Project ID: **NTC2014-SU-R-20**

REVENUE MANAGEMENT AND OPERATIONS OPTIMIZATION FOR HIGH SPEED RAIL

Final Report

by

Cinzia Cirillo

University of Maryland

for

National Transportation Center at Maryland (NTC@Maryland)

1124 Glenn Martin Hall

University of Maryland

College Park, MD 20742

November 2014

ACKNOWLEDGEMENTS

This project was funded by the National Transportation Center (NTC) @ Maryland.

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXCUTIVE SUMMARY

The increasing use of internet as a major ticket distribution channel has resulted in passengers becoming more strategic to fare policy. This potentially induces passengers to book the ticket well in advance in order to obtain a lower fare ticket, and later adjust their ticket when they are sure about trip scheduling. This is especially true in flexible refund markets where ticket cancellation and exchange behavior has been recognized as having major impacts on revenues. Therefore, when modeling this behavior, it is important to account for the characteristic of the passenger that optimally makes decision over time based on trip schedule and fare uncertainty.

In this paper, we propose an inter-temporal choice model of ticket cancellation and exchange for railway passengers where customers are assumed to be forward looking agents. A dynamic discrete choice model (DDCM) is applied to predict the timing in which ticket exchange or cancellation occurs in response to fare and trip schedule uncertainty. Passengers' decisions involve a two step process. First, the passenger decides whether to keep or adjust the ticket. Once the decision to adjust the ticket has been made, the passenger has the choice to cancel the ticket or to change departure time. The problem is formulated as an optimal stopping problem, and a two step look-ahead policy is adopted to approximate the dynamic programming problem.

The approach is applied to simulated and real ticket reservation data for intercity railway trips. Estimations results indicate that the DDCM provides more intuitive results when compared to multinomial logit (MNL) models. In addition, validation results show that DDCM has better prediction capability than MNL. The approach developed here in the context of exchange and refund policies for railway revenue management can be extended and applied to other industries that operate under flexible refund policies.

1.0 INTRODUCTION AND LITERATURE REVIEW

Ticket cancellation and exchange behavior has significant impact on the revenue management (RM) system (Iliescu, 2008). In flexible refund markets, passengers are inclined to book their tickets in advance in order to obtain lower fares, and to exchange/cancel the tickets when changes in their schedule intervene. Moreover, the use of internet as a major ticket distribution channel has affected the behavior of customers who have now better access to fare information, and are becoming more strategic in their choices. Reliable predictions in cancellation and exchange decisions are believed to enable analysts to derive more efficient overbooking and refund/exchange policies. RM applications to air transportation have demonstrated to significantly reduce the number of empty seats on flights for which there is actually a potential demand (Neuling, Riedel et al. 2004).

Existing literatures on choice modeling for revenue management (RM) have mostly ignored temporal effects in individual decision making. Although static models enable analysts to address the dependence of demand on the set of products offered by the provider, they are unable to model forward looking agents, who would typically wait and see before making the final decision. There is an emerging research effort toward dynamic frameworks that account for inter-temporal variability in choice modeling. Existing research on inter-temporal price variation that considers forward-looking consumers includes Stokey (1979), Landsberger and Meilijson (1985), and Besanko and Winston (1990). These papers are based on the assumptions that customers are present in the market throughout the entire season, and that the seller's inventory is practically unlimited. Customers purchase at most one unit during the season, and they optimally select the timing of their purchases so as to maximize individual surplus. Su (2007) studied a model of strategic customer by identifying four customer classes, different from each other in two dimensions: high versus low valuations and strategic (i.e., patient) versus myopic (impatient) behavior. The price path is assumed to be predefined by the seller, and after the specific pricing policy is announced, strategic consumers can weigh the benefits of waiting for a discount (if any is offered). The paper demonstrates that the joint heterogeneity in valuations and in the degree of patience is crucial in explaining the structure of optimal pricing policies.

Behavior of ticket cancellation and exchange is clearly influenced by demand uncertainty over time. Stokey (1979) showed that offering a single price can be optimal when inter-temporal differentiation is feasible, but assumes that consumers have perfect information on the future evolutions of their valuations. In Png's (1989), consumers face both uncertainty in their valuations as well as uncertainty about the capacity. Gale and Holmes (1992) examined advance purchase discounts where a monopoly firm offers two flights at different times and where consumers are assumed to not know their preferred flight in advance. In this study, advance purchase discounts are used to smooth the demand of the consumers with a low cost of time. Gallego and Phillips (2004) used a similar approach in their work on flexible products. Dana (1998) showed that advance purchase discounts may improve the revenues of price-taking firms when consumer demand is uncertain. In this case, firms in competitive markets can improve profits by offering advance purchase discounts. Shugan and Xie (2000) developed an inter-temporal consumer choice model for advance purchase which distinguishes the act of purchasing and consumption. The model accounts for buyer's valuation of services that depends on buyer states at the time of consumption and assumes the product capacity to be unlimited. In a later paper, Xie and Shugan (2001) extended this analysis of advance selling to the finite-capacity case and introduced a refund option. Ringbom and Shy (2004) proposed a model where consumers have the same deterministic valuation (maximum willingness to pay) for a certain service of product but different probabilities of showing up; capacity is assumed to be infinite and prices are endogenously given; results show that by adjusting partial refunds it is possible to endogenize the participation rates. Aviv and Pazgal (2008) considered an optimal pricing problem of a fashion-like seasonal good in the presence of strategic customers (forward-looking characteristics) with a time-varying valuation pattern. Customers have partial information about the availability of the inventory and their arrival is assumed to be time dependent. The system is characterized by a leader follower game under Nash equilibrium where customers select the timing of their purchase so as to maximize individual surplus while the seller maximizes expected revenue. Gallego and Sahin (2010) developed a model of customer purchase decision with evolution of trip schedule valuations over time. This analysis considers partial refundable fare based on a call option approach; each customer updates his/her valuation over time and decides when to issue and when to exercise options in a multi-period temporal horizon.

