Frequency Distribution of Leisure Travel by the Japanese: The Past and Future

Hiromichi YAMAGUCHI a and Makoto OKUMURA b

a JSPS research fellow, Graduate School of Engineering, Tohoku University, Sendai, 980-0845, Japan; E-mail: h-ymgc@plan.civil.tohoku.ac.jp
b International Research Institute of Disaster Science, Tohoku University, Sendai, 980-0845, Japan; E-mail: mokmr@m.tohoku.ac.jp

Abstract: How has the frequency distribution of domestic overnight leisure travel changed in the past few decades? What will the distribution be in the next several decades? In order to answer these questions, this paper analyzes the time series change of travel frequency distribution for the Japanese. We conclude that (1) the average annual travel frequency decreased for 20 years, (2) the zero frequency ratio is increasing, the difference of travel frequency between individuals is also increasing, and (3) these time series changes can be explained by the changes in age composition and differences among birth cohort. Moreover, we forecast the travel frequency distribution until 2060, considering changes in age composition. The results indicates that: (4) by 2060, the total overnight travel volume will decrease to half of that in 2010, (5) the difference between individuals will continue to increase, as a result zero frequency ratio will be 65% by 2060.

Keywords: Population Aging, Travel Frequency, Cohort Model, Travel Demand Forecasting

1. INTRODUCTION

Japan has been experiencing severe population aging in the past few decades, and the scenario is likely to worsen in the next several decades. Moreover, its total population has started to decline. These changes in the country’s population composition are likely to affect its travel generation patterns. It is important to understand and forecast these changes in travel generation patterns in order to maintain or improve intercity travel service levels.

We focus on frequency distribution, a feature of travel generation pattern. This distribution provides important information when considering strategies for demand stimulation. For example, decreasing travel volumes on account of an increasing zero frequency ratio, imply the need to improve services for zero or low frequency travelers. Then, we need to analyze how and why the travel frequency distribution of the Japanese has changed over the past few decades and how distribution change will change in the future. This paper attempts to answer these questions.

Most past research analyzing Japanese domestic leisure travel across many periods focused only on the average number of travelers or total travel volumes (e.g., SUGANUMA et al., 2011 and TSUKAI et al., 2010). Furuya and Zeng (2014) offered a novel prospective in that they analyzed the travel frequency distribution. They analyzed the cross-sectional travel frequency distribution data using the latent class model. They concluded that the demand function could be classified to 3 classes by travel frequency, the ratio of the zero frequency class is almost half that of all individual attribute groups, and the effects of attributes (e.g., income, age and so on) on travel frequency differ among classes. However, their methodology is difficult to adopt when considering a time series analysis because of the features of the latent class model.
Moreover, it is difficult to obtain the long-term data with detailed individual attributes.

Therefore, we propose a simple travel frequency distribution model, which considers the features of travel frequency distribution for analyzing long-term changes. Then, we apply the model to the repeated cross-sectional data of travel frequency distribution for the past 20 years. In particular, we focus on the effects of age composition and birth cohort changes. For estimating these effects we adopt the age/period/birth cohort (A/P/C) model to the three parameters of the travel frequency distribution model. Additionally, future travel frequency distributions are forecast considering changes in the age composition and population decline.

The results of the analyses indicate that (1) the average annual overnight leisure travel frequency decreased for 20 years, (2) the difference of travel frequency between individuals (referred to as “inequality”) is increasing, and (3) these time-series changes can be explained by changes in age composition and differences among birth cohorts. We also forecast the travel frequency distribution until 2060. We arrive at the following conclusions: (4) the total overnight travel volume will decrease to half of that in 2010 by 2060, and (5) the inequality and zero frequency ratio continue to increase, such that the latter will be 65% in 2060.

The rest of this paper is organized as follows. Section 2 presents the travel frequency distribution data and provides an overview of time series changes in the aggregate travel frequency distribution. Section 3 explains the travel frequency distribution model with A/P/C effects and the estimation methodology. Section 4 presents the estimated results of the A/P/C effects. Finally, Section 5 explains the methodology used for the forecasting exercise and presents future travel frequency distributions. Section 6 concludes.

