Transportation Research Part E 61 (2014) 68–83

Contents lists available at ScienceDirect

Transportation Research Part E

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

A latent class choice based model system for railway optimal pricing and seat allocation

Pratt Hetrakul [a,*], Cinzia Cirillo [b,1]

[a] Resource Systems Group, Inc., 55 Railroad Row, White River Junction, VT 05001, USA
[b] Department of Civil and Environmental Engineering, University of Maryland, 1173 Glenn L. Martin Hall, Bldg #088, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 10 November 2012
Received in revised form 13 September 2013
Accepted 17 October 2013

Keywords:
Heterogeneity
Latent class model
Multinomial logit model
Pricing
Revenue management
Seat allocation

ABSTRACT

In this paper, discrete choice methods in the form of multinomial logit and latent class models are proposed to explain ticket purchase timing of passenger railway. The choice model and demand functions are incorporated into a revenue optimization problem which jointly considers pricing and seat allocation. The framework provides insightful policy implications in term of fare and capacity distribution derived from actual passenger behavior. It shows that accepting short-haul demand provides greater revenue than long-haul demand using the same capacity. Revenue improvement ranges from 16.24% to 24.96% in multinomial logit models and from 13.82% to 21.39% in latent class models respectively.

1. Introduction

1.1. Background and motivation

Demand forecasting is an essential component of revenue management (RM) model systems; it also plays a relevant role in seat allocation and optimization problems. With respect to pricing, the operators in charge of the reservation system need to know the distribution of the expected late-booking-high-fare demand in order to protect the right number of seats. For capacity allocation, it is necessary to predict the expected market size for each trip in order to provide efficient seat allocation strategies. Discrete choice analysis (DCA) is a standard approach for determining factors influencing decision making process. Recently, researchers in RM have argued that DCA enables for realistic representation of passenger response to RM policy. Moreover, passengers are usually characterized by a high level of taste heterogeneity (Hetrakul and Cirillo, 2013), that depends both on socio-demographic characteristics and different preferences over scheduling and pricing. Over the past decades considerable progress has been made in the characterization of unobserved taste heterogeneity for travel choice behavior (Wen and Lai, 2010). Latent class (LC) model is considered a valid approach to account for taste heterogeneity (Walker, 2001); LC model is a special case of mixture logit models (Train, 2003) in which the mixing distribution is discrete. This approach segments passengers with similar characteristics into classes of unknown size using a function of observable variables. Accounting for differences on individual taste offers a more realistic representation of passenger choice behavior and might lead to significant revenue improvement. To date a very limited number of studies applies this technique to RM and in particular to the pricing and seat allocation optimization problem for railway systems.

* Corresponding author. Tel.: +1 802 295 4999.
E-mail addresses: Pratt.Hetrakul@rsginc.com (P. Hetrakul), ccirillo@umd.edu (C. Cirillo).
[1] Tel.: +1 301 405 7768; fax: +1 301 405 2585.

1366-5545/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tre.2013.10.005

This paper aims at filling this gap by demonstrating that LC choice models can be effectively used to model railway passenger booking behavior and to represent the underlying taste heterogeneity. The case study based on intercity passenger railway internet booking data demonstrates that gain in revenues can be achieved when those techniques are made operational. The remaining of this paper is organized as follows. Previous studies on the use of demand modeling for revenue management are presented in Section 2. Section 3 describes the conceptual framework which incorporates passenger choice models and demand functions in RM revenue optimization. Section 4 presents a passenger choice model for ticket purchase timing decision. In Section 5, the passenger demand function used to estimate demand volume for each market is estimated. Section 6 describes the mathematical formulation and the optimization procedure adopted to incorporate both multinomial logit and latent class models into RM systems. Section 7 presents numerical results in term of seat allocation and fare strategy, revenue improvement, and capacity redistribution. Finally conclusions and future research directions are given in Section 8.

2. Literature review

2.1. Incorporating heterogeneous choice models in RM problem

Accounting for taste heterogeneity is essential for demand forecasting especially for railway RM because passenger preferences generally vary by departure time of day, day of week, and trip distance. A number of papers have investigated special classes of discrete choice model that accommodates taste heterogeneity. Bhat (1998) estimated an intercity travel mode choice model which accommodates variations in response to level of service measures due to observed and unobserved individual characteristics. The study emphasized the necessity to incorporate systematic and random variation in response to level-of-service variables. Greene and Hensher (2003) compared latent class (LC) with mixed logit (ML) model using stated preference data on long distance travel survey in 2000. Shen (2009) compared the difference between latent class and mixed logit models using two stated choice datasets on mode choice from Osaka, Japan using non-nested test to compare the model fits.

In the RM context, Carrier (2008) analyzed the choice of airline itinerary and fare product based on latent class (LC) model framework. In this model, passenger choice set was constituted from booking data, fare rules, and seat availability data. Instead of segmenting passenger by trip purpose, which is not available in booking data, the author utilizes variables such as frequent flyer membership, ticket distribution channel, and travel day of week for the class membership model. The approach is shown to provide a more distinct and intuitive segmentation across passengers. Teichert et al. (2008) applied the latent class model to explore preferences within airlines segments and analyzed respondents' profiles in terms of individual socioeconomic and trip characteristics. They concluded that the segmentation criterion currently applied by airlines does not adequately mirror the heterogeneity in customer's preference patterns. They suggested product marketing be aligned to passenger attitudes and socio-demographic profiles which are different across passenger segments. Wen and Lai (2010) used latent class model to identify airline passengers' potential segments and preferences for international air carriers using individual socioeconomic and trip characteristics as class membership variables. The latent class model is capable of representing heterogeneity across passenger segments which results in improved prediction accuracy over the multinomial logit model. Specifically, the willingness to pay for service attribute improvements is found to be substantially different across air routes and to vary by traveler segments.

Regardless of the number of research efforts on heterogeneity in choice behavior, the application of latent class choice model in RM problem is still relatively limited. Most of the studies which incorporated choice models in RM problem have assumed that customers are homogeneous in taste preferences. Studies which rely on this assumption include the work of Zhang and Adelman (2009), Topaloglu (2009); and Erdelyi and Topaloglu (2010) who incorporated customer choice models in the network RM pricing. In their setting, the price for each product is chosen from a discrete set, and the demand for each product depends on the price of the product only. However, given that RM relies on the premise that different customers are willing to pay different amounts for a product, accounting for passenger heterogeneity is expected to provide high yield toward RM strategy. More specifically, Garrow (2010) suggested that calibrating models by segments to distinguish between time-sensitive and price-sensitive customers can highly impact demand prediction accuracy and contribute to significant RM system performance when being incorporated in RM optimization problem.

Recently, a limited number of studies which incorporate heterogeneous passenger choice model in RM problem have primarily focused on choice-based deterministic linear programming (CDLP) problem. CDLP is a class of revenue optimization which solves for sets of product to be made available to the customers at different points in time during the sales horizon. In this context, Rusmevichientong et al. (2012) analyzed a model that captures the substitution between the products and preference heterogeneity. Each customer is assumed to belong to a particular class and the demand from each customer class is governed by a multinomial logit choice model with class-dependent parameters. This problem considers a set of different products and maximizes the expected profit across all customer classes. Méndez-Díaz et al. (2012) specified LC model which divides customers into segments based on choice of product alternatives considered by each customer. Their demand model allows product to overlap across segments and the preference parameters for each product alternative in the logit model are assumed to be known in advance. The authors prove that the latent class logit assortment problem is NP-Hard, and solve the choice-based deterministic linear program (CDLP) using branch and