Meanwhile, a number of studies on demand uncertainty have focused on the supply chain management approach. To our knowledge, Spinler et al. (2002, 2003) are among the first in the operations management literature that incorporated consumer's uncertainty in valuations into revenue management, and the first to study partially refundable fares. Other studies on uncertain valuations for traditional revenue management problems include Levin et al. (2009), Yu et al. (2008), and Koenigsberg et al. (2006). There is also an emerging literature that deals with strategic consumers who develop expectations on future prices and product availability based on the observed history of prices and availabilities (e.g. Besanko and Winston 1990, Gallego et al., 2009, Liu and van Ryzin, 2005, Aviv and Pazgal, 2008).

In the context of ticket cancellation and exchange model, a number of papers have been published in the past decade. Garrow and Koppelman (2004a) proposed an airline cancellation and exchange behavior model based on disaggregate passenger data; airline travelers' no-show and standby behavior is modeled using a multinomial logit (MNL) model estimated on domestic US itineraries data. The approach enables the identification of rescheduling behavior based on passenger and itinerary characteristics and supports a broad range of managerial decisions. Variable used to identify passenger rescheduling behavior are traveler characteristics, familiarity to the air transportation system, availability of viable transportation alternatives, and trip characteristics. Garrow and Koppelman (2004b) extended their work by introducing a nested logit structure and demonstrated the benefit of directional itinerary information. The nested logit (NL) tree groups show, early standby, and late standby alternatives in one nest and no show alternative in another nest. The analysis emphasized the superiority of nested logit model structure over multinomial logit model and the importance of distinguishing between outbound and inbound itineraries. Iliescu et al. (2008) further expanded the work of Garrow and Koppelman (2004a, 2004b) by proposing a discrete time proportional odds (DTPO) model to predict the occurrence of ticket cancellation and exchange based on the Airline Reporting Corporation (ARC) data. The cancellation probability is defined as a conditional probability that a purchased ticket will be canceled in a specific time period given it survived up to that point (hazard probability). Results show that the intensity of cancellation is strongly influenced by the time from the ticket purchase and the time before flight departure as well as by other covariates (departure day of week, market, group size, etc.). Specifically, higher cancellation is observed for recently purchased ticket and ticket which associated departure dates are near. Graham et al. (2010) adopted discrete time proportional odds (DTPO) model to investigate when and why travelers make changes to their airline itineraries. Analysis is based on a nine month period panel data of university employees in Atlanta, US. The analysis focused on tickets issued less than 60 days before the outbound departure date. The use of panel data enabled the analysts to study how cancellation behavior differs by frequency of travel as well as by carrier. The deriving empirical analysis identifies the reasons why business travelers exchange their ticket, and concluded that differences exists between outbound and inbound itineraries, between exchange and cancellation rates for frequent and infrequent business travelers, across air carriers and timing when refund and exchange events occur. The results also indicate that the timing of cancellation exhibit a strong pattern, i.e., ticket changes are two to three time more likely to happen within the first week after purchase and are more likely to occur as the departure date approaches.

In summary, while many attempts have been made to understand the impact of choice behavior in revenue management, the issue of passenger uncertainty over trip scheduling has not been extensively explored. Behavior of ticket cancellation and exchange is clearly influenced by the evolution of passenger certainty about trip making over time. Specifically, to date none of the existing studies allows for departure time specific exchange decision in the cancellation and exchange model while accounting for inter-temporal behavior of passengers. Thus, our study aims to fulfill this gap.

In this paper, we propose a dynamic framework based on discrete choice models developed in the context of railway revenue management. Dynamic discrete choice models have been firstly developed in economics and applied to study a variety of problems that include fertility and child mortality Wolpin (1984), occupational choice Miller (1984), patent renewal Pakes (1986), and machine replacement Rust (1987). In dynamic discrete choice structural models, agents are forward looking and maximize expected inter-temporal payoffs; the consumers get to know the rapidly evolving nature of product attributes within a given period of time and different products are supposed to be available on the market. The timing of consumers' purchases is formalized as an optimal stopping problem where the agent (consumer) must decide on the optimal time of purchase (Rust, 1987).

To the authors' knowledge, this is the first attempt to incorporate dynamics in individual choices to revenue management modeling and in particular to formalize tickets' exchange and cancel decisions for railway intercity trips. The railway operator in consideration offers fully refundable fare and provides flexibility in ticket exchange which makes ticket cancellation and exchange decision to be very crucial to the RM system. Passengers are incentivized to purchase ticket early and adjust their ticket later when they are more certain about trip schedules. The model accounts for passengers' trip adjustment choice and explicitly specifies the probability of exchanging ticket as a function of the set of available exchange tickets. The choice set is constituted by all departure times offered by the railway operator between a specific origin destination pair.

The remainder of the paper is organized as follows: in Section 2, we analyze the data used for our model focusing on cancellation and exchange behavior. In Section 3, we formulate a dynamic discrete choice model and we formalize the algorithm used for the dynamic programming problem under study. Section 4 presents numerical results from a simulated experiment. Section 5 demonstrates the superiority of the method proposed for modeling exchange and cancel decisions based on real data. Finally, conclusions drawn from the empirical analysis and future research directions are outlined in Section 6.

2.0 DATA ANALYSIS

The data set used for the analysis has been extracted from intercity railway ticket reservation records registered in March 2009. This data set contains 155,175 individual transactions expressed in terms of ticket purchase, cancellation, and exchange over time prior to departure. Ticket exchange decision is defined as the exchange of the original ticket for a new one and the payment of an additional cost depending on the operator's exchange policy. In our case study, passengers are not charged with exchange fee, but have to pay the difference between the new and the old ticket fare. In the case of ticket exchange, passenger either obtains a new ticket right away or after several time periods (repurchase). Ticket cancellation is defined as the final cancellation of the ticket with the passenger obtaining ticket refund depending on the operator's refund policy.

