

Beyond Reported Trips: A Bayesian Framework for Correcting Household Travel Surveys

Melvyn Ming Xi Li

`melvyn.li@utoronto.ca`

University of Toronto

Kaili Wang

Massachusetts Institute of Technology

Khandker Nurul Habib

University of Toronto

Research Article

Keywords: household travel survey, GPS travel survey, missing trip identification, travel survey correction, Bayesian survey correction

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10968197/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Beyond Reported Trips: A Bayesian Framework for Correcting Household Travel Surveys

Melvyn Ming Xi Li*, Ph.D. Candidate
Department of Civil & Mineral Engineering,
University of Toronto
35 St. George Street, Toronto, ON, Canada, M5S 1A4
ORCID: 0009-0000-1282-884X
Email: melvyn.li@utoronto.ca

Kaili Wang, Ph.D.
Postdoctoral Fellow, Department of Civil & Environmental Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge, MA 02139, USA
ORCID: 0000-0002-7996-5600
Email: kaili668@mit.edu

Khandker Nurul Habib, Ph.D., P.Eng.
Professor, Department of Civil & Mineral Engineering
University of Toronto
35 St. George Street, Toronto, ON M5S 1A4
ORCID: 0000-0003-1007-6706
Email: khandker.nurulhabib@utoronto.ca

Total Number of Pages: 34

*Corresponding Author

ACKNOWLEDGMENTS OpenAI's ChatGPT was used to improve the clarity and flow of parts of the manuscript. All content was reviewed and verified by the authors.

AUTHOR CONTRIBUTIONS The authors confirm their contribution to the paper as follows: Study conception and design: M. Li, K. Wang, K.M.N. Habib; Data collection: K. Wang, M. Li; Analysis and interpretation of results: M. Li, K. Wang; Draft manuscript preparation: M. Li, K. Wang; Overall project supervision: K.M.N. Habib. All authors reviewed the results and approved the final version of the manuscript.

STATEMENTS AND DECLARATIONS The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING This research is funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant.

ABSTRACT

Household travel surveys (HTS) remain essential for transportation planning but are susceptible to recall and proxy-reporting biases that can systematically omit short, discretionary, and intermediate trips. This study develops a Bayesian core–satellite framework that uses insight from smaller, GPS-based smartphone travel surveys to identify and correct missing trips in a large conventional HTS while preserving population coverage and socioeconomic detail. The framework is demonstrated using the 2022 Transportation Tomorrow Survey in the Greater Toronto and Hamilton Area, supplemented by multi-day smartphone surveys conducted in 2023 and 2025. Bayesian models estimated from HTS are treated as priors and sequentially updated using GPS observations. Parametric Bayesian models estimate missing non-skeleton trip counts and distances, while non-parametric Bayesian models reconstruct trip purpose, start time, and mode. Primary and proxy respondents are modelled separately to account for differences in reporting errors. Posterior predictive diagnostics show close agreement between observed and simulated trip-count and distance distributions from the updated models. Among mobile primary respondents, correction increases the mean trip rate from 3.05 to 3.90 trips per person per day, while the proxy rate increases from 2.64 to 3.63. Corrected trip rates agree with rates found in the GPS-based data. The reconstructed trips increase representation of discretionary activities and shorter trips while retaining the underlying temporal, distance, and modal structure of the core survey. The framework provides a probabilistic approach for integrating passive mobility data with conventional HTS datasets to improve travel statistics used in transportation planning and demand modelling.

Keywords: household travel survey, GPS travel survey, missing trip identification, travel survey correction, Bayesian survey correction

1 Introduction

Transportation planning depends on reliable information about how individuals conduct daily activities and travel between them. Household travel surveys (HTS) collect information on daily trips alongside socioeconomic characteristics. This data supports all travel analysis and demand modelling used throughout the transportation sector. Transportation master plans and technical reports regularly use HTS-derived indicators to establish existing conditions and justify proposed policies. Planning agencies use HTS data not only to describe travel demand, but also to set policy targets, evaluate infrastructure strategies, estimate and calibrate travel-demand models, and monitor plan performance.

In the GTHA, the Transportation Tomorrow Survey has been used to establish mode-share and trip-distance baselines for active transportation, transit, parking, and travel-demand-management policies; to estimate regional activity-based and four-step models used to assess road and rapid-transit alternatives; and to monitor trip counts, travel distances, and vehicle-kilometers travelled. For example, the GTHA Regional Transportation Master Plan and the Metrolinx 2041 Regional Transportation Plan use mode shares, peak-period trip counts, and trip-distance distributions to support frequent rapid transit, fare integration, cycling-network, and school-travel policies (Province of Ontario 2025; Metrolinx 2018). Similar applications appear elsewhere. In Canada, the city of Ottawa uses peak-period mode shares to set walking, cycling, and transit targets and the city of Vancouver uses mode shares to guide rapid-transit, bicycle-parking, and automobile-parking policies (City of Vancouver 2012; City of Ottawa 2013). Worldwide, the city of Minneapolis, USA uses mode shares to define its non-automobile travel target and the city of London, UK uses its household survey to monitor active-travel participation and daily trip rates (City of Minneapolis 2020; Mayor of London 2025). **Appendix A** summarizes detailed applications of recurring HTS-derived measures in transportation master plans developed by municipalities globally.

The reliability of self-reported travel diaries has long been a concern in HTS research because several related biases affect the data. Recall bias occurs when respondents do not accurately remember their activities. Proxy-reporting bias arises when one household member reports another person's travel without complete knowledge of intermediate stops or trip details. Non-response bias occurs when survey participants differ systematically from those who do not participate. Sampling bias arises when demographic or geographic groups are overrepresented or underrepresented.

Survey quality is therefore not only a data-collection issue; it directly affects all aspects of transportation planning (Nakamya et al. 2007). Because household travel surveys are commonly used to estimate trip-generation rates, calibrate travel-demand models and monitor mobility trends, reporting errors can propagate directly into the evidence used for transportation planning. Research has shown that trip omission in HTS can underestimate modeled vehicle travel by nearly 30%, while diary fatigue and changes in survey methodology can generate spurious temporal trends and large apparent changes in walking mode share (Meurs et al. 1989; Wolf et al. 2003). Consequently, uncorrected survey data may misrepresent baseline travel demand, modal-split trends and the anticipated effects of transportation investments and policies (Sammer et al. 2018; Buehler and Pucher 2024). Despite the growth of GPS-based and passively generated mobility data, large-scale HTS remains a principal source of jointly observed travel behavior and socio-demographic context. Their value lies in their broad population coverage and ability to relate travel patterns to household and individual characteristics. New mobility datasets should complement these strengths instead of looking to replace them entirely (Wang et al. 2022; Zhao et al. 2025).

This study aims to develop a Bayesian framework for improving the practical use of conventional HTS data by identifying and mitigating missing trips that stem from biases inherent to self-reported HTS. This will be done through updating statistics derived from HTS using supplementary GPS-based smartphone travel survey data. The framework will present self-reported and passively collected travel diaries as imperfect but complementary observations of latent travel behavior. Then, correction strategies that target missing trips will be developed through Bayesian inference. This preserves the broad descriptive value of core HTS while using GPS-based data to support more accurate granular travel observations. An appropriate framework must therefore propagate uncertainty and identify cases in which the supporting evidence is limited (Wang et al. 2022; Zhao et al. 2025; Sfeir et al. 2024). Surveys from the GTHA will be used as initial evidence to construct the framework, including the household Transportation Tomorrow Survey (TTS) and two cycles of GPS-based smartphone survey data.

The remainder of the paper proceeds as follows. Section 2 provides a literature review of current HTS limitations and correction procedures. Section 3 describes the HTS and GPS-based datasets used in the development of the framework. Section 4 explains the methodology and layout of the framework, while section 5 discusses the

estimated models as well as results obtained from the correction framework. Section 6 concludes the study and suggests further avenues of research.

2 Literature Review

2.1 Limitations of Household Travel Survey Data

The reliability of self-reported travel diaries has long been a concern in household travel survey research. Comparisons with GPS traces show that respondents may omit trips or report inaccurate travel times and distances (Wolf et al. 2003; Bricka and Bhat 2006). Conventional HTS are also affected by respondent burden, recall limitations, and declining participation rates (Stopher and Greaves 2007). These errors distort aggregate travel estimates. Some errors are minor or random, while others are systematic and can distort aggregate statistics.

Measurement errors do not affect all trips equally. Routine commute trips are often easier to recall because they structure the day, while shorter trips are more easily forgotten. A self-reported diary may therefore capture the basic structure of daily travel while omitting minor but policy-relevant trips. Validation studies have long shown that conventional diaries understate trip-making and that omissions vary by trip type, traveler, and survey conditions (Clarke et al. 1981; Murakami and Wagner 1999; Stopher et al. 2007; Bricka et al. 2012). Missed active and off-peak trips can lead to underestimated walking and cycling shares and an overstated concentration of travel during AM and PM peak periods (Stopher et al. 2007; Chapleau et al. 2018; Sfeir et al. 2024).

Several related biases affect HTS data. Recall bias occurs when respondents do not accurately remember their activities. Proxy-reporting bias arises when one household member reports another person's travel without complete knowledge of intermediate stops or trip details. Proxy-reported diaries therefore contain fewer trips and simpler activity sequences than directly reported diaries (Badoe and Steuart 2002; Verreault and Morency 2015; Maruyama et al. 2019; Wang et al. 2022). Non-response bias occurs when survey participants differ systematically from those who do not participate. Selective completion may be associated with privacy concerns, time availability, survey familiarity, or the complexity of an individual's schedule (Lutig et al. 2022; Chapleau et al. 2018; Wang et al. 2022). Sampling bias arises when demographic or geographic groups are overrepresented or underrepresented.

The distinction between sampling bias and trip-reporting error is central to this study. Sampling bias concerns who enters the dataset, whereas trip-reporting error concerns the completeness of diaries after participation. Weighting can improve alignment with known population totals, but it cannot automatically recover omitted trips. A sample may therefore be demographically representative while remaining behaviorally incomplete.

2.2 Existing Survey Correction Methods

Survey weighting is commonly used to correct sample imbalance. Iterative proportional fitting, or raking, adjusts respondent weights until selected marginal distributions match known population totals (Deming and Stephan 1940). Post-stratification assigns weights within population cells and calibration weighting satisfies constraints while respecting original design weights (Holt and Smith 1979; Deville and Särndal 1992; Kott and Chang 2010). However, weighting alone does not recover omitted trips.

Imputation methods replace missing information with plausible values derived from observed records. Hot-deck imputation uses values from similar cases, while multiple imputation generates several completed datasets to retain uncertainty in subsequent analysis (Rubin 1987; Andridge and Little 2010). Statistical matching transfers information across datasets through shared variables (Rubin 1986).

Rule-based checks can also identify some incomplete travel sequences. For example, a diary may contain a discontinuity when the destination of one trip differs from the origin of the next, or when a respondent remains away from home without a reported return trip. GPS traces provide a stronger standard for individual trips because they may reveal movements absent from self-reported diaries (Wolf et al. 2003; Bricka and Bhat 2006). GPS processing workflows reconstruct travel diaries through trip-end detection, segmentation, and mode inference (Yang et al. 2020; Xiong et al. 2020; Li et al. 2024).

Model-based procedures are useful when reporting errors that follow systematic patterns. Regression models, discrete-choice models, and machine-learning classifiers are common models for HTS correction. They can help estimate the probability of under-reporting, the expected number of missing trips, or the characteristics of omitted

travel. Similar approaches are used to account for non-response and sample-selection mechanisms when survey inclusion is associated with the outcome of interest (Kott and Chang 2010; Little et al. 2020).

2.3 GPS Data as Supplementary Evidence

GPS technology is increasingly used for travel recording. Unlike self-reported diaries, GPS-enabled devices passively record location and timing as individuals complete their daily activities. These records may reveal trips that are missing from self-reported diaries, particularly short movements and intermediate stops. Passive sensing is most useful when combined with effective algorithmic processing and respondent validation (Bricka and Bhat 2006; Stopher et al. 2007; Servizi et al. 2021).

However, GPS-based surveys should not be treated as complete replacements for conventional HTS. Regional surveys still provide larger samples, broader geographic coverage, and detailed socioeconomic information. Their most effective role is to supplement conventional surveys through explicit linkage, weighting, correction, or data-fusion procedures that acknowledge residual uncertainty (Bayart et al. 2009; Egu and Bonnel 2020; Lugtig et al. 2022; Wang et al. 2022; Zhao et al. 2025).

Early GPS studies relied on dedicated devices that respondents had to borrow and return. Smartphones have reduced these operational barriers. Modern systems can combine passive sensing with web- or app-based interfaces that allow participants to review and correct generated diaries (Cottrill et al. 2013; Kim et al. 2015). Prompted-recall surveys use detected movements as the basis for travel diary generation, asking respondents to verify recorded events rather than report entire trips from memory (Bachu et al. 2001; Greaves et al. 2010; Cottrill et al. 2013).

However, GPS data also contains limitations. Records may be incomplete when devices lose signal, restrict background tracking, or are not carried by respondents. Smartphone systems must balance location accuracy, battery consumption, and participant burden (Cottrill et al. 2013; Servizi et al. 2021). Processing algorithms must also convert raw traces and points into actual trips and diaries. These procedures may unnecessarily split or merge trips, overlook short stops, confuse similar modes, or otherwise misclassify travel. Processing decisions therefore introduce uncertainty into the final diary (Cottrill et al. 2013; Servizi et al. 2021; Yang et al. 2022).

Sample selectivity is another concern. Participation in GPS studies may vary with age, income, digital literacy, and other factors (Keusch et al. 2019). This is particularly important when correction factors estimated from a small satellite sample are transferred to a larger HTS. Weighting, calibration, and uncertainty analysis are required to limit the transfer of selection bias.

Privacy must also be addressed. GPS traces may reveal sensitive locations such as home or workplaces. Even after identifiers are removed, a small number of spatiotemporal observations may distinguish individuals within a large dataset (de Montjoye et al. 2013; Primault et al. 2019). Thus, GPS-based surveys require informed consent, secure storage, and clear access and retention policies.

2.4 Summary

The reviewed literature demonstrates that conventional household travel surveys are affected by multiple sources of error. However, most established correction procedures focus primarily on improving sample representativeness. Methods such as raking, post-stratification, and weighting align the demographic and geographic composition of survey respondents with known population characteristics. Although these methods can reduce distortions associated with who participates in the survey, they do not directly address whether participating respondents provide complete travel diaries. Consequently, an HTS may be statistically representative of the population while still underestimating trips within individual travel diaries.

On the other hand, GPS-based surveys provide an important supplementary source of evidence because they can identify movements that are absent from self-reported diaries due to various biases. Nevertheless, systematic methods for using GPS observations to update and correct large-scale HTS datasets remain scarce. Existing studies commonly use GPS data for linkage and data fusion or aggregate reweighting, but few studies develop correction frameworks that update individual HTS records based on evidence from GPS data. Limited attention has been given to modelling heterogeneous under-reporting while simultaneously accounting for the uncertainty in the different data sources and sparseness of observations for certain demographic cohorts in GPS-based surveys. Thus, this study targets the development of an integrated framework that uses GPS-based evidence to update individual HTS trip records while preserving the socioeconomic and geographic coverage of conventional surveys.

3 Description of Datasets

This study uses datasets collected under a core–satellite survey paradigm in the Greater Toronto and Hamilton Area (GTHA). The core dataset is the 2022 Transportation Tomorrow Survey (TTS), a large regional household travel survey that sampled approximately 5% of the GTHA population. The TTS collected one-day weekday travel diaries through self-reported survey methods and was weighted to align with the 2021 Canadian Census. It provides a benchmark for travel behavior across the region. However, because the survey relies on self-reported single-weekday diaries, recall limitations may lead to the under-reporting of short, discretionary, and non-mandatory trips (Data Management Group 2024).

The core survey is supplemented by two multi-day GPS-based smartphone satellite survey cycles conducted in 2023 and 2025. The Fall 2023 cycle collected seven-day travel diaries from 956 respondents (Li et al. 2024). The survey design and resulting dataset are described in detail by Li et al. (2024). Similarly, the 2025 cycle used a dedicated smartphone application that automatically identified trips which respondents labelled for mode and purpose (Li et al. 2025). Data collection took place in Spring 2025. A total of 999 respondents generated at least two days of travel data and 872 completed at least one week of tracking. Respondents contributed an average of 13.7 days of captured data and 10.7 days of fully labelled data. Further design details and dataset information are described by Li et al. (2025). Both GPS-based smartphone cycles also collected socioeconomic characteristics, work-from-home schedules, and e-shopping activities for the same observation period.

