UTTRI

Research Report

Vehicle for Hire Bylaw Review

Report 2: Comparison of 2016 PTC and 2016 TTS

Gozde Ozonder and Eric J. Miller

April 2019

UTTRI TECHNICAL SUPPORT FOR THE CITY OF TORONTO VEHICLE FOR HIRE BYLAW REVIEW

Report No: 2

Comparison of 2016 PTC and 2016 TTS

Gozde Ozonder, M.Sc.

Eric J. Miller, Ph.D.

April, 2019

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

UTTRI

TABLE OF CONTENTS

List of Figures **Page No.**

1. INTRODUCTION

This technical report presents the work undertaken in support of the City of Toronto's *Vehicle for Hire Bylaw Review* by the University of Toronto Transportation Research Institute (UTTRI) to analyze and compare the patterns in the usage of ride-sourcing services provided by the Private Transportation Companies (PTCs) as recorded by these companies and as recorded in the 2016 Transportation Tomorrow Survey (TTS).

The 2016 TTS conducted in the fall (September-December) 2016 time period collected information on Uber trips as an explicit mode of travel (Ashby, 2018; Miller, et al., 2019). The trip records in the data have socioeconomic attributes of trip-makers (e.g., age, sex, income class, etc.), their household characteristics (household size, number of vehicles owned, etc.) and trip purposes attached (Data Management Group – Reports, n.d.). They, therefore, provide a statistically representative description of the users of the services provided by PTCs and their reasons for travel along with time and start/end locations of the trips. Despite the relatively modest market penetration of PTCs (only Uber at the time) in the City during the fall of 2016, TTS provides a considerably richer description of PTC travel and trip-makers than can be obtained from the PTC records alone, since PTC records do not have user-attributes or trip purposes attached (Miller, et al., 2019). Thus, if it can be shown that the trip records in the 2016 TTS data set can successfully represent the patterns observed in the actual usage of the services, statistical models can be built to impute further attributes associated with the trips recorded by the PTCs, such as the corresponding trip purposes, user attributes, etc. The purpose of this study as documented in this report is to identify the spatiotemporal patterns in both data sets and determine the similarities and/or discrepancies by examining the distributions of the trips mainly considering their time- and space-related attributes.

This report is one of the deliverables by the UTTRI team and it complements Report No. 1 which focusses on the demographics of the trip-makers and their household characteristics as recorded in the 2016 TTS. The rest of the report is organized as follows. Section 2 reports the results of the comparative analyses. Section 3 concludes the report with a summary of findings.

2. COMPARATIVE ANALYSES

This section presents the results of the analyses conducted to compare the spatial and temporal patterns in Uber trips between the 2016 TTS data set and Uber's own data. The data set received from Uber is referred to as the PTC data set in the rest of this document.

To render the comparison compatible the data sets have been cleaned considering several aspects prior to the analyses. First, since the interview period of the 2016 TTS is between September 6th and December 16th, the trips that fall into the same interval are selected from the PTC data set. Second, the TTS collects trip records only for week days, and thus, the trips made during the weekends are removed from the PTC data set. Third, the analyses are restricted to the trips that have both ends (i.e., origins and destinations) within the Greater Toronto and Hamilton Area

(GTHA) in both data sets. The compatible TTS and PTC data sets include over 40,000 and 4,000,000 trips respectively. Due to the large difference in the total number of trips most graphs included in this report are prepared using the percentages to compare the distributions in both data sets on the same graph.

Figure 1 shows the monthly distribution of trips in both data sets. It is seen that in both data sets a larger portion of trips are recorded during October and November. One might anticipate more trips in December considering the approaching Christmas holiday season, however, the observed distribution is probably because not all the days from September and December are included in the survey interview period, and hence, in this analysis. Furthermore, it can be seen that the TTS has recorded lower a portion of trips in September and December when compared to PTC data set as these months are the survey's ramp-up and ramp-down periods, respectively.

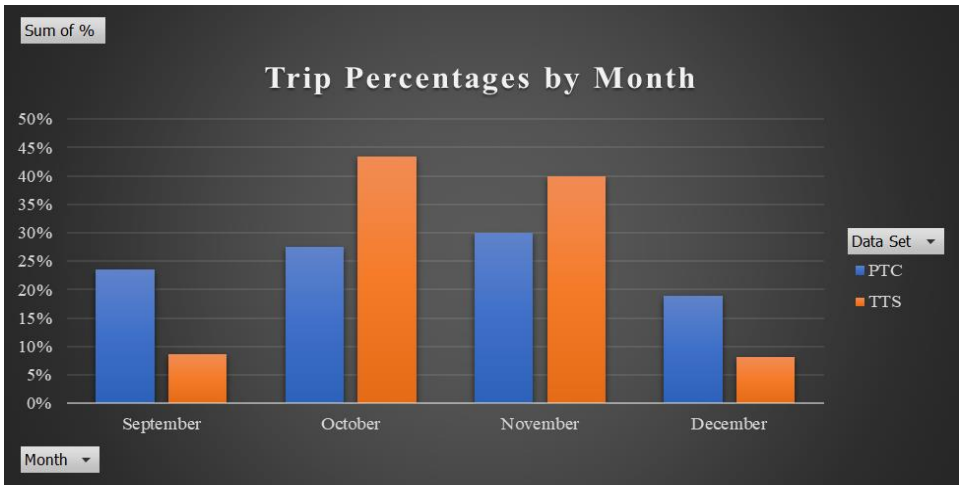

Figure 1. Trip Percentages by Month

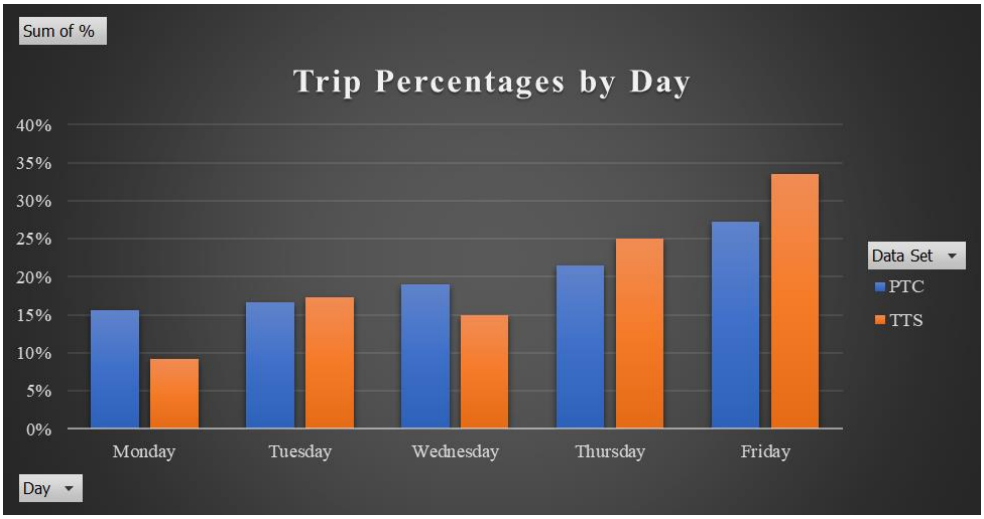

Figure 2. Trip Percentages by Day

The distribution of trip percentages by day is shown in Figure 2. In the PTC data set, the percentage of trips increases from Monday to Friday. Although the TTS data set does not precisely capture this trend, the distribution can be considered quite similar. It is important to note Fridays are oversampled in the TTS since the majority of Uber-users (93% as demonstrated in Report No. 1) have filled the survey online, many of whom carried out the survey task during the weekend, and hence, they have reported their trips for Friday.