Table 1 shows the descriptive statistics derived from the dataset in use. Ticket exchange and cancellation account for 18.22% and 29.75% of the sample respectively. Single exchange and no more than two exchanges account for 80.82% and 95.79% of the exchange ticket respectively (14.73% and 17.46% of the sample). We observe that only 2.26% of the sample make an exchange prior to ticket cancellation; thus in our model, we assume that passenger make ticket adjustment no more than once (either exchange or cancel). Based on this assumption, data are constructed to model the first exchange decision in case of multiple exchange, and model final cancellation in case passenger both exchange and cancel. We do not consider passenger who change origin/destination or reschedule departure day because the share of these population is relatively low accounting for 3.08% and 1.90% (0.91% + 0.99%) of the sample respectively. Consideration of changes in origin/destination and departure day decisions requires the definition of a choice set that is significantly different across passengers and no information is available to construct a realistic choice set for each passenger. This results in the focused sample population to be composed of entire sample (155,175) subtracted by passengers with origin/destination change and departure day change (a, b, and c in Table 1) which results in 147,457 individual ticket reservation records of the sample.

Table 1: Data Overview.

Ticket exchange	**No. reservation**	**% of exchange**	**% of total**
1. Total exchange	**28,280**	**100.00%**	**18.22%**
1.1 Number of exchange			
Exchange (one time)	**22,857**	**80.82%**	**14.73%**
Exchange (one or two times)	**27,088**	**95.79%**	**17.46%**
Exchange (more than 2 times)	**1,193**	**4.22%**	**0.77%**
1.2 Type of exchange			
Change OD (a)	**4,773**	**16.88%**	**3.08%**
No change (either OD or departure)	**7,001**	**24.76%**	**4.51%**

Reschedule departure day (b)	**1,406**	**4.97%**	**0.91%**
Reschedule departure time	**13,565**	**47.97%**	**8.74%**
Reschedule departure day and time (c)	**1,539**	**5.44%**	**0.99%**
Ticket Cancellation	**No. reservation**	**% of cancel**	**% of total**
2. Total final cancellation	**46,158**	**100.00%**	**29.75%**
2.1 Final cancellation after exchanged	**3,506**	**7.60%**	**2.26%**
Total (Northbound, March 2009, Coach Class)	**155,175**		**100.00%**
Effective Sample (Total - (a) - (b) - (c))	**147,457**		**95.03%**

The problem is further simplified by considering only passengers who made weekday trips from south end terminal station to 3 major destinations (named STA1, STA2, and STA3), and purchased the ticket 15 days before departure which results into a time horizon of 16 days for each decision maker (from 15 days before departure until departure day). This results in 696 valid individual passenger records for model estimation.

3.0 PROBLEM FORMULATION

3.1 PASSENGER STOPPING PROBLEM

We consider a passenger set $\mathfrak{T} = \{1, \dots, M\}$ where each passenger $i \in \mathfrak{T}$ can be in one of the two possible states $s_{it} = \{0,1\}$ in time period $t \in \{0,1, \dots, T\}$. Passenger is considered to be in the decision process when $s_{it} = 0$ and out of the decision process when $s_{it} = 1$. In each time period t, passenger i in state $s_{it} = 0$ has two options:

1. To make change to the ticket (either exchange or cancel). Once decided to adjust the ticket, the passenger makes the choice of $j \in \mathfrak{J}_t$ which is composed of exchange (departure time specific exchange decision at time period t) and cancel alternatives and obtain a terminal period payoff u_{ijt}. The utility of exchange is primarily a function of fare difference between the original and the exchange ticket at time period t. The utility of ticket cancellation is primarily a function of trip characteristics, and the refund amount.
2. To keep the original ticket and obtain a one-period payoff U_{ikt}, which is normalized to have a mean of zero before departure day and equal to c on departure day.

The two-step decision process assumes that, at each time period, the passenger decides whether to keep or change the ticket. The optimal time period in which passenger decides to change the ticket is denoted by τ, where the passenger chooses the ticket change alternative j_t^* that maximizes the utility from $\mathfrak{J}$. The passenger decision is the optimal stopping problem at time t:

$$D\left(u_{i1t}, \dots, u_{iJt}, U_{ikt}, t\right) = \max_{\tau} \left\{ \sum_{k=t}^{\tau-1} U_{ikt} + E\left[\max_{j \in \mathfrak{J}} u_{ij\tau}\right] \right\} \tag{1}$$

Let $v_{it} = \max_{j \in \mathfrak{J}} u_{ijt}$. We assume that v_{it} is Gumbel distributed with a scale factor equals to 1. Based on the dynamic programming theory (Rust, 1994), the passenger's decision can be transformed from into:

$$D(v_{it}, U_{ikt}) = \max\left\{v_{it}, U_{ikt} + E\left[D\left(v_{i,t+1}\right)\right]\right\} \tag{2}$$

The reservation utility is defined by function:

$$W_{it} = U_{ikt} + E\left[D\left(v_{i,t+1}\right)\right] \tag{3}$$

And consider the optimal policy:

$$\begin{cases} v_{it} & \text{if } v_{it} \geq W_{it} \\ W_{it} & \text{otherwise} \end{cases} \tag{4}$$

The problem can be simplified as:

$$D(v_{it}) = \max(v_{it}, W_{it}) \tag{5}$$

3.1.1 Keep Ticket Probability

The passenger i will keep the ticket at time t when $W_{it} \geq v_{it}$. Let π_{i0t} denotes the probability of keeping ticket until the next period, which can be written as:

$$\pi_{i0t} = P[\, v_{it} \leq W_{it}] = P[\text{keep}|s_{it} = 0] \tag{6}$$

$$= F_v[W_{it},\, v_{it}] = e^{-e^{-(W_{it} - r_{it})}} \tag{7}$$

Where r_{it} is the mode of the distribution of v_{it} that is:

$$r_{it} = \ln G(e^{V_{i1t}}, \dots, e^{V_{ijt}}) \tag{8}$$

3.1.2 Change Ticket Probability

The probability of ticket change is $P[\text{change}|s_{it} = 0] = 1 - \pi_{i0t}$ and the choice specific ticket change probability is:

$$\pi_{ijt} \quad = P\left[U_{ijt} \geq U_{ilt}, \forall l \neq j, u_{it} \geq W_{it}\right] \tag{9}$$

$$= P\left[U_{ijt} \geq W_{it} \middle| U_{ijt} \geq U_{ilt}, \forall l \neq j\right] P\left[U_{ijt} \geq U_{ilt}, l \neq j\right] \tag{10}$$

$$= (1 - \pi_{i0t}) P\left[U_{ijt} \geq U_{ilt}, l \neq j\right] \tag{11}$$

3.2 OBJECTIVE FUNCTION AND PARAMETERS TO ESTIMATE

The parameter estimation is performed by maximizing the likelihood function:

$$\mathcal{L}(\beta) = \prod_{i=1}^{N} \prod_{t=0}^{T} P_{it}\,[\text{decision}] \tag{12}$$

