SEM ANALYSIS FOR TOURIST EXPENDITURE IN AN EMERGING COUNTRY

ANÁLISIS SEM PARA EL GASTO TURÍSTICO EN UN PAÍS EMERGENTE

ANÁLISE SEM DE GASTOS TURÍSTICOS EM UM PAÍS EMERGENTE

Jorge Guillén Uyen

Abstract
The following study attempts to examine the relationship between consumer preferences and expenditure in a Structural Equation Model (SEM) framework. We use data from a survey applied in an emerging country like Peru. There were 6,772 respondents classified by sex, country of residence, education level, and age. The foreign respondents specified their motives and preferences to travel as well as money spent. The SEM technique permit to find a relationship between the latter variables. This assessment is relevant for tourism policymakers since it will help to find the main drivers of expenditure by foreign visitors within an emerging country. Our study permitted us to identify the motives to visit: leisure, business and education, and others as main drivers of touristic expenditure.

Keywords: SEM; touristic expenses; socioeconomic variables; motive variables

Resumen
El siguiente estudio intenta examinar la relación entre las preferencias del consumidor y el gasto en un marco de Modelo de Ecuación Estructural (SEM, por sus siglas en inglés). Usamos datos de una encuesta aplicada en un país emergente como Perú. Hubo 6,772 encuestados clasificados por sexo, país de residencia, nivel educativo y edad. Los encuestados extranjeros especificaron sus motivos y preferencias para viajar, así como el dinero gastado. La técnica SEM permite encontrar una relación entre las últimas variables. Esta evaluación es relevante para los responsables de la política...
turística, ya que ayudará a encontrar los principales factores de gasto de los visitantes extranjeros dentro de un país emergente. Nuestro estudio permitió identificar los motivos para visitar: ocio, negocios y educación, y otros, como principales impulsores del gasto turístico.

**Palabras clave:** SEM; gastos turísticos; variables socioeconómicas; variables de motivo

---

**Resumo**

O estudo a seguir tenta examinar a relação entre preferências e gastos do consumidor em uma estrutura do Modelo de Equações Estruturais (SEM). Utilizamos dados de uma pesquisa aplicada em um país emergente como o Peru. Havia 6 772 entrevistados classificados por sexo, país de residência, nível de escolaridade e idade. Os entrevistados estrangeiros especificaram seus motivos e preferências para viajar, bem como o dinheiro gasto. A técnica SEM permite encontrar uma relação entre as últimas variáveis. Essa avaliação é relevante para os formuladores de políticas de turismo, pois ajudará a encontrar os principais fatores de despesa dos visitantes estrangeiros em um país emergente. Nosso estudo permitiu identificar os motivos para viajar: lazer, negócios e educação, entre outros como principais motivadores de gastos turísticos.

**Palavras-chaves:** SEM; gastos turísticos; variáveis socioeconómicas; variáveis de motivação

---

**1. Introduction**

Tourism in Latin America has been a relevant economic activity, forecasts suggest continued growth (Correia, Santos and Pestana, 2007). The later situation can be explained because of Latin America’s sheer diversity in scenic beauty, cuisine and cultures have combined with an increasingly sophisticated domestic industry to cater to every kind of traveler. Historical views from Mayas to Incas permitted to attain a variety of motives to visit. In addition, natural landscapes also attract a different kind of tourists.

Since the subprime crisis of 2009, tourism’s direct contribution to Gross Domestic Product (GDP) in Latin America grew significantly to reach an estimated $121 billion in 2016, which represents 3,2% of GDP. Also, the touristic industry contributes to almost 2,9% of the total share of employment data³. It is projected that the touristic contribution to GDP rises to 3,5% in 2027. Forecasts for contribution to employment will be 3,3 % of GDP⁴ (Figure 1).

---

³ The data is updated until 2016.
⁴ The statistic considers all kind of revenue generated by tourism-oriented services such as hotels and airlines, restaurant and leisure industries that attract tourists.
In comparison to other emerging areas, the tourism in North-East Asia and Asia Pacific Area has a contribution to GDP below Latin America (Figure 2), (2.5% and 2.8% respectively). Therefore, the industry of tourism in Latin America is very important. Only in the Caribbean area the contribution to GDP is 4.7% but is only reports an approximate total contribution of US$ 17.9 billion (World Travel and Tourism Council [WTTC], 2016; World Bank, 2016).