2. PREVIOUS AGGREGATE TRAVEL FREQUENCY DISTRIBUTION

This section presents the data for leisure travel frequency and analyzes time series changes in the aggregated distribution. We use the frequency data of overnight leisure travel over 5-year intervals (1991, 1996, 2001, 2006 and 2011). These data are sourced from the National Survey on Time Use and Leisure Activities of Japan, a questionnaire with large-scale samples (approximately 200,000 per period). This survey aggregates annual travel frequency and age group into 9 and 15 groups, respectively (see Tables 1 and 2, respectively).

Figure 1 presents the travel frequency distribution data aggregated for all-age groups for 2011. Figure 1 indicates two features of this travel frequency distribution. First, the component ratios of higher frequencies are basically smaller than that of lower frequencies. Second one, the component ratios of “0” and “over 10-” appear inflated than the decreasing law of others. These features are discussed in detail in the next section. In this section, we discuss the time series changes from two angles: average annual number of travelers and difference in travel frequency (as seen by the Lorenz curve).

First, we consider the time-series change in the average number of travelers. Figure 2 indicates that the average number of travelers decreased from 1.6 to less than 1.2 within a span of 20 years. The average number of travelers $AV_p$ is derived from the following equation:

$$AV_y = \sum_{k \in \{0,9\}} kq_y(k) + 10q_y(\text{"over}10")$$  \hspace{1cm} (1)

where,

$q_y(k)$: component ratio of frequency $k$ in year $y$.
As expressed in equation (1), we assume that the travel frequency of group “over 10” is 10 times/year. Because of data limitation, we can’t discuss about the detail of this group. The component ratio of this group “over 10” is only about 2%, therefore this assumption is not so large effects to following discussions.

Second, Figure 3 presents the Lorenz curves for leisure travel. Point A indicates that the bottom 70% of the population undertook 10% of the total leisure travel in 2011. In other words, only the upper 30% of the population was responsible for 90% of the total leisure travel in 2011. The time series changes in Lorenz curves are apparent: as the curves shift downward and to the right. This shift indicates that the difference (inequality) in travel frequencies has been increasing. Returning to the detail of travel frequency distribution, Figure 4 indicates that increasing difference may be attributed to the increase of zero frequency ratio. In the following sections, we analyze these changes in the time series of the travel frequency distribution using the A/P/C model and we also estimate future time series changes.

<table>
<thead>
<tr>
<th>Frequency (group no. k)</th>
<th>Frequency (times/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>6 or 7</td>
</tr>
<tr>
<td>8</td>
<td>8 or 9</td>
</tr>
<tr>
<td>9</td>
<td>10-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1. Classification by frequency</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Age (group no. a)</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>85-</td>
</tr>
<tr>
<td>2</td>
<td>80-84</td>
</tr>
<tr>
<td>3</td>
<td>75-79</td>
</tr>
<tr>
<td>4</td>
<td>70-74</td>
</tr>
<tr>
<td>5</td>
<td>65-69</td>
</tr>
<tr>
<td>6</td>
<td>60-64</td>
</tr>
<tr>
<td>7</td>
<td>55-59</td>
</tr>
<tr>
<td>8</td>
<td>50-54</td>
</tr>
<tr>
<td>9</td>
<td>45-49</td>
</tr>
<tr>
<td>10</td>
<td>40-44</td>
</tr>
<tr>
<td>11</td>
<td>35-39</td>
</tr>
<tr>
<td>12</td>
<td>30-34</td>
</tr>
<tr>
<td>13</td>
<td>25-29</td>
</tr>
<tr>
<td>14</td>
<td>20-24</td>
</tr>
<tr>
<td>15</td>
<td>15-19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Classification by age group</th>
<th></th>
</tr>
</thead>
</table>
Figure 1. Aggregated data of travel frequency distribution in 2011

Figure 2. Average number of annual travelers

Figure 3. Lorenz curves of leisure travel
3. TRAVEL FREQUENCY DISTRIBUTION MODEL WITH AGE/PERIOD/BIRTH COHORT EFFECTS

3.1 Travel Frequency Distribution Model

In order to analyze the time series changes of travel frequency distribution, we propose a travel frequency distribution model with a small number of parameters. As discussed in section 2, the travel frequency distribution for 2011 showed two features: (1) a monotonic decrease in count data, and (2) inflated data for “0” and “over 10”. In order to express these features, we apply the exponential-mixed Poisson model with zero and high-value inflated parameters.

