ETC データを用いた
首都高速道路における交通行動の日変動特性の分析
Analysis of Daily Variability Characteristics of Travel Behavior on Tokyo Metropolitan Expressway using ETC Data

ティラタナパコン タウィン*・大口 敬**・田中 伸治***
洪 性 俊**・割田 博****
Tawin TIRATANAPAKHOM, Takashi OGUCHI, Shinji TANAKA, Sungjoon HONG and Hiroshi WARITA

要 旨
的确な道路交通の管理・運用のためにはドライバの交通行動の変動を的确に理解しなければならない。従来の研究では日変動を対象とする場合がほとんどであり、個別ドライバの交通行動までは分析できていない。ところが、日変動の相当な部分は個別ドライバの交通行動の変動によるものであることが報告されている。ETC データを利用すれば、複数日の個別ドライバの交通行動データを収集できる。本研究では 2010 〜 2011 年の首都高速道路の ETC データを利用して、交通行動の変動を個別ドライバの変動と個別ドライバ間の違いに分けて分析した。また、その長期間のデータを利用して、曜日別・月別の変動特性についても分析した。

Abstract
An understanding of variability in travel behavior is essential for both analytical and operational processes. Conventional methods rely on interpersonal variability from one-day data. However, it has been found that the intrapersonal variability may constitute a substantial proportion in the total variability. The availability of electronic toll collection (ETC) opens possibilities to collect multiday individual data. This paper uses the multiday ETC data on Tokyo Metropolitan Expressway in 2010-2011 to examine the interpersonal and intrapersonal variability of the daily travel behavior. This long term data can be used to observe the effects of the day and the month to these two types of variability.

1. Introduction

The conventional methods in analyzing and modeling travel demand and travel behavior rely mainly on the one-day data from a common weekday travel survey. With this data, only the variability in travel behavior among individuals (interpersonal) based on one day data can be obtained. Using this one-day data, in general, we can assume that the travel pattern is stable and is sufficient to represent behavioral rhythms on each analysis day. In this case, the existence of day-to-day variability in individual travel behavior (intrapersonal) is commonly ignored. In reality, however, individual travel behavior might not repeat every weekday. Ignoring this variability, the results from analytical process might provide misleading information.

The understanding of day-to-day variability in travel behavior might change the travel theory and provide efficiency in transportation planning and management. Researchers have realized the need of deeper understanding of variability in travel behavior over a span of time. For example, as presented in and , the variability in daily travel behavior is composed of two major components: interpersonal and intrapersonal variability. In , the proportion of intrapersonal variability is found to have an effect on the goodness-of-fit on the trip generation model. The study in also confirms the significance and existence of intrapersonal variability.

The components of variability, particularly, the intrapersonal variability can be measured using the multiday individual travel data (longitudinal data) only. In practice, collecting multiday data is usually expensive and difficult. For example, in traditional travel survey, the respondents might not be willing to
input their travel information for several consecutive days. Therefore, the valid data for analysis usually contains small sample size within the short span of time. Using advanced technology, such as GPS device, offers the possibility to collect multiday travel information more accurately\(^7\). However, the GPS-based method only captures travel information of the participants who volunteer to carry the device. In some cases, particularly for the handy GPS unit, participants might not carry or operate the device every day.

In Tokyo Metropolitan Expressway (MEX), the Electronic Toll Collection (ETC) system is widely implemented throughout the network. Moreover, more than 80% of expressway users pay their toll using this system\(^6\). The ETC system collects toll automatically when vehicles pass the toll gate. At the same time, the system records the passing time, toll gate location, and vehicle ID. In addition, the same vehicle ID will be recognized when the same vehicle re-enters the system. With this characteristic, we can observe the multiday and multi-period travel behavior for the same vehicle on urban expressway from ETC data.

This study shows the use of ETC data to examine the components of variability in travel behavior. The analysis focuses on the day-to-day variability in the number of journeys made by each vehicle between the selected on- and off-ramp sites on metropolitan expressways in Tokyo.

