Tourism Management 43 (2014) 36–45

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Determinants of individual tourist expenditure as a network: Empirical findings from Uruguay

Antonino Abbruzzo [a], Juan Gabriel Brida [b,*], Raffaele Scuderi [b]

[a] Dipartimento di Scienze Economiche, Aziendali e Statistiche, University of Palermo, Viale delle Scienze, ed. 13, Palermo, Italy
[b] School of Economics and Management and TOMTE, Free University of Bozen-Bolzano, Piazza Università 1, Bozen-Bolzano, Italy

HIGHLIGHTS

- We use graphical models to analyze tourist expenditure in Uruguay.
- Trip-related variables are important determinants of spending.
- Socioeconomic variables are not linked to spending directly.
- Each expenditure item is analysed further via logistic regression.

ARTICLE INFO

Article history:
Received 3 December 2012
Accepted 16 January 2014

Keywords:
Tourist expenditure
Log-linear graphical model
Uruguay

ABSTRACT

This paper introduces the use of graphical models for assessing the determinants of individual tourist spending. These models have the advantage of synthesizing and visualizing the relationships occurring within large sets of random variables, through an easy to interpret output. To this end, individual data from a large official survey of international tourists in Uruguay are used. Symmetric conditional independence structures are first investigated. Then subgraphs of each expenditure item's neighbourhood are extracted in order to assess the impact of main effects and interactions through proportional ordinal logistic regression. Results highlight the marginal role of socio-demographic variables and direct importance of accommodation type, destination and length of stay.

1. Introduction and background

In modern economies the attention to tourism has increased noticeably, because of its impact and influence on economic growth (Brida, Lanzillotta, Lionetti and Risso, 2010). Accordingly the interest towards the analysis of demand and its determinants has constantly risen. This is testified by different review papers of theoretical and empirical literature – see among others Divisekera (2014, 2013); Brida and Scuderi (2013); Wang and Davidson (2010); Divisekera (2010); Song and Li (2008); Lim (1997); Crouch (1994). Detection and analysis of the factors influencing individual spending constitutes a significant subset of them. Indeed, assessing the role of a set of variables in affecting individual decisions to consume can be crucial in order to evaluate how likely certain drivers can boost tourism in a destination. All this can have substantial implications on the actions of public and private actors. However papers investigating the determinants of tourist expenditure at individual level have been relatively more limited than the ones on aggregate demand. This regarded both the number of contributions, and heterogeneity of the applied statistical and econometric models (Brida & Scuderi, 2013; Wang & Davidson, 2010). Typically, these approaches have used classic regressions for the sake of quantifying the significance and impact of each regressor from a set on the response variable (Brida & Scuderi, 2013).

This paper contributes to the strand of the literature on the determinants of tourist expenditure. We adopt a different and complementary approach to classic regressions. We introduce the use of graphical models as a way to represent tourism consumption behaviour, seen as a network of variables (Edwards, 2000; Lauritzen, 1996; Whittaker, 1990). We then apply regression models to variables selected from the graph structure, in order to quantify their impact on each expenditure item. The focus of this paper is the analysis of incoming tourists to Uruguay. This country represents an interesting case study for tourism. Uruguay is the second smallest country in South America. However it is ranked first in terms of per capita tourists per year, with tourist flows

* Corresponding author. Tel.: +39 0471 013492; fax: +39 0471 013009.
E-mail addresses: antonino.abbruzzo@unipa.it (A. Abbruzzo), JuanGabriel.Brida@unibz.it, elbrida@gmail.com (J.G. Brida), Raffaele.Scuderi@unibz.it (R. Scuderi).

0261-5177/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tourman.2014.01.014

accounting for about 80% of the total population, and this percentage has increased over time (Brida, Pereyra, & Scuderi, 2014). The analysis was conducted on an official dataset from a direct survey published by Ministerio de Turismo y Deporte (2011).

2. Background

2.1. The analysis of individual tourist expenditure

In the tourism literature, econometric models for the analysis of individual consumption can be classified in two categories (Brida & Scuderi, 2013). The first one is based on the Engel curve representation of the demand for tourism goods. It models expenditure level as function of socio-demographic, trip-related, psychographic variables, plus budget constraints. The use of OLS estimations has been very frequent (Jang, Bai, Hong, & O'Leary, 2004; Kozak, 2001; Marcussen, 2011). However other techniques were used in order to address specific estimation issues. A set of them aimed to properly handle expenditure as zero-censored variable, as well as the self selection bias mechanism between spenders and non spenders. These include Tobit models (Bilgic, Florkowski, Yoder, & Schreiner, 2008; Cai, 1998; Lee, 2001) and two-part approaches (Brida, Disegna, & Scuderi, 2013a; Brida, Fasone, Scuderi, & Zapata-Aguirre, 2013b; Jang & Ham, 2009; Jang, Ham, & Hong, 2007; Nicolau & Más, 2005). More recently, quantile regression was used (Chen & Chang, 2012; Thrane & Farstad, 2011). Other alternative approaches were offered, among others, by Alegre, Mateo, and Pou (2010), who performed switching regressions to study the spending behaviour of first time and repeat visitors; and the recent study by Zhang, Zhang, and Kuwano (2012), who presented an integrated model of tourist expenditure behaviour that mixes information on time use and the influence of destination visit behaviour, and applies a Type-II Tobit model with copula functions.

