International Journal of Transportation
Vol.6, No.3 (2018), pp.11-24
http://dx.doi.org/10.14257/ijt.2018.6.3.02

The Study of Continuously Distributed VOTT with User Heterogeneity

Kai Huang[1], Zhiyuan Liu[2*], Kun An[1] and Xuewu Chen[2]

[1]Department of Civil Engineering, Monash University, Clayton, VIC, 3800, Australia
[2]Jiangsu Key Laboratory of Urban ITS,
[2]Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, Nanjing 210096, China
[*]zhiyuanl@seu.edu.cn; leakeliu@163.com

Abstract

The Value of Travel Time (VOTT) plays an important role in the studies of transport economics and management topics, where two terms in different units are involved: time and money. The VOTT is used as a conversion between these two units for the holistic analysis. It is a case-specific value with regard to the different income level of the travelers as well as their trip emergency, thus VOTT should be considered as a continuously distributed term rather than a fixed average value with user heterogeneity. However, the existing studies do not reach an agreement on its distribution. Hence, this paper addresses a methodology based on the large-scale Revealed Preference (RP) survey and E-map technology to analyze the continuously distributed VOTT. In contrast to the traditional method of using constant/mean values, the VOTT is dealt with by calculating the ratio of the coefficients of the travel cost and the travel time, which is obtained by conducting the Multinomial Logit Regression Analysis and the Nested Logit Regression Analysis respectively. Finally, various VOTTs are used to calculate the Cumulative Distribution Function (CDF) and the Kolmogorov-Smirnov (K-S) test is carried out to analyze the optimizing continuously distributed function. In addition, some interesting issues like the travel frequency, the travel purpose, and the departure time are discussed further.

Keywords: *Continuously Distributed VOTT, User Heterogeneity, Logit Regression Analysis; K-S Test*

1. Introduction

Faced with the increasing traffic problems, especially for the traffic congestion, the Value of Travel Time (VOTT) is attracting more and more attention (Li *et al.*, 2010; Carrion and Levinson, 2012). It is mainly because the VOTT plays an important role in the area of transport economics, which can help the transport managers in conducting in-depth analysis and taking proper decisions (Liu *et al.*, 2018; Liu *et al.*, 2013b; 2016). For example, the congestion charge at peak time can help to reduce the trip demand of travelers with low VOTTs, which can alleviate the traffic congestion (Liu *et al.*, 2013a; 2014; 2017).

Studies show that the VOTT is similar for most travelers who are in the same group (Wu and Huang, 2015). The groupings are primarily based on the similarity of income

Received (May 10, 2017), Review Result (April 26, 2018), Accepted (August 11, 2018)
[*] Corresponding Author

ISSN: 2287-7940 IJT

level. With the change in the population divided by the income, the VOTT can be seen as a continuous distribution (Nie and Liu, 2010). Meng et al. (2012) supposed that the VOTT based on the change of different income groups follows Normal Distribution. However, Berg and Verhoef (2011) raised that the VOTT with user heterogeneity follows Log-normal Distribution. Normally, researchers utilized the hypothetical continuously distributed VOTT in the traffic management and control.

In this study, income is used to represent the earning power of travelers. Hence, an assumption is made that the VOTT should also be continuously distributed with the change of the population divided by the income. Furthermore, the specific distribution of the VOTT for different groupings with different incomes are explored.

1.1. Literature Review

In order to study the methodology of continuously distributed VOTT with user heterogeneity based on the existing research, this section mainly reviews two main aspects: the method of calculating the VOTT and its continuous distribution with user heterogeneity.

Studies show that there are two main methods of calculating the VOTT: calculating the income of per working hour and second, from the ratio of the marginal utility of the travel cost and the marginal utility of the travel time (Fosgerau, 2006; Santos and Bhakar, 2006; Huang *et al.*, 2018). The two calculating methods are absolutely different, however, the results are similar (Paleti *et al.*, 2015). The first one is easier to obtain as it involves the income of travelers. For example, the VOTT equals the monthly income divided by the working hours in one month. Compared to the first method, the method of calculating the ratio is more complicated but accurate. Normally, the marginal utility is replaced by the coefficient of the travel cost and the travel time that is obtained from the regression analysis of the factors influencing the mode choice (Harline and Burris, 2014). In the second method, calculation of the VOTT requires the study of travelers' model choice that is generally carried out using the Logit regression analysis (Kumar, *et al.*, 2004; Alvarez *et al.*, 2007).

The logit regression analysis that is widely used includes the Binary Logit (BL), the Multinomial Logit (MNL), the Nested Logit (NL), and the Ordered Logit (OL) models. In 1997, Abdel-Aty *et al.*, used the BL model to study the drivers' route choice with the VOTT. Then some scholars used the MNL model to analyze the multiple factors influencing travelers' choice (Hensher, 2001; Bhat and Sardesai, 2006). Their survey methods are more complicated, like the RP-SP survey. In addition, the NL and OL models are the extensions of the MNL and can study the interaction of several factors during the analysis of VOTT (Antoniou *et al.*, 2007).