We set the following two assumptions for this mixed distribution. First we assume that the annual travel frequency \( k_i \) follows the Poisson distribution, which is formulated as follows:

\[
p(k_i | \lambda_i) = \frac{\lambda_i^k}{k!} e^{-\lambda_i},
\]

where, \( \lambda_i \) : expected travel frequency (continuous variable) of individual \( i \).

Second, we assume that the expected travel frequency \( \lambda_i \) differs among individuals and follows the exponential distribution, which is formulated as follows:

\[
p(\lambda | \mu) = \frac{1}{\mu} e^{-\frac{\lambda}{\mu}},
\]

where, \( \mu \) : parameter of exponential distribution (expected value of \( \lambda_i \)).

Mixing the two distributions, the observed ratio of a sample with frequency \( k \) is formulated
as follows:

\[ p(k \mid \mu) = \int p(k \mid \lambda) p(\lambda \mid \mu) d\lambda = \frac{\mu^k}{(\mu + 1)^{k+1}}. \] (4)

Equation (4) indicates that the exponential and Poisson mixed model assumes the following relation:

\[ \frac{p(k + 1 \mid \mu)}{p(k \mid \mu)} = \frac{p(k + 2 \mid \mu)}{p(k + 1 \mid \mu)} = \frac{\mu}{\mu + 1} = s(\text{constant}) \quad \forall k \geq 0. \] (5)

The validity of this assumption for the travel frequency data is verified in section 3.3.

In order to explain the zero and high-value inflated data, we use the “with zero model”, which was proposed by Mullahy (1986) for analyzing the zero-inflated data. Thus this model includes a parameter for explaining inflated zero data. Following this approach, we add two parameters to the exponential-mixed Poisson model: the zero-inflated parameter \( z \) and the “over 10”-inflated parameter \( h \). As a result, the travel frequency distribution model is formulated as follows:

\[ p(k \mid s, z, h) = \begin{cases} 
(z + (1-s)(1-z))(1-h) & k = 0 \\
(1-s)s^k (1-z)(1-h) & 0 < k < 10 \\
h + (1-h)(1-z)\int_{10}^{\infty} (1-s)s^k dk & k = "over10"
\end{cases}. \] (6)

This model includes only three parameters: decline ratio \( s \), zero-inflated parameter \( z \) and “over 10”-inflated parameter \( h \). In sub-section 3.2, we formulate the model for analyzing the differences in above-mentioned parameters between age groups, periods, and birth cohorts.

### 3.2 A/P/C model

In this sub-section, we formulate a model for dividing the time series changes in the travel frequency distribution into three effects: age, period and birth cohort. This model is called the A/P/C model of cohort model. In this model, we assume that \( s_a, z_y, h_c \) parameters which characterize the \( a \)-th age group in the \( y \)-th survey period, can be decomposed into the following form:

\[ s_a = s_0 + s_a^A + s_y^P + s_c^C, \quad \sum_a s_a^A = \sum_y s_y^P = \sum_c s_c^C = 0 \]

\[ z_y = z_0 + z_a^A + z_y^P + z_c^C, \quad \sum_a z_a^A = \sum_y z_y^P = \sum_c z_c^C = 0 \]

\[ h_c = h_0 + h_a^A + h_y^P + h_c^C, \quad \sum_a h_a^A = \sum_y h_y^P = \sum_c h_c^C = 0 \] (7)

where \( s_0 \) is the grand mean effect, \( s_a^A \) is the effect of age-group \( a \), \( s_y^P \) is the effect of survey period \( y \), and \( s_c^C \) is the effect of birth cohort \( c \).