The decision probability is presented as:

$$P_{it}[\text{decision}] = P_{it}[\text{decision}, s_{it} = 0\,] + P_{it}[\text{decision}, s_{it} = 1] \tag{13}$$

$$= P_{it}[\text{decision}|\, s_{it} = 0\,] P[s_{it} = 0\,] + P_{it}[\text{decision}|\, s_{it} = 1] P[s_{it} = 1\,] \tag{14}$$

The state s_{it} is observed in the data set, if the passenger has not changed the ticket, $P[s_{it} = 0\,] = 1$ and $P[s_{it} = 1\,] = 0$. Once the passenger changes the ticket, the passenger is considered to be out of the decision process, therefore $P[s_{it} = 0\,] = 0$ and $[s_{it} = 1\,] = 1$. As a result, the complete likelihood function in this problem is:

$$\mathcal{L}(\beta) = \prod_{(i,t)\in V} P_{it}[\text{decision}, s_{it} = 0\,] \tag{15}$$

Where $V = \{(i,t)|i \in \{1, \dots, M\}, t \in \{1, \dots, T\}$ and $s_{it} = 0\}$. The decisions include keeping the ticket and ticket change specific choice. Thus $P_{it}[\text{decision}, s_{it} = 0\,] = \{\pi_{i0t}, \pi_{ijt}\}$.

3.3 DYNAMIC ESTIMATION PROCESS

The estimation process is done with maximum likelihood estimation method. First π_{i0t} must be obtained in order to calculate π_{ijt}. The probability π_{i0t}, depends on W_{it} which can be calculated from : $W_{it} = U_{ikt} + E\left[D\left(v_{i,t+1}\right)\right]$, assuming that r_{it} is the mode of the distribution of v_t.

W_{it} is composed of two parts: the utility of the current ticket attributes (U_{ikt}) and the expected utility in the next time period $\left(E\left[D\left(v_{i,t+1}\right)\right]\right)$. At each time period, the passenger is assumed to have a perception about the future scenarios, which are characterized by the alternative attributes changing over time. The expectation utility accounts for the possible market conditions in the passenger's perceived scenario; in our specification, the fare of each departure time specific exchange decision has been selected as independent variable in the utility specification. Passenger is assumed to have a perception of future attributes on a limited number of time periods, denoted by T. At time period t, the passenger faces two alternatives, keeping the ticket or changing the ticket. The passenger will continue the decision process into the period $t + 1$ only if he had decided to keep the ticket in time period t. Therefore, the decision process can be characterized by a scenario tree with a unique pattern (shown in Figure1). This scenario tree constitutes the base for the expected utility calculation. The following steps describe the procedure to calculate $\pi_{i0,0}$ and $E[D(v_{i1})]$ which will be indicated by $E[D_1]$ because all the expectations in the example are for individual i.

The procedure for calculating the expected utility will be described in detail as follows:

- First, we assume that the passenger has the expectation over a limited number of future time periods, which is limited to two in order to reduce the number of leaves in the scenario tree. At time period $t = 0$, the passenger can anticipate the future ticket characteristics (i.e. fare) from time period $t = 1$ and $t = 2$. The terminal time period expected utility $E[D_3] = 0$ because the passenger knows nothing for time period 3 when being at time period 0.
- Calculate $E[D_1]$. In order to obtain $\pi_{i0,0}$ from equation (6), the reservation utility (W_{i0}) is required. The reservation utility (W_{i0}) can be obtained from equation (3) $W_{i0} = U_{ik0} + E[D_1]$ which requires the calculation of $E[D_1]$. At time 0, the passenger has two alternatives for successive time 1, keep the ticket or change the ticket. The second term at the right hand side of the function $E[D_1] = \mathrm{E}\{\max[v_1, U_{ik1} + E[D_2]]\}$ represents the utility of keeping alternative; therefore when calculating $E[D_2]$, it is necessary that the term corresponded to the left leave of the tree be obtained (indicated by dash line in Figure 1). The calculation $E[D_2] = \mathrm{E}\{\max[v_2, U_{ik2} + E[D_3]]\}$ demands the same function to be calculated for time period 3 ($E[D_3]$) which is assumed to be zero according to the above assumption. The process of calculating $E[D_1]$ is recursive with known utility at the

end of the perspective horizon (assumed to two periods in this formulation). After $E[D_1]$ is calculated, reservation utility at time 0 (W_{i0}) can be obtained.
- This calculation procedure can be repeated to calculate $\pi_{i0,1}$ with the assumption that respondent can anticipate characteristics for time period 3 and $E[D_4] = 0$.

The reason that a terminal value for the expected utility has to be fixed at zero is because it is difficult to predict a particular value for the individual's perspective when future time period is far beyond his knowledge of information. This means that in the long term, the individual has not enough information to predict the future; passengers cannot anticipate the utility of keeping or cancelling the ticket. With this approach, after a limited number of time periods, information on future ticket fare attribute is just ignored.