The regression method can be used to obtain several VOTTs with different income groups. The study of discrete VOTTs with user heterogeneity has a long development history. Beesly studied the value of time spent in the travel in the 1960s. In 1963, Fastor and Beesly used the London underground railway data to analyze the VOTT. Their work points out one important research direction: studying the diverse VOTTs for multi-users can bring more findings in the future (Hensher, 2001). Besides, for the study of road charging, discrete VOTTs with user heterogeneity are paid more attention. In 1987, Cohen studied the fine tolling with two user groups of different VOTTs. In 1999, Cheung divided the respondents into several groups according to VOTTS and studied the change in travel demand after charging the congestion fee. Later, an increasing number of scholars studied the discrete VOTTs with user heterogeneity (Berg and Verhoef, 2011).

However, in terms of the complete sample, the discrete VOTTs will approach the continuous distribution. In 2004, Yang and Huang proposed that the discrete VOTTs belong to the continuous distribution by dividing the users into more and more groups with different incomes. Furthermore, existing studies show that the continuously distributed VOTT is Log-normal or Normal Distribution (Ben-Akiva *et al.*, 1993; Nie and Liu, 2010). For example, Nie and Liu (2010) discussed the function of Log-normal Distribution in a static congestion pricing model. Then, in 2012, Meng *et al.*, proposed a methodology of toll-charge function by assuming that the continuously distributed VOTT is normally distributed.

1.2. Objectives and Contributions

Studies show that the continuously distributed VOTT with user heterogeneity is supposed as the Log-normal or Normal Distribution in the research of congestion charging and travel behaviour. However, there are no persuasive results to support these assumptions. Hence, to overcome these issues, this paper mainly focuses on the following three aspects:

(a) Giving a clear methodology for analysing the VOTT. Apart from the method of calculating the VOTT with the ratio of the coefficients of the travel cost and the travel time by using the Logit Regression Analysis model, the data collection measures are also introduced in detail.
(b) Calibrating and analysing the continuously distributed VOTT with user heterogeneity. In this paper, a continuously distributed VOTT with the population divided by the income is studied.
(c) Utilizing E-map data to fill the gaps of RP survey. Normally, in the questionnaire survey, respondents provide an approximate travel distance. However, it is not accurate as most values are estimated based on individual judgement, especially for the walking distance and the riding distance. Hence, to solve this problem, E-map data is raised to overcome these limitations.

The remaining part of this paper is organized as follows: Section 2 clearly describes the methodology of analysing the VOTT. The data collected from the field survey and E-map is introduced in Section 3. In Section 4, the analysis results and some findings of travelers' behaviours are presented. Section 5 provides suggestions based on this study. The conclusions are drawn in Section 6.

2. Methodology

To get the continuously distributed VOTT based on the ratio of coefficients of the travel time (β_t) and the travel cost (β_c), the key procedures include data collection, regression analysis of factors influencing travelers' bahaviours, calculation of VOTT, and distribution estimation. In this section, the methodologies mainly including the following four aspects are discussed.

2.1. Data Collection

The Revealed Preference (RP) survey and the Stated Preference (SP) survey are considered as the basic survey methods (Lam and Small, 2001; Tirachini *et al.*, 2016). The RP survey focuses on the data which is generated before conducting the survey (Boyle, 2003). The SP survey can help the investigator to get the expected data, which is based on supposition (Zhang *et al.*, 2014). In this paper, the survey focuses on the mode choice of the past trip and the data is collected from RP survey.

The RP questionnaire includes citizens' personal information (travel time, travel cost, travel purpose, travel frequency, educational level, income, age, and job) and the past mode choice (the choice of travel modes). However, the travel distance was not included in the survey because most travelers do not know their accurate travel distance, particularly riding a motorcycle and taking a bus. Hence, E-Map is introduced, which can calculate the distance between any origin and destination shown in the electronic map like the Google Map. Furthermore, because of the linear relationship between travel distance and travel cost, the latter can be easily obtained.

2.2. Logit Regression Analysis Model

To analyze travelers' behaviors, especially the mode choice, the Logit regression analysis model is widely applied (Li *et al.*, 2016). In this section, the MNL model is introduced to study travelers' choice from four modes: motorcycle, bus, private vehicle, and taxi. When conducting the MNL regression, the Independence of Irrelevant Alternatives (IIA) is considered. However, when choosing a mode from several options, the IIA assumes that the odds of choosing option A over B do not depend on whether A's alternative C is present or absent. Hence, the NL model is used to conduct the regression analysis. Finally, the Hausman Test is used to analyze whether the IIA assumptions should be rejected (the NL regression analysis is better than the MNL regression analysis).