Average trip rates differ substantially across the three datasets. The self-reported 2022 TTS records the lowest weekday rate, at 2.06 trips per person per day, whereas the passively collected GPS surveys report 3.16 weekday trips in 2023 and 3.83 in 2025. The GPS cycles also record weekend rates of 3.95 and 3.30 trips per person per day, respectively. The lower TTS rate is consistent with recall and proxy-reporting biases in conventional household travel surveys, which may omit short, discretionary trips. A summary of the calculated trip rates across datasets is in **Table 1**.

Table 1. Comparison of Trip Rates Across the 2022 TTS and 2023 and 2025 GPS Cycles

Day of Week	2022 TTS	2023 GPS	2025 GPS
Weekday Average	2.06	3.16	3.83
<i>Monday</i>	1.98	2.94	3.62
<i>Tuesday</i>	2.08	3.10	3.94
<i>Wednesday</i>	2.12	3.14	3.66
<i>Thursday</i>	2.06	3.24	3.56
<i>Friday</i>	2.02	3.39	3.77
Weekend Average	-	3.95	3.30
<i>Saturday</i>	-	4.23	3.64
<i>Sunday</i>	-	3.58	2.96

Li et al. (2025) provide an in-depth comparison of socioeconomic attributes and descriptive statistics between the three datasets. Together, the datasets represent the population coverage of the TTS with the temporal detail and improved trip detection of GPS-based surveys. The satellite samples are smaller and are not fully representative of the entire GTHA population. These differences should be addressed through sample weighting or model-based adjustment when estimates are transferred to the broader population. Nevertheless, the repeated GPS cycles provide a basis for identifying systematic differences between self-reported and passively recorded travel diaries and for estimating corrections to the core survey.

4 Methodology

4.1 Framework Overview

A Bayesian framework to improve descriptive travel statistics derived from a conventional HTS using GPS-based satellite surveys is defined. The proposed approach operates at the respondent level. Models estimated for trip rates and attributes using the core HTS are updated through Bayesian inference using GPS data as likelihood. These models are then applied to each core survey respondent to estimate the number of trips that would be expected under the posterior. Where the posterior prediction exceeds the number of observed trips, the difference is treated as a probabilistic estimate of missing travel. Attributes are generated for each missing trip and inserted into a copy of the

core dataset. The methodology combines parametric and non-parametric Bayesian models. The two components are linked through a sequential trip-reconstruction procedure. An overview of the estimation and application of the framework is found in **Figure 1**.

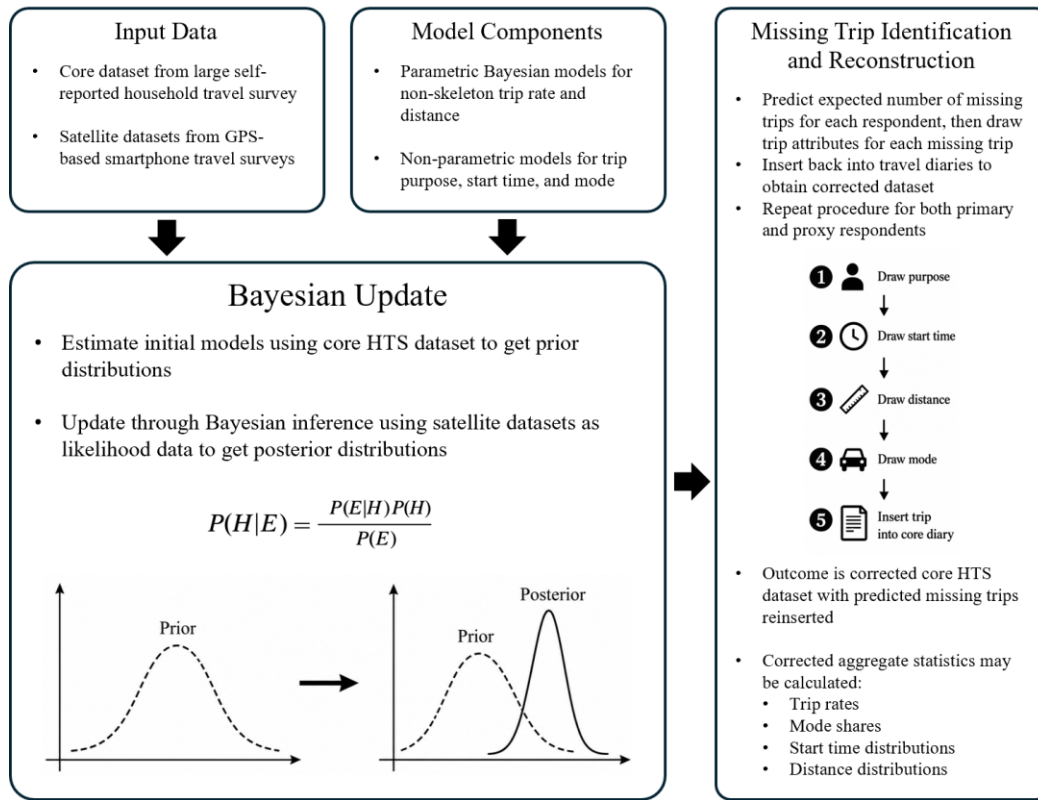


Figure 1. Schematic of Bayesian Core-Satellite Correction Framework

4.2 Bayesian Statistics and Inference

Bayesian statistics provide a formal framework for combining multiple evidence sources while explicitly representing and propagating uncertainty. Through Bayesian inference, prior distributions can be updated into posteriors by incorporating new data as the likelihood. Self-reported and GPS-derived trip counts can be treated as separate measurements of a latent true count, with Bayesian statistics serving as a tool to utilize both. This approach is consistent with Bayesian evidence-synthesis methods that combine heterogeneous sources while preserving their individual limitations (Ades and Sutton 2006; Richardson and Gilks 1993; Muff et al. 2015).

Bayesian models may be divided broadly into parametric and non-parametric approaches. A parametric regression-based model assumes that the observed data may be expressed as a function of a finite number of explanatory variables following a specified distribution. Non-parametric Bayesian methods, by contrast, provide greater flexibility. The model complexity is not restricted to a finite-dimensional form and can adapt as additional data becomes available. This is useful when distributions are complex, multimodal, or difficult to specify in advance. Results are influenced by choices such as base distributions, concentration parameters, and kernels (Ferguson, 1973; Müller and Mitra, 2013).

The two methods need not be treated as mutually exclusive. A practical framework may utilize both types of models. For example, parametric models can be used to estimate a set of outcomes adequately represented by standard parametric distributions while non-parametric models can be used to estimate the remaining values that are not easily described parametrically. This combined approach preserves interpretability while allowing complexity where required.

4.3 Data Harmonization and Sample Adjustment

The core and satellite datasets must first be harmonized, as they stem from different surveys. Common definitions will be established for trip purpose, mode, trip start time, trip distance, respondent characteristics, and reporting type. GPS observations will be restricted to comparable weekdays because the core HTS records travel for previous weekdays only. Records with incomplete trip details or validation will be identified and removed through quality checks before model estimation.

4.4 Parametric Bayesian Models

Parametric Bayesian regression models are used for trip parameters that can be represented using standard parametric distributions. Candidate model specifications will be compared through posterior predictive checks and model-comparison criteria. The final selection will reflect the empirical distribution of each parameter while retaining a clear interpretation due to its parametric definition.

4.4.1 Skeleton and Non-Skeleton Trip Rates

Trips are separated into skeleton and non-skeleton categories before model estimation. Skeleton trips are structural travel associated with work, school, and home activities. These activities tend to anchor the daily schedule and are generally easier to remember. Non-skeleton trips are flexible or discretionary travel, including shopping, errands, recreation, and other purposes. Trip rates are non-negative integer outcomes. For respondent i and trip category c , where c denotes skeleton or non-skeleton travel, the initial specification is a negative binomial model:

$$\begin{aligned} N_{ic} &\sim \text{NegBin}(\mu_{ic}) & (1) \\ \log(\mu_{ic}) &= \beta_{0c} + x_i^T \beta_c & (2) \end{aligned}$$

where N_{ic} is the trip count, μ_{ic} is the expected count, x_i is a vector of respondent characteristics, and β_c is a coefficient vector. Candidate explanatory variables include age, household size, number of household vehicles, employment status, driver's license status, transit-pass ownership, dwelling type, household region, workplace arrangement, etc.

A Poisson model will be considered as a simpler alternative. However, a negative binomial specification is expected to be more suitable when the variance exceeds the mean. Hurdle models will also be evaluated if the core dataset contains a substantial number of zero-trip records. These alternatives distinguish between the probability of observing an immobile day and the expected number of trips among respondents with reported travel.

Each trip-rate model is estimated using the core HTS first. The resulting coefficient distributions form the HTS-informed priors. The same model is then updated using GPS-based satellite observations. The posterior predictive distributions are applied to core survey respondents to estimate their expected number of skeleton and non-skeleton trips.

The analysis primarily models non-skeleton trip rates since skeleton trips are more accurately reported. Skeleton trips are structural travel associated with work, school, and home activities. These activities tend to anchor the daily schedule and are generally easier to remember. Non-skeleton trips are flexible or discretionary travel, including shopping, errands, recreation, and other purposes. These trips are more likely to occur as intermediate stops or short movements and may be more susceptible to biases and omission (Hassounah et al. 1993; Bricka and Bhat 2006; Verreault and Morency 2015; Aschauer et al. 2018). Preliminary results show that the difference between estimated and observed skeleton trips is distributed around zero with relatively small variance. Thus, there is little evidence of systematic under-reporting in skeleton trips compared to non-skeleton trips.

4.4.2 Trip Distance

Trip distance is continuous, non-negative, and typically right skewed. A log-normal model is used as the initial specification:

$$\log(d_{ij}) \sim \mathcal{N}(\gamma_0 + z_{ij}^T \gamma, \sigma^2) \quad (3)$$

where d_{ij} is the distance of trip j made by respondent i , z_{ij} is a vector of trip and respondent characteristics, γ is a coefficient vector, and σ^2 is the residual variance. Candidate explanatory variables include trip purpose, start time, respondent characteristics, and geographic area.

4.5 Non-Parametric Bayesian Models

Several travel attributes have empirical distributions that may not be represented adequately by a single parametric family. Trip purposes are distributed into discrete categories. Trip start times can exhibit multiple peaks across the day. Mode choice depends on distance and trip purpose. These attributes will be modelled using non-parametric Bayesian inference.

For attribute a , let G_a denote an empirical base distribution derived from the core HTS. A Dirichlet-process prior is specified as:

$$G_a \sim \text{DP}(\alpha_a, G_{H,a}) \quad (4)$$

where α_a is a concentration parameter controlling the influence of the HTS-informed base distribution. After observing values $y_{a1}, \dots, y_{an}$ from the GPS-based satellite survey, the posterior becomes:

$$G_a \mid y_{a1}, \dots, y_{an} \sim \text{DP}\left(\alpha_a + n_a, \frac{\alpha_a G_{H,a} + \sum_{j=1}^{n_a} \delta_{y_{aj}}}{\alpha_a + n_a}\right) \quad (5)$$

where $\delta_{y_{aj}}$ denotes a point mass at GPS observation y_{aj} . The posterior distribution combines the broad coverage of the HTS with the additional behavioral detail observed in the satellite survey.

For categorical outcomes with a finite set of alternatives, a Dirichlet-multinomial formulation provides a simpler implementation. Let $q_{H,a}$ denote the vector of proportions for attribute a in the core HTS, let κ_a denote an effective prior sample size, and let $n_{G,a}$ denote the GPS counts. The prior and posterior are expressed as:

$$\begin{aligned} \pi_a &\sim \text{DP}(\kappa_a, q_{H,a}) \\ \pi_a \mid D_G &\sim \text{DP}(\kappa_a, q_{H,a} + n_{G,a}) \end{aligned} \quad (6) \quad (7)$$

4.5.1 Trip Purpose

Trip purpose is estimated based on the posterior distribution of non-skeleton purposes. The core HTS provides the initial purpose distribution, while the GPS observations update the probability assigned to each purpose. The purpose for each missing trip is drawn first as it can be considered the most independent attribute of the trip. Intuitively, one would have an objective in mind for each out-of-home trip before deciding other details such as time of day and mode.

$$\Pr(P_{ij} = p \mid D_H, D_G) \quad (8)$$

where P_{ij} is trip purpose, D_H is the core HTS dataset, and D_G is the GPS satellite dataset.

4.5.2 Trip Start Time

Trip start time is estimated using posterior distributions over time-of-day intervals based on trip purpose. A binned formulation is transparent and directly supports aggregate planning statistics. For an imputed trip j made by respondent i , the relevant posterior distribution is:

$$\Pr(T_{ij} = t \mid P_{ij}, D_H, D_G) \quad (9)$$

where T_{ij} is start time and P_{ij} is trip purpose.

4.5.3 Mode Choice

Travel mode is estimated after distance and purpose have been assigned. The relevant posterior distribution is:

$$\Pr(M_{ij} = m \mid d_{ij}, T_{ij}, D_H, D_G) \quad (10)$$

where M_{ij} is the selected travel mode, d_{ij} is the trip distance, and T_{ij} is start time. The model accounts for the relationship between travel distance and the feasibility of different modes. Short trips may receive a higher probability of walking or cycling, while longer trips may be more likely to use automobile or transit.

4.6 Survey Correction Procedure

4.6.1 Sequential Identification and Imputation of Missing Trips

After the parametric and non-parametric models have been estimated, the posterior predictive distributions are drawn for each core survey respondent.

For respondent i , the predicted trip count is sampled as:

$$N_i \sim p(N_i | \mathbf{x}_i, D_H, D_G) \quad (11)$$

The number of expected missing trips is then calculated as:

$$R_i = \max(0, N_i - O_i) \quad (12)$$

where R_i is the number of missing trips to be inserted for respondent i and O_i is the observed number of trips in the core. A positive value indicates that the respondent may have omitted one or more trips while a value of zero indicates that no additional trips are inserted. Cases in which the observed number exceeds the posterior prediction are retained for diagnostic analysis but do not lead to the removal of reported trips in the core dataset.

Attributes are generated sequentially for each expected missing trip. First, trip purpose is drawn. Second, trip start time is sampled from the posterior start-time distribution conditional on the sampled purpose. Third, distance is generated from the posterior predictive distance model conditional on purpose and respondent characteristics. Finally, travel mode is drawn from the posterior mode-choice distribution conditional on distance and start time:

$$\begin{aligned} P_{ij}^* &\sim p(P | \mathbf{x}_i, D_H, D_G) \\ T_{ij}^* &\sim p(T | P_{ij}^*, \mathbf{x}_i, D_H, D_G) \\ d_{ij}^* &\sim p(d | P_{ij}^*, \mathbf{x}_i, D_H, D_G) \\ M_{ij}^* &\sim p(M | d_{ij}^*, T_{ij}^*, \mathbf{x}_i, D_H, D_G) \end{aligned} \quad (13)$$

The superscript asterisk denotes an imputed attribute. This methodology recovers the missing trips and associated attributes required for aggregate planning statistics. However, more granular imputations such as determination of origins and destinations as well as specific routes are infeasible with simple regression or distribution-based models. Where applications require complete activity-trip chains, additional location choice and routing models can be introduced after the initial missing trip imputation.

4.6.2 Construction of Corrected Core Datasets

Imputed trips are inserted into the corresponding diaries of each respondent that is missing trips. This may be defined as:

$$\tilde{D}_i = D_i \cup R \quad (14)$$

where D_i is the original diary of each respondent and R is the set of imputed missing trips for the respondent, generated by drawing from the Bayesian posteriors. Then, the corrected copy of the HTS dataset can be defined as:

$$\tilde{H} = \sum_{i=1}^n \tilde{D}_i \quad (15)$$

where the corrected HTS dataset $\tilde{H}$ is the set of all corrected diaries $\tilde{D}_i$ and n is the total number of respondents in the original HTS. Aggregate statistics are calculated using the corrected dataset. The outputs include trip rates by purpose and mode, mode shares, time-of-day distributions, average trip distances, and vehicle-kilometers travelled. Posterior means provide the expected corrected values. Credible intervals communicate the uncertainty associated with model estimation, satellite-sample size, sample selectivity, and missing-trip reconstruction.