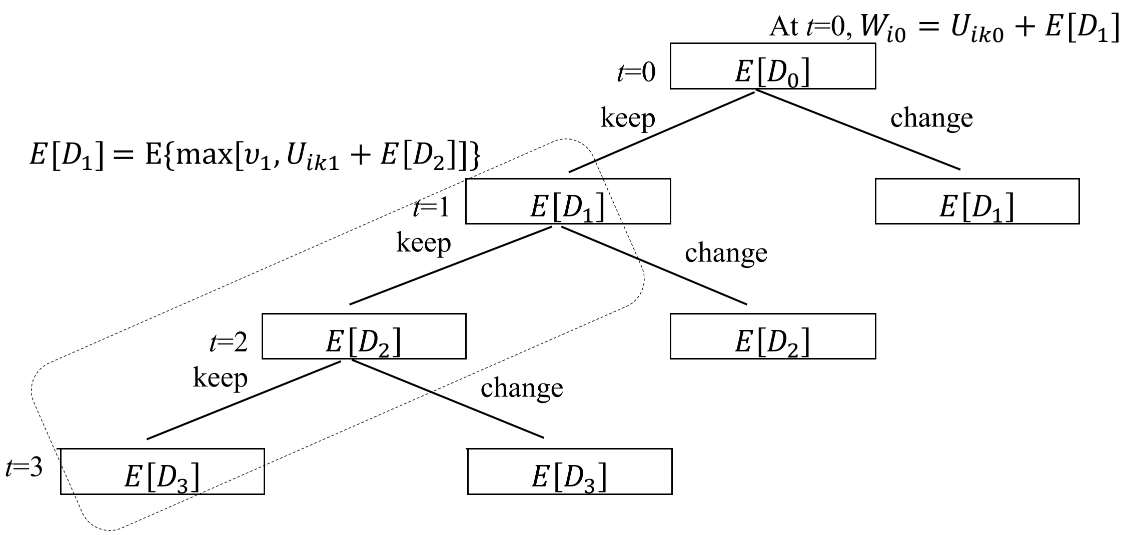

Figure 1: Scenario Tree.

4.0 EXPERIMENT WITH SIMULATED DATA

Synthetic ticket reservation data over time periods are simulated to validate the proposed dynamic discrete choice formulation. The data is created assuming that the characteristics of choice behavior is known; by adopting this procedure it is possible to test the ability of the dynamic discrete choice model to recover the true value of the parameters used to generate the data and to reproduce observed choice of individuals over time. Comparisons with static models, in the form of multinomial logit, are also presented.

4.1 DATA CONSTRUCTION

The simulated data is partially simulated from the real individuals' record, which characteristics are described in Section 2. Synthetic data assume that passengers have the same origin and destination as the real data, while individual characteristics, departure day of week, and departure time, vary from the real data. Concerning individual characteristics, the group size variable is generated from a uniform distribution and varies between 1 and 3 persons. Departure day of week is assumed to be uniformly distributed across the weekdays, while departure time is assumed to be uniformly distributed on discrete hour clock time between 5:00 AM and 7:00 PM. Ticket fare of the original departure time and other departure times within the same departure day are constructed for each day over the decision horizon based on historical data; the constructed fares vary by departure day of week and time of day.

Each individual is supposed to provide responses over a 16 day time period starting from 15 days before departure until the departure day. A total of (16×696) observations are then generated. There are 17 alternatives in the choice set, the first 15 alternatives refer to departure time specific exchange decisions (5:00 AM to 7:00 PM), the 16^{th} alternative is cancel, and the 17^{th} alternative is keeping the ticket. An important assumption in the data construction process is that if at one period in time the passenger decides to make change to his ticket, then this passenger will no longer be part of the decision process in the next time period (he is out of the market). True value parameters have been used to determine individual choices. Synthetic observations are then used to estimate both the static multinomial logit (MNL) and the dynamic discrete choice model (DDCM).

4.2 MODEL SPECIFICATION

The model specification considers 16 discrete time horizon defined by $t \in \{0,1, \dots ,15\}$ where t also represents the number of day from original ticket purchase. The first time period is the day when original ticket is purchased ($t = 0$), (day1). The last time period is departure day ($t = 15$), (day16). The utility specification is defined as follows:

$$\left.\begin{aligned} &U_{i5t=}\beta_{cost}(f_{5t} - f_{b0}) + \varepsilon_{i5t} + \varepsilon_i \\ &\quad \dots \dots \dots \dots \dots \dots \dots \dots . \\ &U_{ijt=}\beta_{cost}\left(f_{jt} - f_{b0}\right) + \varepsilon_{ijt} + \varepsilon_i \\ &\quad \dots \dots \dots \dots \dots \dots \dots \dots \dots . \\ &U_{i19t=}\ \beta_{cost}(f_{19t} - f_{b0}) + \varepsilon_{i19t} + \varepsilon_i \end{aligned}\right\} \text{15 Exchange alternatives (5:00 – 19:00).}$$

$$U_{ict=}\ ASC_{cnl} + \beta_{ev}ev + \beta_{Fri}Fri + \beta_{STA3}STA3 + \beta_{ref}f_{b0} + \varepsilon_{ict} + \varepsilon_i \tag{16}$$

$$U_{ikt=}\begin{cases} c + \varepsilon_{ikt} + \varepsilon_i & \text{if } t = 15 \\ \varepsilon_{ikt} + \varepsilon_i & \text{if } t < 15 \end{cases}$$

The utility of individual i on alternative j is denoted by U_{ijt}. For ticket exchange decision, the index j indicate 15 exchange departure times (5:00 AM to 7:00 PM). The utility of exchange (U_{ijt}) includes exchange cost defined as the difference between the original fare (f_{b0}) and the new fare (f_{jt}) at time t. The model allows passengers to exchange ticket for the same departure time as in the original ticket; transactions of this type are observed from the real data. This decision will result in passenger paying the difference between the original cost ($t = 0$) and the cost at time t. The utility of cancel (U_{ict}) includes alternative specific constant (ASC), refund, dummy of original departure in the evening (3:00-7:00 PM.), dummy of original departure on Friday, and dummy of STA3 destination. The utility of keep (U_{ikt}) has two different specifications. In the last time period ($t = 15$) passengers deciding to keep the ticket obtain the utility which includes the constant term referring to utility of traveling with the original ticket. In other time periods ($t < 15$) the systematic term of the keep utility is normalized to zero. ε_{ijt} is the random error term for each alternative at a given time period. ε_i is the individual error term which is assumed to be constant across all observations produced by the same respondent.