2.2.1. MNL Model

The dependent variable is the choice of travel modes. The independent variables are shown in Table 1.

Table 1. Independent Variables

Name	Variable	Property	Parameter
Travel time	T	Continuous variable	β_1
Travel cost	C	Continuous variable	β_2
Travel purpose	P	Discrete variable	β_3
Travel frequency	F	Continuous variable	β_4
Educational level	E	Discrete variable	β_5
Income	I	Continuous variable	β_6
Age	A	Discrete variable	β_7
Job	J	Discrete variable	β_8

The utility function is given in Eq.(1), U is the utility, V is the systematic component, ε is the random component and β_0 is the constant term. In Eq.(2), $P(n)$ is the probability of choosing the n th trip mode.

$$U = V + \varepsilon = \beta_0 + \beta_1 T + \beta_2 C + \beta_3 P + \beta_4 F + \beta_5 E + \beta_6 I + \beta_7 A + \beta_8 J + \varepsilon \tag{1}$$

$$P(n) = \frac{exp(U)}{\sum_{n=1}^{M} exp(U)} \tag{2}$$

To analyze the correlation between independent variables, the Maximum Likelihood Estimation (MLE) method is used. Finally, the Stata software is used to conduct the MNL regression analysis, which gives the goodness-of-fit (R^2) and the coefficients of influencing factors. P-value can be taken as one main factor to estimate the significance of each independent variable.

2.2.2. NL Model

The NL model uses the independent variables presented in MNL model in Section 2.1.1. Suppose n travel modes are divided into m branches (nests), $P(n)$ is the probability of choosing the nth trip mode, $P(m)$ is the probability of choosing the mth branch, and V_n is the systematic component of the nth trip mode.

$$P(n|m) = \frac{exp(\mu_m V_n)}{\sum_{n=1}^{N_m} exp(\mu_m V_n)} \tag{3}$$

$$I_m = \frac{1}{\mu_m} ln \sum_{n=1}^{N_m} exp(\mu_m V_n) \tag{4}$$

$$P(m) = \frac{exp(\mu I_m)}{\sum_{m=1}^{M} exp(\mu I_m)} \tag{5}$$

$$P(n) = P(n|m) \times P(m) \tag{6}$$

where μ_m and μ are the correlations of the random components ($0 \le \mu_m, \mu \le 1$). The MLE method is proposed to estimate the unknown parameters (including the β, μ_m and μ). The Stata software is used to conduct NL regression analysis.

2.3. Method of Calculating the VOTT

To get the different values of different user groups, the recommended method is utilizing the coefficients of the independent variables as shown in Eq.(7),

$$VOTT = \frac{\beta_t}{\beta_c} \tag{7}$$

Where β_t is the coefficients of travel time and β_c is the coefficients of travel cost.

The value of VOTT varies with the different grouping methods. Normally, the VOTT is continuously distributed with the population. In the following section, to test these hypotheses, the VOTTs are divided into different values according to the user heterogeneity.

2.4. Continuously Distributed VOTT Estimation

In terms of the different income levels, the target population is divided into 6 groups based on the total income per year. The income groups are; less than ￥19999, ￥20000-￥39999, ￥40000-￥59999, ￥60000-￥79999, ￥80000-￥119999 and larger than ￥120000 respectively. For the above 6 groups, the number of citizens in each group can be obtained from the Statistics Yearbook 2014 of Huzhou City. Based on the individual VOTT of 6 groups and the proportion of the population that these income groups represent, we can find the befitting Cumulative Distribution Function (CDF).

One-sample Kolmogorov-Smirnov (K-S) test for one independent sample is used to determine the distribution function. K-S test quantifies the maximum distance between the distribution function based on the statistics (samples) and the CDF of the supposed distribution function.

In Eq.(8), the $F_x(x)$ is the CDF based on the supposed distribution function and $S_n(x)$ is the CDF based on the statistics. D is the maximum difference value between $F_x(x)$ and $S_n(x)$.

$$D = Max\left|S_n(x) - F_x(x)\right| \tag{8}$$

Taking $D(n,\alpha)$ as the critical value at the significance level α of n statistics, if $D(n,\alpha)$ is larger than D, the supposed distribution function is acceptable. However, if $D(n,\alpha)$ no larger than D, the supposed distribution function is not acceptable.

Hence, for the continuous VOTT estimation, K-S test method can be used to analyze the continuous distribution function.

3. Data

The data survey was conducted in Huzhou, China in 2015. The city consists of 16 sub-districts and 62 communities. It was divided into 55 traffic zones (as showing in Figure 1). To collect citizen's information related to the mode choice, the Revealed Preference (RP) survey was conducted.