The inserted trips are probabilistic estimates rather than directly observed records. Their intended use is to improve aggregate statistics and characterize the range of plausible corrections. This distinction is important when the corrected data is interpreted or transferred into downstream applications.

4.6.3 Estimation of Proxy-Reporting Bias

The correction procedure is applied separately to primary and proxy respondents. Primary respondents report their own travel and may omit trips due to recall limitations. In proxy records, one household member reports another member's travel. These proxy trip records contain recall errors as well as additional omissions caused by incomplete knowledge of another traveler's activities.

Separate HTS-informed priors and GPS-updated posterior distributions are estimated for primary and proxy records. The first model set estimates corrections for directly reported diaries. The second model set estimates corrections for proxy reported diaries. Differences between the resulting posterior corrections provide an estimate of the additional error associated with proxy reporting.

5 Results and Discussion

The final posterior estimates and evaluation of the resulting dataset produced by the Bayesian correction framework are discussed. The results are organized into three parts. The final parametric models are assessed, and the non-parametric posterior distributions are reviewed. Then, the corrected dataset is compared against the original TTS and the two GPS-based satellite cycles. The correction framework was estimated separately for primary and proxy respondents. Separation is essential because the two reporting mechanisms have different error structures. This distinguishes between errors in self-reporting and additional proxy-reporting biases and errors.

5.1 Non-Skeleton Trip Count Model

The non-skeleton trip count model uses a hurdle Poisson specification. It was selected because the outcome is a non-negative integer with a substantial number of zero-count observations. The explanatory variables in the final model include the number of household persons, number of household vehicles, household region, household income, dwelling type, workplace arrangement, student status, and number of skeleton trips. For the categorical variable of dwelling type, detached house was kept as the reference. Below \$15,000 is the reference bracket for income, Toronto is the reference for household region, unknown status is the reference for student status, and not a worker is the reference for workplace arrangement. The prior model from the TTS data was updated sequentially, first with the larger 2025 GPS dataset and then the smaller 2023 GPS dataset.

The inclusion of skeleton trips is based on preliminary comparison between the TTS and satellite data. Skeleton activity structures the day and is expected to significantly affect non-skeleton trip counts. Respondents with more skeleton trips may also have more intermediate stops or linked discretionary activities. For this reason, skeleton trip count was retained as an explanatory variable rather than treated as a correction target.

Figures 2 and 3 compare the TTS-informed prior coefficient distributions with the GPS-updated posterior distributions for primary and proxy respondents. The posterior distributions show how the satellite evidence modifies the prior relationships learned from the TTS sample. In cases where the posterior remains close to the prior, the satellite data is consistent with the self-reported TTS relationship. In cases where the posterior shifts away from the prior, the satellite data indicates a systematic difference in travel. The proxy model is particularly informative because it captures reporting differences that are not present in primary self-reported diaries alone.

The initial intercept shows a negative value while the posterior has a mean near zero. This means in the posterior model the trip count is better explained by the parameters. With respect to dwelling type, both models expect slightly fewer trips from apartments with the posterior reporting a smaller effect. For the category of household income, both prior and posterior found that a higher income correlates to higher non-skeleton trip count. This effect is the same or less prominent in the posterior categories under \$125,000, but stronger over \$125,000. Greater number of household members appears to link with very slightly less individual trips in the prior, but not in the posterior. On the other hand, greater number of vehicles has the same slightly positive effect on all models. The effect of region is similar in both models with slight differences for the regions of Durham and York. Both prior and posterior models correlate greater number of skeleton trips with more non-skeleton trips, with the influence in the posterior being reduced. Student status and workplace arrangement are comparable in all models, with usual workplace outside of home having a slightly smaller effect on trip rate in the posterior. All posterior coefficient distributions are narrower and taller in shape than the priors, indicating that updating with GPS-based data has narrowed the variance of each.

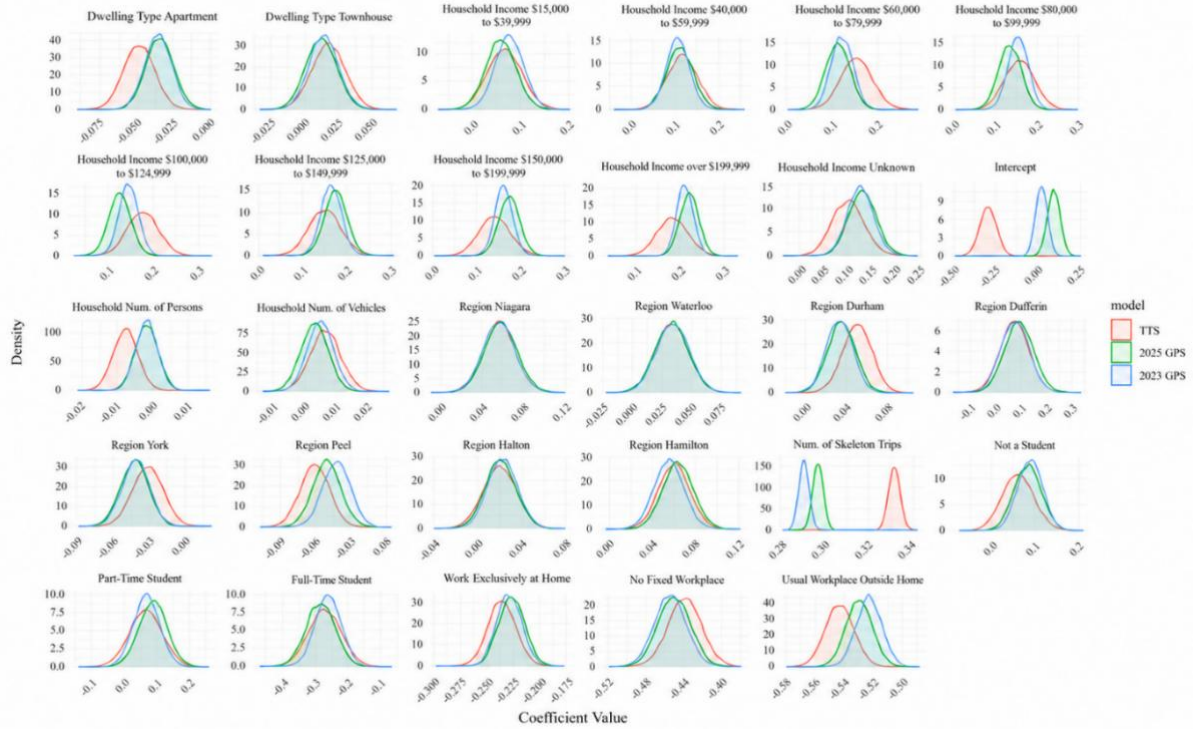


Figure 2. Primary Respondent Prior vs. Posterior Coefficient Distribution

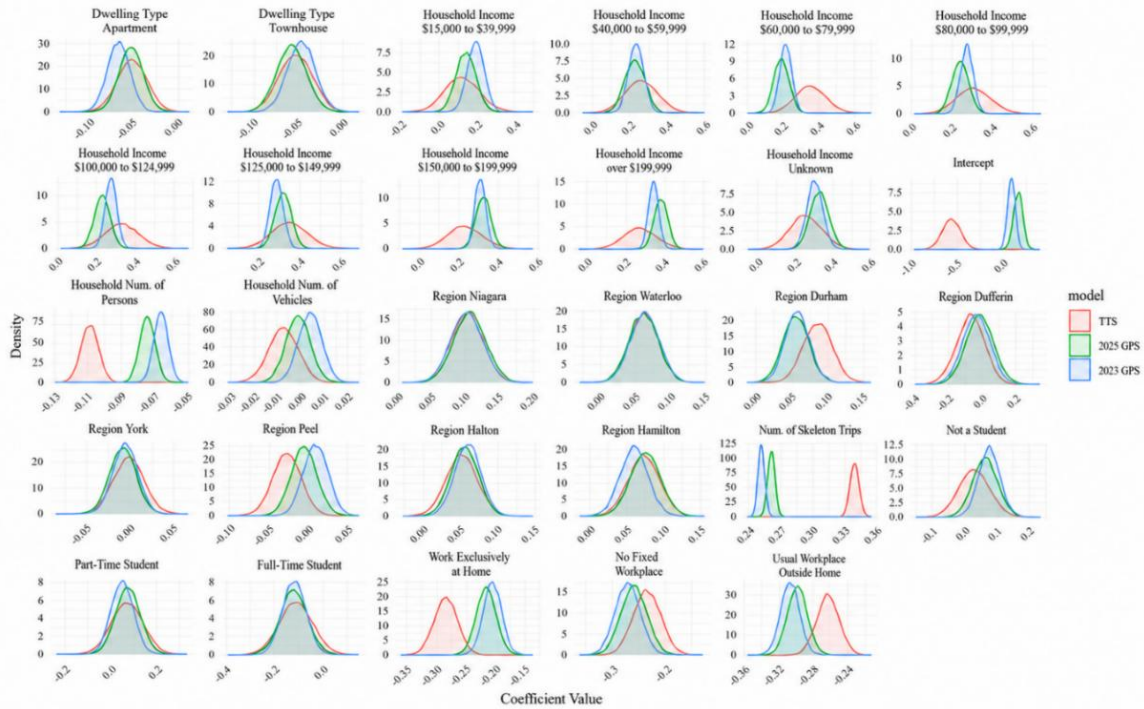
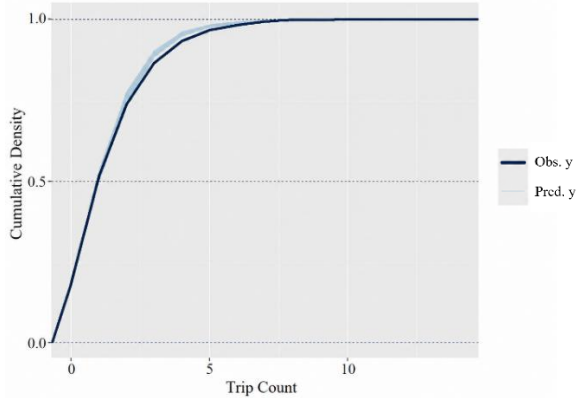


Figure 3. Proxy Respondent Prior vs. Posterior Coefficient Distribution

Model fit was assessed using posterior predictive diagnostics. The empirical cumulative distribution overlays in **Figure 4** show close agreement between observed and simulated counts. As well, the Bayesian R-squared is 0.156 and 0.110 for the primary and proxy models respectively. These results support the use of the final models for missing non-skeleton trip estimation.

Primary Respondent Predicted vs. Observed Trip Count Cumulative Density



Proxy Respondent Predicted vs. Observed Trip Count Cumulative Density

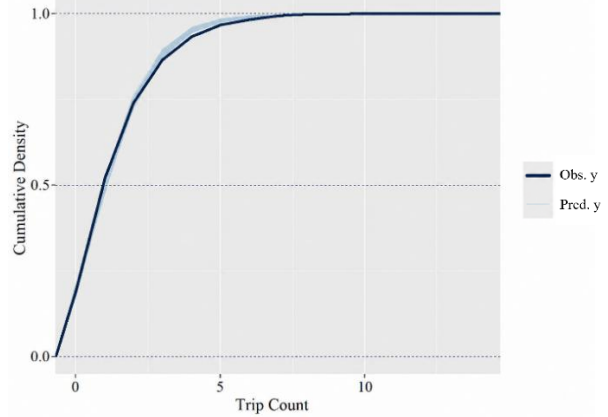


Figure 4. Non-Skeleton Trip Rate Model Posterior Predictive Diagnostics

5.2 Trip Distance by Purpose Model

Trip distance was modelled using a lognormal specification. This form is appropriate because distance is positive and right skewed. The final model estimates distance in meters using trip purpose and household region as explanatory variables. Purpose is included because each has distinct distance profiles. Household region is included to account for differences in urban density and existing transportation network across the GTHA. Similar to the trip count model, the prior model using the TTS was updated sequentially with the 2025 GPS data then the 2023 data. The distance model is applied after the number of missing trips has been determined and a purpose has been drawn for each trip.

Figures 5 and 6 show the prior and posterior coefficient distributions for the distance models estimated for primary and proxy respondents. The GPS-updated posterior distributions retain the structure of the TTS distance relationships while allowing the satellite surveys to influence the estimated distance profiles. The posterior intercept shifts to indicate a slightly greater distance on average, with the shape of the distribution being shorter and wider. This suggests GPS-based data indicates a greater variance in trip distances overall. Compared to the reference trip purpose category of work, other purposes show a lower distance on average. School, shopping, worship, passenger dropoff, and home trips have a less negative effect on distance compared to the prior, while recreation and other purpose trips are greater. With respect to region, most regions are associated with longer trips on average than the reference of Toronto.

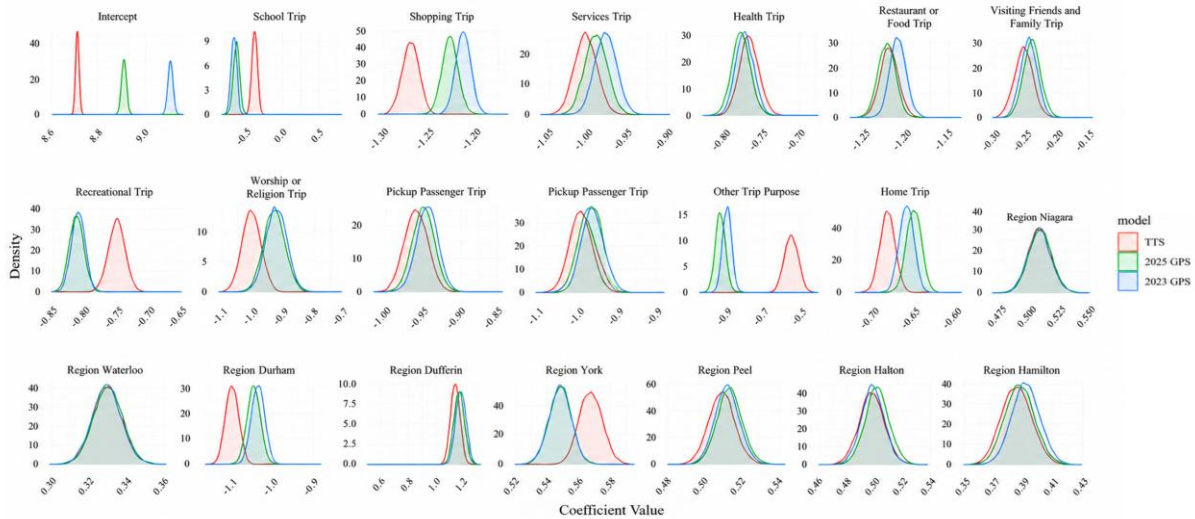


Figure 5. Primary Respondent Trip Distance Prior vs. Posterior Coefficient Distribution

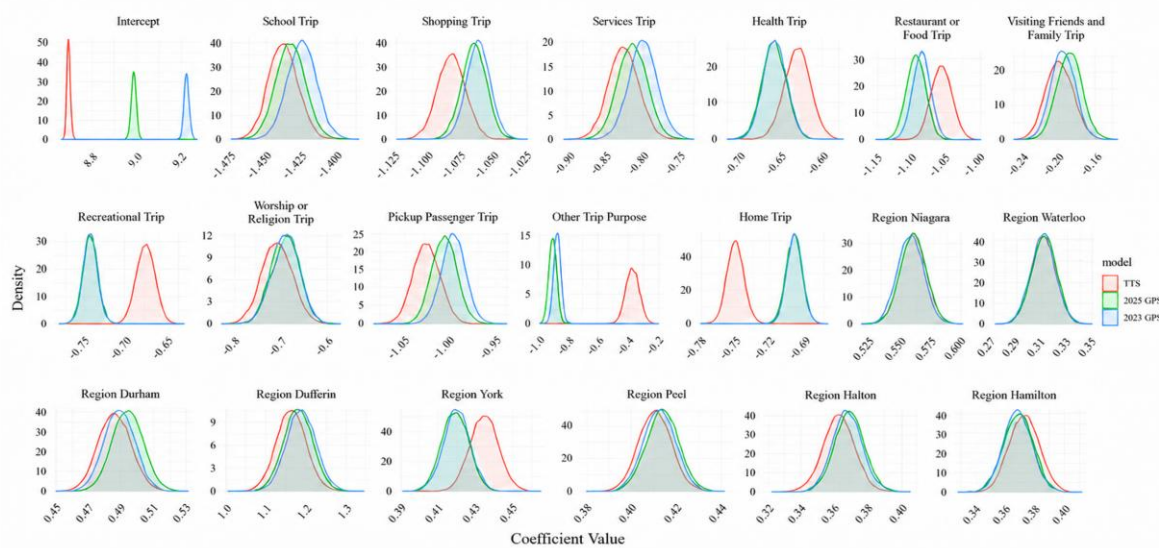


Figure 6. Proxy Respondent Trip Distance Prior vs. Posterior Coefficient Distribution

Posterior predictive checks for the trip distance models are presented in **Figure 7**. The empirical cumulative distribution overlays show close agreement across most of the distance range. The Bayesian R-squared is 0.143 for the primary model and 0.145 for the proxy. This validates the use of the model for generating distances of imputed trips.