When we estimate these three effects, there is an identification problem in that these three effects cannot be separated without some prior information (Mason & Fienberg
1985). Rogers (1982), Nakamura (1986), and Yang and Land (2006) proposed some solutions to overcome this problem. We adopt the classical approach of Rogers (1982). As we need to assume more than one constraint for some effects, we assume equal constraints for old and later birth cohorts (see Tables 3 and 4). As the data for elders over 85 years old are aggregated in our case, we cannot separate the effects of this birth cohort. The birth cohort effects of later cohorts is the most important for the forecasting. Therefore, it is better to conduct the estimation by aggregating the data from several periods for certain birth cohorts.

\[
(s^*, z^*, h^*) = \arg \max_{(s,z,h)} \log(L(s,z,h)) = \arg \max_{(s,z,h)} \left\{ \sum \log \left( \frac{p(k_i | s,z,h)}{L_p k \sum_s} \right) \right\}. 
\]

We solve equation (8) using the methodology proposed by Lambert (1992). Lambert (1992) adopted an expectation-maximization (EM) algorithm technique to estimate a zero-inflated Poisson model. In this algorithm, we repeat the alternate estimation of inflated parameters \(z, h\) (M-step) and decline ratio \(s\) (parameters of the exponential distribution \(\mu\), as shown in equation (5); E-step).

### 3.3 Estimation Methodology

We estimate all parameters simultaneously using the Maximum Likelihood Estimation (MLE) as follows:

\[
(s^*, z^*, h^*) = \arg \max_{(s,z,h)} \log(L(s,z,h)) = \arg \max_{(s,z,h)} \left\{ \sum \log \left( \frac{p(k_i | s,z,h)}{L_p k \sum_s} \right) \right\}. 
\]

### 3.4 Fitness of the Model

In this sub-section, we discuss the fitness of the model used to analyze the travel frequency data. First, we discuss the fitness of the travel frequency distribution model (equation (6)) used to analyze the distribution of each period and age group. Second, we refer to the fitness of the A/P/C model.
In order to check the first type of fitness, we estimate the parameters of all survey periods and age groups \((s_{ay}^*, z_{ay}^*, h_{ay}^*)\). The fitness is evaluated using the coefficient of determination:

\[
R^2_{ay} = \frac{\sum_{k \in K} (q_{ay}(k) - p(k | s_{ay}^*, z_{ay}^*, h_{ay}^*))^2}{\text{Var}(q_{ay}(k \in K))}.
\]  

Some statistical indexes of the estimated values of \(R^2_{ay}\) are shown in Table 5. The values of the mean and S.D. in Table 5 indicate that \(R^2_{ay}\) (except minimum data) is around 0.89. This confirms the goodness of fit of the model used for the observed travel frequency distribution namely, equation (5), and the validity of the associated assumption (see section 3.1). Additionally, Figures 5 and 6 show both the observed and the estimated frequency distributions. These figures also verify the goodness of fit of the travel frequency distribution model, as seen by the small value for \(R^2_{ay}\) (Figure 6).

Regarding the second type of fitness, Table 6 shows the null deviance and residual deviance of the whole model. In order to discuss the fitness of the A/P/C model, the null model is defined as a model that adopts only grand mean effect \(s_0, z_0, h_0\). Table 6 indicates a low value for residual deviance compared with null deviance. This means that the differences in age and survey periods are explained by the three estimated effects.

### Table 5. Statistical indexes of coefficients of determination \(R^2_{ay}\)

<table>
<thead>
<tr>
<th>Mean</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
<th>Number of distributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.889</td>
<td>0.120</td>
<td>0.361</td>
<td>0.999</td>
<td>75</td>
</tr>
</tbody>
</table>

\(\text{Min.} = \text{“20-24,”; 1991} \quad \text{Max.} = \text{“85-,”; 2001}\)

![Figure 5. Fitness of the travel frequency distribution model using 2011 data. \(R^2=0.971\)](image)

![Figure 6. Fitness of the travel frequency distribution model using 2011, 20-24 years old data. \(R^2=0.361\)](image)