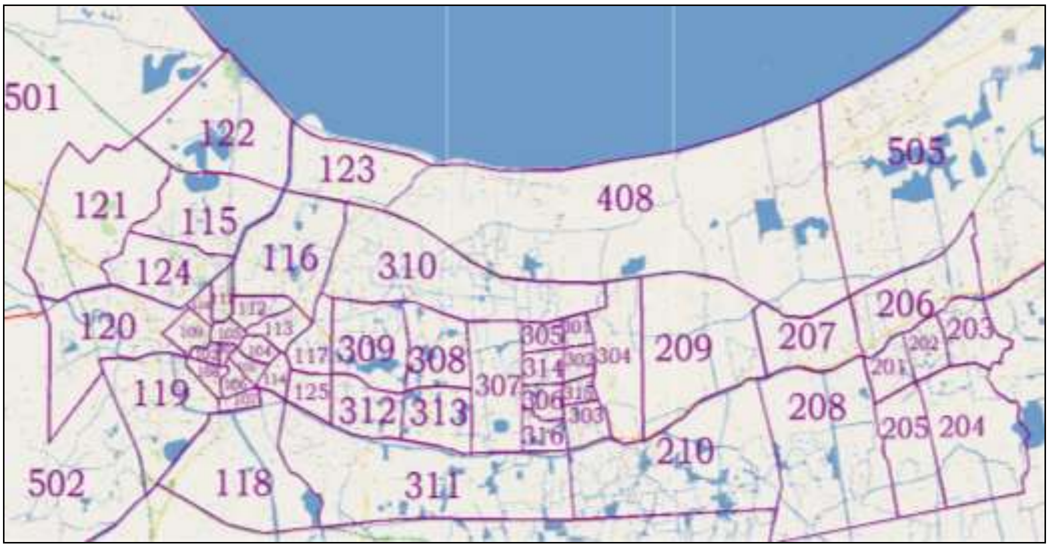

Figure 1. 55 Traffic Zones in Huzhou

The partition of traffic zones is mainly based on the population density. The trip demands in a zone are directly proportional to the number of people residing in it i.e. more demand is associated in zones with more people. Hence, a traffic zone will be smaller if there are more travelers. As shown in Figure 1, due to the high population density, the areas from Zone 101 to Zone 114 are smaller than others. For the zones like Zone 408, 501, 502 and 505, the fewer trip demand result due to fewer citizens.

3.1. RP Survey Data

The respondents' age, income, gender, and residence are reasonably distributed. The percentage of middle-aged respondents is slightly higher than other age groups while the respondents' income level is subjected to a normal distribution. The gender ratio between males and females is nearly 1:1 (50.43%:49.57%). The respondents' residents distribute over the 55 traffic zones equitably. 1544 valuable questionnaires are collected in which each sample contains 22 types of data. The main data used in this paper includes age, area type, educational level, gender, job, income, travel cost, travel frequency, travel purpose, travel time and trip mode. The following Table 2 shows the data from the RP survey.

Table 2. Variables of RP Survey

Variable	Description	Statistic
Age	Discrete variable	The numbers 1, 2, 3, 4, 5, 6, 7 and 8 indicates the age group ranging from 6 to 14, 15 to 19, 20 to 24, 25 to 29, 30 to 39, 40 to 49, 50 to 59 and over 60 respectively.
Area type	Discrete variable	The numbers 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 indicate the area type namely residence, school, administration, business, culture and recreation, sport, hospital, factory, park, and others respectively.
Educational level	Discrete variable	The numbers 1, 2, 3 and 4 indicate the educational level; primary school and junior high school, senior high school, college and higher respectively.
Gender	Discrete variable	The numbers 1 and 2 indicates the gender male and female respectively.
Job	Discrete variable	The numbers 1, 2, 3, 4, 5, 6, 7 and 8 indicates the job of the respondent; student, workman, office-bearer, enterprise staff, business owner, retiree, jobless and others respectively.
Income	Discrete variable	The numbers 1, 2, 3, 4, 5 and 6 indicates the sub-categories of annual total income; lower than 20000, from 20000 to 40000, from 400000 to 60000, from 60000 to 80000, from 80000 to 120000 and over 120000 respectively.
Travel frequency	Continuous variable	The total number of the trips on a random day.
Travel purpose	Discrete variable	The numbers 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 indicate the purpose of travel; working, studying, official business, shopping, recreation, picking up others, going to a hospital, going home, going back to work and others respectively.
Travel time	Continuous variable	The travel time in one trip.
Trip mode	Discrete variable	The numbers 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 indicate the modes of travel; walking, bicycle, electric bike, motorcycle, private vehicle, bus, taxi and others respectively.