Looking at the individual country’s performance, there is some heterogeneity. Economic growth and government policy of tourism led to a stunning 50 percent growth in direct GDP contribution in Peru, but a decline of over 40 percent in Suriname Brazil, the region’s largest country, is also its tourism power player, here the contribution to GDP is 3.2% with a maximum of 4.2% in 2003. In the case of Venezuela, the contribution is not
doing an outstanding performance. Since the latter country has political problems, the
tourism contribution varies from 3.6% in 2009 to 3.2% in 2016 (WTTC, 2016; World
Bank 2016).

We will explore relationships between tourism expenditure, motivation to visit the
country as well as the socioeconomic characteristic of foreign travelers for an emerging
economy that belongs to the Latin American Region: Peru. According to the data shown
above, this country is relevant in the area in the touristic industry.

This research has the aim and objective of developing a SEM of the Peruvian tourism
demand as well as using this model for explaining tourists’ behavior. The paper is organized
as follows: the next section will have a literature review for the touristic demand, then
we propose the SEM. Section 4 explains the results and finally, we conclude.

2. Literature review of demand for tourism

In the existing literature, there are not several studies that pay attention to traveler’s
expenditure. Mitchell and Reid (2001) find some socioeconomic benefits in Taquile
Island for most residents due to touristic activity in an integrated community.

Also, for Desforges (2000) Peru is a good case study of touristic reform after the neoliberal
structural adjustment. The author reports changes in the sector since the market open
in the country in 1990. These changes focus on the role of government and the private
sector as well. The latter strategy has benefited Peruvian tourism since it increased
significantly. The Peruvian touristic model cannot be compared yet on an international
scale.

Moreover, Divino and MacAleer (2010) show for an empirical result that international
tourist arrivals and their growth rates are stationary. The authors run a Dickey Fuller
test to proceed with their hypothesis, estimated symmetric and asymmetric conditional
volatility models all fit the data extremely well and the behavior of the touristic arrivals
behave like financial variables.

Our study explores the determinants of expenditure like above, but the approach is
different since we use a technique that captures qualitative data in the SEM framework.
In this case, there is not any seminal paper that hits this point and from the mentioned
approach. In the literature, there is an estimation of determinants of traveler’s
expenditure for countries like Brazil, Mexico, South Africa, USA, and European Countries
by using panel data models.

Chaitip, Chaiboonsri, Kovács, and Balogh (2010) used the SEM to analyze touristic
demand for the case of Greece. In this study, a survey containing Likert scale questions
was conducted to collect data from 100 tourists who had traveled to Greece’s tourist
destination. With the help of factor analysis, four dimensions were identified for scales
used in the study: travel cost satisfaction, tourism product, tourism product attributes,
and tourism product management.
Also, Dong-Wan and Stewart (2002) tests the structural equation model between residents’ perceived tourism impacts and attitudes toward the host community. The study is run for residents of Cheju Island, Korea. The construct community satisfaction was influenced by the perception of tourism impacts.

Besides motive variables, sociodemographic variables have been widely used to predict the level of travel or recreation expenditure. Dardis, Derrick, Lehfeld, and Wolfe (1981) utilized 1972-73 Consumer Expenditure Survey conducted by the US Bureau of Labor Statistics to investigate the impact of various socio-demographic characteristics on recreation expenditures in the US.

Dardis, Soberon-Ferrer, and Patro (1994) found that income was again found to be an influencing factor in explaining variations of household expenditures on three categories of recreation items including visiting museums, attending sports events, and other recreation and entertainment. Mok and Lam (1997) found that the Taiwanese tourists concentrate more on shopping than any other category, this behavior can be attributed to a cultural value of keeping relationships through giving a gift.

Saayman and Saayman (2014) incorporated socio-economic variables and income variables as well. They also include the length of stay as a determinant of tourist expenditure. For the same authors, the variable sex is also incorporated but they found it irrelevant. For other authors like Lehto, O’Leary, and Morrison (2004) women spend more than men but for Jones, Wood, Catlin, and Norman (2009) the result is just the opposite. The latter is interesting for touristic planning segmentation across gender.

Age is an important variable that may explain touristic demand. Younger people tend to be interested in adventurous activities such as wildlife viewing and beach holidays. With advances in age, individuals search for holiday activities that can give them more knowledge and understanding of issues outside their usual environment of life, particularly cultural activities (Van der Merwe, Slabbert, and Saayman, 2011).

Besides socio-demographic characteristics, Kruger (2009) added cultural differences that may explain the touristic demand.

The next section will explain the SEM model that runs the assessment of determinants of the demand for tourism.