### Table 6. Fitness of A/P/C model

<table>
<thead>
<tr>
<th>residual deviance</th>
<th>null deviance</th>
<th>((b-a)/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.516 \times 10^4)</td>
<td>(4.851 \times 10^4)</td>
<td>0.687</td>
</tr>
</tbody>
</table>
4. ESTIMATED RESULTS OF AGE/PERIOD/BIRTH COHORT EFFECTS

4.1 Decline Ratio

This section, explains the estimated A/P/C effects of each three parameters. First let us analyze the effects of A/P/C on the decline ratio parameter $s$. The estimated grand mean effect of the decline ratio is $s_0 = 0.579$. Figure 7 shows the estimated parameters of $s^A, s^P, s^C$. From the definition of parameter $s$ (equations (3) and (5)), we can derive the expected travel frequency using the base exponential-mixed Poisson distribution as follows:

$$E(\lambda) = \mu = \frac{s}{1-s}, \quad 0 < s < 1. \tag{10}$$

Equation (10) indicates that the larger the decline parameter $s$, the higher the expected travel frequency $\lambda$ of the base distribution.

Figure 7(A) indicates that the decline ratios of the younger age groups are basically larger than those of the elderly. In other words, the expected travel frequency of the base distribution is large for the younger age. We observe a small peak around 60-64 years old, which can be attributed to the effect of retirement.

Figure 7(P) indicates that the effect of the period on the decline ratio is small compared to the other two effects. Figure 7(C) indicates that the decline ratio of the later birth cohort is basically smaller than that of the former birth cohort. In other words, the expected travel frequency of the base distribution continues to decrease for the birth cohort.

![Figure 7. Estimated A/P/C effect on the decline ratio](image)

4.2 Zero-Inflated Ratio

Second, we check the A/P/C effects on the zero-inflated ratio parameter $z$. The estimated grand mean effect of the zero-inflated ratio is $z_0 = 0.210$.

Figure 8 shows the estimated parameters of $z^A, z^P, z^C$. Comparing the max absolute value of the three effects, the age effect is larger than the other effects. According Figure 8(A), the zero-inflated ratios of elderly age groups are larger than those of the younger age groups. Notably, for the group older than 85 years old, the zero frequency ratio is approximately 80% (derived by $z_0 + z^{A}_{85-}$). This result can be explained by the health issues commonly experienced by this age group.
4.3 Over10-Inflated Ratio

Third, we check the A/P/C effects on the over10-inflated ratio parameter $h$. The estimated grand mean effect of the over10-inflated ratio is $h = 0.017$. This ratio is very small compared with the zero-inflated ratio.

Figure 9 shows the estimated parameters of $h^A, h^P, h^C$. The over10-inflated ratios of younger age groups (those in their 20s) are larger than the ratios for the other age group (see Figure 9(A)). Moreover, unlike the other parameters, the effect of “85-” is not so different from those for the other age groups.

The effect of the period (Figure 9(P)) indicates that the over10-inflated ratios in the 2000s were larger than the corresponding levels in the 1990s. The birth cohort effect (Figure 9(C)) is not as large compared with the other two effects.

5. FUTURE TRAVEL FREQUENCY DISTRIBUTION

5.1 Forecast Methodology

Using the estimated age and birth cohort effects, we forecast the future travel frequency distributions. We begin by forecasting the three parameters $s, z, h$ for every age groups for a future period using the A/P/C model. We assume the following for future A/P/C effects: period effects are omitted, the estimated age effects are the same for future periods, and the
effects of birth cohorts born after 1995 are assumed to be the same as those of cohort number 13 (those born between 1981 and 1995). The last assumption means that the decline in Figure 7(C) is assumed to stop in the future. Then, using the three parameters, we derive the future travel frequency distribution with equation (6).

Next, we obtain the aggregated distribution for all age groups using the forecasted population by age group. For the forecasted population data, we adopt the estimation results of the middle projection made by the National Institute of Population and Social Security Research (2012).