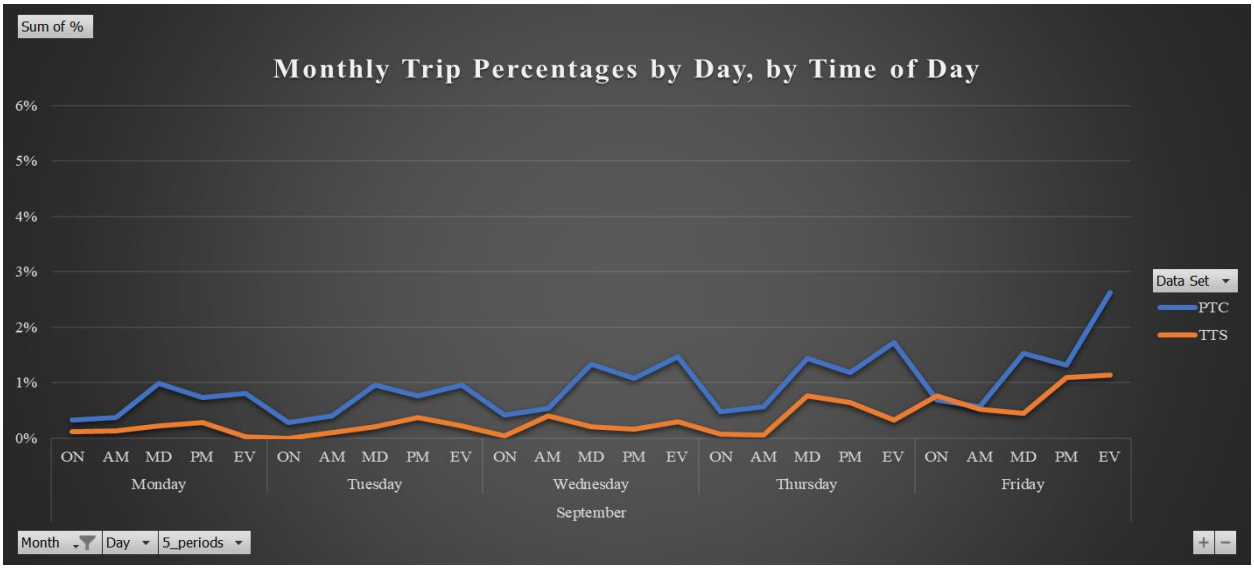

Figure 3. Trip Percentages in September by Day, by Time of Day

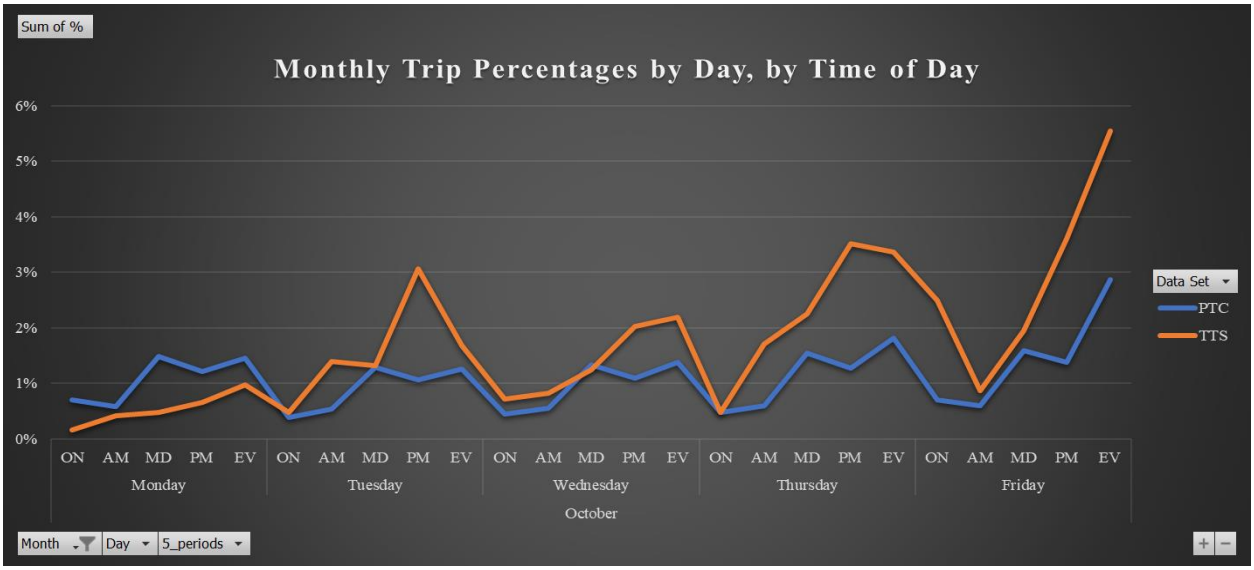

Figure 4. Trip Percentages in October by Day, by Time of Day

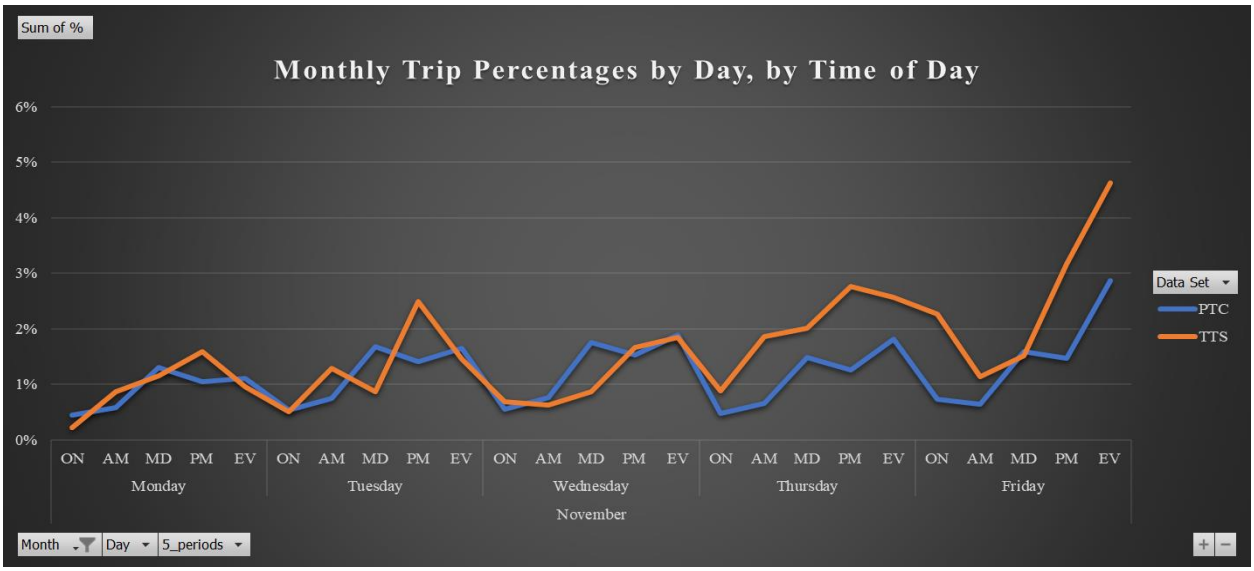

Figure 5. Trip Percentages in November by Day, by Time of Day

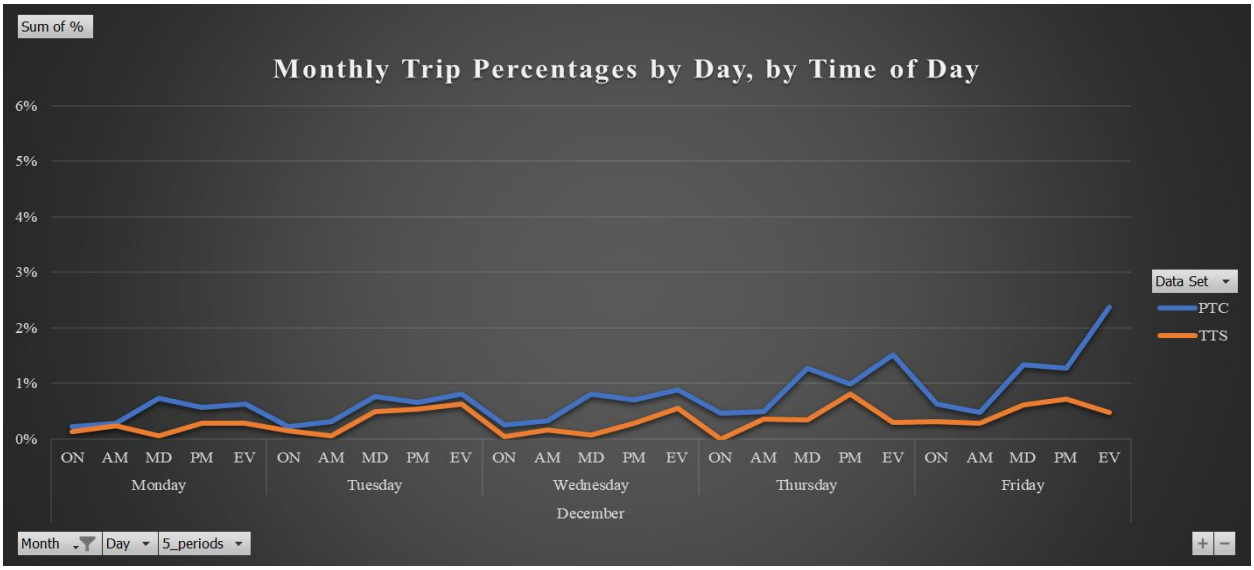

Figure 6. Trip Percentages in December by Day, by Time of Day

Figures 3-6 present the distributions of trip percentages by day and by time of day for four months: September, October, November and December, where a day is divided into five time periods: "ON" period starts at 12 am and continues until 6 am, "AM" period starts at 6 am in the morning, continues until 9 am; "MD" period starts at 9 am continues until 3 pm; "PM" period starts at 3 pm, continues until 7 pm; lastly, "EV" period starts at 7 pm, continues until 12 am.

