with telephones on a typical autumn 2011 weekday.

[8] Sampling error computed for response proportions of 50%, the circumstance for which the maximum variation of survey results is usually observed. For response proportions approaching 0% or 100%, the sampling error is less than the maximum sampling error computed here.