Photo or figure (optional)

Stochastic frontier estimation of holiday budgets for Multiple Discrete-Continuous Extreme Value model: An application to tourist expenditure analysis

Andrea Pellegrini, Università della Svizzera italiana
Igor Sarman, Università della Svizzera italiana

Conference paper STRC 2017

STRC

17th Swiss Transport Research Conference
Monte Verità / Ascona, May 17-19, 2017

Stochastic frontier estimation of holiday budgets for Multiple Discrete-Continuous Extreme Value model: An application to tourist expenditure analysis

Andrea Pellegrini
Università della Svizzera
italiana
city

Phone: +41 58 666 47 85
Fax: +41 58 666 46 62
email:
andrea.pellegrini@usi.ch

May 2016

Abstract

Understanding the dynamics of visitor expenditures represents one of the main strengths that a tourist destination may hold and this is particularly true considering the evolution of the tourism industry characterized by cheap flights, do-it-yourself experience and sharing of information and experiences on social media.

Traditional analyses on this subject focus on two approaches: the first one does not distinguish among expenditure categories; in the second case, when this distinction is considered, the approach only focuses on one particular category at a time.

Given that tourists take multiple, non-mutually exclusive decisions of expenditure allocation, we implement a Multiple Discrete-Continuous Extreme Value (MDCEV) model to understand tourist expenditure behavior. Specifically, we propose an empirical framework able to model simultaneously both the tourists' decisions to distribute their expenditures on different activities and products and the amount of money spent on each category of expenditure.

We also contribute to the existing literature by deploying a stochastic frontier approach to estimate holiday budgets for the MDCEV model. This approach is essential when the underlying money budgets addressing a situation choice are unobserved, but the expenditures on the choice alternatives of interest are observed.

Keywords

Stochastic Frontier – Multiple Discrete Continuous Model – Tourism consumption

1. Introduction

The analysis of tourist expenditures has captured the interest of researchers and analysists in recent years. Understanding the dynamics of expenditures represents one of the main strengths that a tourist destination may hold and this is particularly true in the modern life of the tourism industry, which is living an age of evolution characterized by cheap flights, do-it-yourself planning and sharing of information and experiences on social media. Tourism literature that considers the topic of visitors' expenditure is varied in terms of approaches, frameworks and evidences; it is usually grounded on microeconomic theories and microeconometric practices and has evolved over the years, taking into account the high complexity of the tourist product.
If tourist decision-making is framed as a sequence of choices, one recognizes that the phase of "expenditure" is just one of a set of interdependent decisions that a tourist takes. If we dig a bit deeper, we can also assess that the expenditure phase is a collection of decisions, all of which are aimed at allocating money to diverse activities (accommodation, food and beverage, transportation and the like). This article contributes to the stream of literature aimed at analyzing tourist expenditures from a microeconometric perspective and responds to the necessity called by other authors to adopt appropriate methodologies (Brida and Scuderi, 2013; Rashidi and Koo, 2016). Many scholars assessed the fragilities inherent to simplistic models (i.e. ordinary linear regression and the like) when analysing expenditure choices and, more in general, individual behavior that implies an interaction between two or more decisions. We adopt a MCDEV model to assess two types of decisions which are interrelated and sequential: the decision concerning the money allocation on a category of expenditure and the decision concerning the amount of money to allocate on the same category of expenditure.
One further important aspect touched in this article is the consideration of the tourist budget. Money spent on tourist activities comes from individual monetary availability and it is commonly assumed that an agent (an individual or a group of individuals) sets an expense budget and, during the course of the holiday, this budget is allocated to purchase activities, services and products. In this setting, an issue arises as it is complicated (if not impossible) to observe what is the actual maximum amount of money an agent predetermines to spend and hence the tourist budget remains latent. In this work an application of a Stochastic Frontier (SF) estimation of travel budget is proposed to identify an unbosrved budget frontier and quantify what can be considered the maximum possible amount of money an individual is willing to allocate in tourism activities.
The consideration of the *frontier* is strictly connected to the analysis of the multiple discrete-continuous choices an individual takes when allocating money to tourist activities. The analysis of tourist expenditures proposed in this article follows a two-stage design: first, a stochastic frontier regression model is adopted in order to estimate the *frontier* and, in a second moment, a MCDEV model is formulated in order to analyse the determinants of expenditure decisions

and budget allocation over different expense categories. To the best of authors' knowledge, this is the first application of a MCDEV model enriched by a SF estimation of travel budget in tourism studies.

The remainder of the article presents an extensive overview of the SF modelling and the formalization of the MCDEV model. Next, the empirical context of the present research is presented with a description of data and data collection process. It follows a description of results emerging for empirical analysis and the discussion and, finally, the concluding chapter provides a summary of the study and some solutions for future research.