It can be seen that in the PTC data set in all five week days in all four months trip distribution peaks at "MD" and "EV" periods, i.e. during middays and evenings. TTS data cannot capture this trend very well, which can be attributed to under-reporting in the survey during these periods, which is demonstrated in Figures 7 and 8. Nevertheless, it is observed that the survey can capture the "EV" peaks on Fridays in October and November very well. In addition, for statistical model building purposes, it might be useful to filter the trips on Tuesdays in December (within the aforementioned interval, i.e., between December 1st and December 16th), where both peaks in the PTC data are successfully captured in the TTS data.

However, these distributions might be misleading in terms of highlighting when the hourly peaks in the usage are observed due to aggregation bias, as the "MD" and "EV" periods are longer than "AM" and "PM" periods, which correspond to usual commuting hours, thus, it is not surprising that a greater portion of the trips fall into those longer periods ("MD" and "EV"). Although "ON" is also a long period, it corresponds to the night period, and since in this analysis only week day trips are analyzed, the portion of overnight trips are low.

When the hourly distributions of total number of trips in the data sets are plotted (refer to Figure 7 for the PTC data, and Figure 8 for the TTS data), the under-reporting periods of the trips in the survey can be observed, as previously discussed. Comparing the two graphs, one can spot the discrepancy in the "MD" and "EV" periods. Figure 7 shows that the largest number of trips are made around 7 pm on week days. It makes sense that the number of trips reduces from 1 am to 4 am, and then starts increasing in the morning. Hourly plots provide a better sense of when the actual hourly peaks happen, which in fact fall into the "AM" and "PM" periods according to the aforementioned categorization.

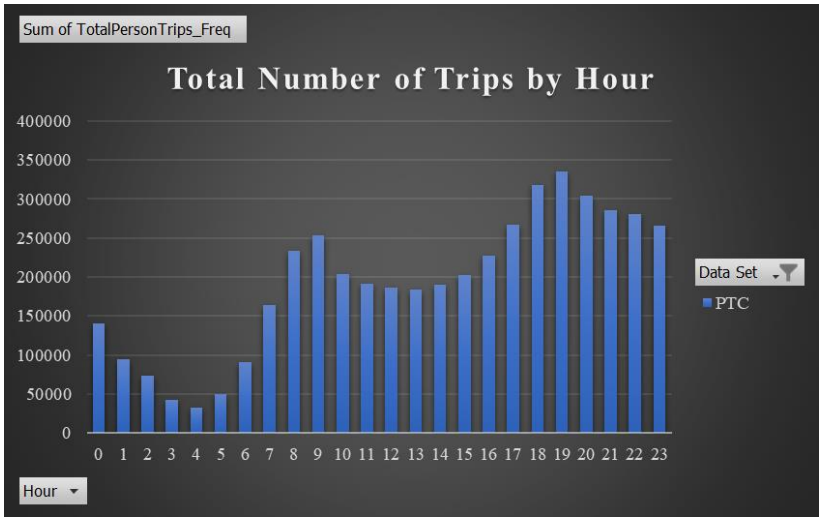

Figure 7. Total Number of PTC Trips by Hour

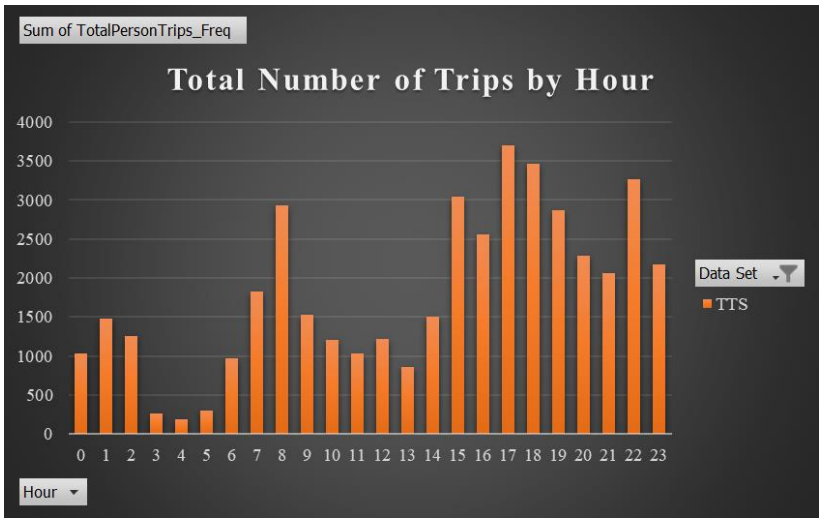

Figure 8. Total Number of TTS Trips by Hour

The trip length frequency distributions for the data sets are shown separately in Figures 9 and 10, where the trip length is in kilometres and is aggregated into ten categories. Both distributions display similar skewed trends. It is seen that Uber trips are usually short trips that are less than ten kilometres, which makes sense as the trips are likely to get more costly with an increase in the distance, similar to taxi fares[1], in which case the individuals might prefer alternative transportation modes. Trips between five to ten kilometres seem to be oversampled in the TTS compared to the distribution of Uber trips in the PTC data set.

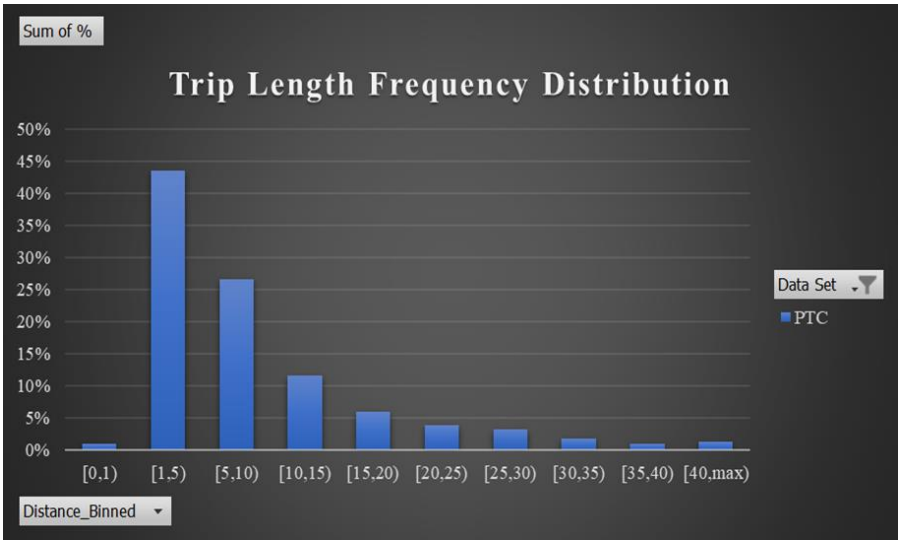

Figure 9. Trip Length Frequency Distribution – PTC Data

[1] This is just an assumption as the pricing methods of Uber is unknown to the UTTRI researchers.

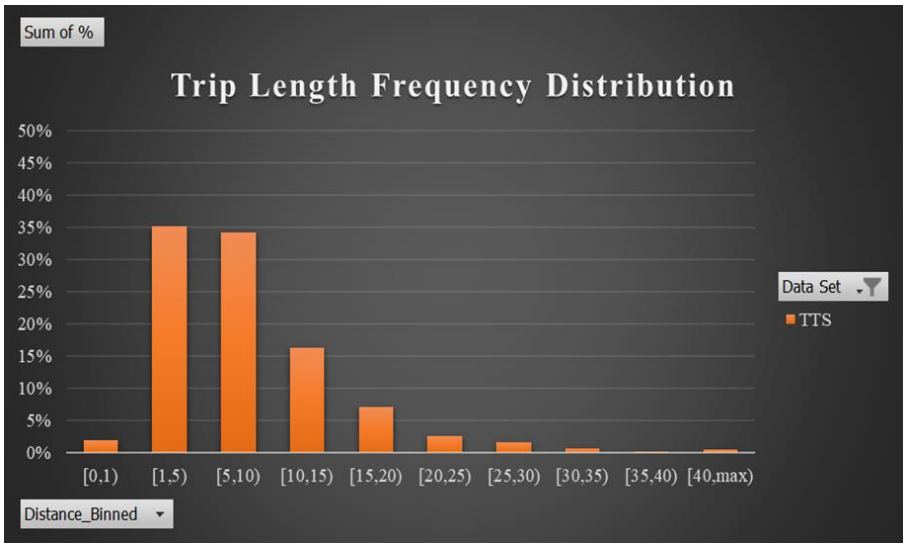

Figure 10. Trip Length Frequency Distribution – TTS Data

Figure 11 shows a more disaggregate distribution of trip length frequencies, illustrating the distributions by time of day. The trends in both data sets in all time periods seem to match quite well. For further disaggregation please refer to Figure A1 in the Appendix.