To evaluate the ability to recover the true value of the model, the root mean square deviation (RMSD) is adopted as measure of differences between the true values and the estimated coefficient values. The RMSD is defined as:

$$RMSD(\hat{\theta}) = \sqrt{E[(\hat{\theta}-\theta)^2]} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{\theta}_i-\theta_i)^2}{n}} \tag{17}$$

Where n is the number of parameters. Using the simulated data with the utility specification defined above two models are estimated: dynamic discrete choice model (DDCM) and static multinomial logit (MNL) model. In the static model, the attribute of the future ticket characteristic (fare) are not considered when making decision at each time period. The model is simply formulated as a traditional MNL model with 17 alternatives (15 exchange decisions, 1 cancel, and 1 keep). The dynamic model with the algorithm defined in the formulation is coded in C language and make use of optimization tools available in AMLET (Another Mixed Logit Estimation Tool), (Bastin, 2011).The static model is estimated using ALOGIT (ALOGIT, 2007).

The utilities specifications are assumed to be the same for the static and the dynamic model; the deriving estimation results are compared in Table 2. All parameters in both MNL and DDCM models are statistically significant at 5% significance level. The RMSD value obtained with the dynamic model is lower when compared to the MNL model (0.82 compared to 3.93); this indicates that the dynamic model outperforms MNL model in recovering the true value of the parameters.

Table 2: Estimation Result: Simulated Data.

	Exchange	Cancel	Keep	True Value	MNL Est	MNL T-stat		Dynamic (2-SL) Est	Dynamic (2-SL) T-stat	
ASC cancel		x		-5.00	-12.730	-14.2	*	-5.434	6.7	*
Orig Deptt 3-7 pm		x		2.50	1.392	5.2	*	3.927	8.1	*
Depart Friday		x		-2.00	-1.456	-3.8	*	-2.018	3.3	*
STA3 destination		x		4.00	2.517	10.1	*	5.252	10.8	*
Exchange cost	x			-0.02	-0.040	-77.8	*	-0.020	12.1	*
Refund		x		0.03	0.032	8.4	*	0.030	5.5	*
Keep (day 16)			x	-7.00	-6.477	-8.9	*	-5.868	27.0	*
Cancel day1		x		3.00	6.609	9.1	*	2.275	13.2	*
Exchange day16	x			1.50	-6.144	-8.4	*	2.343	8.5	*
Early exchange	x			-2.00	-6.366	-41.4	*	-2.598	13.6	*
Log-likelihood (0)						-28,896			-2,908	
Log-likelihood (final)						-17,875			-852	
Likelihood ratio index						0.38			0.70	
RMSD						3.93			0.82	
R-square wrt 0						0.3814				
No. individual						696				
No. observations						10,199				

* Statistically significant at 5% significance level.

4.3 MODEL VALIDATION

The results obtained from the estimation are used to validate how well the models reproduce the observed simulated decisions (Table 3).

Table 3: Validation Result: Simulated Data.

Choice		Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
Exchange	Pred	MNL	3	3	3	3	2	2	2	3	3	2	3	2	3	2	3	3
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
2	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
Exchange	Pred	MNL	4	4	4	4	4	6	7	4	4	5	5	5	5	3	3	2
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
3	No.	Observed	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
Exchange	Pred	MNL	4	5	7	6	6	8	6	6	6	7	4	3	3	2	2	2
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
4	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10
Exchange	Pred	MNL	15	10	11	11	12	11	11	12	11	13	13	13	13	12	10	7
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
5	No.	Observed	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	27
Exchange	Pred	MNL	31	31	30	22	23	20	21	21	20	17	14	15	23	19	19	17
	Pred	DDCM	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	20
6	No.	Observed	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	18
Exchange	Pred	MNL	44	36	30	34	30	27	27	31	29	33	29	27	26	25	21	17
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
7	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27
Exchange	Pred	MNL	50	53	35	33	36	32	28	39	38	44	43	46	47	33	30	33
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25
8	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26
Exchange	Pred	MNL	48	51	35	34	37	38	42	34	40	24	27	26	35	38	33	33
	Pred	DDCM	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	22
9	No.	Observed	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	27
Exchange	Pred	MNL	44	46	37	37	38	32	45	37	38	43	40	39	37	35	29	22
	Pred	DDCM	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	29
10	No.	Observed	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	28
Exchange	Pred	MNL	17	16	17	15	14	19	17	20	17	21	21	23	19	16	17	14
	Pred	DDCM	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	30
11	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Exchange	Pred	MNL	16	17	17	21	19	17	26	21	21	19	17	19	19	23	17	13
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	19

Choice		Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
12	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28
Exchange	Pred	MNL	12	18	22	19	20	17	20	19	18	20	16	19	14	15	14	15
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28
13	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32
Exchange	Pred	MNL	23	27	38	43	37	32	36	29	31	39	32	29	28	28	25	22
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	37
14	No.	Observed	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43
Exchange	Pred	MNL	59	77	69	79	67	109	64	61	66	56	58	73	47	50	51	41
	Pred	DDCM	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	42
15	No.	Observed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	76
Exchange	Pred	MNL	183	151	132	122	135	113	127	137	137	110	132	104	108	107	113	113
	Pred	DDCM	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	77
16	No.	Observed	54	0	1	0	0	0	0	1	0	0	0	0	0	0	0	94
Cancel	Pred	MNL	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	67
	Pred	DDCM	42	0	2	0	2	0	1	3	0	0	0	0	1	3	0	88
17	No.	Observed	638	637	636	635	634	634	633	632	632	632	632	632	632	632	632	148
Keep	Pred	MNL	90	93	150	153	154	152	155	157	153	179	180	188	205	224	246	210
	Pred	DDCM	653	637	634	636	632	633	631	629	632	631	632	630	631	629	632	172
Total	No.	Observed	4	1	0	1	1	0	1	0	0	0	0	0	0	0	0	390
Exchange	Pred	MNL	552	545	487	483	481	482	479	476	479	453	452	444	427	408	386	355
(1-15)	Pred	DDCM	1	1	1	0	1	1	2	1	0	1	0	2	0	0	0	372
		Total	696	638	637	636	635	634	634	633	632	632	632	632	632	632	632	632

The validation in Table 3 indicates that the major drawback of MNL model is the over-prediction of exchange decisions especially exchange decision choice 15 (exchange to 7 PM); which is characterized by low fare. MNL model predicts the cancel decision in the first time period (day1) precisely; however, its prediction on cancel decision in the last time period (day16) is not as good as the one produced by DDCM. More importantly, in the last time period (day16) with a high number of exchange decision, the DDCM model clearly outperforms MNL. We selected the first (day1) and last time period (day16) of cancel decision and exchange decisions to show in detail the prediction capability of the DDCM over the MNL (Figure 2).