---

5 Also, Vargas-Sánchez, Porras-Bueno, and Plaza-Mejía (2011) tested attitudes towards tourism development for the Spanish province of Huelva. They found that “behavior of tourists”, “density of tourists” and “level of tourism development perceived by the resident” are relevant for resident attitude.

Seiler, Hsieh, Seiler, and Hsieh (2002) used the SEM model to explain the determinants of expenditure for the Taiwanese tourist visiting the USA.

6 Mok and Armstrong (1995) found that age, occupation and income permit to have a different pattern of touristic demand. More educated people prefer specific activities. People with more income spend more, mainly in high-quality facilities.
3. Structural Equation Model of the demand for tourism

Structural Equation Model\(^7\) (SEM) is well known in the business literature (Jöreskog and Sörbom, 1982) it may capture relationship between qualitative variables and quantitative variables as well. In the literature of tourism Chaiboonsri and Chaitip (2010) run an analysis between tourist travel motivations (travel cost satisfaction and tourist demographics) and tourist destination (tourism product, tourism product attributes, and tourism product management).

Results for latter authors indicated that the travel cost satisfaction of international tourists had a positive influenced on tourism product attributes. Also, the travel cost satisfaction had a positive influence on tourism product management. The tourist demographics had a positive influenced on tourism product and tourist demographics, which had a positive influenced on tourism management.

Swanson and Horridge (2004) studied the Structural Equation Modeling to test the causal relationships between tourist travel motivations (travel activities and tourist demographics) and souvenir consumption (souvenir products, product attributes, and store attributes). The results indicated that the travel activities of tourists had positive correlations with souvenir consumption, while tourist demographics had no correlation.

Our model considers socio-economic variables and motive variables as well. We classify the motives into three different categories: Leisure, Health and Education and Other. The latter can include up to religious motives to travel such as conferences or events. Tables 1 and 2 describe some statistics of our variables.

Table 1

<table>
<thead>
<tr>
<th>Sex</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4293</td>
<td>63,39</td>
<td>63,39</td>
</tr>
<tr>
<td>2</td>
<td>2479</td>
<td>36,61</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>6772</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Source: Survey Promperu.

\(^7\) See also Bollen, Tueller, and Oberski (2013).
Table 2
Descriptive Statistics of Motives to Travel

<table>
<thead>
<tr>
<th>Motives</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vacation and leisure</td>
<td>4240</td>
<td>62,61</td>
<td>62,61</td>
</tr>
<tr>
<td>2. Visit friends</td>
<td>787</td>
<td>11,62</td>
<td>74,23</td>
</tr>
<tr>
<td>3. Business</td>
<td>1200</td>
<td>17,72</td>
<td>91,95</td>
</tr>
<tr>
<td>4. Seminars</td>
<td>257</td>
<td>3,8</td>
<td>95,75</td>
</tr>
<tr>
<td>5. Education</td>
<td>92</td>
<td>1,36</td>
<td>97,11</td>
</tr>
<tr>
<td>6. Health</td>
<td>49</td>
<td>0,72</td>
<td>97,83</td>
</tr>
<tr>
<td>7. Work</td>
<td>104</td>
<td>1,54</td>
<td>99,37</td>
</tr>
<tr>
<td>8. Mission/Religion</td>
<td>43</td>
<td>0,63</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6772</strong></td>
<td><strong>100</strong></td>
<td></td>
</tr>
</tbody>
</table>

Note: Motive 1, Leisure: 1, 2, and 3; Motive 2, Education and Health: 4, 5, and 6; Motive 3, Other: 7 and 8
Source: Survey Promperu.

Basically, the theoretical model is shown in Figure 3. We can see that we attempt to capture socio-economic variables and motives with touristic expenditure. We control for correlations between the variables. Among the different categories for the Structural Equation Model, we consider Path Model Analysis since we are not able to construct latent variables. The model in Figure 3 considers indicators of socio-economic, motive, and touristic expenditure variables. Since we have a mix of dummies, continue and ordinal variables then we must apply weighted Maximum Likelihood Estimation.

Figure 3. The Model of Touristic Expenditure
Source: Own elaboration

The next section will describe the data before introducing the results of the model presented here.
4. Data analysis

According to Tables 1 and 2, most of the people are males in the sample conducted by Promperu which is the main regulator of tourism in Peru. We have 6 772 respondents. Table 2 contains the motive variables which are classified into three different categories: Leisure, Health and Education and Other.