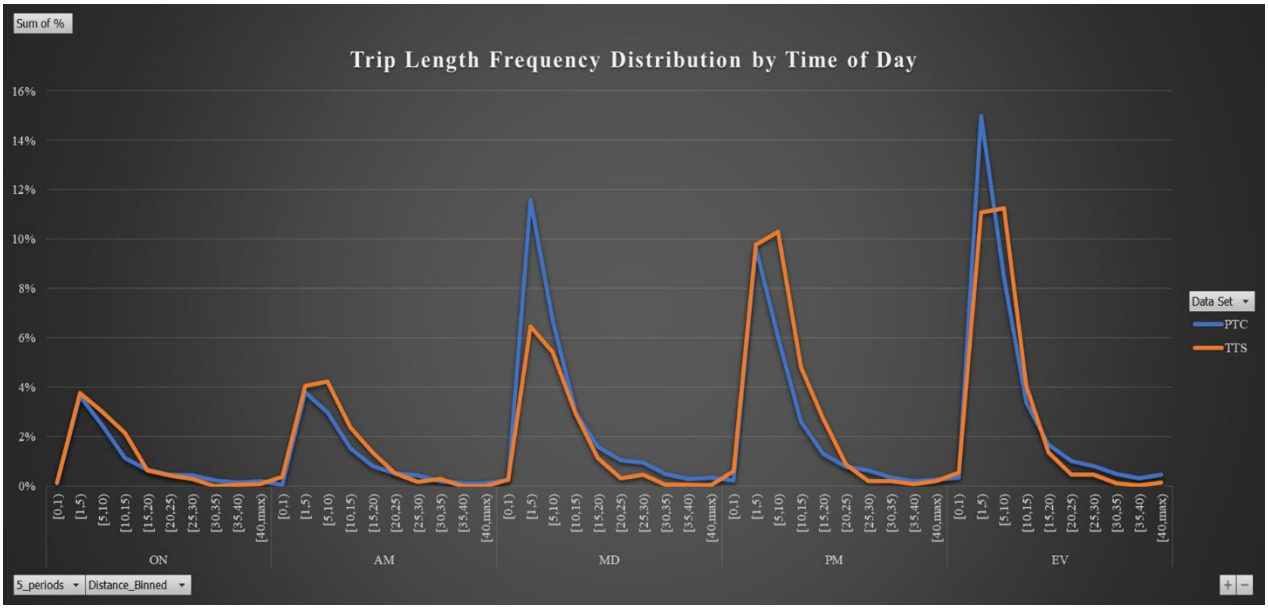

Figure 11. Trip Length Frequency Distribution by Time of Day

The trip duration frequency distributions for the data sets are shown separately in Figures 12 and 13, where the trip duration is in minutes. Duration is also aggregated into ten categories. Trip durations are mostly less than 20 minutes. Although the distributions are not very different, more discrepancy is observed when the duration distributions are studied compared to the trip lengths. This can be attributed to the fact that trip durations are not recorded in the survey, they are attached to the TTS trip records based on the zone-to-zone travel times received from Emme4 network assignment model (Travel Modelling Group, 2019). These durations do not reflect the point-to-point durations of the trips, they are at a more aggregate level, unlike the durations reported for the trips in the PTC data set.

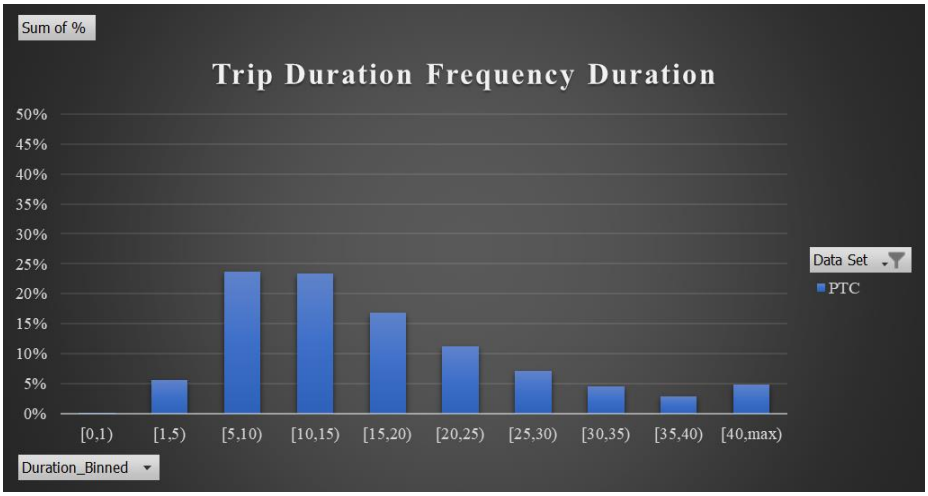

Figure 12. Trip Duration Frequency Distribution – PTC Data

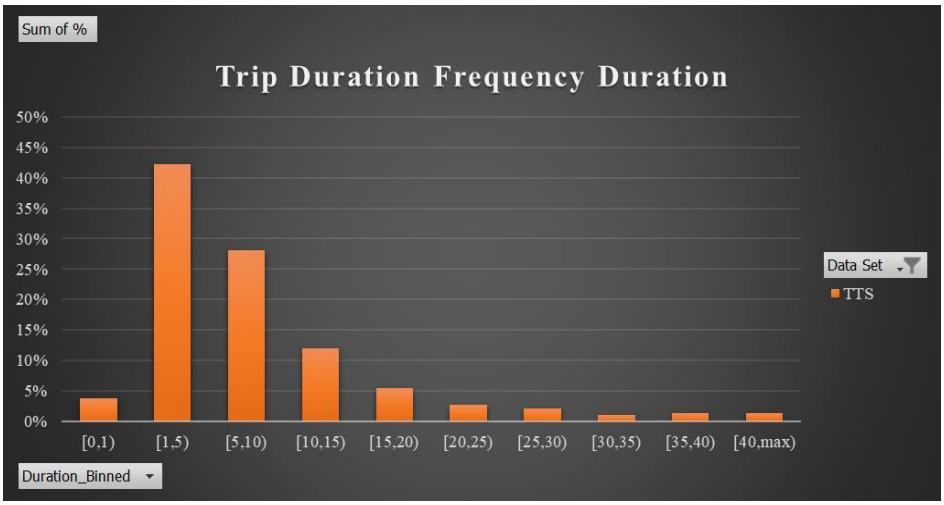

Figure 13. Trip Duration Frequency Distribution – TTS Data

Figure 14 shows a more disaggregate distribution of trip duration frequencies, illustrating the distributions by time of day. As discussed above, the trends do not display a one-to-one match. Nonetheless, the orange line can be thought as a lagged version of the blue line where the distribution shapes are similar. For further disaggregation please refer to Figure A2 in the Appendix.