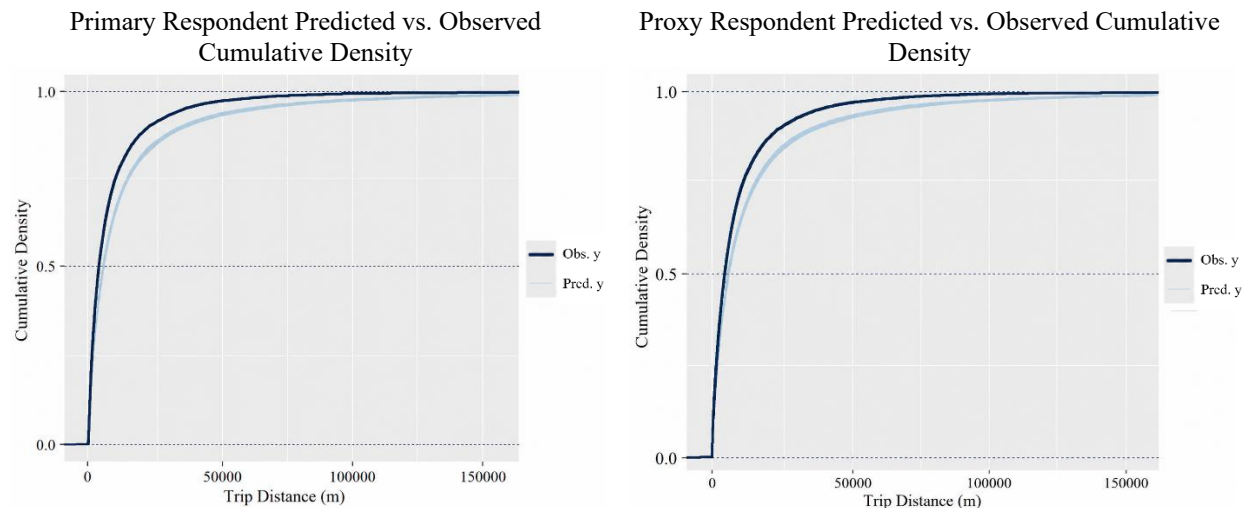


Figure 7. Trip Distance Model Posterior Predictive Diagnostics

5.3 Final Non-Parametric Posteriors

Non-parametric Bayesian updating was used for trip attributes that are difficult to represent with a single parametric form. These include trip purpose, trip start time, and trip mode. In each case, the TTS was used as the prior distribution, and the satellite data was used to update the posterior.

5.3.1 Posterior Trip Purpose Distribution

The first non-parametric update estimates the purpose distribution for missing non-skeleton trips. The purpose posterior is central to the correction process because trip attributes are highly conditional on purpose. A missing recreation trip should not be assigned the same start time, distance, or mode profile as a missing home or shopping trip. The updated purpose distribution ensures that the missing trips inserted are consistent with the composition of the diary. **Table 2** compares the relative frequencies of prior and posterior trip purposes for primary and proxy respondents. The posterior distributions show a slight reduction in the relative share of home purposes and a

corresponding increase in discretionary purposes, including recreation, shopping, and other non-mandatory activities. This pattern is consistent with the expected purposes of missing travel in self-reported surveys being non-essential travel that may be forgotten more easily.

Table 2. Prior vs. Posterior Trip Purpose Relative Frequencies

Trip Purpose	TTS Prior	2023+2025 GPS	Posterior
Primary Respondents			
<i>Work</i>	11.66%	12.05%	11.99%
<i>School</i>	0.42%	1.10%	1.00%
<i>Shopping</i>	15.23%	15.00%	15.03%
<i>Services</i>	3.14%	1.52%	1.76%
<i>Health</i>	4.14%	3.58%	3.67%
<i>Restaurant</i>	4.22%	7.91%	7.35%
<i>Visiting Friends</i>	3.62%	5.32%	5.06%
<i>Recreation</i>	6.03%	10.53%	9.85%
<i>Worship and Religion</i>	0.52%	0.68%	0.65%
<i>Pickup Passenger</i>	3.07%	2.54%	2.62%
<i>Dropoff Passenger</i>	4.28%	2.48%	2.76%
<i>Other</i>	0.35%	3.69%	3.19%
<i>Home</i>	43.31%	33.59%	35.06%
Proxy Respondents			
<i>Work</i>	16.13%	12.05%	12.67%
<i>School</i>	1.58%	1.10%	1.17%
<i>Shopping</i>	11.43%	15.00%	14.46%
<i>Services</i>	1.88%	1.52%	1.57%
<i>Health</i>	3.37%	3.58%	3.55%
<i>Restaurant</i>	3.54%	7.91%	7.25%
<i>Visiting Friends</i>	3.08%	5.32%	4.98%
<i>Recreation</i>	5.21%	10.53%	9.72%
<i>Worship and Religion</i>	0.48%	0.68%	0.65%
<i>Pickup Passenger</i>	2.84%	2.54%	2.59%
<i>Dropoff Passenger</i>	3.60%	2.48%	2.65%
<i>Other</i>	0.36%	3.69%	3.19%
<i>Home</i>	46.50%	33.59%	35.55%

5.3.2 Posterior Start Time Distribution

Start time distributions were also estimated for each unique trip purpose. Separate purpose-specific distributions were used because travel timing differs substantially by activity type. **Figures 8 and 9** compare the TTS prior and GPS-updated posterior start time distributions by purpose for primary and proxy respondents. The posterior distributions generally preserve the main temporal structure observed in the TTS. For most purposes, the morning and afternoon peaks remain visible. However, the satellite data adds modest trips during the midday and evening periods. The effect is most apparent for discretionary purposes and return-home trips. Work and school trips tend to concentrate around conventional peak periods, while shopping, recreation, and other discretionary trips are more dispersed across the day. These shifts are consistent with the GPS datasets, which capture more inter-peak travel and a wider spread of evening activity than the self-reported TTS.

The results indicate that the correction procedure does not impose a fundamentally different daily rhythm. Instead, it adjusts the prior distributions in places where GPS evidence suggests under-reporting. This is important because many analyses rely on time-of-day profiles. A correction that increases trip rates without considering timing could distort peak-period estimates. The updated start time distributions help allocate missing trips in a way that remains consistent with observed behavior.

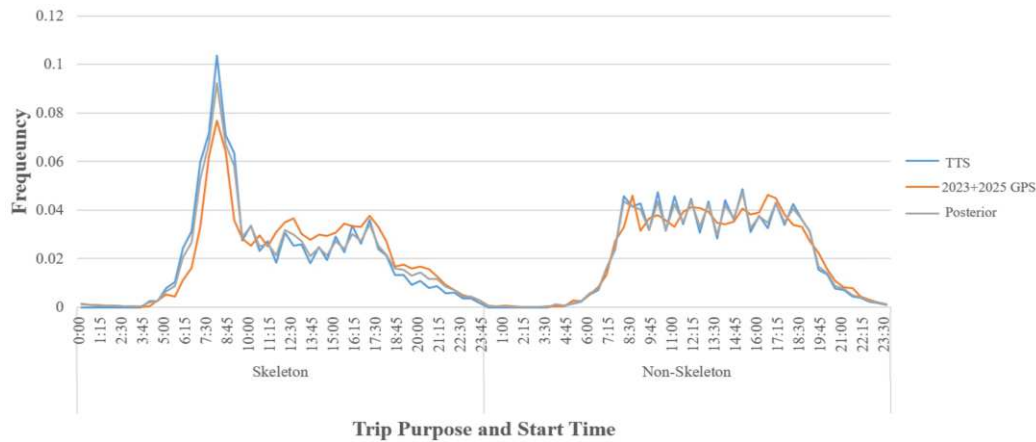


Figure 8. Prior vs. Posterior Primary Respondent Trip Start Time by Purpose

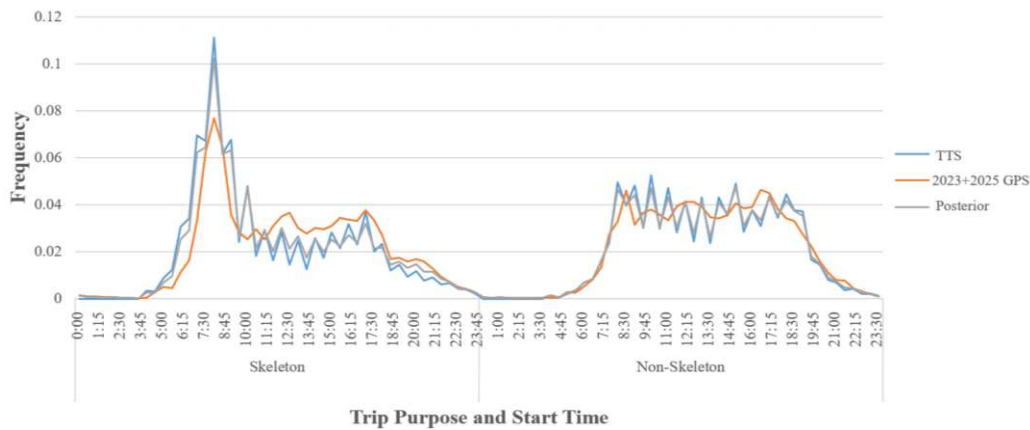


Figure 9. Prior vs. Posterior Proxy Respondent Trip Start Time by Purpose

5.3.3 Posterior Mode Distribution

Finally, mode distributions by distance category are updated. This reflects the strong relationship between distance and mode feasibility. Walking is most common for short trips, whereas cycling is concentrated in medium-distance ranges. Automobiles and transit (both local transit (LT) and regional transit (RT)) become more prominent as distance increases.

Figures 10 and 11 compare the prior and posterior mode distributions by distance category for primary and proxy respondents. The general distance-mode relationship is stable across datasets. Walking has the highest share for the shortest trips and declines sharply after around 2 kilometers. Automobile and transit shares increase with distance. The satellite data introduces a slightly higher walking share among short trips, especially for trips under two kilometers. This reflects the ability of the GPS-based satellite data in capturing shorter non-skeleton trips that are commonly associated with walking, often missed in self-reported surveys.

The distance-based mode posterior ensures that inserted trips are behaviorally plausible. A correction based only on aggregate mode shares could assign too many walking trips to long distances or too many automobile trips to very short distances. Conditioning mode on distance avoids this problem and preserves the observed distance sensitivity of mode choice.

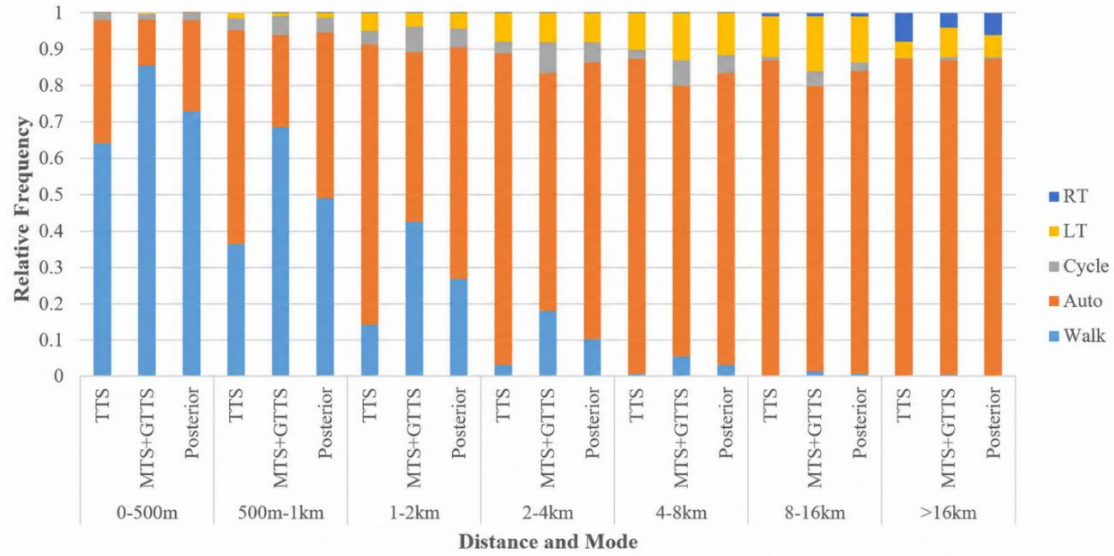


Figure 10. Prior vs. Posterior Primary Respondent Trip Mode by Distance Distributions

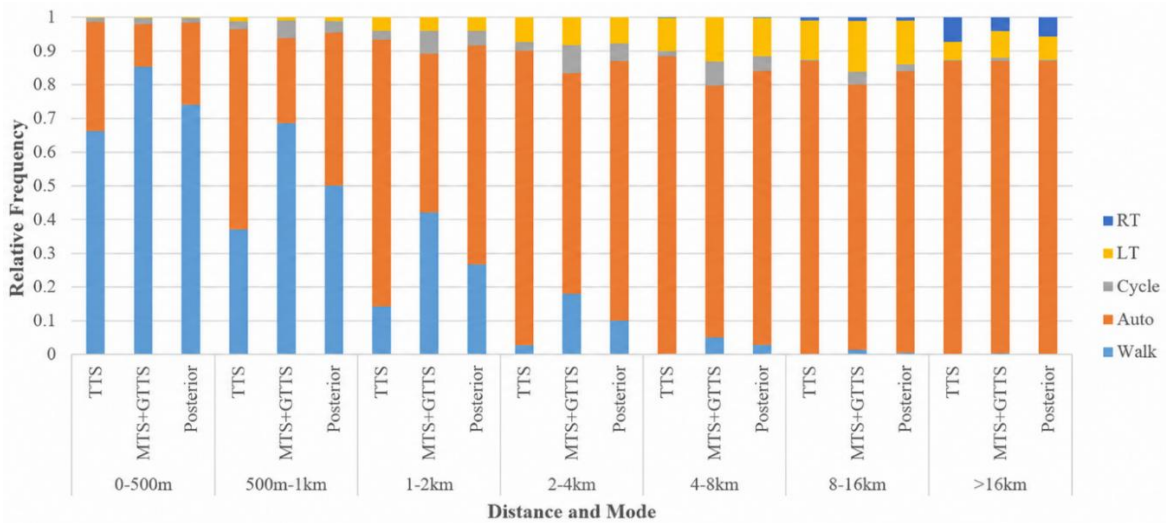


Figure 11. Prior vs. Posterior Proxy Respondent Trip Mode by Distance Distributions

5.4 Comparison of Corrected Trip Rates

Table 3 compares average skeleton, non-skeleton, and combined trip rates per person across the original TTS primary and proxy respondents, the 2023 GPS cycle, the 2025 GPS cycle, and the corrected TTS. The analysis only considers mobile respondents for all datasets, that is, respondents who reported at least one trip. The comparison separates skeleton and non-skeleton travel because the correction framework is designed to target non-skeleton trips, the portion of the diary most likely to be incomplete.