### 2. Field data

In this paper, the travel information from ETC data of urban expressway in Tokyo was used. The selected study area serves inbound traffic between the west and central part of Tokyo. Between these locations, ETC data can be obtained from the entry gate at Yoga toll station and the exit gate at Kasumigaseki toll station. The individual daily journey frequency is extracted from multiday ETC data collected between August 2010 and March 2011. The analysis mainly considers the travel information on common workdays (163 days) of small and medium vehicles classified based on MEX vehicle payment type.

Between the on- and off-ramp sites, around 2,000 vehicles have been observed daily. In addition, more than 100,000 IDs were recorded travelling along this route. Among these IDs, some vehicles have been found to travel only few times during the observation period (e.g. 1-2 times in eight months). In the analysis, to maintain the same data set, the sample was selected based on the number of the vehicle's appearance in each month. A vehicle will be included in the analysis, if it appears to travel in the selected study area every month (i.e. at least one time in a month). As a result, the number of vehicles considered in this paper is 2,362 vehicles.

For the whole observation period, the total number of journeys per vehicle in each month constitutes various patterns of individual travel behavior. From this characteristic, users can be split into different groups by using the clustering method.

The dendrogram in Figure 1 provides the image of clustering results. To choose the number of clusters, using a simple approach, the dendrogram is cut at the large changes in distances. As a study case, the sample was partitioned into three clusters by cutting across the branches at the distance of around 300. The results from clustering method provide the composition of clusters as summarized in Table 1.

![Figure 1. Clustering results](image)

### Table 1. Summary of the clustering results

<table>
<thead>
<tr>
<th>Cluster</th>
<th>N</th>
<th>%N</th>
<th>Ave. number of travel day [%]</th>
<th>Mean</th>
<th>SD</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1,795</td>
<td>75.99</td>
<td>15.13</td>
<td>0.157</td>
<td>0.381</td>
<td>3</td>
</tr>
<tr>
<td>C2</td>
<td>376</td>
<td>15.92</td>
<td>40.32</td>
<td>0.418</td>
<td>0.524</td>
<td>4</td>
</tr>
<tr>
<td>C3</td>
<td>191</td>
<td>8.09</td>
<td>75.92</td>
<td>0.817</td>
<td>0.517</td>
<td>3</td>
</tr>
<tr>
<td>All</td>
<td>2,362</td>
<td>100</td>
<td>24.05</td>
<td>0.252</td>
<td>0.461</td>
<td>4</td>
</tr>
</tbody>
</table>

As shown in Table 1, cluster 1 (C1) constitutes the majority of sample size while clusters 2 (C2) and 3 (C3) make up a medium and small group of this sample, respectively.

Conversely, the number of days used and the daily journey rate indicate that C3 travels along this route more frequent than C2 and C1. This data provides travel information in one traffic direction. In general, the observed maximum number of daily journeys is less than four journeys per day.

From the clustering results, the analysis can be made by comparing the variability in travel behavior of different groups of users. Based on the number of journeys that each individual travels during the observation period, we define C1 as infrequent user, and C2 and C3 as moderate and frequent users, respectively.

The table also shows the information of total variability in daily journey rate of each cluster through the standard deviation (SD). In this case, drivers in category C1 appears to have low
level of variability while variability in C3 is higher than C1 but slightly lower than C2. The components of this total variability will be examined in section 4.

3. Methodology

As discussed in\(^1\) and\(^2\), interpersonal and intrapersonal variability constitute the total variability in travel behavior. Different individuals have different needs and desires for travel. This can be considered as interpersonal variability. On the other hand, the need for travel of each individual also varies from day to day. This type of variability can be considered as intrapersonal variability.

Day-to-day variability in travel behavior can be observed in various travel behavior measures such as trip frequency, travel time, departure/arrival time, etc. The present study focuses only on the day-to-day variability in the number of journeys made by each vehicle. The daily number of journeys in this study is defined as the total number of movements of one particular vehicle from entry to exit gate in the study area within one day.