2. The consumer behaviour theory

Several tourist decisions are characterized by the simultaneous choice of multiple alternatives. An example of multiple discreteness is a situation in which a traveller might decide to either participate in multiple leisure activities within a given time or allocate his/her holiday budget to various expenditure categories, such as, accommodation, transportation, souvenirs and food & beverage. In addition to such multiple discrete choices, a consumer normally makes continuous decisions on how much of each selected alternative to consume.

The multiple discrete continuous (MDC) choices have received an increasing attention for the analysis of consumer behaviour in a variety of social sciences, including marketing, psychology and transportation economy. The most popular approach is based on the classical microeconomic consumer theory of utility maximization. Such approach assumes that the direct utility function $U(\boldsymbol{q})$ is quasi-concave, increasing and continuously differentiable with respect to the consumption quantities $\boldsymbol{q} = (q_1, q_2, \ldots, q_K)$. Consumers maximise their utility function over a set of non negative quantities $\boldsymbol{q}$ subject to a linear budget constraint, as below:

$$\max U(\boldsymbol{q}) \; such\ that \sum_{k=1}^{K} p_k q_k = y \; and \; q_k \geq 0 \; \forall k = 1,2,..K$$

Wales and Woodland (1983) propose a particularly attractive approach for deriving the demand function from the above equation, which is based on the Kuhn Tucher or KT (1951) first-order conditions for constrained random utility maximization. With this approach, they assume that the utility function $U(\boldsymbol{q})$ is randomly distributed over the population and then derive the optimal consumption quantities $\boldsymbol{q}^*$ for the random utility model by applying the KT conditions. When the utility function is assumed to be randomly distributed, the KT conditions become stochastic and provide the framework for obtaining the probability expression of the consumption patterns. Nevertheless, its applicability has been rare until recently because of a complicated likelihood function that implies an intensive data exercise and a multidimensional integration. Bhat (2005) addresses this issue by developing the MDCEV model whose probability function is tractable and can be entailed even for situations with a broader set of discrete consumption

alternatives. In addition, such empirical model includes explicitly a way to deal with the diminishing marginal returns when an alternative is consumed.

Despite the improvements in the use of KT approach for analysing the MDC choices, only recently the literature in the area of consumer behaviour has started to deeply investigate the role of budget in the constrained random utility maximization. More specifically, almost all studies that adopt the KT approach, including the MDCEV, consider the available budget for total expenditure to be fixed for each consumer. When the budget is fixed, an eventual variation due to socio-demographic characteristics or changes in the alternative specific attributes can only lead to a reallocation of the budget among diverse choice alternatives. Hence, the KT formulation does not allow the total available budget to either increase or decrease. For example, in the context of households' vehicle ownership and utilization (Bhat, 2009), the observed total annual mileage is assumed to be the available budget for each household. Therefore, any change in fuel price or in household features can only lead to a reallocation of the observed total budget among the transportation alternatives.
Similarly, the sum of diverse expenditure categories (such as, food, transportation, accommodation, souvenirs) is assumed to be equal to the total disposable budget. The KT framework does not allow any variation in the total budget due to changes in price of tourist expenditure categories. In such a situation, the researcher can only observe the alternatives chosen by the decision-maker together with the amount consumed of each of those but cannot observe the maximum amount of expenditure the consumer is willing to allocate to each alternative. One understands that assuming that the actual total tourist expenditure or the total annual mileage represent the entire budget available to a consumer is not realistic.

To overcome this issue, the literature considers the two-stage budgeting approach along with the assumption of separability of preferences across a specific number of consumption categories and homothetic preferences within each category (Chintagunta and Nair, 2010)

The two stage budgeting approach allows, in the first stage, to allocate the total expenditure to broad groups of consumption goods, and in the second stage to distribute the group expenditures to single commodities. The allocation process needs that its result is equal to what would happen if the allocation was made in one stage only. The main strength of this approach is its flexibility in a sense that it can be extended (more than two stages) and modified in several ways.
With the separability assumption one considers the final utility as a function of as many subutilities as the number of groups of consumption goods. This can be only done when the different bundles within each consumption group can be ranked independently of consumption levels observed for the other groups.

2.1 The unobserved budget

In this study, we propose to estimate the unobserved budget by applying the stochastic frontier regression approach. The stochastic frontier regression is normally adopted in firm production economy to identify the maximum production capacity that can be obtained from any given input (Aigner et all, 1977, Kumbhakar and Lovell, 2000). Typically, we observe the production level of a firm given a vector of inputs, but unobserved production frontier is assumed to exist. Such frontier represents the maximum production level that a firm is capable to achieve given the observed inputs. In the context of out-of-home activities and time-use (Augustin, 2014), the stochastic frontier is performed on the observed total out-of-activity time to determine the unobserved out-of-home time frontier. Such frontier represents the maximum amount of time that people are willing to allocate to out-of-home activities in a day.