The corrected TTS shows a substantial increase in non-skeleton trip rates relative to the uncorrected TTS. The resulting trip rates move closer to those observed in the smartphone-based satellite surveys, while still retaining the trip characteristics of the TTS. This indicates that the correction procedure preserves the underlying structure of the data rather than uniformly inflating all trip categories. This improvement is especially clear for proxy respondents, whose original trip rates are lower than both primary TTS respondents and GPS survey respondents. The result is consistent with the expectation that proxy diaries omit more intermediate and discretionary travel.

Table 3. Comparison of Mean and Standard Deviation of Trip Rates Across All Datasets

Dataset	Mean	Standard Dev.
TTS Primary (Core)		
<i>Skeleton</i>	1.73	0.940
<i>Non-Skeleton</i>	1.32	1.307
<i>Combined</i>	3.05	1.658
TTS Proxy (Core)		
<i>Skeleton</i>	1.87	0.804
<i>Non-Skeleton</i>	0.77	1.001
<i>Combined</i>	2.64	1.261
2023 GPS Cycle (Satellite)		
<i>Skeleton</i>	2.09	1.183
<i>Non-Skeleton</i>	1.81	1.610
<i>Combined</i>	3.90	2.125
2025 GPS Cycle (Satellite)		
<i>Skeleton</i>	2.10	1.34
<i>Non-Skeleton</i>	1.90	1.69
<i>Combined</i>	4.00	2.187
Corrected TTS Primary		
<i>Skeleton</i>	1.73	0.940
<i>Non-Skeleton</i>	2.17	1.429
<i>Combined</i>	3.90	1.898
Corrected TTS Proxy		
<i>Skeleton</i>	1.87	0.804
<i>Non-Skeleton</i>	1.76	1.172
<i>Combined</i>	3.63	1.708

Figure 12 presents the predicted-minus-observed non-skeleton trip counts for primary and proxy respondents. Positive values indicate respondents for whom the posterior predicts more non-skeleton trips than were recorded in the original TTS diary. These positive values form the basis for missing-trip insertion. The distributions show that the correction is not applied evenly across the sample. Many respondents receive no additional trips, while a smaller set receives one or more imputed non-skeleton trips. Rather than a uniform adjustment factor, it is recognized that not all diaries are equally incomplete.

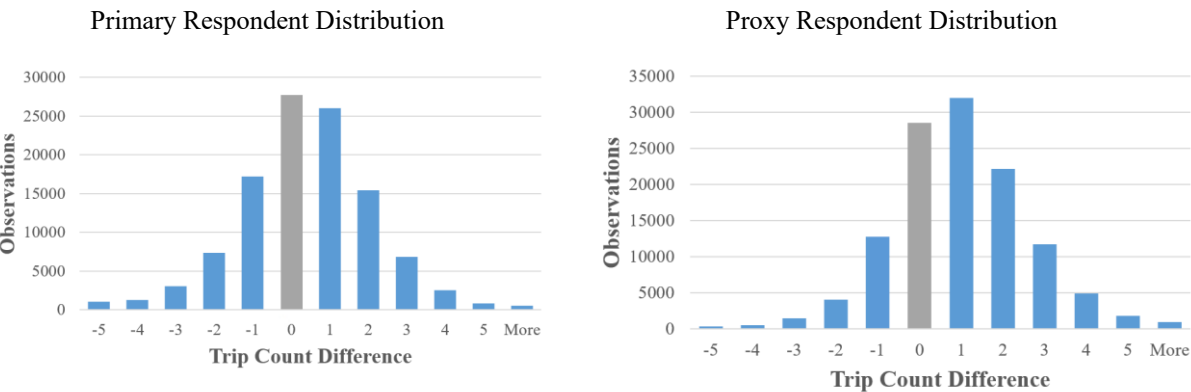


Figure 12. Predicted Minus Observed Non-Skeleton Trip Count Distributions

The correction is larger for proxy respondents than for primary respondents. This is consistent with the presence of proxy-reporting bias in such trips. The corrected dataset produces a larger trip-rate adjustment for proxy records, where the expected reporting error is greater.

Figure 13 shows the distribution of predicted missing non-skeleton trips inserted into the corrected dataset. For primary respondents, most corrections are modest and concentrated at low counts. For proxy respondents, the distribution contains a larger share of cases with multiple missing trips. This difference supports the decision to estimate separate correction models for primary and proxy diaries.

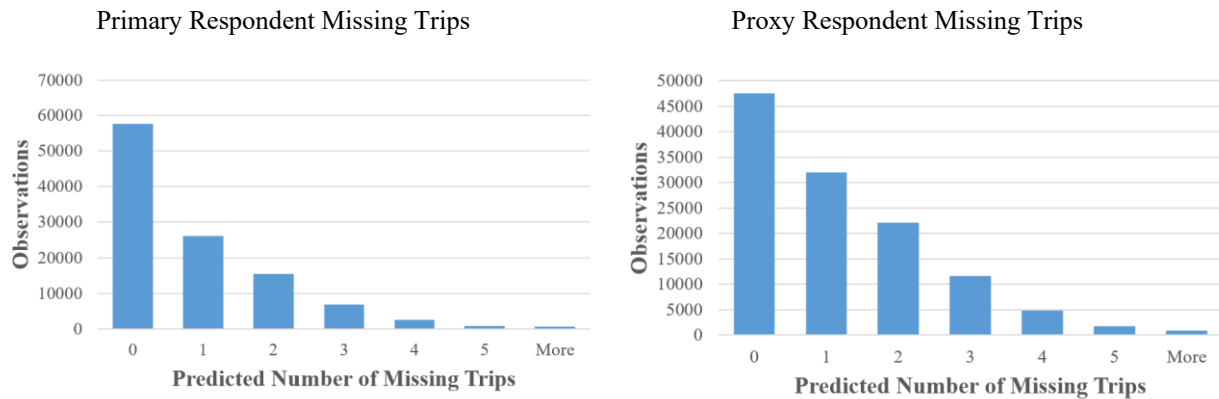


Figure 13. Predicted Number of Missing Non-Skeleton Trips

5.5 Comparison of Corrected Trip Attributes

Attributes for each missing trip were drawn from the posterior distributions. Thus, the aggregate statistics of the corrected dataset closely follow the estimated posteriors. The corrected dataset was compared against the TTS, 2023 GPS cycle, and 2025 GPS cycle using trip purpose, distance, start time, and mode distributions. These comparisons assess whether the correction procedure generates plausible trip attributes after increasing trip rates. This step is necessary because a corrected trip count alone is not sufficient for planning applications.

5.5.1 Trip Purpose Distributions

Figures 14 and 15 compare trip purpose distributions across the TTS, 2023 and 2025 GPS cycles, and corrected TTS for primary and proxy respondents. The corrected TTS shifts in the expected direction. The share of home and work trips declines slightly, while discretionary purposes increase. The distribution remains closer to the original TTS than to either satellite dataset. The proxy-corrected distribution shows a stronger adjustment than the primary-corrected distribution. This again reflects the larger missing-trip problem in proxy diaries. The correction adds discretionary trips that are unlikely to be fully reported by another household member.

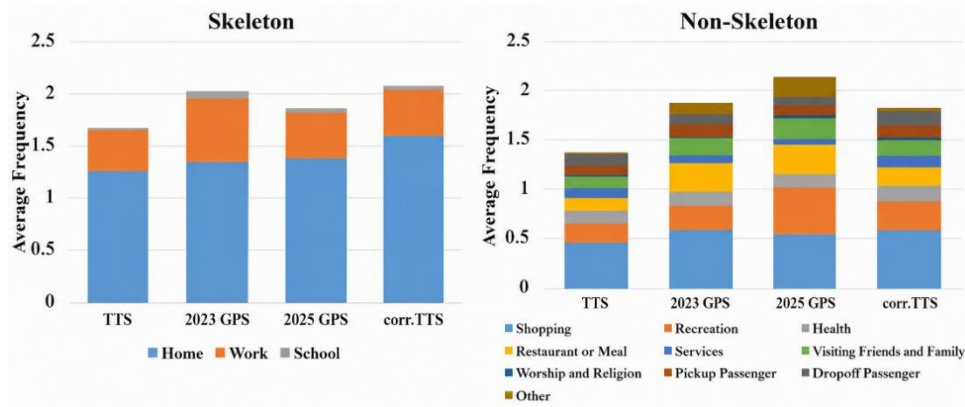


Figure 14. Comparison of Primary Trip Purpose Between Corrected and Observed Datasets

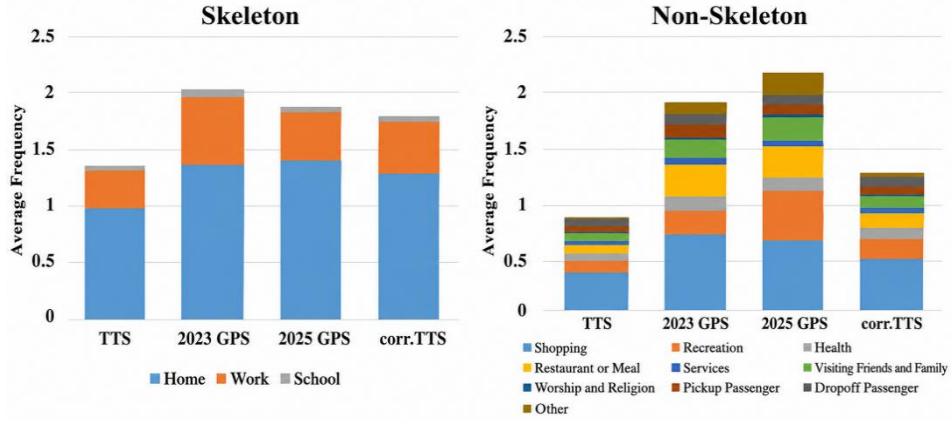


Figure 15. Comparison of Proxy Trip Purpose Between Corrected and Observed Datasets

These results are significant because the correction improves not only the total trip rate but also the purpose composition. Trip omission in HTS is reported as purpose selective. Short discretionary and non-home-based trips are reported less completely than regular work and school trips, and proxy reporting further suppresses these discretionary activities (Hassounah et al. 1993; Badoe and Steuart 2002). Kumar and Levinson (1993) similarly found that household surveys substantially underreported short and irregular shopping trips and warned that missing or misclassified intermediate stops can misspecify trip-generation models. Identifying and reinserting these trips reduces the artificial dominance of home and work trips and restores more credible shares of discretionary activities. Because the distribution of corrected purposes remains anchored to the original survey, the adjustment improves behavioral accuracy without discarding the broader population structure of core HTS.

5.5.2 Distance Distributions

Figures 16 and 17 compare trip distance distributions by purpose for the four datasets. Across most purposes, the corrected TTS remains close to the base TTS distribution. This indicates that the imputation process does not substantially alter the distance structure of the core survey. At the same time, the final dataset shows a modest influence from the satellite data, particularly in the short-distance ranges where GPS-based surveys record more trips.

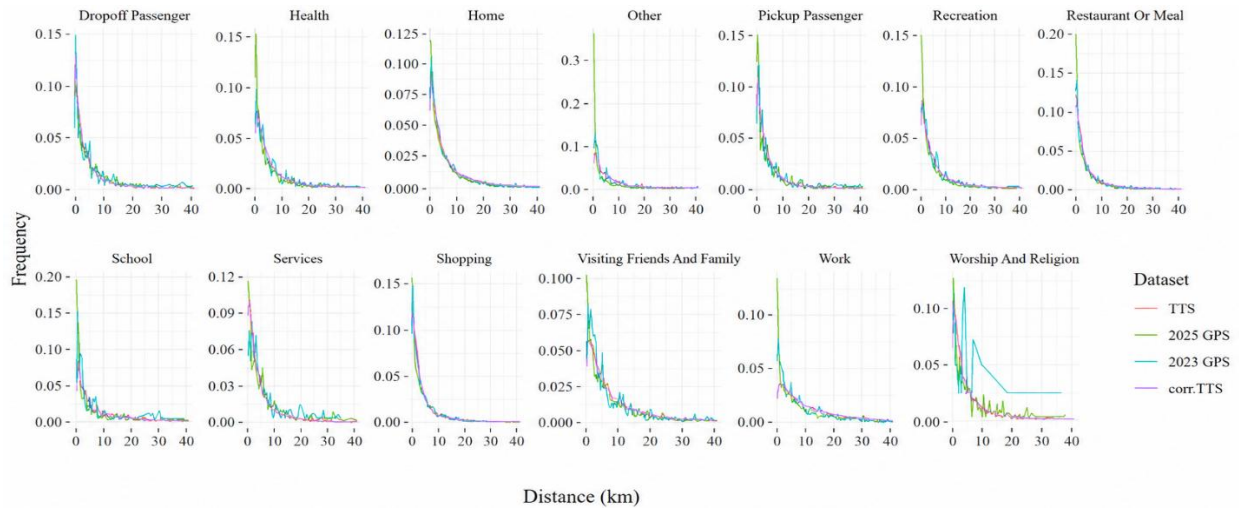


Figure 16. Comparison of Primary Trip Distance Between Corrected and Observed Datasets

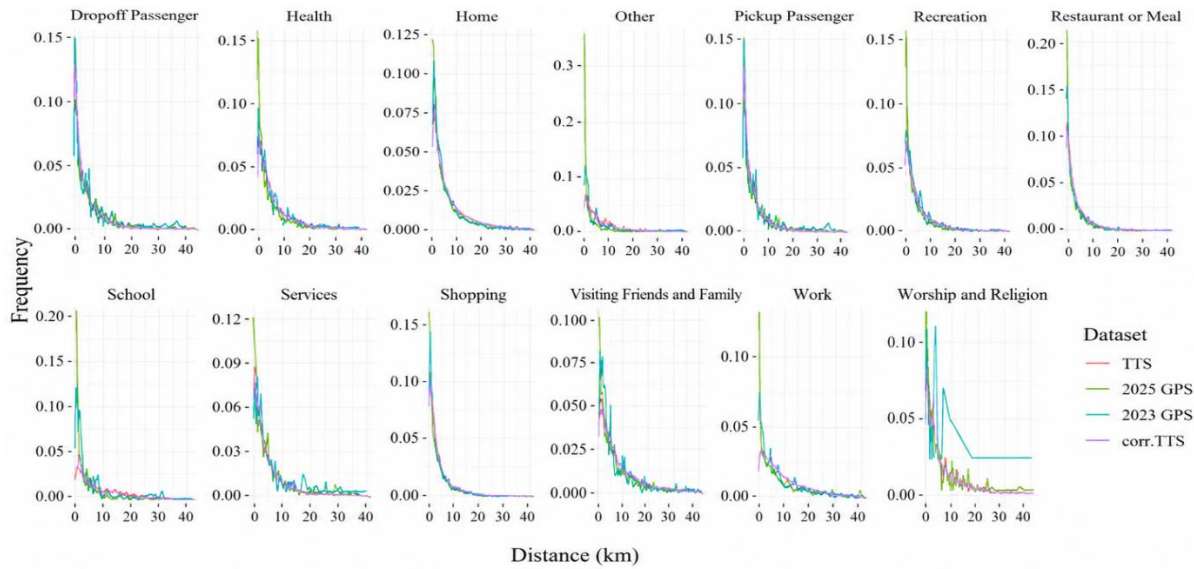


Figure 17. Comparison of Proxy Trip Distance Between Corrected and Observed Datasets

5.5.3 Start Time Distributions

Figures 18 and 19 compare trip start time distributions by purpose across the datasets. For most purposes, the start time distributions are similar across datasets. This reflects the strong temporal regularity of many activity types. However, the satellite data introduces a modest increase in midday travel and a wider spread of evening homebound trips. These changes are consistent with the GPS-based datasets, which capture more inter-peak activity than the self-reported TTS. The correction therefore improves temporal coverage without removing the core survey's main trip start time structure.