The components of variability in travel behavior can be measured and quantified following the methodology presented in\(^1\) and\(^2\). The idea of this method is somewhat similar to the calculation used in analysis of variance (ANOVA). Each component of variability is quantified by calculating the sum of squares. The total variability (TSS) can be partitioned into two major components as presented in equation (1).

\[
TSS = BPSS + WPSS
\]

where TSS is the total sum of squares representing the total variability while BPSS and WPSS represent the interpersonal or between-person and intrapersonal or within-person variability, respectively.

From equation (1), the total sum of squares TSS is calculated by subtracting each individual travel behavior measure (i.e. daily journey frequency) from the grand mean as shown in equation (2).

\[
TSS = \sum_i \sum_k (t_{ik} - \bar{t})^2
\]

where \(t_{ik}\) is the number of journeys made by vehicle \(i\) on day \(k\), and \(\bar{t}\) is the grand mean or overall mean number of journeys made per vehicle per day.

The BPSS and WPSS are defined in equations (3) and (4), respectively.

\[
BPSS = \sum_i K (\bar{\ell}_i - \bar{\ell})^2
\]

\[
WPSS = \sum_i \sum_k (t_{ik} - \bar{\ell}_i)^2
\]

where \(\bar{\ell}_i\) is the mean number of journeys that vehicle \(i\) made per day and \(K\) is the number of days in the observation period.

4. Empirical results

This section presents the empirical results of measuring the components of day-to-day variability and investigates the effects of day and of month on the change in level of variability.

In the latter part of this section, the level of contribution of each cluster to the components of variability will be examined.

4.1 Day of the week effect on level of variability

Table 2 illustrates the components of variability according to the different types of day in a week. Since the sample sizes (\(N\)) are different in each case, to make a comparison, the sum of squares value of each case are divided by \(N\).

<table>
<thead>
<tr>
<th>Day</th>
<th>(N)</th>
<th>Number of travel days [%]</th>
<th>TSS</th>
<th>BPSS</th>
<th>WPSS</th>
<th>(\frac{WPSS}{TSS}) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>1,941</td>
<td>26.03</td>
<td>5.80</td>
<td>2.12</td>
<td>3.68</td>
<td>63.44</td>
</tr>
<tr>
<td>Tuesday</td>
<td>2,622</td>
<td>24.90</td>
<td>6.35</td>
<td>1.52</td>
<td>4.83</td>
<td>76.02</td>
</tr>
<tr>
<td>Wednesday</td>
<td>2,295</td>
<td>26.33</td>
<td>6.68</td>
<td>1.44</td>
<td>5.24</td>
<td>78.39</td>
</tr>
<tr>
<td>Thursday</td>
<td>2,288</td>
<td>25.82</td>
<td>6.33</td>
<td>1.34</td>
<td>4.99</td>
<td>78.86</td>
</tr>
<tr>
<td>Friday</td>
<td>2,271</td>
<td>25.36</td>
<td>6.25</td>
<td>1.28</td>
<td>4.97</td>
<td>79.48</td>
</tr>
<tr>
<td>Saturday</td>
<td>1,602</td>
<td>15.49</td>
<td>4.66</td>
<td>0.50</td>
<td>4.16</td>
<td>89.20</td>
</tr>
<tr>
<td>Sunday</td>
<td>617</td>
<td>10.05</td>
<td>3.00</td>
<td>0.55</td>
<td>2.45</td>
<td>81.82</td>
</tr>
</tbody>
</table>

For weekdays, overall, the level of total variability is approximately the same on Tuesday, Thursday and Friday. The highest variability is observed on Wednesday while the lowest is found on Monday.

The level of intrapersonal variability (WPSS) is generally larger than the interpersonal variability (BPSS). The percentage of WPSS is about 63 percent of TSS in the case of Monday. For other common weekdays, the levels of WPSS reach up to 78 percent. Moreover, it is observed that the WPSS gradually increase from Tuesday to Friday (76 to about 79 percent). This indicates that drivers are more regular in making journeys on Monday. In addition, the drivers tend to reduce their regularity in making journey from beginning to the end of the weekday period.

