Transport Policy 18 (2011) 347–357

Contents lists available at ScienceDirect

Transport Policy

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

An exploratory analysis of relationships between socioeconomic, land use, activity participation variables and travel patterns

C.S. Pitombo [a,*], E. Kawamoto [b], A.J. Sousa [a]

[a] CERENA, Mining Engineering Department of Instituto Superior Técnico, Technical University of Lisbon (TULisbon), Avenida Rovisco Pais, 1049-001 Lisbon, Portugal
[b] Department of Transportation, Engineering School of São Carlos, University of São Paulo, Avenida Trabalhador SãoCarlense, 400 Centro, CEP 13566-590, São Carlos, São Paulo, Brazil

ARTICLE INFO

Available online 10 December 2010

Keywords:
Exploratory Multivariate Data Analysis techniques
Land use
Socioeconomic characteristics
Activity participation

ABSTRACT

This paper aims to find relations between the socioeconomic characteristics, activity participation, land use patterns and travel behavior of the residents in the São Paulo Metropolitan Area (SPMA) by using Exploratory Multivariate Data Analysis (EMDA) techniques. The variables influencing travel pattern choices are investigated using: (a) Cluster Analysis (CA), grouping and characterizing the Traffic Zones (TZ), proposing the independent variable called Origin Cluster and, (b) Decision Tree (DT) to find a priori unknown relations among socioeconomic characteristics, land use attributes of the origin TZ and destination choices. The analysis was based on the origin–destination home-interview survey carried out in SPMA in 1997. The DT application revealed the variables of greatest influence on the travel pattern choice. The most important independent variable considered by DT is car ownership, followed by the Use of Transportation "credits" for Transit tariff, and, finally, activity participation variables and Origin Cluster. With these results, it was possible to analyze the influence of a family income, car ownership, position of the individual in the family, use of transportation "credits" for transit tariff (mainly for travel mode sequence choice), activities participation (activity sequence choice) and Origin Cluster (destination/travel distance choice).

1. Introduction

One of the most important topics in travel demand analysis concerns the relation between city configuration, the personal needs to fulfill the dispersed activities in the urban environment, the place where people live, individual and household characteristics and travel behavior (Lu and Pas, 1999; Bowman and Ben-Akiva, 2000; Bhat and Singh, 2000).

People generate complex travel patterns in course of their daily activities, depending on the spatial urban distribution and socioeconomic variables. Recognizing the importance of the contextual characteristics in the generation of trip-chaining patterns, it was suggested that the activity-based travel demand approach must be rooted in the urban characteristics, besides individual attributes (Kwan, 2000; Sun et al., 1998).

The idea that the individual characteristics, household structure and activity participation can affect travel patterns choices has been developed in recent years (Arentze et al., 2000; Balasubramaniam and Goulias, 1999). The affirmation that travel behavior can be influenced by variables such as income, person role in the family, household structure, gender and household size predominates in the literature (Srinivasan and Athuru, 2005; Simma and Axhausen, 2001).

Land use variables form a part of a factors set that affect individual decisions and travel schedule. The complexity of the interrelation between the land use attributes and travel patterns has been examined in a large body of work, generating inconclusive results (Kitamura et al., 1997; Giuliano, 1995; Cervero and Radisch, 1996; Boarnet and Sarmiento, 1996). The incorporation of such variables to the study of travel behavior involves considerable difficulties such as the following:

(A) How to represent or measure the land use? There is a wide range of parameters such as density, diversity and design among the others. The main difficulty is related to the finding of an adequate way to represent the independent land use variables (numeric, categorical, dummy), searching the relations for the dependent variables/travel behavior (Van Acker et al., 2007).
(B) How can these variables be used for the estimating models? It is essential to represent the land use variables in a reasonable way and to use the appropriate statistical methods to find these relationships. An inadequate approach can generate different unbased findings about land use and travel behavior relations. Thus, the utility of land use variables in estimating models depends on the methodological decisions. Taking into account that the urban characteristics influence travel choices, the incorporation of such variables into trip generation models or travel forecast models can improve the travel behavior analysis (Purvis, 1998; Boarnet and Crane, 2001).

* Correponding author.
E-mail address: cirapitombo@gmail.com (C.S. Pitombo).

0967-070X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tranpol.2010.10.010

The main aim of this work is to find relations between socioeconomic characteristics, activity participation, land use and travel patterns of the residents in the São Paulo Metropolitan Area (SPMA) by application of Exploratory Multivariate Data Analysis (EMDA) techniques. The EMDA techniques used in this paper were: (A) Cluster Analysis (CA), grouping and characterizing the Traffic Zones in SPMA, proposing the independent variable called Origin Cluster and (B) Decision Tree (DT) to find *a priori* unknown relations between the socioeconomic characteristics, land use attributes and destination choices.