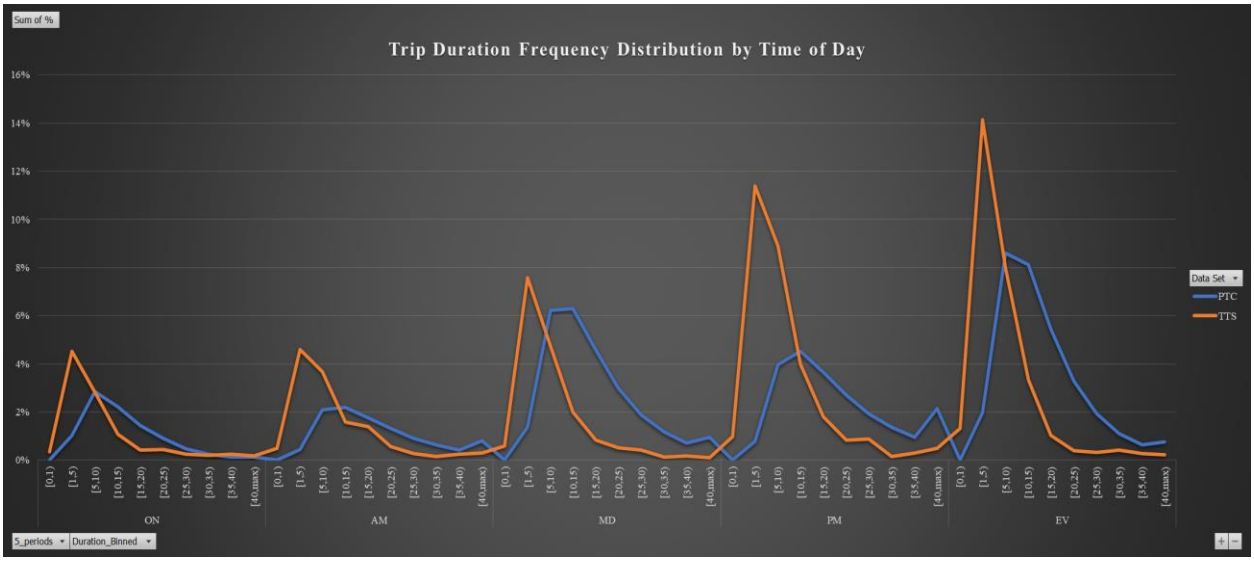

Figure 14. Trip Duration Frequency Distribution by Time of Day

Figures 15-19 show the daily trip percentages by time of day in both data sets that originate from the City of Toronto categorized by the Planning Districts (PDs)[2]. Each figure shows the distribution for one week day, starting from Monday (Figure 15). These figures show that the TTS data has been able to capture the spatial distribution within the City quite well. Most of the trips originate from the downtown area in the City, PD1 in all days and in all time periods. Some neighbouring PDs (e.g., PD2, PD4, PD6) also have relatively higher percentage of trips.

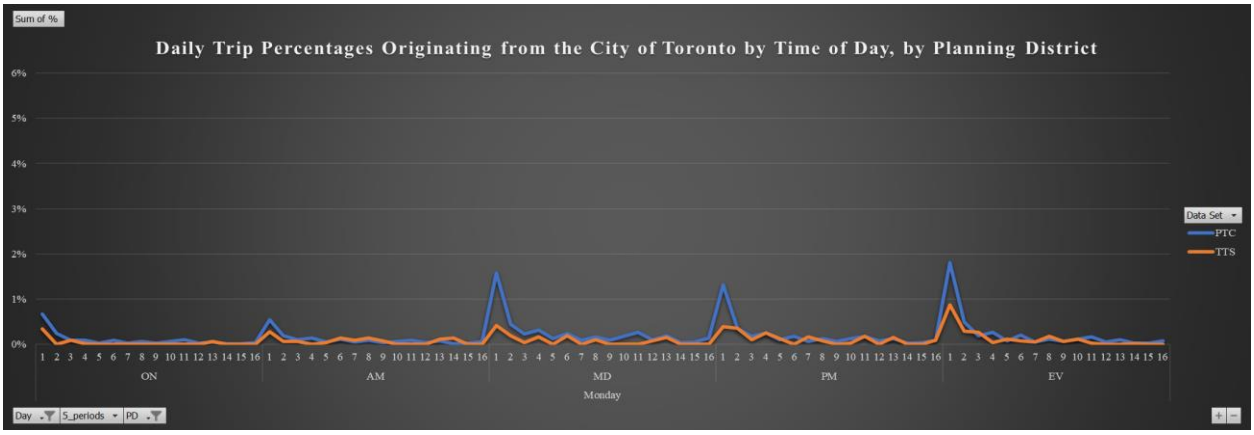

Figure 15. Daily Trip Percentages Originating from the City of Toronto by Time of Day, by Planning District – Monday

[2] A map of the PDs is shown in Figure 20 for reference.

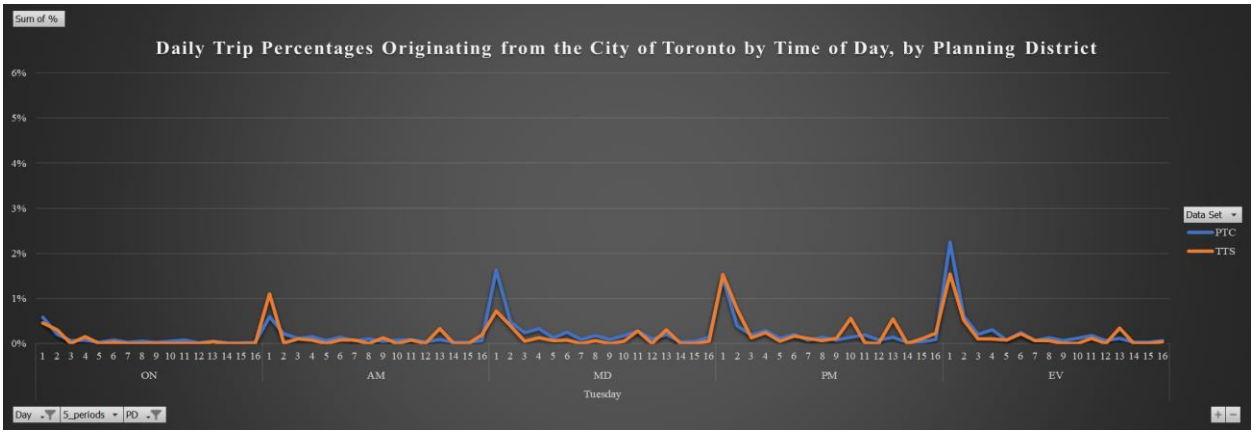

Figure 16. Daily Trip Percentages Originating from the City of Toronto by Time of Day, by Planning District – Tuesday

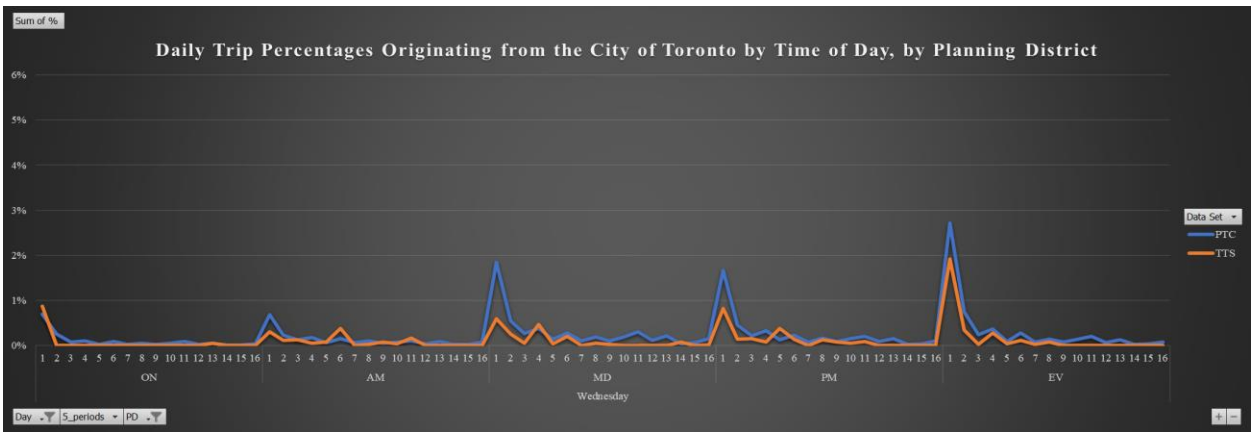

Figure 17. Daily Trip Percentages Originating from the City of Toronto by Time of Day, by Planning District – Wednesday

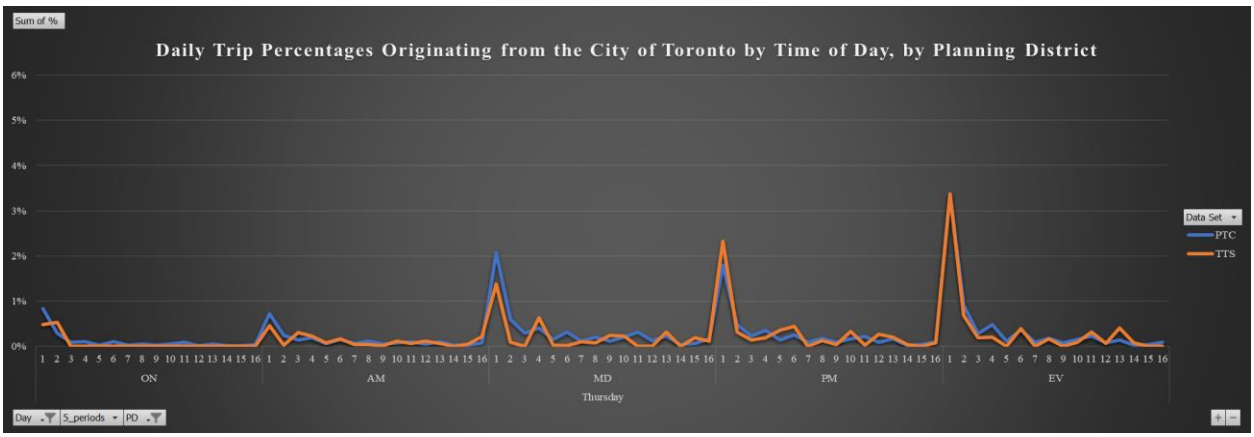

Figure 18. Daily Trip Percentages Originating from the City of Toronto by Time of Day, by Planning District – Thursday

