A Two-stage Model of Tourists’ Multi-destination Movement

Yang Yang  
*Department of Geography, University of Florida*

Timothy Fik  
*Department of Geography, University of Florida*

Jie Zhang  
*Institute of Tourism Studies, Nanjing University*

Follow this and additional works at: [https://scholarworks.umass.edu/ttra](https://scholarworks.umass.edu/ttra)
A Two-stage Model of Tourists’ Multi-destination Movement

Yang Yang
Department of Geography
University of Florida

Timothy Fik
Department of Geography
University of Florida

and

Jie Zhang
Institute of Tourism Studies
Nanjing University

ABSTRACT

To overcome the shortcoming within the traditional micro-model and incorporate tourists’ multi-destination movement, this paper proposes an alternative two-stage micro-model of tourists’ multi-destination travel. In the first stage, tourists in a destination decide whether to go to the next destination or terminate traveling by returning home, while in the second stage, those opting to continue must then choose the second destination. By comparing the sign and magnitude of several configuration parameters embedded in the model, the empirical results confirmed different traveling patterns of tourists with different motivation. Finally, several implications are provided for tourism marketing strategies and tactics.

Keywords: micro-model, sample selection, multi-destination travel.

INTRODUCTION

Among previous studies on tourist flow/movement analysis, both micro- and macro-models have been introduced to explain tourists’ spatial movements and destination choices (Ryan 2003; Smith 1983). A major drawback of various macro-models, which have been tailored to fit the aggregate data of tourist movement, comes from their failure to consider individual heterogeneity of tourists (Uysal and Crompton 1985; Eugenio-Martin 2003). On the other hand, the discrete choice model (DCM) has been overwhelmingly popular as a micro-model to explain the movement of individual tourists among a set of alternative destinations (Nicolau and Más 2008; Seddighi and Theocharous 2002). However, a major problem associated with the DCM is the limited number of destinations in the choice set. In the DCM’s specifications, the outcome of the model could only incorporate a handful of destination alternatives due to computational considerations (Cameron and Trivedi 2005).

To overcome the shortcoming within the traditional micro-model and incorporate tourists’ multi-destination movement, this paper applies an alternative two-stage micro-model
of tourists’ multi-destination movement. In the first stage, tourists in a destination decide whether to go to the next destination or terminate traveling by returning home, while in the second stage, those opting to continue must then choose the second destination.

**MODEL AND DATA**

We propose the model as a two-stage sample selection model (Greene 2007; Becken and Schiff 2011). The first-stage model captures the decision of whether to travel to a second destination \((y = 1\) if choose to continue\) as a probit model. The second-stage model is a linear regression model, whose dependent variable, \(\ln D\), is the log distance between the first and second destinations. The model is specified as follows:

\[
\begin{align*}
Y &= \begin{cases} 
1 & \text{if } y = 1 \\
0 & \text{if } y = 2
\end{cases}, \\
Y &= X_1 \beta_1 + \varepsilon_1; \\
\ln D &= \begin{cases} 
\ln D^* & \text{if } y = 1 \\
\ln D^* & \text{if } y = 2
\end{cases}, \\
\ln D^* &= \alpha_2 \ln distrd1 + \beta_2 \ln distrd2 + X_2 \beta + \varepsilon_2
\end{align*}
\]