In this methodology, some other effects not related to the age effects or birth cohort effects are omitted for forecasting. For example, the effects of 3.11 East Japan Earthquake (2011) or several change of economic conditions are account in the period effects. Because these affect to all age-groups in certain period. But these effects are not so strong comparing age effects and birth cohort effects as shown in Figure 7-9. Therefore this forecasting result is reliable if there were no drastic change of surrounding conditions or travel trends outside of aging effect comparing the 1991 – 2011.

5.2 Forecast Results

Figure 10 presents the forecasted results of the average annual travel frequency. This figure indicates that average annual travel frequency will decrease to 0.8 (times/year) until 2060 because of age and cohort effects. Additionally, the comparison between the estimated and observed average travel frequency shows that the estimated value is quite close to the observed data up to 2011. Thus the observed time series changes can be explained by changes in the age composition and differences between birth cohorts. Needless to say, the estimated average travel frequency derived from the null model (which refers to the grand mean effects only) is constant.

Figure 11 plots the annual total travel volume. These values are derived by multiplying forecasted population and estimated average travel frequency. This means that the change in the null model result comes solely from the change in population. Therefore, the result from the null model turns to decrease in 2010. However, the forecasted volume will decrease at a faster pace earlier than the pace of population decline. As a result, the total travel volume will decrease to one-third of that in 1990 (or half of that in 2010) by 2060.

The time series changes in the component ratio of frequencies (Figure 12), show that the zero-frequency ratio will increase to 65% until 2060. As a result of the time series changes, the Lorenz curve shifts downward and to the right (Figure 13). The curve for 2060 indicates that 90% of total leisure travel will be undertaken by the upper 25% of the population in 2060.
Figure 10. Forecasted average annual travel frequency

Figure 11. Forecasted total leisure travel volume

Figure 12. Forecasted cumulative ratio of each frequency
6. CONCLUSION

In this paper, we focus on the time series changes in domestic overnight leisure travel frequency distribution in Japan. We analyze the past travel frequency distribution for 20 years and conclude that that the average annual travel frequency has decreased and the difference in travel frequency between individuals has increased. These results are mainly explained by the changes in age composition (population aging). Moreover, we forecast the travel frequency distribution until 2060 after considering forecasted population aging and decline. In conclusion, the total overnight travel volume will decrease to half of 2010 in 2060, and the zero-frequency ratio will continue to increase, till it reaches 65% in 2060.

These results suggest that Japan will experience a large decrease in domestic overnight leisure travel within the next fifty years. This will be serious problem for Japan’s tourism industries.

One countermeasure for this problem is to encourage the international travel to Japan. However this countermeasure have limitations for keeping the travel volume. The volume of international travel to Japan was only 13 million in 2014 (JNTO, 2015). In order to compensate the decrease of domestic travel by encouraging the international travels to Japan, the international travel volume should be increased to 60 million (about 5 times) for fifty years. It seems unrealistic, even if we account economic growth of Asian countries.

It is also required to encourage the domestic leisure travel and to change or ease the decreasing trend. Especially, zero travel frequency groups is important target. More than half of population are classified this group and the component ratio of this group will increase in the forecast. Therefore it is urgently required to investigate the features of zero travel frequency groups by using the more detailed dis-aggregate data.

Additionally, policy makers need to account the shrinking travel volumes and must avoid heavy investments to infrastructure assets for larger volumes of travelers. Then it may more important to invest for improving the quantities of the existent destinations and give more satisfaction to the present travelers, and encourage their repeated visits.

At last, this paper also suffers from some limitations. In future studies, it would be prudent to identify the cause of each effects by using time allocation models (as in LaMondia
et al., 2008) with more detailed data. For example, the decline in the travel frequency of the latter birth generation (observed in Figure 7C) may be explained by their increased time allocation for communication via information technology or change of disposable incomes. It would also be worthwhile to study the temporal change in the effects and to find the effective measure to change this trend.

ACKNOWLEDGEMENT

This research was supported in part by a grant from JSPS scientific research fund (No. 15J03532).

REFERENCES


http://www.ipss.go.jp/syoushika/tohkei/Mainmenu.asp (last access: 2015/02/23).