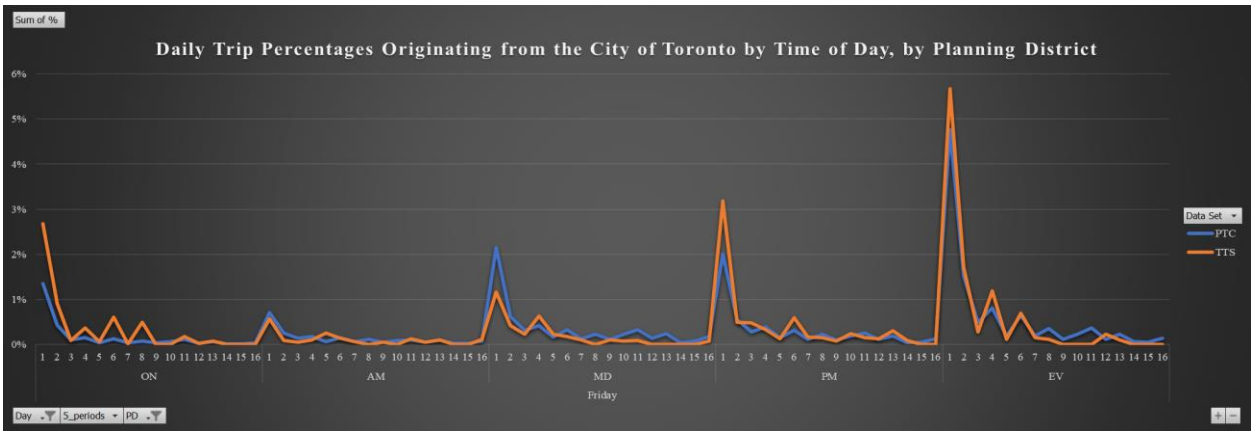

Figure 19. Daily Trip Percentages Originating from the City of Toronto by Time of Day, by Planning District – Friday

In the Appendix, additional maps are provided to demonstrate the hot spots of trip origins within the GTHA (Figures A3-A7). Each map shows the distribution of trip origins in a certain data set, on a given time period, in a given day.

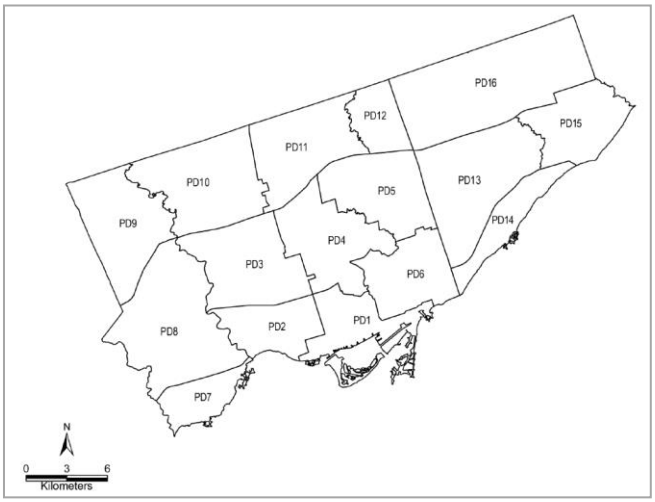

Figure 20. Map of Planning Districts in the City of Toronto - Source: (Ashby, 2018)

3. CONCLUSIONS

Spatiotemporal attributes of week day trips as recorded by Uber and as recorded in the 2016 Transportation Tomorrow Survey (2016 TTS) are compared in this report considering the trips that originate from/terminate in the Greater Toronto and Hamilton Area (GTHA) during the fall of 2016. Monthly, daily, hourly, etc. comparisons are demonstrated. The descriptive analyses documented in this report are crucial because unlike the PTC data set, the TTS data set includes socio-demographic information regarding the users of the service which enables building statistical models that can identify the factors that influence the demand for the ride-sourcing services.

Despite the small number of Uber trips in the 2016 TTS, it has been shown that in general the TTS data set captures the distributions in the PTC data set successfully. There are slight discrepancies, such as under-reporting of trips during the midday ("MD") and evening ("EV") periods in the survey. However, it has been shown that the general trends in trip length distribution, spatial distribution, etc. all support the use of the TTS data to impute further attributes for the trips recorded in the PTC data set.

It is useful to note that the small number of Uber trips reported in the 2016 TTS is not surprising since the interview period corresponds to the early days of Uber usage in the City of Toronto. As underlined in Report No. 1, it is expected that the next cycle of the TTS, the 2021 TTS, will capture the trends in Uber usage better, providing more information to the researchers and planners.

Acknowledgements & Disclaimer

This report was prepared as part of the project *UTTRI Technical Support for the City of Toronto Vehicle for Hire Bylaw Review*, funded by the City of Toronto. The findings and opinions expressed in this report are those of the authors alone and do not necessarily reflect those of the City of Toronto.

REFERENCES

Ashby, B. (2018). Transportation Tomorrow Survey 2016: Data Guide. [Technical Report]. Available at http://dmg.utoronto.ca/pdf/tts/2016/2016TTS_DataGuide.pdf.

Data Management Group [University of Toronto]. (n.d.). Reports. Available at http://dmg.utoronto.ca/transportation-tomorrow-survey/tts-reports.

Miller, E. J., Habib, K. M. N., Shalaby, A. (2019). UTTRI Technical Support for the City of Toronto Vehicle for Hire Bylaw Review. Project Work Plan.

Travel Modelling Group [University of Toronto]. (2019). GTAModel V4.0 – Network Assignment