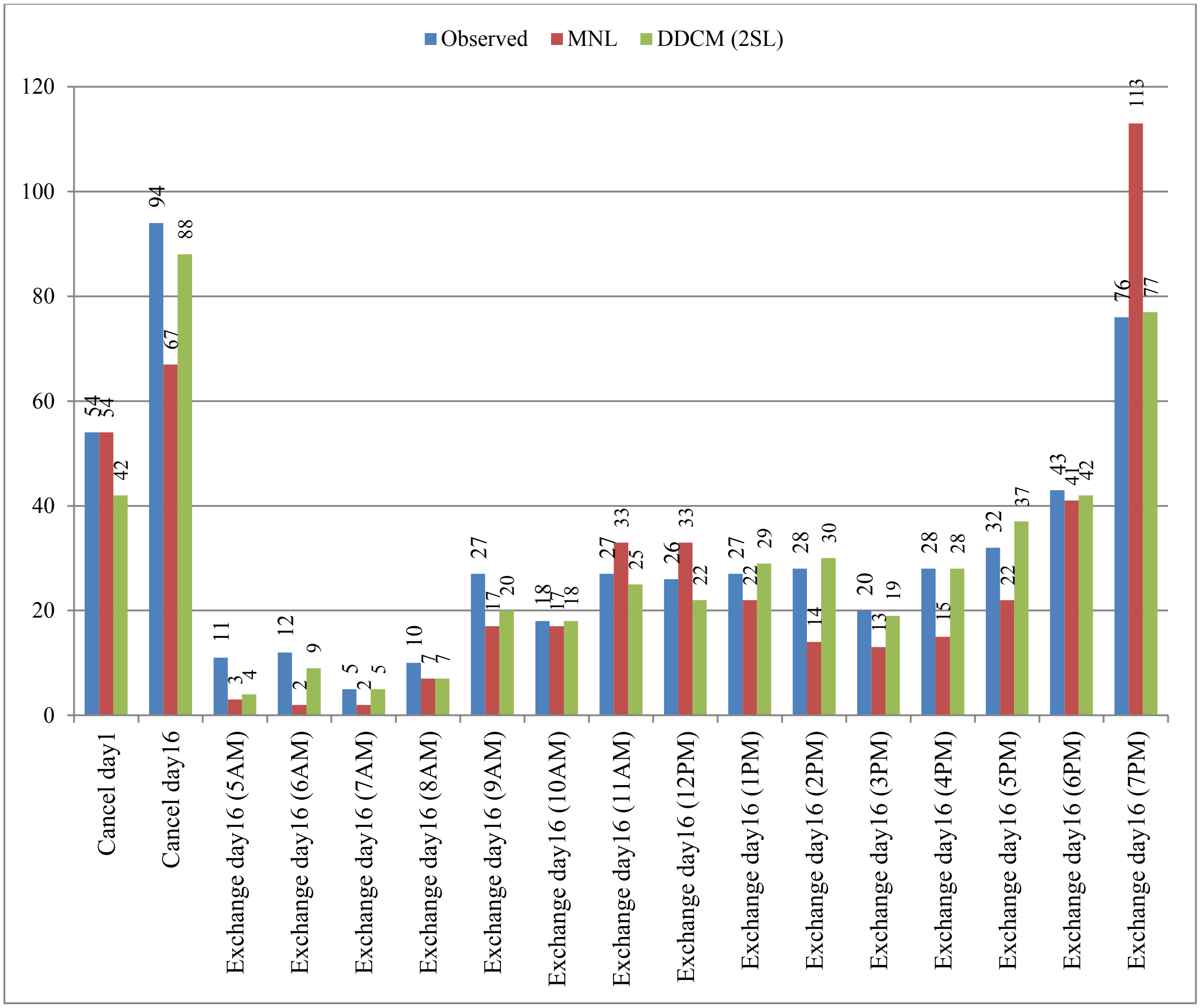

Figure 2: Simulated Data Validation: Departure time specific exchange and cancel decision.

Figure 3 to Figure 5 briefly summarize the predictions over different time periods (day) where exchange decision is aggregated over all departure times.

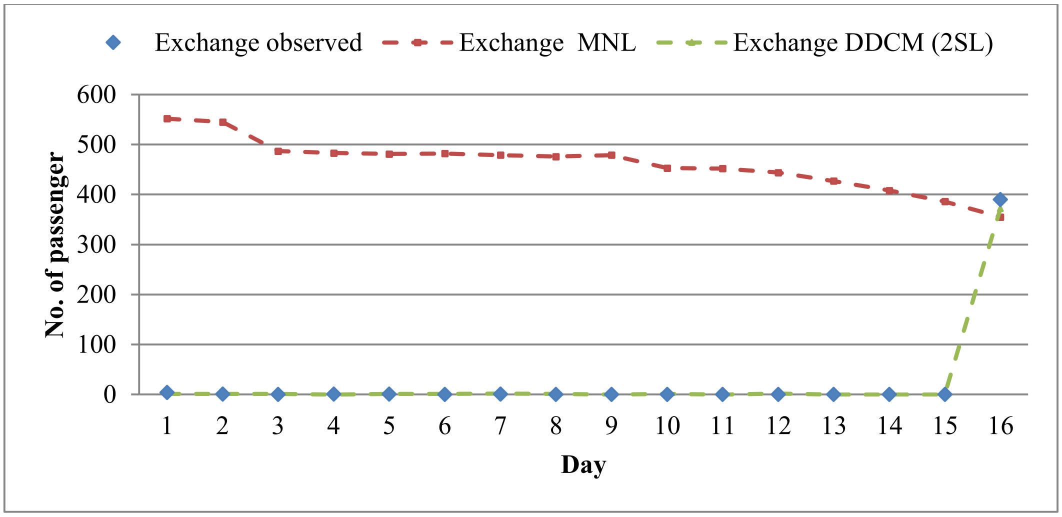

Figure 3: Validation of Exchange Decision: Simulated Data.