where \(X_1\) and \(X_2\) denote the vector of explanatory variables specified in the two models, respectively. In particular, \(\ln distrd1\) denotes the log distance between the residence and the first destination, while \(\ln distrd2\) represents the log distance between the residence and the second destination. Moreover, the error terms of two equations, \(\varepsilon_1\) and \(\varepsilon_2\), are jointly distributed, which suggests that the two stages of decision-making are inter-connected. To estimate the proposed model, we use the individual tourist survey data from a province-wide domestic tourist survey in Jiangsu Province of China, which comprises the information of 27,417 Chinese domestic tourists’ movements after visiting a particular destination in Jiangsu Province. The definition of other explanatory variables is presented in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>The age of tourist</td>
<td>1=age 14 and below, 2=15-24, 3=25-44, 4=45-64, 5=age 65 and above</td>
</tr>
<tr>
<td>night</td>
<td>The nights tourist spend in the 1st destination</td>
<td></td>
</tr>
<tr>
<td>mode</td>
<td>The traveling pattern of tourists</td>
<td>1=organized by affiliations, 2=with friends and relatives, 3=package tour, 4=alone</td>
</tr>
<tr>
<td>pastvisit</td>
<td>Number of previous visits to the first destination</td>
<td></td>
</tr>
<tr>
<td>distrd1</td>
<td>Distance between the residence and the first destination</td>
<td></td>
</tr>
<tr>
<td>purpose</td>
<td>The major purpose of the trip</td>
<td>1=leisure/vacation, 2=sightseeing, 3=VFR, 4=business/conference, 5=others</td>
</tr>
</tbody>
</table>
**Table 1**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>distrd2</td>
<td>Distance between the residence and second destination</td>
</tr>
<tr>
<td>A4</td>
<td>The number of AAAA scenic spots in the second destination</td>
</tr>
<tr>
<td>NationalPark</td>
<td>The number of National Parks in the second destination</td>
</tr>
<tr>
<td>CD</td>
<td>The competition destination effect Using Fik, Amey, and Mulligan (1992)'s hierarchical index, which composes two levels</td>
</tr>
<tr>
<td>IO</td>
<td>The intervening opportunity effect Using Fik, Amey, and Mulligan (1992)'s hierarchical index, which composes two levels</td>
</tr>
</tbody>
</table>

$\delta_1$ and $\delta_2$ are of particular interest to illustrate the pattern of multi-destination travel. $\delta_1$ indicates the relative length between the travel distance from residence to the first destination and that from the first to the second destination. If $\delta_1 > 0$, it is suggested that the travel distance from the first to the second destination is longer than that from residence to the first destination. In the same manner, $\delta_2$ compares the travel distance from residence to the second destination and the distance from the first to the second destination. $\delta_2 > 0$ suggests that tourists are choosing a second destination that is further away from residence than the first destination. Figure 1 demonstrates the four possible spatial patterns of movement based on these two coefficients.

**Figure 1**

**Different spatial patterns of tourists’ movement indicated by model coefficients**

<table>
<thead>
<tr>
<th>$\delta_1 &gt; 0$</th>
<th>$\delta_1 &lt; 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_2 &gt; 0$</td>
<td>$\delta_2 &lt; 0$</td>
</tr>
</tbody>
</table>

![Diagram showing different spatial patterns of tourists' movement](attachment:image.png)
RESULTS

Table 2 presents the estimation results of empirical models. In Model 1 including all observations, the results from the first stage model suggest that tourists who stay shorter in a destination, travel alone, and visit relatives and friends are more likely to continue their travel to a second destination. Regarding the results from the second-stage model, $\beta_2$ is estimated to be negative, while $\beta_1$ is positive, which suggests the general pattern of spatial movement as shown in the upper right graph of Figure 1. Moreover, other significant coefficients in Model 1 suggest that tourists with business/conference purposes and organized by affiliations travel longer to the second destination. The number of national parks in the second destination seems to be a striking attraction to tourists as indicated by the significant and positive coefficient of NationalPark, and the competition destination effect (CD) and intervening opportunity effect (IO) are found to be statistically significant.

To further look into tourists with different purposes, we estimate separate models to unveil their multi-destination travel patterns. Models 2 to 6 in Table 2 present these results. For the coefficients in the first-stage model, it is found that frequent tourists with leisure/vacation purpose are more likely to visit a second destination, but their counterparts with VFR and business/conference purposes are less likely to travel to a subsequent destination. As shown by the estimated coefficients of $\beta_1$ and $\beta_2$, the spatial movement pattern of leisure/vacation and sightseeing tourists are similar as the general pattern depicted in Model 1.

Table 2

<table>
<thead>
<tr>
<th>Estimation results of empirical models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model1</td>
</tr>
<tr>
<td>All data</td>
</tr>
<tr>
<td>age</td>
</tr>
</tbody>
</table>
night  
-0.0516*** -0.0597*** -0.0773*** -0.0721*** -0.0122 -0.00428
(0.0062) (0.0138) (0.0102) (0.0184) (0.0139) (0.0176)

mode = 2  
0.0454 0.226*** -0.105* -0.381*** 0.305*** -0.106
(0.0232) (0.0494) (0.0434) (0.0825) (0.0701) (0.0656)

mode = 3  
0.295*** 0.303*** 0.190*** 0.206 0.0678 0.588***
(0.0306) (0.0607) (0.0502) (0.1480) (0.1250) (0.1010)

mode = 4  
0.178*** 0.323*** 0.027 -0.253** 0.306*** 0.0877
(0.0221) (0.0525) (0.0470) (0.0844) (0.0396) (0.0512)

pastvisit  
-0.00401 0.0498*** 0.0283 -0.0763*** -0.0717*** 0.00156
(0.0081) (0.0193) (0.0146) (0.0215) (0.0176) (0.0211)