During the RP survey, for the respondents who have several trips in a day, they should answer multiple times. Some of the variables are same for all the respondents like age, educational level, gender, job, income, and travel frequency. The other variables vary with the change in travel purpose, which includes area type, travel purpose, travel time, and trip purpose.

3.2. E-map Data

For the RP survey data collected in Huzhou, there are two variables required to be added in the 1544 questionnaires. We calculate the travel distance and travel cost via E-map firstly and then add them to each collected RP data. In terms of the travel distance, it is not accurate for travel modes except the motorcycle, private vehicle, and taxi. To get those values, E-map called AMAP is used, which is an advanced E-map provider with the path-finding, fixed position, and secondary development services. Firstly, the TransCAD is used to find out the centroid of each zone. Through the function of the secondary development of E-map, the optimal travel route and travel time for different travel modes is extracted. The following code shows the partial pseudo code of E-map.

Pseudo code for travel distance with E-map

```
<!DOCTYPE HTML>
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
<title>RoutingPlaning</title>
<link rel="stylesheet" type="text/css" href="zero.css" />
<script                                                    language="javascript"
src="http://webapi.amap.com/maps?v=1.2&key=caaa086bdf5666322fba3baf5a6a2c03"></scr
ipt>
</head>
<body onLoad="mapInit()">
  <div id="iCenter"></div>
  <div id="iControlbox">
    <ul>
      <li>
        <button onclick="javascript:driving_route();">Driving</button>
      </li>
      <li>
        <button onclick="javascript:transfer_route();">Public Transportation</button>
      </li>
      <li>
        <button onclick="javascript:walking_route();">Walking</button>
      </li>
      <li>
        <button onclick="javascript:clearMap();">Clear</button>
      </li>
    </ul>
  </div>
  <div id="result"></div>
</body>
<script language="javascript">
var mapObj;
function mapInit(){
        mapObj = new AMap.Map("iCenter",{
        center:new AMap.LngLat(120.1000, 30.8766),
        level:11
        });
}
function clearMap(){
        mapObj.clearMap();
        document.getElementById("result").innerHTML = ' ';              }
var ad = new Array();
```



```
ad[1]=new AMap.LngLat(120.10266,30.867054)
ad[2]=new AMap.LngLat(120.09536,30.866726)
ad[3]=new AMap.LngLat(120.10358,30.872642)
ad[4]=new AMap.LngLat(120.11806,30.868422)
ad[5]=new AMap.LngLat(120.114014,30.862691)
ad[6]=new AMap.LngLat(120.104868,30.856874)
ad[7]=new AMap.LngLat(120.108024,30.851056)
ad[8]=new AMap.LngLat(120.094796,30.859468)
ad[9]=new AMap.LngLat(120.085625,30.873472)
ad[10]=new AMap.LngLat(120.092396,30.88211)
……
```

The second data required is the travel cost. Normally, it is not easy to know the cost spent on the private vehicle or a motorcycle. The following are the methods to obtain the travel cost.

$$C_{Walk} = C_{Bike} = C_{Electric\ bike} = 0 \tag{9}$$

$$C_{Bus} = 2 \tag{10}$$

$$C_{Taxi} = \begin{cases} 7 & 0 < l \leq 2.5 \\ 7 + 2.5 \cdot (l - 2.5) & 2.5 < l \leq 7 \\ 20.75 + 2.5 \cdot (l - 7) \times 1.5 & 7 < l \end{cases} \tag{11}$$

$$C_{Private\ vehicle} = 0.424 \cdot l \tag{12}$$

$$C_{Motor\ bike} = 0.104 \cdot l \tag{13}$$

where C is the travel cost and l is the travel distance.

In Eq.(9) and Eq.(10), the charges for walking, riding a bike, riding an electric bike and taking a bus are fixed values (the first three equals 0 while the last equals ￥2). In Eq.(11), the charge rate is ￥7 within 2.5 kilometers, the charge rate is ￥2.5 per kilometer from 2.5 to 7 kilometers and the charge rate is ￥3.75 per kilometer over 7 kilometers. In Eq.(12) and Eq.(13), the charge is based on the fuel consumption and the oil price.

4. Estimation Result and Discussion

This section mainly focuses on calculating and analyzing the continuously distributed VOTT with user heterogeneity. Besides, some interesting findings on travelers' behaviors are also discussed.

4.1. Continuously Distributed VOTT with User Heterogeneity

After the RP survey and the secondary development of E-map, the fundamental trip information and travellers' personal information are obtained, especially the travel cost and travel time. Then the MNL model and NL model are established to get the coefficients of the travel time and travel cost. With the individual VOTTs and proportion of the population in each group, the continuous distribution function is tested. However, due to the low precision of the collected data, this paper is unable to determine persuasive continuously distributed functions.