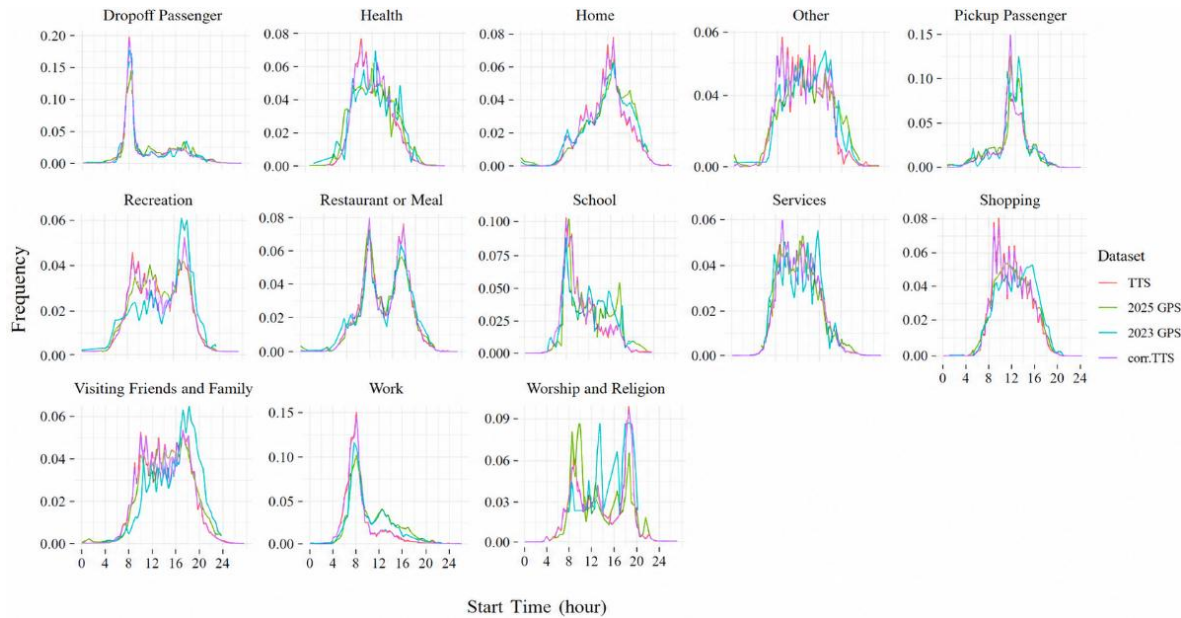


Figure 18. Comparison of Primary Trip Start Time Between Corrected and Observed Datasets

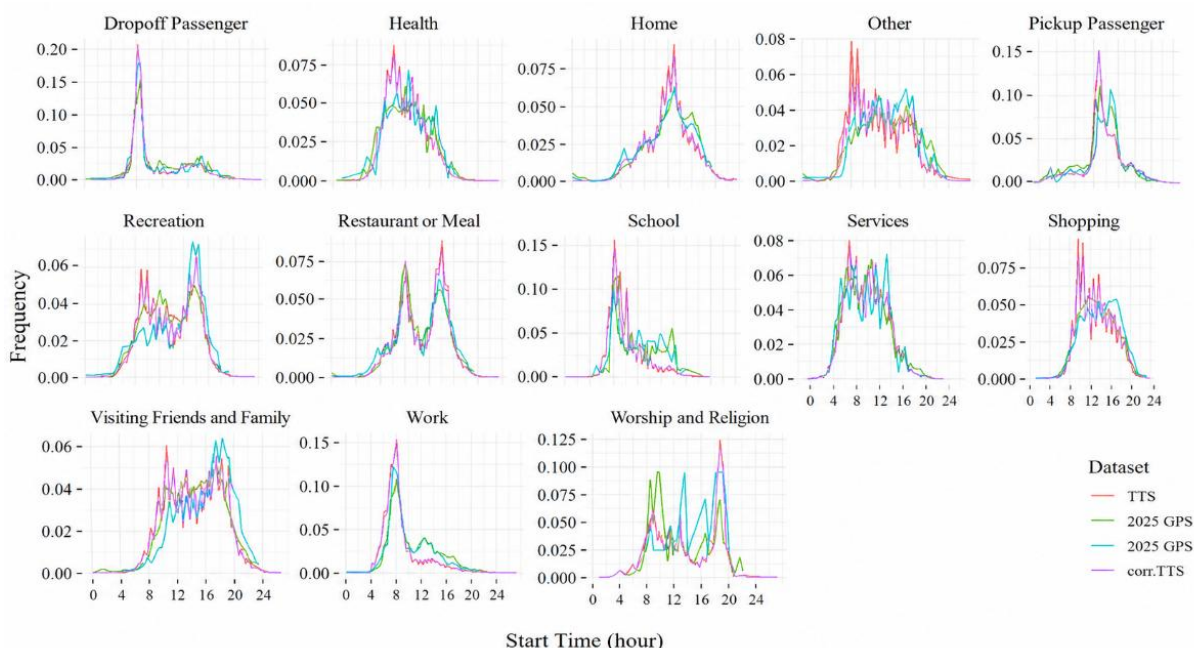


Figure 19. Comparison of Proxy Trip Start Time Between Corrected and Observed Datasets

5.5.4 Mode Share Distributions

Figures 20 and 21 compare mode distributions by distance category. The corrected TTS follows the base TTS closely but begins to reflect the satellite datasets in the shortest distance categories. In particular, the walking share increases slightly for trips under 500 meters and for trips between 500 meters and 1 kilometer. This is consistent with the expected under-reporting of short walking trips in self-reported surveys. The mode correction is concentrated where the evidence of missed travel is strongest.

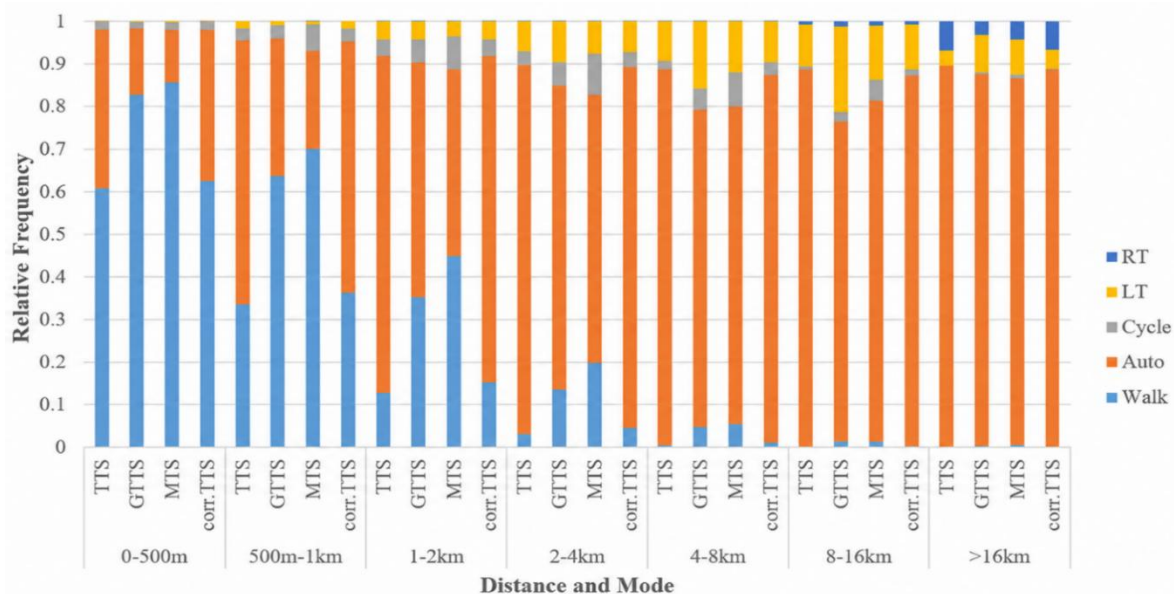


Figure 20. Comparison of Primary Trip Mode Between Corrected and Observed Datasets

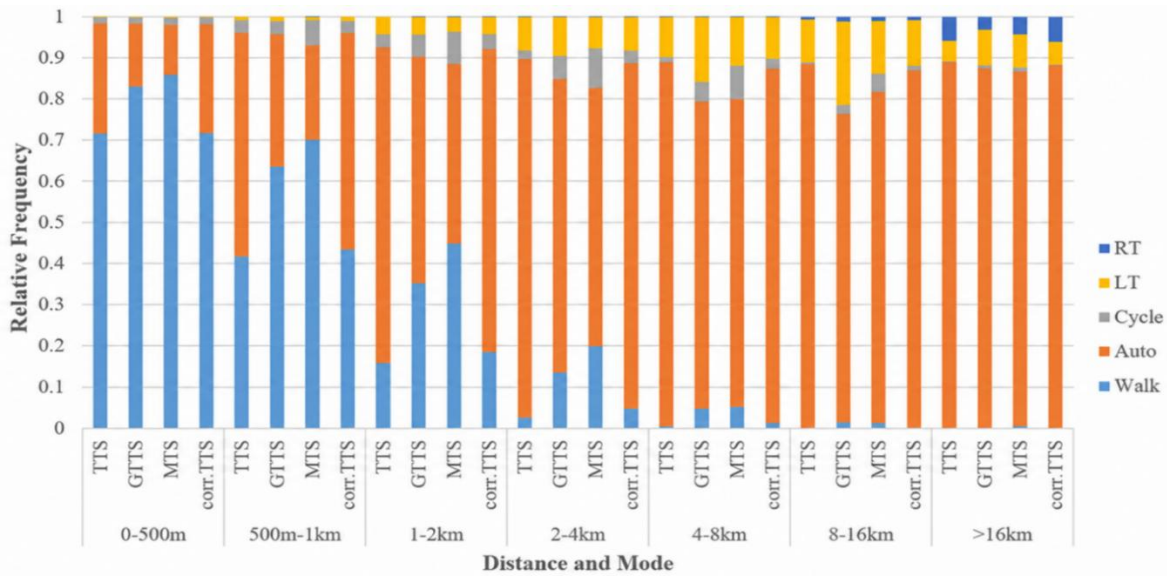


Figure 21. Comparison of Proxy Trip Mode Between Corrected and Observed Datasets

These results correct a systematic weakness in household travel surveys. Short walking trips are more likely to be forgotten or excluded, causing HTS to understate walking and overstate the relative shares of automobile and other motorized modes (Buehler and Pucher 2024). GPS-diary comparisons further show that trip underreporting varies with travel intensity and trip-chain complexity, meaning that uncorrected mode shares can misrepresent actual behavior rather than simply reduce total trip counts (Bricka and Bhat 2006). Restoring these trips produces more credible walk-mode shares and short-distance distributions, improving the calibration of trip-generation, mode-choice, and assignment models and reducing errors that can propagate into estimates of mode share and travel time (Wolf et al. 2003).

6 Conclusion and Further Research

This study develops a Bayesian core-satellite correction framework to improve descriptive statistics derived from the 2022 Transportation Tomorrow Survey by systematically identifying and recreating possible missing trips. The approach responds to a common limitation of conventional household travel surveys: self-reported one-day diaries can understate travel activity, particularly for short, discretionary, and proxy-reported trips. The 2023 and 2025 GPS-based satellite surveys provided supplementary evidence for identifying these reporting gaps while retaining the population coverage, socioeconomic detail, and policy relevance of the TTS.

The framework treats the TTS and GPS surveys as complementary sources of evidence. The TTS provides the base structure of regional travel behavior while the satellite datasets provide more complete observations of individual trips. Bayesian updating was used to combine these sources without assuming that the GPS data is error-free. Parametric models were applied to non-skeleton trip counts and trip distances. Non-parametric Bayesian updating was used for trip purpose, start time, and mode distributions. Primary and proxy respondents were modelled separately to reflect differences in reporting quality.

A central finding is that skeleton trips were comparatively stable across the self-reported and GPS-based datasets. Work, school, and home trips appear to be reported with less error because they are structural to any weekday and thus easier to recall (Hassounah et al. 1993; Bricka and Bhat 2006; Verreault and Morency 2015; Aschauer et al. 2018). Therefore, the correction focuses on non-skeleton travel including shopping, errands, recreation, and other discretionary purposes. This distinction prevents missing discretionary travel from being incorrectly represented as additional work or school trips.

The corrected dataset substantially improves trip rates while preserving the original structure of the TTS. Among mobile primary respondents, the combined trip rate increased from 3.05 to 3.90 trips per person per day, matching the 2023 and 2025 GPS satellite rates of 3.90 and 4.00. The increase was driven by non-skeleton trips, which rose from 1.32 to 2.17. The correction was larger for proxy respondents. The combined proxy trip rate increased from

2.64 to 3.63, with non-skeleton trips rising from 0.77 to 1.76. This confirms that proxy-reported diaries are more likely to omit non-mandatory and intermediate movements.

The inserted trips also produced plausible trip-level distributions. The corrected purpose distribution shifts toward discretionary travel but remains anchored to the TTS. Distance distributions continue to follow the core survey closely, with modest adjustments in short-distance ranges. Start time distributions retain the main weekday pattern of the TTS but show more midday travel and a wider evening spread for return-home trips. Mode shares by distance remain behaviorally consistent, with walking concentrated among short trips and motorized modes increasing with distance. The correction is consequently more than a naïve trip-rate adjustment. It inserts missing trips on the respondent level with realistic purpose, timing, distance, and mode attributes.

Several limitations should be noted. The inserted trips are posterior predictions, not directly observed records. They should be interpreted as plausible corrections for aggregate analysis rather than as exact reconstructions of daily schedules. The present framework also does not assign origins, destinations, or detailed routes to the missing trips. These attributes would require additional location choice and routing models. In its current form, the corrected dataset is best suited for descriptive statistics and distributional analysis rather than full activity-chain reconstruction.

Further research should extend the framework to spatial imputation so that missing trips can be assigned plausible origins, destinations, and routes. This would allow corrected HTS data to support network assignment, accessibility analysis, and activity-based model estimation. Future work should also incorporate schedule-consistency constraints so inserted trips form coherent daily activity chains.

Hierarchical model structures may also be explored. Partial pooling across regions, age groups, employment types, and household categories would improve estimates for sparsely observed groups while preserving variation in reporting errors. The framework can also be tested with additional validation sources, including simultaneous HTS-GPS data, smart-card records, traffic counts, and other passive mobility datasets. These sources could provide external checks on corrected mode shares, time-of-day distributions, and distances.

Finally, future applications should examine how corrected HTS data affects planning outcomes. Missing trips may change estimates of active mode shares, discretionary travel, off-peak demand, and vehicle-kilometers travelled. These changes can influence mode-shift targets, emissions estimates, transit planning, and travel demand model calibration. The proposed framework provides a practical way to strengthen the empirical base used in transportation policy and planning, but care should be exercised in verifying the efficacy of the results.

References

- Ades, Anthony E., and Alex J. Sutton. 2006. "Multiparameter evidence synthesis in epidemiology and medical decision-making: Current approaches." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 169 (1): 5–35. <https://doi.org/10.1111/j.1467-985X.2005.00377.x>.
- Andridge, Rebecca R., and Roderick J. A. Little. 2010. "A review of hot deck imputation for survey non-response." *International Statistical Review*. 78 (1): 40–64. <https://doi.org/10.1111/j.1751-5823.2010.00103.x>.
- Aschauer, Florian, Reinhard Hössinger, Kay W. Axhausen, Basil Schmid, and Regine Gerike. 2018. "Implications of survey methods on travel and non-travel activities: A comparison of the Austrian national travel survey and an innovative mobility-activity-expenditure diary (MAED)." *European Journal of Transport and Infrastructure Research*. 18 (1): 4–35. <https://doi.org/10.18757/ejtir.2018.18.1.3217>.
- Bachu, Prashanth K., Trisha Dudala, and Sirisha M. Kothuri. 2001. "Prompted recall in global positioning system survey: Proof-of-concept study." *Transportation Research Record*. 1768 (1): 106–13. <https://doi.org/10.3141/1768-13>.
- Badoe, Daniel A., and Gerald N. Steuart. 2002. "Impact of interviewing by proxy in travel survey conducted by telephone." *Journal of Advanced Transportation*. 36 (1): 43–62. <https://doi.org/10.1002/atr.5670360104>.
- Bayart, Caroline, Patrick Bonnel, and Catherine Morency. 2009. "Survey mode integration and data fusion: Methods and challenges." *Transport Survey Methods*. 587–611. Edited by Patrick Bonnel, Martin Lee-Gosselin, Johanna Zmud, and Jean-Loup Madre. Bingley, UK: Emerald Group Publishing. <https://doi.org/10.1108/9781848558458-034>.
- Bricka, Stacey G., and Chandra R. Bhat. 2006. "Comparative analysis of global positioning system-based and travel survey-based data." *Transportation Research Record*. 1972 (1): 9–20. <https://doi.org/10.1177/0361198106197200102>.
- Bricka, Stacey G., Sudeshna Sen, Rajesh Paleti, and Chandra R. Bhat. 2012. "An analysis of the factors influencing differences in survey-reported and GPS-recorded trips." *Transportation Research Part C: Emerging Technologies*. 21 (1): 67–88. <https://doi.org/10.1016/j.trc.2011.09.005>.
- Buehler, Ralph, and John Pucher. 2024. "The challenge of measuring walk trips in travel surveys: Problems of undercounting and incomparability among countries and over time." *Transport Reviews*. 44 (5): 937–43. <https://doi.org/10.1080/01441647.2024.2319415>.
- Chapleau, Robert, Philippe Gaudette, and Tim Spurr. 2018. "Strict and deep comparison of revealed transit trip structure between computer-assisted telephone interview household travel survey and smart cards." *Transportation Research Record*. 2672 (42): 13–22. <https://doi.org/10.1177/0361198118758297>.
- City of Edmonton. 2009. "The Way We Move: Transportation Master Plan." https://www.edmonton.ca/sites/default/files/public-files/documents/land_sales/TransportationMasterPlan.pdf.
- City of Minneapolis. 2020. "Minneapolis Transportation Action Plan: Minneapolis Streets 2030." <https://go.minneapolismn.gov/minneapolis-streets-2030>.
- City of Mississauga. 2021. "Transportation Master Plan." <https://www.mississauga.ca/publication/transportation-master-plan/>.
- City of Ottawa. 2013. "2013 Transportation Master Plan." <https://documents.ottawa.ca/en/node/5836>.
- City of Richmond Hill. 2023. "Transportation Master Plan Update." <https://www.richmondhill.ca/en/learn-more/transportation-master-plan.aspx>.
- City of Ryde. 2025. "Integrated Transport Strategy 2045." <https://www.ryde.nsw.gov.au/files/assets/public/v/1/publications/operations/transport/final-integrated-transport-strategy-its-2045-november-2025.pdf>.