\( lndistrd1 \)  
0.000925*** 0.0012*** 0.0010*** 0.00078*** 0.00078*** 0.00068***
(0.0000) (0.0000) (0.0000) (0.0000) (0.0000) (0.0000)

\( purpose = 2 \)  
-0.0286
(0.0218)

\( purpose = 3 \)  
-0.0971***
(0.0292)

\( purpose = 4 \)  
-0.0452
(0.0271)

\( purpose = 5 \)  
-0.0395
(0.0293)

\( constant \)  
-0.436*** -0.745*** -0.404*** 0.168 -0.422*** -0.291*
(0.0494) (0.1070) (0.0807) (0.1400) (0.1240) (0.1330)

**Second-stage Model**

\( lndistrd1(2) \)  
-0.0742*** -0.125*** -0.145*** 0.0138 0.0407 -0.0715
(0.0143) (0.0332) (0.0224) (0.0328) (0.0306) (0.0396)

\( lndistrd2(2) \)  
0.154*** 0.274*** 0.224*** 0.0655** 0.0481* 0.0865***
(0.0096) (0.0218) (0.0169) (0.0251) (0.0191) (0.0254)

\( A4 \)  
-0.0805*** -0.0686*** -0.0781*** -0.0882*** -0.0838*** -0.0902***
(0.0018) (0.0044) (0.0031) (0.0048) (0.0039) (0.0047)

\( A4\_square \)  
0.00252*** 0.0022*** 0.0024*** 0.00281*** 0.00253*** 0.00276***
(0.0001) (0.0001) (0.0001) (0.0001) (0.0001) (0.0001)

\( NationalPark \)  
0.0772*** 0.0807*** 0.0496*** 0.0977*** 0.142*** 0.047*
(0.0070) (0.0172) (0.0110) (0.0190) (0.0147) (0.0210)

\( CD \)  
0.00678*** 0.0068*** 0.0074*** 0.00610*** 0.00597*** 0.00673***
(0.0001) (0.0001) (0.0001) (0.0002) (0.0001) (0.0002)

\( IO \)  
(0.0380) (0.0944) (0.0620) (0.0986) (0.0789) (0.1110)

\( purpose = 2 \)  
0.0595***
(0.0139)

\( purpose = 3 \)  
0.0271
(0.0187)
| purpose=4  | 0.101*** | (0.0171) |
| purpose=5  | 0.0666*** | (0.0186) |
| mode=2     | -0.0480** | 0.115**  | -0.128*** | 0.0313  | 0.0263  | -0.150** | (0.0152) | (0.0398) | (0.0274) | (0.0552) | (0.0442) | (0.0495) |
| mode=3     | -0.347*** | -0.333*** | -0.327*** | -0.152  | -0.0032 | -0.563*** | (0.0204) | (0.0471) | (0.0306) | (0.0778) | (0.0761) | (0.0816) |
| mode=4     | -0.0314*  | 0.128**  | -0.0462  | -0.0489 | -0.0549 | -0.0424  | (0.0151) | (0.0436) | (0.0287) | (0.0509) | (0.0316) | (0.0364) |
| constant   | 10.72***  | 9.453*** | 10.72*** | 10.89*** | 10.80*** | 12.18*** | (0.228)  | (0.536)  | (0.334)  | (0.480)  | (0.498)  | (0.655)  |
| Mil's lambda| 0.0303  | 0.293**  | 0.0797  | -0.0139 | -0.0078 | -0.493** | (0.0504) | (0.106)  | (0.0702) | (0.121)  | (0.113)  | (0.182)  |
| observations| 27411  | 5626  | 9954  | 3408  | 5024  | 3399 |

(Notes: standard error in parenthesis, * indicates significant at the 0.10 level, ** indicates significant at the 0.05 level, *** indicates significant at the 0.01 level.)

REFERENCES