4.2. Resident Trip Characteristics

Some interesting findings are observed during the data analysis procedure like the distribution of the departure time, the travel frequencies and the travel purposes. The following Figure 2 shows the findings of the characteristics of citizens' trip. In Figure 2(a), the numbers from 1 to 12 in the horizontal ordinate indicate the number of travel trips observed during a day. In Figure 2(b), the numbers from 1 to 10 in the horizontal ordinate represent the travel purposes: working, studying, official business, shopping, recreation, picking up others, going to a hospital, going home, going back to work and others. In Figure 2(c), the numbers from 1 to 10 in the horizontal ordinate represent the time from 0:00 to 20:00 for 20 hours.

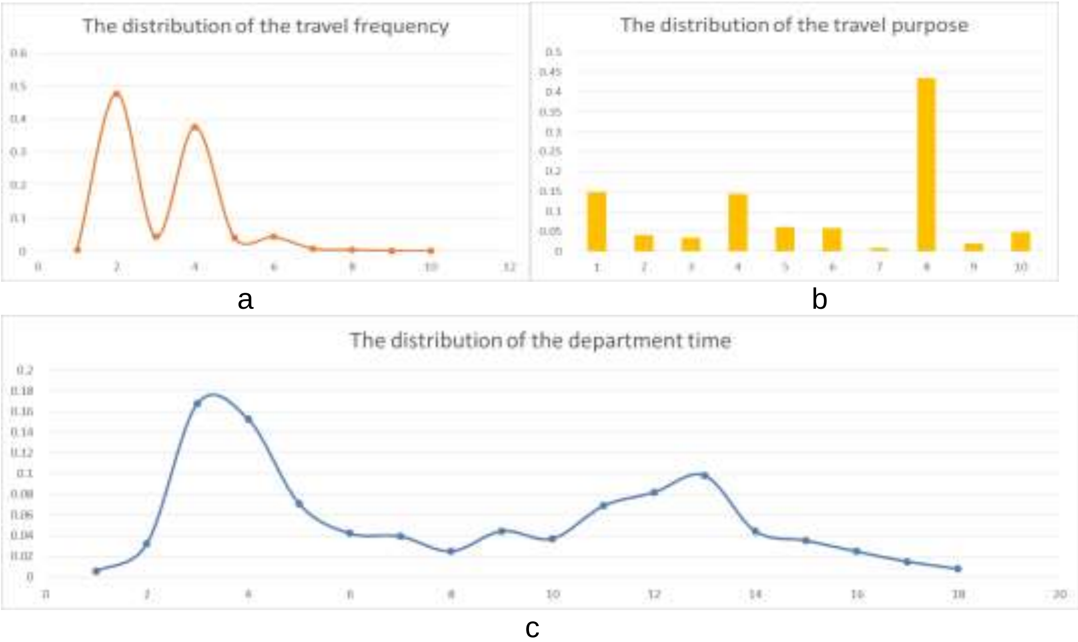

Figure 2. The Distribution of the Travel Frequency, the Travel Purpose, and the Departure Time

For the travel frequency, most of the travelers make 2 to 4 trips in a day. The statistical data shows that the frequency of 2 to 4 trips is about ten times larger than 5 to 6 trips. For the travel purpose, the percentage of going home is the largest. It is mainly because that each traveler has to go back home after each trip like going shopping, going to work, etc. For the departure time, two peak hours from 7:30 to 9:30 and from 16:00 to 18:30 can be clearly observed. This reveals that the collected data can represent the characteristic of the trip.

5. Suggestions

Based on the study of this paper, two proposals are raised in this paper. The first proposal is to improve the data collection method. In the survey of mode choice, the respondents' preferences for alternatives which do not exist are also needed to estimate. Due to the advantages of incorporating hypothetical situations, SP survey can provide more data. However, the large number of hypothetical situations used to describe the choices could increase the complexity of the survey and make it difficult for the respondents to consider all the situation while making their choices. It is recommended that the number of situations in SP survey should be no more than 4 to 6.

Another proposal for the estimation of continuously distributed VOTT is to segment the sample into more income groups in order to have sufficient data points for the estimation of the CDF. However, one of the issue with this is that when there are more income groups, the sample size required for the survey is larger. The study should take this into account when aiming to estimate the VOT distribution and balance the required sample size and the number of income groups to be estimated. Moreover, when collecting data from different income groups, it is inevitable that there will be a lack of access to income groups of high income and should be considered when designing the survey. In the future studies, these challenges should be taken into account to achieve higher accuracy while estimating continuously distributed VOTT.

6. Conclusion

The aim of this paper is to study the continuously distributed VOTT with user heterogeneity. It clearly describes the methodology of analyzing the VOTT followed by a case study to test it. Besides, some interesting findings are proposed, which reveal the characteristics of travelers' behaviors. The travel frequency, the travel purpose, and the departure time are reasonably distributed.