- 1 City of Toronto. 2019. "Background reports: ConsumersNext." [https://www.toronto.ca/city-government/planning-](https://www.toronto.ca/city-government/planning-development/planning-studies-initiatives/consumersnext/background-reports-consumersnext/)
- 2 [development/planning-studies-initiatives/consumersnext/background-reports-consumersnext/](https://www.toronto.ca/city-government/planning-development/planning-studies-initiatives/consumersnext/background-reports-consumersnext/).
- 3 City of Toronto. 2022. "Yonge Street North planning study." [https://www.toronto.ca/city-government/planning-](https://www.toronto.ca/city-government/planning-development/planning-studies-initiatives/yonge-street-north-planning-study/)
- 4 [development/planning-studies-initiatives/yonge-street-north-planning-study/](https://www.toronto.ca/city-government/planning-development/planning-studies-initiatives/yonge-street-north-planning-study/).
- 5 City of Vancouver. 2012. "Transportation 2040 Plan." [https://vancouver.ca/streets-transportation/transportation-](https://vancouver.ca/streets-transportation/transportation-2040.aspx)
- 6 [2040.aspx](https://vancouver.ca/streets-transportation/transportation-2040.aspx).
- 7 Clarke, Mike, Martin Dix, and Peter Jones. 1981. "Error and uncertainty in travel surveys." *Transportation*. 10 (2):
- 8 105–26. <https://doi.org/10.1007/BF00165261>.
- 9 Cottrill, Caitlin D., Francisco Camara Pereira, Fang Zhao, Ines Ferreira Dias, Hock Beng Lim, Moshe E. Ben-Akiva,
- 10 and Christopher Zegras. 2013. "Future Mobility Survey: Experience in developing a smartphone-based
- 11 travel survey in Singapore." *Transportation Research Record*. 2354 (1): 59–67.
- 12 <https://doi.org/10.3141/2354-07>.
- 13 Data Management Group. 2024. "2022 TTS: Data Guide." [https://dmg.utoronto.ca/wp-](https://dmg.utoronto.ca/wp-content/uploads/2024/12/2022TTS_Data_Guide-1.pdf)
- 14 [content/uploads/2024/12/2022TTS_Data_Guide-1.pdf](https://dmg.utoronto.ca/wp-content/uploads/2024/12/2022TTS_Data_Guide-1.pdf).
- 15 de Montjoye, Yves-Alexandre, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. 2013. "Unique in the
- 16 crowd: The privacy bounds of human mobility." *Scientific Reports*. 3: 1376.
- 17 <https://doi.org/10.1038/srep01376>.
- 18 Deming, William Edwards, and Frederick F. Stephan. 1940. "On a least squares adjustment of a sampled frequency
- 19 table when the expected marginal totals are known." *The Annals of Mathematical Statistics*. 11 (4): 427–44.
- 20 <https://doi.org/10.1214/aoms/1177731829>.
- 21 Deville, Jean-Claude, and Carl-Erik Särndal. 1992. "Calibration estimators in survey sampling." *Journal of the*
- 22 *American Statistical Association*. 87 (418): 376–82. <https://doi.org/10.1080/01621459.1992.10475217>.
- 23 Egu, Oscar, and Patrick Bonnel. 2020. "How comparable are origin-destination matrices estimated from automatic
- 24 fare collection, origin-destination surveys and household travel survey? An empirical investigation in
- 25 Lyon." *Transportation Research Part A: Policy and Practice*. 138: 267–82.
- 26 <https://doi.org/10.1016/j.tra.2020.05.021>.
- 27 Ferguson, Thomas S. 1973. "A Bayesian analysis of some nonparametric problems." *The Annals of Statistics*. 1 (2):
- 28 209–30. <https://doi.org/10.1214/aos/1176342360>.
- 29 Greaves, Stephen P., Simon Fifer, Richard B. Ellison, and George Germanos. 2010. "Development of a global
- 30 positioning system web-based prompted recall solution for longitudinal travel surveys." *Transportation*
- 31 *Research Record*. 2183 (1): 69–77. <https://doi.org/10.3141/2183-08>.
- 32 Hassounah, Mazen I., Loy-Sai Cheah, and Gerald N. Steuart. 1993. "Underreporting of trips in telephone interview
- 33 travel surveys." *Transportation Research Record*. 1412: 90–94.
- 34 Holt, David, and Terence Michael Frederick Smith. 1979. "Post stratification." *Journal of the Royal Statistical*
- 35 *Society: Series A (General)*. 142 (1): 33–46. <https://doi.org/10.2307/2344652>.
- 36 Keusch, Florian, Bella Struminskaya, Christopher Antoun, Mick P. Couper, and Frauke Kreuter. 2019. "Willingness
- 37 to participate in passive mobile data collection." *Public Opinion Quarterly*. 83 (S1): 210–35.
- 38 <https://doi.org/10.1093/poq/nfz007>.
- 39 Kim, Youngsung, Francisco C. Pereira, Fang Zhao, Ajinkya Ghorpade, Christopher Zegras, and Moshe Ben-Akiva.
- 40 2015. "Activity recognition for a smartphone and web-based travel survey." arXiv. arXiv:1502.03634.
- 41 <https://doi.org/10.48550/arXiv.1502.03634>.
- 42 Kott, Phillip S., and Ted Chang. 2010. "Using calibration weighting to adjust for nonignorable unit nonresponse."
- 43 *Journal of the American Statistical Association*. 105 (491): 1265–75.
- 44 <https://doi.org/10.1198/jasa.2010.tm09016>.

- 1 Kumar, Ajay, and David Levinson. 1993. "Specifying, estimating, and validating a new trip generation model: Case
2 study in Montgomery County, Maryland." *Transportation Research Record* 1413: 107–13.
- 3 Li, Melvyn, Kaili Wang, Yicong Liu, and Khandker Nurul Habib. 2024. "Deriving weeklong activity-travel diary
4 from Google Location History: Survey tool development and a field test in Toronto." *Transportation*.
5 <https://doi.org/10.1007/s11116-024-10523-3>.
- 6 Li, Melvyn, Sumaiya Afrose Suma, Felita Ong, Kaili Wang, and Khandker Nurul Habib. 2025. "Beyond the typical
7 day: Evaluating multi-day travel behaviour through a core–satellite survey approach in the Greater Toronto
8 and Hamilton Area." Research Square. <https://doi.org/10.21203/rs.3.rs-7273201/v1>.
- 9 Little, Roderick J. A., Brady T. West, Philip S. Boonstra, and Jingwei Hu. 2020. "Measures of the degree of
10 departure from ignorable sample selection." *Journal of Survey Statistics and Methodology*. 8 (5): 932–64.
11 <https://doi.org/10.1093/jssam/smz023>.
- 12 Lugtig, Peter, Katie Roth, and Barry Schouten. 2022. "Nonresponse analysis in a longitudinal smartphone-based
13 travel study." *Survey Research Methods*. 16 (1): 13–27. <https://doi.org/10.18148/srm/2022.v16i1.7835>.
- 14 Maruyama, Takuya, Kenta Hosotani, and Tomoki Kawano. 2019. "Inferring proxy response in household travel
15 surveys with unknown completer using a group-based choice model." *Transportation*.
16 <https://doi.org/10.1007/s11116-019-10053-3>.
- 17 Mayor of London. 2025. "Delivering the Mayor's Transport Strategy 2024/25." *Transport for London*.
18 <https://tfl.gov.uk/cdn/static/cms/documents/delivering-the-mayors-transport-strategy-2024-25-acc.pdf>.
- 19 Meurs, Henk, Leo van Wissen, and Jacqueline Visser. 1989. "Measurement biases in panel data." *Transportation*. 16
20 (2): 175–94. <https://doi.org/10.1007/BF00163114>.
- 21 Metrolinx. 2018. "About the 2041 regional transportation plan." [https://www.metrolinx.com/en/projects-and-](https://www.metrolinx.com/en/projects-and-programs/regional-transportation-plan)
22 [programs/regional-transportation-plan](https://www.metrolinx.com/en/projects-and-programs/regional-transportation-plan).
- 23 Muff, Stefanie, Andrea Riebler, Leonhard Held, Håvard Rue, and Philippe Saner. 2015. "Bayesian analysis of
24 measurement error models using integrated nested Laplace approximations." *Journal of the Royal*
25 *Statistical Society: Series C (Applied Statistics)*. 64 (2): 231–52. <https://doi.org/10.1111/rssc.12069>.
- 26 Müller, Peter, and Riten Mitra. 2013. "Bayesian nonparametric inference—why and how." *Bayesian Analysis*. 8 (2):
27 269–302. <https://doi.org/10.1214/13-BA811>.
- 28 Murakami, Elaine, and David P. Wagner. 1999. "Can using global positioning system (GPS) improve trip
29 reporting?" *Transportation Research Part C: Emerging Technologies*. 7 (2–3): 149–65.
30 [https://doi.org/10.1016/S0968-090X\(99\)00017-0](https://doi.org/10.1016/S0968-090X(99)00017-0).
- 31 Nakamya, Juliet, Elke Moons, Suzana Koelet, and Geert Wets. 2007. "Impact of data integration on some important
32 travel behavior indicators." *Transportation Research Record*. 1993 (1): 89–94.
33 <https://doi.org/10.3141/1993-12>.
- 34 Peel Region. 2019. "Long Range Transportation Plan." [https://peelregion.ca/transportation-roads/plans-studies/long-](https://peelregion.ca/transportation-roads/plans-studies/long-range-transportation-plan)
35 [range-transportation-plan](https://peelregion.ca/transportation-roads/plans-studies/long-range-transportation-plan).
- 36 Primault, Vincent, Antoine Boutet, Sonia Ben Mokhtar, and Lionel Brunie. 2019. "The long road to computational
37 location privacy: A survey." *IEEE Communications Surveys & Tutorials*. 21 (3): 2772–93.
38 <https://doi.org/10.1109/COMST.2018.2873950>.
- 39 Province of Ontario. 2025. "Connecting the GGH: A Transportation Plan for the Greater Golden Horseshoe."
40 <https://www.ontario.ca/page/connecting-ghh-transportation-plan-greater-golden-horseshoe>.
- 41 Region of Durham. 2017. "Transportation Master Plan." [https://www.durham.ca/en/living-](https://www.durham.ca/en/living-here/transportationmasterplan.aspx)
42 [here/transportationmasterplan.aspx](https://www.durham.ca/en/living-here/transportationmasterplan.aspx).

- Richardson, Sylvia, and Walter R. Gilks. 1993. "A Bayesian approach to measurement error problems in epidemiology using conditional independence models." *American Journal of Epidemiology*. 138 (6): 430–42. <https://doi.org/10.1093/oxfordjournals.aje.a116875>.
- Rubin, Donald B. 1986. "Statistical matching using file concatenation with adjusted weights and multiple imputations." *Journal of Business & Economic Statistics*. 4 (1): 87–94. <https://doi.org/10.1080/07350015.1986.10509497>.
- Rubin, Donald B. 1987. "Multiple Imputation for Nonresponse in Surveys." New York: Wiley.
- Sammer, Gerd, Christian Gruber, Gerald Roeschel, Rupert Tomschy, and Max Herry. 2018. "The dilemma of systematic underreporting of travel behavior when conducting travel diary surveys: A meta-analysis and methodological considerations to solve the problem." *Transportation Research Procedia*. 32: 649–58. <https://doi.org/10.1016/j.trpro.2018.10.006>.
- San Diego Association of Governments. 2021. "Final Amended 2021 Regional Plan." <https://www.sandag.org/regional-plan/2021-regional-plan/final-2021-regional-plan>.
- Servizi, Valentino, Francisco C. Pereira, Marie K. Anderson, and Otto A. Nielsen. 2021. "Transport behavior-mining from smartphones: A review." *European Transport Research Review*. 13: 57. <https://doi.org/10.1186/s12544-021-00516-z>.
- Sfeir, Georges, Filipe Rodrigues, Maya Abou-Zeid, and Francisco Camara Pereira. 2024. "Analyzing the reporting error of public transport trips in the Danish national travel survey using smart card data." *Transportation*. <https://doi.org/10.1007/s11116-024-10535-z>.
- Stopher, Peter R., Camden FitzGerald, and Min Xu. 2007. "Assessing the accuracy of the Sydney Household Travel Survey with GPS." *Transportation*. 34 (6): 723–41. <https://doi.org/10.1007/s11116-007-9126-8>.
- Stopher, Peter R., and Stephen P. Greaves. 2007. "Household travel surveys: Where are we going?" *Transportation Research Part A: Policy and Practice*. 41 (5): 367–81. <https://doi.org/10.1016/j.tra.2006.09.005>.
- Verreault, Hubert, and Catherine Morency. 2015. "What about proxy respondent bias over time?" CIRRELT. Report No. CIRRELT-2015-55. <https://www.cirrelt.ca/documentstravail/cirrelt-2015-55.pdf>.
- Wang, Kaili, Sanjana Hossain, and Khandker Nurul Habib. 2022. "A hybrid data fusion methodology for household travel surveys to reduce proxy biases and under-representation of specific sub-group of population." *Transportation*. 49 (6): 1801–36. <https://doi.org/10.1007/s11116-021-10228-x>.
- Wolf, Jean, Marcelo Oliveira, and Miriam Thompson. 2003. "Impact of underreporting on mileage and travel time estimates: Results from global positioning system-enhanced household travel survey." *Transportation Research Record*. 1854 (1): 189–98. <https://doi.org/10.3141/1854-21>.
- Xiong, Chenfeng, Aref Darzi, Yixuan Pan, Sepehr Ghader, and Lei Zhang. 2020. "A data-driven analytical framework of estimating multimodal travel demand patterns using mobile device location data." arXiv. arXiv:2012.04776. <https://doi.org/10.48550/arXiv.2012.04776>.
- Yang, Mofeng, Yixuan Pan, Aref Darzi, Sepehr Ghader, Chenfeng Xiong, and Lei Zhang. 2020. "A data-driven travel mode share estimation framework based on mobile device location data." arXiv. arXiv:2006.10036. <https://doi.org/10.48550/arXiv.2006.10036>.
- Yang, Yitao, Bin Jia, Xiao-Yong Yan, Jiangtao Li, Zhenzhen Yang, and Ziyong Gao. 2022. "Identifying intercity freight trip ends of heavy trucks from GPS data." *Transportation Research Part E: Logistics and Transportation Review*. 157: 102590. <https://doi.org/10.1016/j.tre.2021.102590>.
- York Region. 2022. "Transportation Master Plan." <https://www.york.ca/york-region/plans-reports-and-strategies/transportation-master-plan>.
- Zhao, Guangchen, Mohammad B. Al-Khasawneh, Tiziana Tuoto, and Cinzia Cirillo. 2025. "Data fusion for travel analysis: Linking travel survey and mobile device location data." *Transportation*. <https://doi.org/10.1007/s11116-025-10666-x>.