The answers ‘vacation and leisure’, ‘visit friends’ and ‘business seminars’ are considered as Leisure motive. The answers ‘education and health’ are Education and health motive, and ‘work mission/religion is considered as another motive.

According to the literature previously discussed, Mitchell and Reid (2001) and Dardis, Derrick, Lehfeld, and Wolfe (1981) found the significant and relevant influence of socio-economic variables in the touristic sector but the result contrasts with Swanson and Horridge (2004). Also, motive is widely discussed with a positive significant effect on tourism expenditure as well.

The available information of the survey run by Promperu allows us to consider the latter variables plus the touristic expenditure for our analysis.

5. Results of the Model

Table 3 shows some relevant results in terms of Fitness of the Model, the indicator of Standardized Root Mean Residual was below 0,10, so the fitness of the model8 describe above is almost perfect. Also, the Coefficient of determination was 0,19, low but we are not predicting the touristic expenditure but find its determinants. The model is estimated using a weighted maximum likelihood, so we have just these two tests to verify the fitness of the model. Since there is a mix of dichotomic and continuous variables, we use the later estimator.

Table 3
Fitness of the Model

<table>
<thead>
<tr>
<th>Fit statistic</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of residuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRMR</td>
<td>0,005</td>
<td>Standardized root mean squared residual</td>
</tr>
<tr>
<td>CD</td>
<td>0,193</td>
<td>Coefficient of determination</td>
</tr>
</tbody>
</table>

Note: The Model was fit with vce (robust); only stats (residuals) valid.
Source: Own elaboration

---

8 Adjustment of the Structural Model. We do not have many parameters of fitness since it is an estimation of dichotomic variables, and we used Maximum Likelihood estimation.
Also, Table 4 and Figure 4, shows the results of the structural model set up in the previous section of the paper. From Table 1 we can infer that almost all the variables are significant except for the constant where motive 3 is endogenous.

Figure 4 contains the weights for each of the relationships among the variables. The variables are not latent but indicators, so the model to be estimated in a Path Analysis Model instead of a pure Structural Model. All these weights are significant at 99% of confidence. As we have discussed before, the results are robust since the indicator of Standardized Root Mean Squared Residuals (SRMR)\(^9\) was low, which make us trust the result obtained in Table 4.

Touristic expenditure is explained by the three different motives: Leisure, Education and Health and Other. So foreign visitors may come following either one of the three models explained. The coefficient Leisure motive was the highest followed by Education & Health and Other Motive. This result goes along the results found by Chaitip et al. (2010) for a different sample set.

Also, as the people get older the expenditure is higher and the number of nights influences positively the touristic expenditure as well. The result goes along the line for a different sample set Dardis et al. (1981; 1994). There is more than one direction of causality in the model and the goodness of the Path Analysis conducted here is that we can capture all positive correlations and relationships. The Fitness of the model described permits us to trust this final specification as the best we can conduct among several different combinations.

The socio-economic variables ‘sex’ and ‘night’ explain the motives Leisure and Other, while age explains Education and Health and other motives to travel. Being male influences positively the three different motives to travel. Being older motive the motive Education and Health but there is a negative relationship with respect to other motives. The number of nights influence positively the variable Other Motive but negatively Leisure motive. Besides the relationship between motive variables and socioeconomic variables, it is interesting to look at the influence between socioeconomic and motive variables. There is not a homogeneous relationship between socioeconomic and motive to travel variables.

In addition, motives are correlated with each other negatively and significantly, so if the visitor comes for Leisure, it is also not necessarily expected than they come for a Business or Religious motive at the same time. The latter is an interesting result for marketing segmentation in order to find a better demand for visitors.

---

\(^9\) The SRMR is an absolute measure of fit and is defined as the standardized difference between the observed correlation and the predicted correlation. We have a sample bigger than 200 which is good for a Structural Equation or Path Analysis Model in our case (see Hu and Bentler, 1999, for a detailed explanation of Model Fitness). It is well used in non-complex models also, which is our case since we have a few available variables to perform the assessment.
### Table 4
**Structural Model Results**