To understand the methodology of analyzing continuously distributed VOTT with user heterogeneity, there are four main steps. (i) RP survey is conducted to collect travelers' personal information and travel information, and E-map is used to make up for the lack of accurate travel time and travel cost. (ii) MNL regression analysis model and NL regression analysis model is established to find out the coefficients of travel time and travel cost, which can be taken as the marginal utility. (iii) The marginal utilities of travel time and travel cost are used to calculate the VOTT. (iv) The one-sample K-S test is used to estimate the continuously distributed function.

There are three main findings from our study. The first is using E-map data to enrich the RP data provides some interesting findings. The second is proposing a methodology to study the continuous VOTT distribution with user behaviours. The last is to show the trip characteristics of travelers. The travel frequency, the travel purpose and the departure time are found to be reasonably distributed.

However, there are still some limitations that need to be addressed in the future studies. No doubt, the accuracy of the collected data should be improved further. In theory, the regression analysis of travel time and travel cost is absolutely significant. However, our results are not ideal. In addition, the method of supplementing data with E-map should be improved further. The RP data used in this study is the historical data. However, the secondary development of E-map provides the real-time data. Hence, to enhance its accuracy, some improved measures should be carried out.

Acknowledgment

This study is supported by the Youth Program (No. 71501038), Surface Project (No. 71771050), Key Projects (No. 51638004) of the National Natural Science Foundation of China, and the Natural Science Foundation of Jiangsu Province, China (BK20150603).

References

[1] M. A. Abdel-Aty, R. Kitamura and P. P. Jovanis, "Using stated preference data for studying the effect of advanced traffic information on drivers' route choice", Transportation Research Part C, vol. 5, no. 1, **(1997)**, pp. 39-50.

[2] O. Alvarez, P. Cantos and L. Garcia, "The value of time and transport policies in a parallel road network", Transport Policy, vol. 14, pp. 366-376.

[3] C. Antoniou, E. Matsoukis and P. Roussi, "A methodology for the estimation of value-of-time using state-of-the-art econometric models", Journal of Public Transportation, vol. 10, no. 3, **(2007)**, pp. 1-19.

[4] M. Ben-Akiva, D. Bolduc and M. Bradley, "Estimation of travel choice models with randomly distributed values of time", Transportation Research Record, vol. 1413, **(1993)**, pp. 88-97.

[5] V. V. D. Berg and E. T. Verhoef, "Congestion tolling in the bottleneck model with heterogeneous values of time", Transportation Research Part B, vol. 45, no. 1, pp. 60-78.

[6] C. R. Bhat and R. Sardesai, "The impact of stop-making and travel time reliability on commute mode choice", Transportation Research Part B, vol. 40, no. 9, **(2006)**, pp. 709-730.

[7] K. J. Boyle, "Introduction to revealed preference methods", In A primer on nonmarket valuation, **(2003)**, pp. 259-267.