1

Appendix A. Applications of HTS-derived Indicators in Transportation Master Plans and Technical Reports

2 Table A1. Applications of household travel survey data in transportation master plans and technical reports from the Greater Toronto Area

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
York Region, Ontario	York Region Transportation Master Plan (York Region 2022)	Integrated active-transportation network	Expand bike lanes, trails, and multi-use paths so walking/cycling are safer, more direct, and better connected to destinations and transit.	Total mode share	Transportation Tomorrow Survey (2016, with comparisons to 2011, 2006, and 2001)
York Region, Ontario	York Region Transportation Master Plan (York Region 2022)	Rapid transit corridors	Identify and protect rapid transit corridors to support planned growth to 2051.	Estimation of ABM model using HTS	Transportation Tomorrow Survey (2016, with trend comparison back to 2001)
York Region, Ontario	York Region Transportation Master Plan (York Region 2022)	Optimized road network	Manage congestion and improve access by making best use of existing roads, lanes, and intersections, and by targeting road improvements where they are most needed.	Estimation of ABM model using HTS	Transportation Tomorrow Survey (2016)
City of Toronto	Yonge Street North Transportation Master Plan (City of Toronto 2022)	Parking management strategy	Eliminate minimum parking requirements and set maximum parking supply using a formula tied to the plan's desired mode share.	Estimation of ABM model using HTS	Transportation Tomorrow Survey (TTS), especially 2016 TTS for comparable Toronto areas
City of Toronto	Yonge Street North Transportation Master Plan (City of Toronto 2022)	Reduce automobile reliance / 2041 target mode split	Shifting the study area from auto dependence toward transit and active transportation, use that target in transportation analysis for the network.	Total mode share	Transportation Tomorrow Survey (TTS) 2006, 2011, 2016, plus 2016 TTS for comparable transit-oriented Toronto areas
City of Toronto	Yonge Street North Transportation Master Plan (City of Toronto 2022)	Cycling and short-trip active transportation	Increase cycling for short trips and first/last-mile access to transit through cycling network and related active transportation improvements.	Trip distance distributions by mode	Transportation Tomorrow Survey (TTS) 2006, 2011, 2016 for the Focused Study Area
City of Toronto	ConsumersNext Transportation Master Plan Final Report (City of Toronto 2019)	Transit mode-shift strategy tied to Sheppard LRT	"LRT as is" strategy assumes higher transit use as part of the preferred transportation solution, uses TTS trends to justify increasing transit mode share.	Total mode share	Transportation Tomorrow Survey (TTS)
City of Toronto	ConsumersNext Transportation Master Plan Final Report (City of Toronto 2019)	Pedestrian and cycling improvements	The active transportation package, including cycling and pedestrian improvements, is explicitly justified by TTS evidence that many short trips still made by car could shift to walking or cycling.	Mode share by distance	Transportation Tomorrow Survey (TTS), 2011
City of Toronto	ConsumersNext Transportation Master Plan Final Report (City of Toronto 2019)	Higher-occupancy vehicle strategy	Encourage higher-occupancy vehicles and reduce peak-period auto dependence as part of the HOV-transit lane strategy.	Total mode share	Transportation Tomorrow Survey (TTS), 2006 and 2011

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
City of Toronto	ConsumersNext Transportation Master Plan Final Report (City of Toronto 2019)	Regional transit hub and improved GO connections	Propose a regional transit hub and GO bus rerouting to attract longer-distance trips away from auto use.	AM peak mode share	Transportation Tomorrow Survey (TTS), primarily 2011
Regional Municipality of Durham, Ontario	Durham Transportation Master Plan (Region of Durham 2017)	Action 109: support ongoing data collection	Support new and on-going data collection initiatives critical to monitoring transportation conditions and predicting future needs, including the Transportation Tomorrow Survey.	Total mode share, trip distance distributions by mode, estimation of model using HTS	Transportation Tomorrow Survey (TTS)
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	50% sustainable mode-share target	Core policy target to shift half of trips to sustainable modes by 2041, implemented through the Sustainable Transportation Strategy (STS).	Total mode share	Transportation Tomorrow Survey (TTS), especially 1996, 2001, 2006, 2011 household trip data
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	Preferred alternative: road and active-transportation improvements	Selected alternative combining planned road improvements with active transportation infrastructure to achieve the 50% sustainable mode share.	Total mode share, AM peak trip counts	TTS / STS household mode-share baseline
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	TDM: carpooling and vanpooling	Promote ridesharing and new carpool lots to reduce driving and improve network efficiency.	Total mode share	Transportation Tomorrow Survey (TTS) household trip mode-share data, primarily 2011 baseline
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	Walking and cycling infrastructure	Expand pedestrian improvement areas and cycling facilities to make walking/cycling viable travel choices and support first/last-mile access.	Total mode share	Transportation Tomorrow Survey (TTS) household trip and trip-distance data, mainly 1996-2011 trend/target work
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	Transit enhancements on Regional roads	Transit is treated as the largest required mode-share increase within the 50% sustainable mobility objective; the plan supports transit enhancement on regional roads.	Total mode share	Transportation Tomorrow Survey (TTS)
Region of Peel, Ontario	Let's Move Peel: Long Range Transportation Plan (Peel Region 2019)	Kitchener GO Line advocacy	Advocate for two-way, all-day, 15-minute GO service on the Kitchener Line from Union Station to Mount Pleasant GO Station.	Total mode share	Transportation Tomorrow Survey (TTS)
Greater Toronto and Hamilton Area (regional plan)	Metrolinx 2041 Regional Transportation Plan (Metrolinx 2018)	Priority Action 2.1: Frequent Rapid Transit Network	Build a region-wide Frequent Rapid Transit Network, including subway/BRT/LRT/GO and Priority Bus services to provide a stronger alternative to driving.	Total mode share	Transportation Tomorrow Survey (TTS), including 2006, 2011, and 2016 travel-market mode shares
Greater Toronto and Hamilton Area (regional plan)	Metrolinx 2041 Regional Transportation Plan (Metrolinx 2018)	Priority Action 3.1: integrated transit services and fares	Remove fare barriers and create an integrated fare system so transit travel across municipal boundaries is seamless and more attractive.	AM peak mode share, AM peak trip counts	Transportation Tomorrow Survey (TTS)

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
Greater Toronto and Hamilton Area (regional plan)	Metrolinx 2041 Regional Transportation Plan (Metrolinx 2018)	Strategy 4: walking/cycling-supportive design and Regional Cycling Network	Align transportation and land use so communities and road networks support transit, cycling, and walking; establish a Regional Cycling Network to connect cyclists to rapid transit and key destinations.	Trip distance distributions by mode	Transportation Tomorrow Survey (TTS)-based household trip-length and mode-share analysis
Greater Toronto and Hamilton Area (regional plan)	Metrolinx 2041 Regional Transportation Plan (Metrolinx 2018)	Priority Action 4.9: K-12 school travel programs	Coordinate with partners to advance active and sustainable school travel tailored to each community, prioritizing active/sustainable travel by children and youth.	Total mode share	Transportation Tomorrow Survey (TTS), especially 2006, 2011, and 2016 school-trip mode-share analysis
City of Mississauga	Mississauga Transportation Master Plan (City of Mississauga 2021)	Sustainable mode-share target and Action 11	Set a city-wide policy direction that half of trips to, from, and within the city should be taken by sustainable modes. Action 11 breaks down policy into more specific targets by mode, geography, land use type, or other segments.	Total mode share	Transportation Tomorrow Survey (TTS) 2016
City of Mississauga	Mississauga Transportation Master Plan (City of Mississauga 2021)	Pedestrian network and short-trip walking	Identify gaps and make it safe, easy, and comfortable for people to make short trips on foot.	Trip distance distributions by mode	Transportation Tomorrow Survey (TTS) 2016
City of Mississauga	Mississauga Transportation Master Plan (City of Mississauga 2021)	Long-term cycling network	Implement long-term cycling network so cycling becomes a safe, connected, comfortable, and convenient option.	Trip distance distributions by mode	Transportation Tomorrow Survey (TTS) 2016
City of Mississauga	Mississauga Transportation Master Plan (City of Mississauga 2021)	School Walking Routes program	Expand the School Walking Routes program to encourage walking and cycling by students, alongside ensuring no-cost access options where walking distance is unreasonable.	Total mode share	Transportation Tomorrow Survey (TTS) 2011
City of Mississauga	Mississauga Transportation Master Plan (City of Mississauga 2021)	Action 73: Milton GO two-way, all-day service	Advocate for all-day, two-way GO train service on the Milton line.	Total mode share, total trip counts	Transportation Tomorrow Survey (TTS), 1986-2016 trends
Greater Golden Horseshoe (Ontario regional plan)	Connecting the GGH: A Transportation Plan for the Greater Golden Horseshoe (Province of Ontario 2025)	Travel demand management	Reduce or shift the need to travel, especially during peak hours while promoting telecommuting, flexible work hours, and alternative modes.	Total mode share	Transportation Tomorrow Survey (TTS) 2016; TTS 2011 and 2016
Greater Golden Horseshoe (Ontario regional plan)	Connecting the GGH: A Transportation Plan for the Greater Golden Horseshoe (Province of Ontario 2025)	Connected region-wide transit system	Transform the region from a downtown-focused radial network into a connected grid with integrated fares and services and better local connections.	Total mode share	Transportation Tomorrow Survey (TTS) 2016; TTS 2011 and 2016
Greater Golden Horseshoe (Ontario regional plan)	Connecting the GGH: A Transportation Plan for the Greater Golden Horseshoe (Province of Ontario 2025)	Active transportation for short trips	Make walking and cycling more attractive and widely available for short trips.	Total mode share	Transportation Tomorrow Survey (TTS) 2016

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
City of Richmond Hill	Richmond Hill Transportation Master Plan (City of Richmond Hill 2023)	Scenario 2 - City Improvements+	Recommend a multi-modal preferred alternative combining city road improvements with Regional and Provincial improvements to support growth to 2051.	Estimation of 4-step model using HTS	Transportation Tomorrow Survey (TTS) 2006 and 2016
City of Richmond Hill	Richmond Hill Transportation Master Plan (City of Richmond Hill 2023)	Prioritize active transportation and localized improvements	Emphasize active transportation and other sustainable modes instead of relying only on traditional road expansion.	Estimation of 4-step model using HTS	Transportation Tomorrow Survey (TTS) via the calibrated York Region model
City of Richmond Hill	Richmond Hill Transportation Master Plan (City of Richmond Hill 2023)	Prioritize local improvements and non-auto modes	Set a city-wide mode-shift direction for future policy updates.	Total mode share	Transportation Tomorrow Survey (TTS) 2016 baseline, with future monitoring tied to TTS-based indicators
City of Richmond Hill	Richmond Hill Transportation Master Plan (City of Richmond Hill 2023)	Monitoring plan: multi-modal data collection	Use recurring multi-modal indicators to monitor Transportation Master Plan performance over time.	Total trip counts/distances (VKT)	Transportation Tomorrow Survey (TTS)

1

2

3

4

5

6

7

8

9

10

11

1 Table A2. Applications of household travel survey data in Canadian transportation master plans and technical reports outside the Greater Toronto Area

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
Ottawa	Transportation Master Plan (City of Ottawa 2013)	Action 4-1: Implement the Ottawa Pedestrian Plan	Action implements the updated Pedestrian Plan and explicitly ties it to a walking target: plan includes Section 2.4 which sets a 2031 morning peak-period walking mode share target of 10% city-wide.	AM peak mode share	National Capital Region household travel survey (latest survey in fall 2011; formal survey title not specified in the plan).
Ottawa	Transportation Master Plan (City of Ottawa 2013)	Action 5-1: Implement the Ottawa Cycling Plan	Action implements the updated Cycling Plan and explicitly ties it to cycling targets set in Section 2.4. The chapter states a 2031 morning peak-period cycling mode share target of 5% city-wide, plus a refined 8% target for travel inside the Greenbelt.	AM peak mode share	National Capital Region household travel survey (latest survey in fall 2011; formal survey title not specified in the plan).
Ottawa	Transportation Master Plan (City of Ottawa 2013)	Action 6-1: Implement the Affordable RTTP Network by 2031	Transit chapter states that OC Transpo morning peak-period travel share was just over 22% in 2011, the plan aims to increase it to 26% by 2031. Action 6-1 is the main implementation for the rapid transit and transit-priority network.	AM peak mode share	National Capital Region household travel survey (latest survey in fall 2011; formal survey title not specified in the plan).
Vancouver	Transportation 2040 (City of Vancouver 2012)	T 1.1. Advance new and improved rapid transit	Policy states that fast, frequent, reliable, high-capacity transit is essential to attract riders and meet mode share targets.	Total mode share	TransLink trip diary surveys (the plan also mentions census data in the same trend analysis; formal survey title not specified).
Vancouver	Transportation 2040 (City of Vancouver 2012)	C 2.1.1. Periodically review policies for new developments to ensure abundant and conveniently located secure bicycle parking and end-of-trip facilities	Action states bicycle parking minimums should support long-term mode share targets and ownership levels.	Total mode share	TransLink trip diary surveys (plus census in the trend analysis; formal survey title not specified).
Vancouver	Transportation 2040 (City of Vancouver 2012)	M 2.1.1. Develop and implement a strategy to revise minimum off-street parking requirements elsewhere based on target mode shares	Parking action ties off-street parking minimums to target mode shares.	Total mode share	TransLink trip diary surveys (plus census in the trend analysis; formal survey title not specified).
Vancouver	Transportation 2040 (City of Vancouver 2012)	M 2.1.2. Introduce parking maximum allowances throughout the city based upon current ownership levels or existing mode share	Action explicitly bases parking maximums on existing mode share.	Total mode share	TransLink trip diary surveys (plus census in the trend analysis; formal survey title not specified).
Edmonton	The Way We Move: Transportation Master Plan (City of Edmonton 2009)	Strategic Objective 5.1 - The City will pursue expansion of the LRT to all sectors of the city with a goal to increase transit ridership and transit mode split	Expand the LRT to all sectors of the city to increase transit use and support compact urban communities. The plan links objective to existing household travel survey evidence on current mode shares.	Total mode share	2005 Edmonton Household Travel Survey.

1 Table A3. Applications of household travel survey data in international transportation master plans and technical reports

City / region	Plan name and year	Policy / action	Brief description	General HTS metric	HTS name
San Diego region (CA, US)	San Diego Forward: The 2021 Regional Plan, Appendix M (San Diego Association of Governments 2021)	Update regional model assumptions using HTS data	Conduct the San Diego Regional Transportation Study every ten years to update assumptions in the ABM2+ activity-based regional model.	Model estimation / calibration inputs	San Diego Regional Transportation Study (2016-2017)
Minneapolis (MN, US)	Transportation Action Plan (City of Minneapolis 2020)	Mode-shift goal: 60% of trips by non-auto modes	Target 60% of trips by modes other than private automobile, including 35% by walking and cycling and 25% by transit.	Total mode share	Travel Behavior Inventory (2019)
London (UK)	Delivering the Mayor's Transport Strategy 2024/25 (Mayor of London 2025)	Monitor the active-travel target	Use the London Travel Demand Survey to monitor the share of residents achieving at least 20 minutes of active travel per day.	Active-travel participation rate	London Travel Demand Survey (LTDS), 2010/11-2023/24
London (UK)	Delivering the Mayor's Transport Strategy 2024/25 (Mayor of London 2025)	Monitor trips per person per day	Track average daily trip rates across recent LTDS cycles to monitor changes in travel activity.	Trip rate by day	London Travel Demand Survey (LTDS), 2005/06-2023/24
City of Ryde (Sydney, NSW, AU)	City of Ryde Integrated Transport Strategy 2045 (City of Ryde 2025)	Use modal split as an existing-conditions benchmark	Use the household travel survey modal split for an average workday to characterize current travel patterns and support the transport strategy.	Total mode share	Transport for NSW Household Travel Survey (HTS), 2022/23

2

3

4

5