APPENDIX

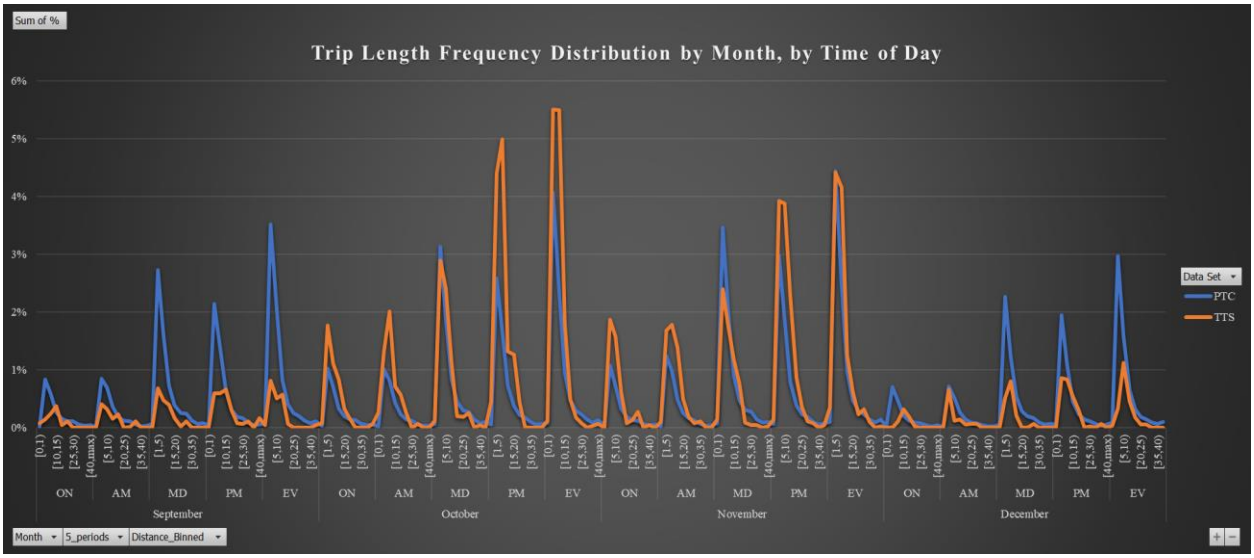

Figure A1. Trip Length Frequency Distribution by Month, by Time of Day

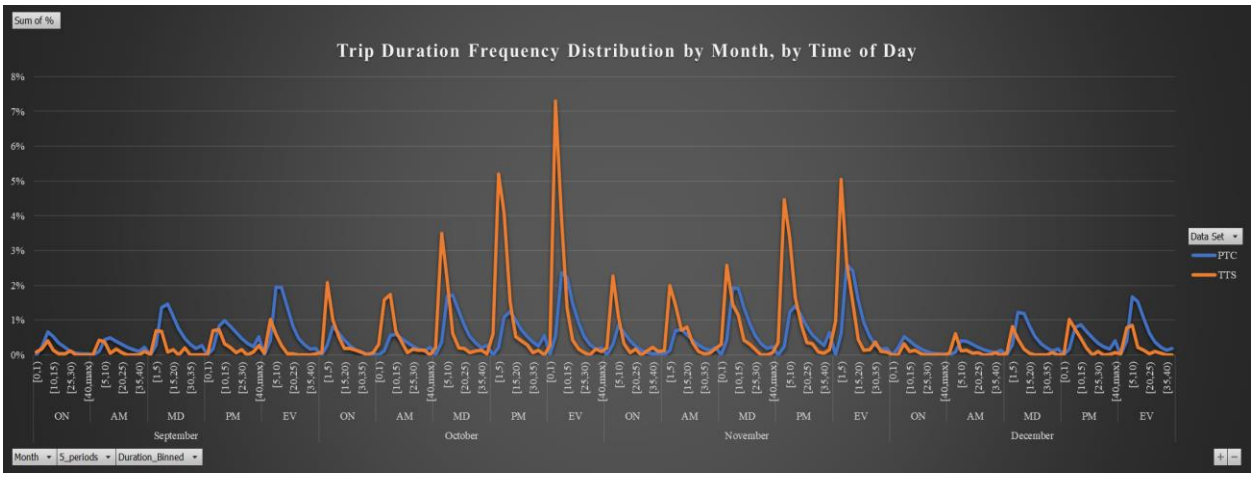

Figure A2. Trip Duration Frequency Distribution by Month, by Time of Day

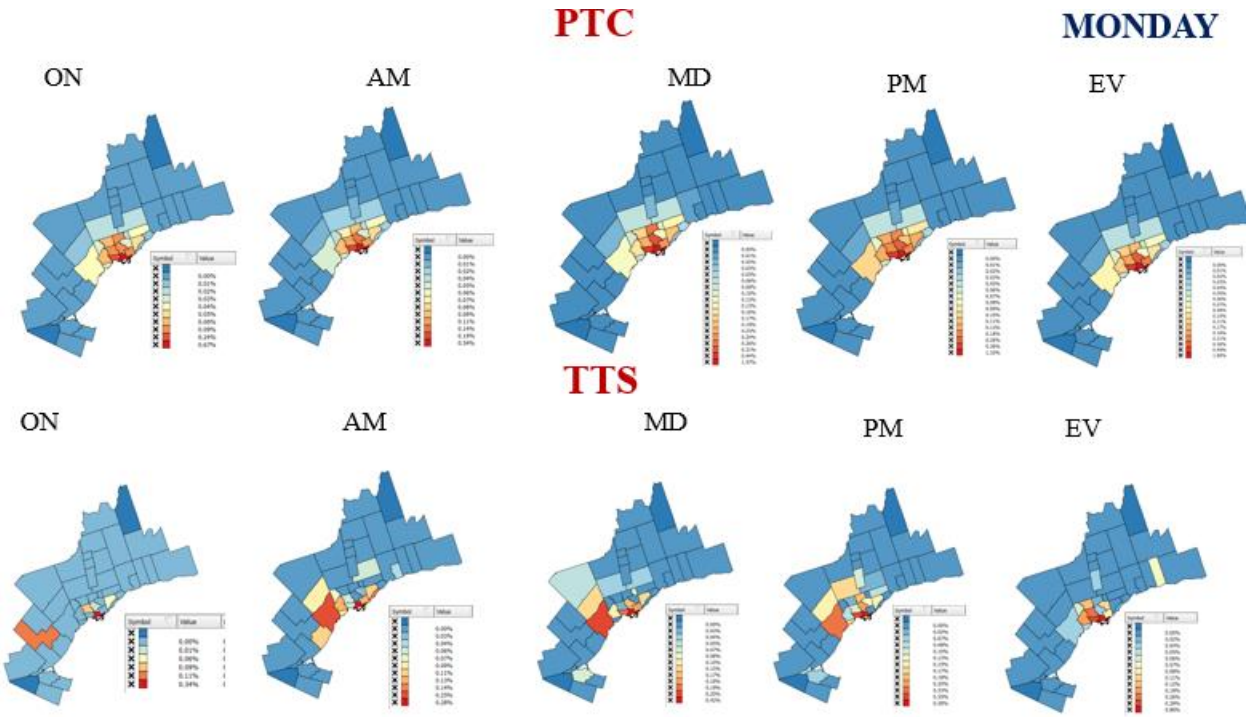

Figure A3. Spatial Distribution of Daily Trip Origins by Time of Day – Monday

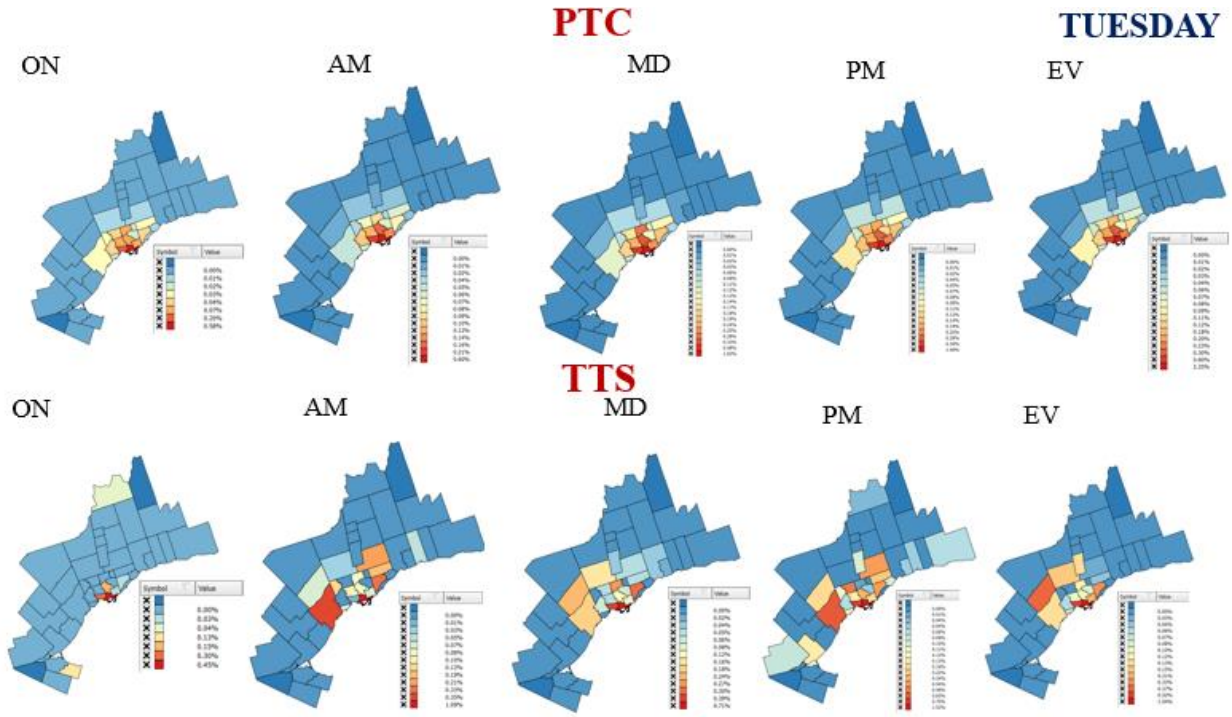