[8] G. E. Cantarella and M. Binetti, "Stochastic Equilibrium Traffic Assignment with Value of time Distributed Among Users", International Transactions in Operational Research, vol. 5, no. 6, **(1998)**, pp. 541-553.
[9] C. Carrion and D. Levinson, "Value of travel time reliability: A review of current evidence", Transportation research part A, vol. 46, no. 4, **(2012)**, pp. 720-741.
[10] W. M. Cheung, H. Yang and W. H. Tang, "Modeling private toll roads with heterogeneous users", Proceedings of the Fourth Conference of the Hong Kong Society for Transportation Studies, Hong Kong, **(1999)** December 4, pp. 189-198.
[11] Y. Cohen, "Commuter welfare under peak-period congestion tolls: who gains and who loses?", International Journal of Transport Economics, vol. 14, no. 3, **(1987)**, pp. 239-266.
[12] M. Fosgerau, "Investigating the distribution of the value of travel time savings", Transportation Research Part B, vol. 40, **(2006)**, pp. 688-707.
[13] C. D. Foster and M. E. Beesley, "Estimating the social benefit of constructing an underground railway in London", Journal of the Royal Statistical Society, vol. 126, no. 1, pp. 46-93.
[14] C. E. Harline and M. W. Burris, "Impact of traffic images on route choice and the value of time estimates in stated preference surveys", Journal of the Transportation Research Board, vol. 2405, no. 04, **(2014)**, pp. 24-32.
[15] D. A. Hensher, "Measurement of the Valuation of Travel Time Savings", Journal of Transport Economics and Policy, vol. 35, no. 1, **(2001)**, pp. 71-9.
[16] K. Huang, G. H. de Almeida Correia and K. An, "Solving the station-based one-way carsharing network planning problem with relocations and non-linear demand", Transportation Research Part C, vol. 90, **(2018)**, pp. 1-17.
[17] E. P. Kroes and R. J. Sheldon, "Stated preference methods: An introduction", Journal of Transport Economics and Policy, vol. 22, **(1988)**, pp. 11-25.
[18] C. Kumar, D. Basu, B. Maitra and I. Kharagpur, "Modeling generalized cost of travel for rural bus users: a case study", Journal of Public Transportation, vol. 7, **(2004)**, pp. 59-72.
[19] T. C. Lam and K. A. Small, "The value of time and reliability: measurement from a value pricing experiment", Transportation Research Part E, vol. 37, no. 2, **(2001)**, pp. 231-251.
[20] D. Li, T. Miwa, T. Morikawa and P. Liu, "Incorporating observed and unobserved heterogeneity in route choice analysis with sampled choice sets", Transportation Research Part C, vol. 67, **(2016)**, pp. 31-46.
[21] Z. Li, D. A. Hensher and J. M. Rose, "Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence", Transportation research part E, vol. 46, no. 3, **(2010)**, pp. 384-403.
[22] Z. Liu, Q. Meng and S. Wang, "Speed-based toll design for cordon-based congestion pricing scheme", Transportation Research Part C, vol. 31, **(2013a)**, pp. 83-98.
[23] Z. Liu, S. Wang, W. Chen and Y. Zheng, "Willingness to board: a novel concept for modeling queuing up passengers", Transportation Research Part B, vol. 90, **(2016)**, pp. 70-82.
[24] Z. Liu, S. Wang and Q. Meng, "Optimal joint distance and time toll for cordon-based congestion pricing", Transportation Research Part B, vol. 69, **(2014)**, pp. 81-97.
[25] Z. Liu, S. Wang, K. Huang, J. Chen and Y. Fu, "Practical taxi sharing schemes at large transport terminals", Transportmetrica B, **(2018)**, pp. 1-21.
[26] Z. Liu, S. Wang, B. Zhou and Q. Cheng, "Robust optimization of distance-based tolls in a network considering stochastic day to day dynamics", Transportation Research Part C, vol. 79, **(2017)**, pp. 58-72.
[27] Z. Liu, Y. Yan, X. Qu and Y. Zhang, "Bus stop-skipping scheme with random travel time", Transportation Research Part C, vol. 35, **(2013b)**, pp. 46-56.
[28] Q. Meng, Z. Liu and S. Wang, "Optimal distance tolls under congestion pricing and continuously distributed value of time", Transportation Research Part E, vol. 48, no. 5, **(2012)**, pp. 937-957.
[29] R. Paleti, P. Vovsha, D. Givon and Y. Birotker, "Impact of individual daily travel pattern on value of time", Transportation, vol. 42, no. 6, **(2015)**, pp. 1003-1017.
[30] G. Santos and J. Bhakar, "The impact of the London congestion charging scheme on the generalised cost of car commuters to the city of London from a value of travel time savings perspective", Transport Policy, vol. 13, no. 1, **(2006)**, pp. 22-33.
[31] A. Tirachini, L. Sun, A. Erath and A. Chakirov, "Valuation of sitting and standing in metro trains using revealed preferences", Transport Policy, vol. 47, **(2016)**, pp. 94-104.
[32] W. X. Wu and H. J. Huang, "An ordinary differential equation formulation of the bottleneck model with user heterogeneity", Transportation Research Part B, vol. 81, **(2015)**, pp. 34-58.
[33] H. Yang and H. J. Huang, "The multi-class, multi-criteria traffic network equilibrium and systems optimum problem", Transportation Research Part B, vol. 38, no. 1, **(2004)**, pp. 1-15.
[34] G. Zhang, Z. Wang, K. R. Persad and C. M. Walton, "Enhanced traffic information dissemination to facilitate toll road utilization: a nested logit model of a stated preference survey in Texas", Transportation, vol. 41, no. 2, **(2014)**, pp. 231-249.

Authors

Kai Huang, he received his master degree in transportation engineering in 2016 from southeast university, where he is currently working toward the Ph.D. degree. His main research interests include sharing transportation and transportation data analysis.

Zhiyuan Liu, he received his Ph.D degree in Transportation Engineering in 2011 from National University of Singapore. He is a Professor in the School of Transportation at Southeast University. His research interests include transportation big data analysis and modelling, and transportation network planning and management.

Kun An, she received her Ph.D degree in Transportation Engineering in 2014 from Hong Kong University of Science and Technology. She is a lecturer in the Institute of Transport Studies at Monash University. Her research interests include reliable transit network design with stochastic demand, logistic and supply chain design, robust sensor location and traffic operation and control.

Xuewu Chen, she received her Ph.D degree in Transportation Engineering in 2002 from Southeast University. She is a Professor in the School of Transportation at Southeast University. Her research interests include sustainable transportation system and public transport optimization.