Adopting a two stage budgeting approach, in the first stage, a stochastic frontier regression is performed on the observed total expenditure to estimate the maximum expenditure level that tourists are willing to allocate. Specifically, the estimated budget is used to determine the *outside good*, that is, the difference between the *frontier* and the actual total expenditure on all the *inside goods*. In the second stage, we perform a MDCEV in which the expenditure categories are all the choice alternatives of interest (*inside goods*) plus the *outside good*. The role of the *outside good* is to either supply the additional resources for further purchase or save the unspent portion of the latent budget.

The decision to adopt the stochastic frontier over the classical empirical model, such as, the linear or log linear regression model, is due to the fact that the *frontier*, by definition, is greater than the observed total expenditure. Therefore, the estimated budget allows the total expenditure to either increase or decrease because of changes in alternative specific attributes or tourist characteristics.

3. Data and sample description

Data adopted in this work comprises the tourist expenditure categories as well as the type of the travel and the socio-demographic characteristics of traveller conducted in 2012 in the Canton Ticino, Switzerland. Expenditure information collected in the survey is used to determine nine different categories – accommodation, transportation (e.g. fuel, public transport tickets, rental car cost), purchase of souvenirs, clothes and accessories, boat trip, events, sport activities, mountain transportation (e.g. funicular railway, cable railway) , food & beverage and a residual one.

We found 282 eligible tourists which reported at least one of the expenditure categories. The vast majority of respondents are men (70%) and the average age is around 47 years. 54% of the

sample chose to get to their destination in Canton Ticino by private transportation, including car, rental car and camper whereas 46% decided for a public service such as airplane, taxi, bus and train. On average, tourists in the sample allocated money in 4.5 expenditure categories; 18% allocate their budget in three expenditure categories, 32% in four categories and 44% in at least five categories. The majority of individual spent on accommodation and food & beverage, followed by transportation, souvenirs, clothes and accessories, mountain transportation, boat trip, sport activities, events and other. These considerations call for the use of the multiple discrete-continuous modelling approach for investigating tourist expenditure behaviour.

4. Preliminary results

This section reports preliminary results obtained from applying the MDCEV model to Swiss tourist expenses. The MDCEV model is characterized by two vectors of parameters to be estimated: delta and gamma. Specifically, delta parameter reflects the baseline marginal utility, or the marginal utility at the point of zero consumption. Between two choice alternatives, the alternative with greater baseline marginal utility is more likely to be chosen. By contrast, gamma parameter allows corner solution (the chance of not selecting an alternative) and differential satiation effects (diminishing marginal returns with increasing consumption) for diverse expenditure categories.

All gamma parameters are significantly different from one, implying that tourist decisions on how to allocate their budget simultaneously consider different categories of expenditure. The results indicate that tourists have higher propensities to purchase souvenirs, clothes, accessories and food & beverage. On the other hand, they are less willing to allocate their budget to sport activities and mountain transportation. Among the household sociodemographic characteristics, the estimates show female are more likely to allocate budget to shopping and souvenirs compared to men. With respect to delta parameter, the coefficient related to the duration of the journey indicates that the higher the number of nights at destination the greater the probability of allocating a larger portion of budget to accommodation, transportation and food & beverage categories. The purpose of trip results to be statistically significant for both discrete and continuous stage of the model. More specifically, tourists who travel for a holiday trip are less inclined to spend money on accommodation than those who travel for business.

5. Conclusions and future advances

This article aims at enhancing our understanding of tourist expenditure behavior by overcoming the above limitations of the MDCEV model. In particular, the empirical results demonstrate the importance of estimating an unobserved budget through the application of the stochastic frontier

regression. Such approach could be applied to estimate budgets for many empirical applications involving MDCEV choice analysis, including travel mode choice and usage, or tourist activities and free time allocation. In terms of policy implications, our findings provide evidences about the determinants driving tourist budget allocation. A tourist destination might design suitable packages for specific tourist segments by combining, for example, transportation and accommodation, or accommodation and activities in order to incentivize them to spend a larger part of their holiday budget. Finally, a broader knowledge on tourist expenditure behavior could also guide tourist destinations in marketing and promotion activities for specific expenditure categories.