Figure A4. Spatial Distribution of Daily Trip Origins by Time of Day – Tuesday

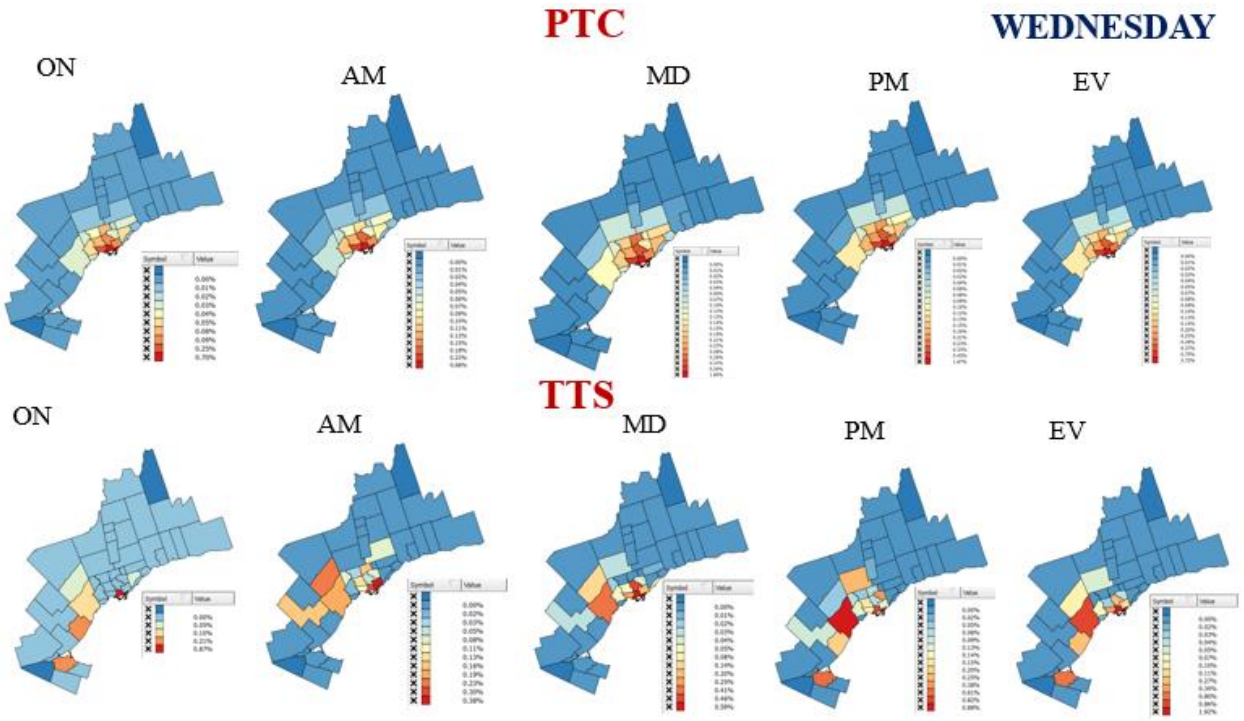

Figure A5. Spatial Distribution of Daily Trip Origins by Time of Day – Wednesday

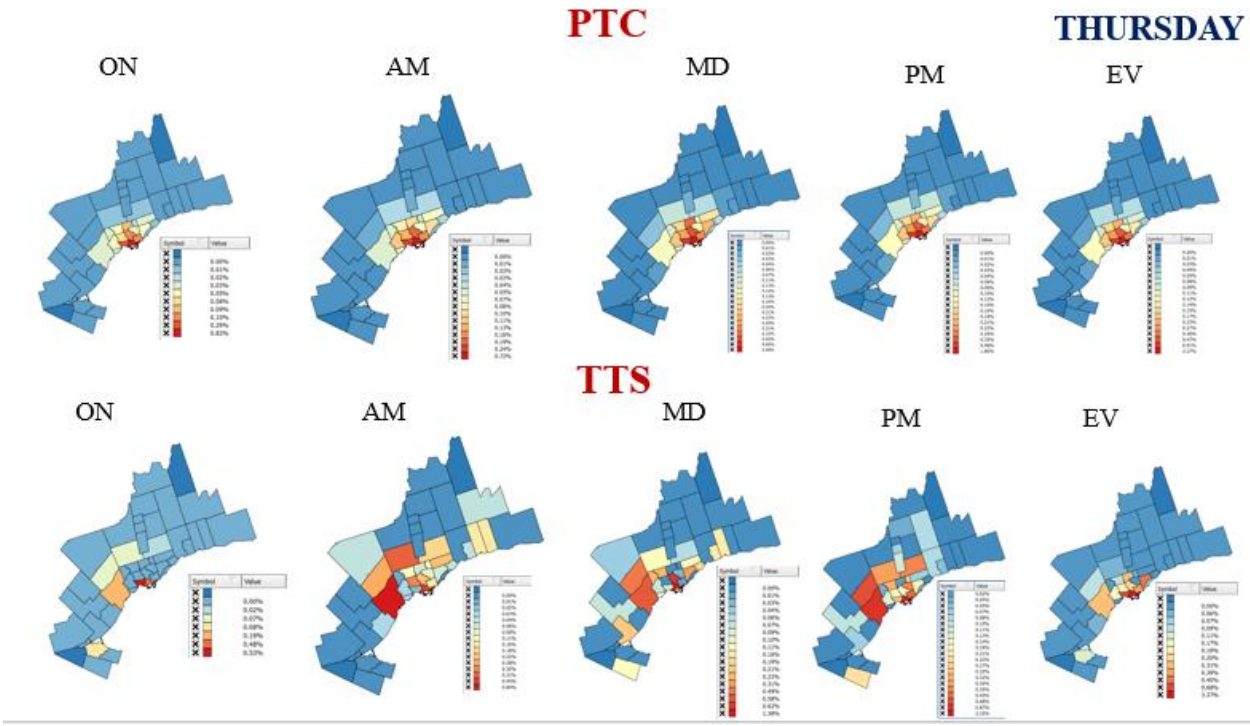

Figure A6. Spatial Distribution of Daily Trip Origins by Time of Day – Thursday

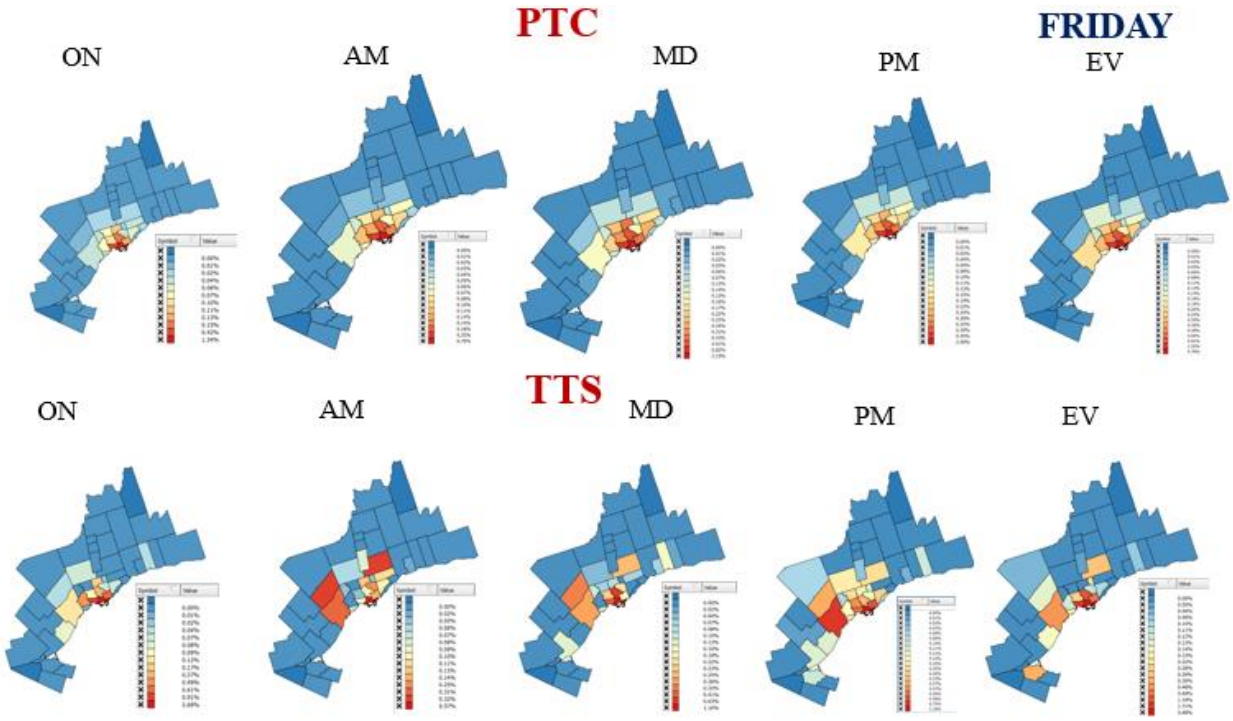

Figure A7. Spatial Distribution of Daily Trip Origins by Time of Day – Friday