A second category of works is based on random utility models. They assess the characteristics that likely influence the probability of an individual to consume tourism goods, in presence of selection process between tourism and other goods. Logistic regression models (Alegre et al., 2010; Saayman & Saayman, 2006) have been the most frequently used ones. A couple of contributors applied also multinomial models, as in the choice experiment of Crouch et al. (2007) on tourism as discretionary spending.

2.2. Graphical models and their main characteristics

Graphical models allow to identify the patterns of conditional independences within a wide set of variables, that is whether significant interrelationships exist. They are particularly suitable to handle large sets of random variables via an easy to interpret output: a graph displays whether variables, represented as *nodes*, are connected each other through *links*. The missing link between a given couple of nodes is also useful in order to understand the role of certain variables relatively to others, from their location and links to certain nodes of the network. In other words, we can deduce whether direct or indirect connections are likely to occur between variables, and identify those nodes that 'mediate' the direct relationship between couples of variables. In addition, similarly to multivariate regressions different response variables can be handled simultaneously. In this paper, such feature allows to investigate whether each category of spending is related to specific nodes, conditionally to other types of expenditure.

Compared to regression models they serve different objectives. Regressions aim to 'quantify' whether and to what extent variables and their interactions affect the dependent variable, conditionally to others. Instead the goal of graphical models is to assess 'if' interrelationships between couples of nodes exist conditionally to the entire network of nodes. The great majority of papers on tourism expenditure rarely put into account the assessment of interrelationships. To this end, only very few studies included selected second order interaction terms as regressors (Roehl & Fesenmaier, 1995; Tyrrell & Johnston, 2003). After all, selecting interactions is everything but a trivial matter in presence of a relatively large dataset, and much is left to the researcher's choice. In fact the inclusion of all possible couples of second order terms from K regressors would imply adding a number of $[K(K-1)]/2$ right-hand variables. Instead, graphical models do not require to exclude any interaction, but they allow to represent only significant ones. In this sense they can be very useful while selecting regressors and interaction. In a few words, using graphs to represent models has the effect of shifting the emphasis from parameters estimation towards assessment of the model structure. Also, unlike regressions where estimation problems may arise when including interactions even in a moderately large dataset, graphical models are computationally feasible.

Another difference with regressions resides in the 'objects' each of the approaches handles. The nodes of a graph are a generalization of a regression's variables. Think for instance about categorical variables with a set of modalities greater than two. In regressions, in order to avoid collinearity one must use a number of dummy variables that is equal to the number of modalities, minus one reference category. Instead, a node of a graphical model can handle a whole politomic variable, with no need to separate its modalities, nor select one reference category.

Unlike many econometric models, basically graphical models have no underpinning behavioural theoretical foundation. The interpretation of results is limited to mere reporting of conditional probabilities as graphs. Nevertheless, they may emphasize the complexities underlying the consumption behaviour, seen as a system of explicit interactions.

3. Dataset description

The Ministry for Tourism and Sport of Uruguay conducts a quarterly survey on non-resident visitors entering the country for purposes different than migration and working. Data are recorded by trained pollsters at the end of tourists' journey in the main exit points of Uruguay. The sampling adopts a stratified design according to the *apriori* available information at each exit point. The present study considers the four 2010 surveys (Ministerio de Turismo y Deporte, 2011). The overall number of used individual records is 6498. The complete list of socio-demographic, trip-related, and expenditure items is reported in Appendix A.

The sample is mainly composed of people residing in Argentina (71.98%) and Brazil (14,39%), who self-organized their trip (96.32%). Many tourists repeated their visit (Brida et al., 2014): only 6.15% came for the first time, whereas 48.42% declared they have already visited Uruguay six times or more, including the one of the interview. The main mean of transportation to enter the country was maritime (43.13%), which once more supports the idea of the importance of tourism from neighbour countries, and in particular of Argentinians crossing Rio de la Plata river. The most frequent length of the visit ranged from 4 to 7 days (40.86%), with a notable percentage of interviewees (about 25%) who spent more than 7 days in the country. The main place of visiting was Montevideo (37,04%), followed by Punta del Este (26.06%) and a set of inland places (15.79%). Groups of tourists were most frequently formed by two members (45.98%), three (14.15%) and four (15.00%). Leisure (57.31%) was the main motivation of the trip, although visitors came prevalently also for the sake of meeting their families (22.85%) and doing business (10.16%). Hotels (44.85%) and relatives'