| Variable                        | Coefficient | Std. Err. | z   | P>|z|  | [95% Conf.] |
|--------------------------------|-------------|-----------|-----|-----|----------------|
| **Motive Leisure (M1) <-**      |             |           |     |     |                 |
| Sex                            | 0.63        | 0.01      | 5.36| 0.00| *** 0.040 0.086 |
| Age                            | -0.03       | 0.02      | -1.31| 0.19| *** -0.063 0.013 |
| Nights                         | -0.06       | 0.02      | -3.74| 0.00| *** -0.097 -0.030 |
| Constant                       | 2.60        | 0.09      | 28.78| 0.00| *** 2.430 2.786 |
| **Tourist Expenditure <-**     |             |           |     |     |                 |
| Motive Leisure (M1)            | 0.23        | 0.02      | 9.26| 0.00| *** 0.178 0.273 |
| Educat. & Health (M2)          | 0.11        | 0.02      | 5.65| 0.00| *** 0.072 0.149 |
| Other (M3)                     | 0.06        | 0.02      | 2.87| 0.00| *** 0.020 0.105 |
| Age                            | 0.17        | 0.02      | 10.10| 0.00| *** 0.139 0.207 |
| Nights                         | 0.33        | 0.03      | 11.13| 0.00| *** 0.279 0.398 |
| Constant                       | -0.52       | 0.09      | -6.07| 0.00| *** -0.687 -0.352 |
| **Educat. & Health Motive (M2) <-** |             |           |     |     |                 |
| Age                            | 0.02        | 0.02      | 1.65| 0.10| * 0.005 0.063  |
| Constant                       | 0.19        | 0.05      | 3.81| 0.00| *** 0.094 0.292 |
| **Other Motive (M3) <-**       |             |           |     |     |                 |
| Sex                            | 0.04        | 0.02      | 2.51| 0.01| *** 0.009 0.071 |
| Age                            | -0.07       | 0.02      | -4.66| 0.00| *** -0.107 -0.044 |
| Nights                         | 0.18        | 0.04      | 4.55| 0.00| 0.105 0.264    |
| Constant                       | 0.10        | 0.06      | 1.63| 0.10| -0.021 0.227   |
| **Media <-**                   |             |           |     |     |                 |
| Sex                            | 2.84        | 0.00      | 651.52| 0.00| *** 2.830 2.847 |
| Age                            | 2.90        | 0.03      | 107.33| 0.00| *** 2.845 2.951 |
| Nights                         | 0.74        | 0.04      | 17.60| 0.00| *** 0.657 0.821 |
| **Variance <-**                |             |           |     |     |                 |
| Motive Leisure (M1)            | 0.99        | 0.00      |     |    | 0.987 0.997    |
| Tourist Expenditure            | 0.85        | 0.02      |     |    | 0.809 0.896    |
| Educat. & Health (M2)          | 1.00        | 0.00      |     |    | 0.997 1.001    |
| Other (M3)                     | 0.96        | 0.02      |     |    | 0.924 0.987    |
| Variable       | Coefficient | Std. Err. | z   | P>|z| | [95% Conf.] |
|----------------|-------------|-----------|-----|-----|-------------|
| Covariance <-  |             |           |     |     |             |
| M1 & M2        | -0.74       | 0.01      | -59.85 | 0.00 | -0.765 -0.716 |
| M1&M3           | -0.36       | 0.02      | -18.70 | 0.00 | -0.400 -0.324 |
| M2&M3           | -0.03       | 0.00      | -7.31  | 0.00 | -0.041 -0.024 |
| Age & Sex       | -0.02       | 0.02      | -1.41  | 0.16 | -0.059 0.010  |
| Sex&Nights      | 0.04        | 0.02      | 2.78   | 0.01 | 0.013 0.074  |
| Age&Nights      | -0.10       | 0.01      | -6.67  | 0.00 | -0.128 -0.070 |

Source: Own elaboration

**Figure 4. Structural Model**

Source: Own elaboration

**6. Conclusions**

This research provides some feedbacks for policymakers in the area of touristic segmentation. This is an interesting model brand new in the literature of Tourism Economics which can find the determinants of touristic demand in terms of socio-economic variables but considering the idea of several relationships in a conceptual map model.

Structural Equation Model is not widely used in this area, we have estimated a Path Analysis which is a variety of a Structural Equation Model. Despite there is a lack of information, the results obtained in the research permit us to make some inference...
about the relationship between the described variables. The goodness of fitness was clearly robust, and we can infer that motives of leisure, business and other influence expenditure of the tourist. More weight can be attributed to leisure and socio-economic variables like age, the number of nights and sex also influence the expenditure of the tourist as well.

This paper can be used for marketing segmentation\(^{10}\) in our database of foreigners visiting Peru for the most recent period. Further investigation may compare if there is a structural change among these periods of time. Also, our limitation is that the database considered a survey for foreigners but not domestic visitors.

7. References


Finding the relevant motives and socio-economic variables is important for touristic policies.