The lower travel demand is observed during weekends. Consequently, the amount of TSS is found to be lower than on weekdays. On Sunday, particularly, the level of total variability drops by more than half of the weekdays. The high percentage of WPSS on weekends indicates low level of regularity. This might be due to the small number of working trips observed.
4.2 The effect of month to level of variability

Using methodology described earlier, the results in Table 3 compare daily variability in journey frequency from month to month over the eight-month period. This makes possibility to observe the seasonal effect on the change in level of variability.

Overall, the percentages of WPSS in the daily journey rate of all groups account for around 75 percent of the total variability. The levels of WPSS in August and December are higher than other months. This may be because August is commonly the summer vacation for school students in Japan and December is the year end season. From these results it can be implied that during these two seasons drivers are more likely to change their regular travel patterns.

Furthermore, it can be observed that the shares of WPSS gradually increase from September to November. After New Year season, the share of WPSS drops to around 73 percent and maintains almost the similar level until the end of observation period.

This means that drivers make more regular journeys during the beginning of the year. On the opposite, they tend to decrease their journey regularity during the end of the year.

Table 3. Interpersonal and intrapersonal variability for different types of month

<table>
<thead>
<tr>
<th>Month</th>
<th>TSS</th>
<th>BPSS</th>
<th>WPSS</th>
<th>WPSS [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>August</td>
<td>9.55</td>
<td>2.10</td>
<td>7.45</td>
<td>78.03</td>
</tr>
<tr>
<td>September</td>
<td>10.20</td>
<td>2.57</td>
<td>7.63</td>
<td>74.79</td>
</tr>
<tr>
<td>October</td>
<td>10.61</td>
<td>2.61</td>
<td>8.00</td>
<td>75.37</td>
</tr>
<tr>
<td>November</td>
<td>9.95</td>
<td>2.39</td>
<td>7.56</td>
<td>76.00</td>
</tr>
<tr>
<td>December</td>
<td>10.76</td>
<td>2.27</td>
<td>8.49</td>
<td>78.91</td>
</tr>
<tr>
<td>January</td>
<td>9.88</td>
<td>2.66</td>
<td>7.22</td>
<td>73.06</td>
</tr>
<tr>
<td>February</td>
<td>10.07</td>
<td>2.68</td>
<td>7.39</td>
<td>73.35</td>
</tr>
<tr>
<td>March</td>
<td>10.79</td>
<td>2.95</td>
<td>7.84</td>
<td>72.64</td>
</tr>
</tbody>
</table>

Table 4. Interpersonal and intrapersonal variability of each cluster group

<table>
<thead>
<tr>
<th>Cluster</th>
<th>N</th>
<th>TSS</th>
<th>BPSS</th>
<th>WPSS</th>
<th>WPSS [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1795</td>
<td>45.15</td>
<td>3.54</td>
<td>41.61</td>
<td>50.78</td>
</tr>
<tr>
<td>C2</td>
<td>376</td>
<td>18.54</td>
<td>2.24</td>
<td>16.30</td>
<td>19.89</td>
</tr>
<tr>
<td>C3</td>
<td>191</td>
<td>18.25</td>
<td>10.87</td>
<td>7.38</td>
<td>9.01</td>
</tr>
<tr>
<td>All</td>
<td>2362</td>
<td>81.94</td>
<td>16.65</td>
<td>65.29</td>
<td>79.68</td>
</tr>
</tbody>
</table>

Table 5. Average interpersonal and intrapersonal variability of each cluster group

<table>
<thead>
<tr>
<th>Cluster</th>
<th>N</th>
<th>TSS</th>
<th>BPSS</th>
<th>WPSS</th>
<th>WPSS [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1795</td>
<td>25.15</td>
<td>1.97</td>
<td>23.18</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>376</td>
<td>49.30</td>
<td>5.96</td>
<td>43.34</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>191</td>
<td>95.55</td>
<td>56.91</td>
<td>38.64</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>2362</td>
<td>34.69</td>
<td>7.05</td>
<td>27.64</td>
<td></td>
</tr>
</tbody>
</table>

4.3 Variability of each cluster group

To examine how each cluster contributes to each component of variability, the sum of squares terms in equations (2) through (4) are distributed into three portions as shown in Equation (5), (6) and (7).