6. References

Aigner, Dennis, Lovell, C.A.Knox, Schmidt, Peter, "Formulation and estimation of stochastic frontier production function models", Journal of Econometrics, vol.6

Augstin, B., A.R. Pinjari, A. Faghih Imani, V. Sivaraman, N. Eluru, and R.M. Pendyala. Stochastic Frontier Estimation of Budgets for Kuhn-Tucker Demand Systems: Application to Activity Time-Use Analysis. Forthcoming, Transportation Research Part A.

Ben Akiva, Moshe, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire et al. (2002) Hybrid choice models: progress and challenges, Marketing Letters, 13, 163-175

Bhat, C.R. (2005), "A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions," Transportation Research Part B, Vol. 39, No. 8, pp. 679-707

Bhat, C.R. (2008), "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions," Transportation Research Part B, Vol. 42, No. 3, pp. 274-303

Bhat, C.R., and N. Eluru (2009), "A Copula-Based Approach to Accommodate Residential Self-Selection Effects in Travel Behavior Modeling," Transportation Research Part B, Vol. 43, No. 7, pp. 749-765

Brida J.G., Scuderi R. (2013), 'Determinants of tourist expenditure: A review of microeconometric models', Tourism Management Perspectives, Vol 6, pp 28–40.

Duncan, Gregory M. (1980), Formulation and Statistical Analysis of the Mixed, Continuous/Discrete Dependent Variable Model in Classical Production Theory, Econometrica, Vol 48, No. 4, pp. 839-852.

Gate, Martin. F.m Gurgand, Marc (2007), Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons, Journal of economic surveys. Vol. 21, No.1

Hanemann, W. Michael (1984) Discrete/Continuous Models of Consumer Demand, Econometrica, Vol. 52, No. 3, pp. 541-561.

Heckman, James J. (1979), Sample Selection Bias as a Specification Error, Econometrica, Vol. 47, No. 1, pp. 153-161.

Hossein Rashidi T., Koo TT R., 2016, 'An analysis on travel party composition and expenditure: A discrete-continuous model', Annals of Tourism Research, vol. 56, pp. 48 - 64

Lee, Lung-Fei, Maddala, G.S., Trost, R.P. (1980) Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Model with Selectivity, Econometrica, Vol 48, No.2, pp. 491-503.

Lingling, Wu, Junyi Zhang, Fujiwara, Akimasa (2013), Tourism Participation and expenditure behavior: analysis using a scobit based discrete-continuous choice model, Annals of Tourism Research, Vol.40, pp- 1-7.

Lovell CAK and SC Kumbhakar (2000). Stochastic Frontier Analysis. NY, USA, Cambridge University Press.

Morikawa, Taka, Moshe Ben-Akiva and Daniel McFadden (2002) Discrete choice models incorporating revealed preferences and psychometric data, Econometric Models in Marketing, 16, 29-55

Oppermann, M. (1997), First-time and repeat visitors to New Zealand, Tourism Management, 18, 177-181.

Pinjari, A.R., and C.R. Bhat (2010), "A Multiple Discrete-Continuous Nested Extreme Value (MDCNEV) Model: Formulation and Application to Non-Worker Activity Time-Use and Timing Behavior on Weekdays," Transportation Research Part B, Vol. 44, No. 4, pp. 562-583

Pinjari, A.R., C.R. Bhat, and D.A. Hensher (2009), "Residential Self-Selection Effects in an Activity Time-use Behavior Model," Transportation Research Part B, Vol. 43, No. 7, pp. 729-748

Pradeep K. Chintagunta & Harikesh S. Nair, 2011. "Structural Workshop Paper --Discrete-Choice Models of Consumer Demand in Marketing," Marketing Science, INFORMS, vol. 30(6), pages 977-996, November.

Song, H., Li, G., (2010). Tourism demand modelling and forecasting: how should demand be measured? Tourism Economics, 61: 63-81.

Srinivasan, S., and C.R. Bhat (2006), "A Multiple Discrete-Continuous Model for Independent- and Joint- Discretionary-Activity Participation Decisions," Transportation, 2006 TRB Special Issue, Vol. 33, No. 5, pp. 497-515

Sung, H. H., Morrison, A. M., Hong, G.-S., & O'Leary, J.T. (2001), The effects of household and trip characteristics on trip types: A consumer behavioral approach for segmenting the US domestic leisure travel market. Journal of Hospitality & Tourism Research, 2581), 46-68.

Train, Kenneth E. (1986), Qualitative choice Analysis: Theory Econometrics, and an Application to Automobile Demand, MIT Press, Cambridge.

Train, Kenneth E. (2003), Discrete Choice Methods with Simulation, MIT Press, Cambridge.

Wales, T. J. & Woodland, A. D., 1983. "Estimation of consumer demand systems with binding non-negativity constraints," Journal of Econometrics, Elsevier, vol. 21(3), pages 263-285, April.