It was decided to use exploratory techniques with objective to discover hidden patterns and structures in the database used. These techniques do not estimate parameters; however, they allow exploring analytical models that are not necessarily based on the inference statistics. Therefore, the application of the EMDA techniques is justified by the proposal of the present work dedicated to the exploratory analysis of relations between the variables.

Considering the aforementioned, people seem to choose different destinations depending on their personal and household characteristics and the attributes of origin Traffic Zone (TZ) and other zones—near and distant from their residence. Therefore, groups of people who live in different places, with diverse individual and household characteristics, can make different travel patterns. This study will attempt to improve the understanding of these factors by answering some general and specific research questions:

General questions:

- Which is the predominant travel pattern associated with a determined group of individuals?
- How is the trip sequence related to the land use patterns of the place where people live and to socioeconomic and activity participation variables?

Specific questions:

- Which is the individual's travel trend considering destination choices and travel distance—to fit all the trips in the same TZ (shortest distances), or to vary the destination TZ considering the accessibility or configuration of urban activities?
- Is there any need for the people who live in TZ of high commercial or industrial activity level to change their TZ to accomplish different activities? Is it true that the individuals living in TZ with low activity level are generally going for work to the different region?

As a background, Section 2 organizes and reviews the pertinent literature. Section 3 describes the EMDA techniques used in this work. The application of CA is described in Section 4. Section 5 summarizes the data used in this paper. In Section 6 the travel patterns are coded considering the activities sequence, the travel mode sequence and the destination/travel distances sequence. Finally, the following sections present the applications of DT, discussion and conclusions.

2. Linking individual and household characteristics, land use and travel patterns

Socioeconomic characteristics are strongly related to the individual behavior in a general way. Some variables (for example, family income) provide an appropriate support to investigate individual behavior, in particular related to the travel. Some success was achieved in relating the individual's sociodemographic and location characteristics to his or her travel pattern (Hanson and Hanson, 1981; Mitchell and Town, 1977; Pas, 1984). Many literature sources report on the complexity and variety of travel patterns in accordance with individual and household characteristics, showing the important relations between the household structure, gender, car ownership, family income and travel behavior (Mcguckin and Murakami, 1999; Sarmiento, 1996; Bhat and Koppelman, 1993).

However, it must be pointed out that the socioeconomic characteristics and activity participation are only a part of a variable set determining the travel behavior. Car ownership can predominantly define travel mode sequence, and the activity level in Traffic Zones can explain destination choices.

Recognizing the importance of including the variables considering the urban configuration into the travel patterns choices, some recent publications paid special attention to the land use variables searching for the most realistic representation of trip-chaining behavior. Despite the discrepancies between their findings, the investigation of the connection between land use and travel patterns assists in the construction of adjusted models for transportation demand forecast (Ewing, 1995; Golob and Bownstone, 2005).

Kitamura (1985) examined the relations between the trends of the individuals towards the chain trips and characteristics of a hypothetical linear city. The analysis showed that chain trips greatly depend on the utility of opportunity set and the type of activity to be performed. Srinivasan (2000) investigated the influence of neighborhood characteristics (land use, transportation network and accessibility measures, quantified with a GIS) on travel behavior related to modal choice and trip-linking.

Based on the literature, it is reasonable to assume that travel pattern choice is a function of the people's need to participate in dispersed activities in the urban environment, individual and household characteristics, including a set of options and constrains. Fig. 1 represents a framework proposal of a variable set that influences travel patterns choices.

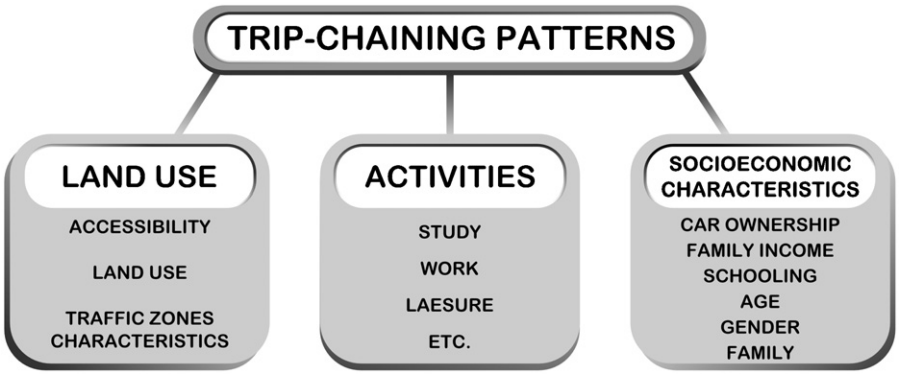

Fig. 1. Suggested set of variables influencing the trip-chaining patterns choice.