$$TSS = \sum_{i\in C1} \sum_{k} (t_{ik} - \bar{t})^2 + \sum_{i\in C2} \sum_{k} (t_{ik} - \bar{t})^2 + \sum_{i\in C3} \sum_{k} (t_{ik} - \bar{t})^2 +$$

$$BPSS = \sum_{i\in C1} K(\bar{t}_i - \bar{t})^2 + \sum_{i\in C2} K(\bar{t}_i - \bar{t})^2 +$$

$$WPSS = \sum_{i\in C1} \sum_{k} (t_{ik} - \bar{t})^2 + \sum_{i\in C2} \sum_{k} (t_{ik} - \bar{t})^2 + \sum_{i\in C3} \sum_{k} (t_{ik} - \bar{t})^2$$

Using the entire eight-month travel information to calculate each type of variability, the results in Table 4 show that around 50 percent of the overall total variability is devoted to WPSS for infrequent users (C1). In addition, the share of WPSS for frequent users (C3) is minimal while for the moderate users (C2), it is between C1 and C3.

Considering the sample size of each cluster, it was observed that the larger the sample size, the higher the amount of WPSS. This means that WPSS accumulates with increasing number of samples. Conversely, the amount of BPSS of C1 is found to be significantly lower than C3 but slightly larger than C2. This indicates that although infrequent users constitute the large portion of WPSS, on average, the different behaviors across drivers are not significant. This is reasonable since the large number of non-travel behavior of infrequent users represents the similar behavior across drivers.

Further comparison can be made by dividing each sum of square value by the sample size. Table 5 shows that on average...
moderate users (C2) generate the largest amount of WPSS per individual. As can be expected, infrequent users make no journey regularly while frequent users are more likely to make the similar journey patterns. Hence, both C1 and C3 generate lower levels of WPSS than C2.

4.4 Individual intrapersonal variability and daily journey rate

To examine the relationship between the mean and variance of intrapersonal variability in the rate of journeys, the intrapersonal variability of each individual can be quantified by using the basic formula similar to the calculation of variance as shown in equation (8).

$$s_i^2 = \frac{1}{K-1} \sum_{k=1}^{K} (t_{ik} - \bar{t}_i)^2$$

where $s_i^2$ represents the intrapersonal variability in journey frequency made by vehicle $i$.

The scatter plot in Figure 2 shows the relationship between intrapersonal variability and average journey rate. Lower intrapersonal variability can be observed from a person who travels with low trip rate and a person who travels regularly everyday (e.g. one journey per day).

As expected, low level of variability can be observed among infrequent users (C1) who make no journey in most days and frequent users (C3) who travel the journey rate of once a day. On the other hand, moderate users (C2) who travel around 0.5 journeys per day and those who travel more than one journey per day tend to have higher level of intrapersonal variability. This tendency is in line with the results shown in Table 5.

5. Conclusion

This paper focuses on the examination of day-to-day variability in daily journey rate using the multiday ETC data collected between Yoga and Kasumigaseki inbound direction on Tokyo metropolitan expressway in 2010-2011.

The paper provides several information of variability in journey rates. In general, the results indicate that intrapersonal variability accounts for a substantial portion of total variability. Investigations in types of day and of month reveal that drivers are more regular in making their journey during the beginning of the week (e.g. Monday) and the beginning of the year.

The results also report that in the observation period, during August and December, drivers are likely to vary their behavior in making journeys more than other months.

Grouping drivers into three clusters, the results report that infrequent users constitute the major source of variability since their population is greater than the other groups. However, the levels of intrapersonal variability per individual indicate that moderate users are likely to have greater intrapersonal variability than the other groups.

In conclusion, the paper confirms the existence of intrapersonal variability in trip making behavior. In addition, it was also found that the multiday ETC data can be used to obtain the knowledge of variability in personal level, particularly for long term period.

6. Acknowledgement

The authors gratefully acknowledge Metropolitan Expressway Company Limited in Tokyo for providing the data used in the research.

(Manuscript received. January 22, 2013)

References

6) www.shutoko.jp/english/technology/operationcontrol
7) http://www.fhwa.dot.gov/ohim/gps/importance.htm