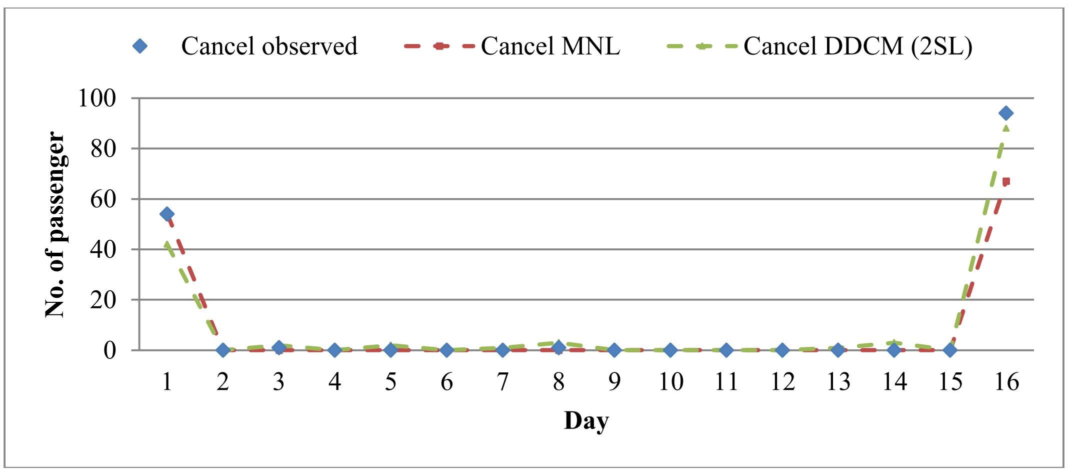

Figure 4: Validation of Cancel Decision: Simulated Data.

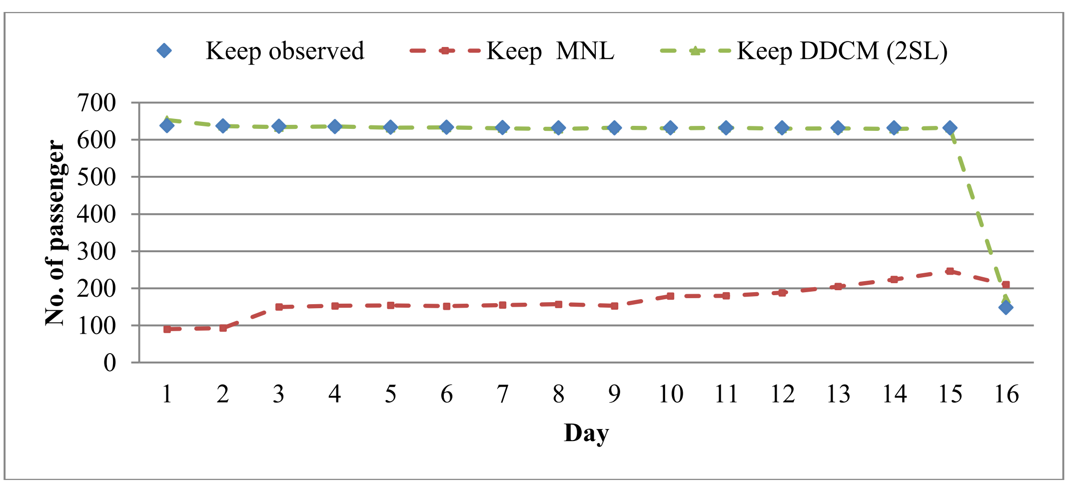

Figure 5: Validation of Keep Decision: Simulated Data.

The choice probability for each alternative observed and predicted together with a measure of errors for the simulated data experiment is reported in Table 4. The absolute error D is used to represent measure of error which is defined as:

$$D = |M_{pred} - M_{obs}| \tag{18}$$

Where D is error norm; M_{pred} is a vector of predicted choice probability, and M_{obs} is a vector of observed choice probability. The D value obtained from the dynamic model is significantly smaller than the corresponding value in the MNL model (0.176 compared to 22.214) indicating a better prediction capability of the dynamic model over the MNL model.

Table 4: Model Validation: Choice Probability of Simulated Data Experiment.

Alternative	**Observed**	**Predicted (Static)**	**Predicted (Dynamic)**
Exchange day1	**0.0057**	**0.7931**	**0.0014**
Exchange day2	**0.0016**	**0.8542**	**0.0016**
Exchange day3	**0.0000**	**0.7645**	**0.0016**
Exchange day4	**0.0016**	**0.7594**	**0.0000**
Exchange day5	**0.0016**	**0.7575**	**0.0016**
Exchange day6	**0.0000**	**0.7603**	**0.0016**
Exchange day7	**0.0016**	**0.7555**	**0.0032**
Exchange day8	**0.0000**	**0.7520**	**0.0016**
Exchange day9	**0.0000**	**0.7579**	**0.0000**
Exchange day10	**0.0000**	**0.7168**	**0.0016**
Exchange day11	**0.0000**	**0.7152**	**0.0000**
Exchange day12	**0.0000**	**0.7025**	**0.0032**
Exchange day13	**0.0000**	**0.6756**	**0.0000**
Exchange day14	**0.0000**	**0.6456**	**0.0000**
Exchange day15	**0.0000**	**0.6108**	**0.0000**
Exchange day16	**0.6171**	**0.5617**	**0.5886**
Cancel day1	**0.0776**	**0.0776**	**0.0603**
Cancel day2	**0.0000**	**0.0000**	**0.0000**
Cancel day3	**0.0016**	**0.0000**	**0.0031**
Cancel day4	**0.0000**	**0.0000**	**0.0000**
Cancel day5	**0.0000**	**0.0000**	**0.0031**
Cancel day6	**0.0000**	**0.0000**	**0.0000**
Cancel day7	**0.0000**	**0.0000**	**0.0016**
Cancel day8	**0.0016**	**0.0000**	**0.0047**
Cancel day9	**0.0000**	**0.0000**	**0.0000**
Cancel day10	**0.0000**	**0.0000**	**0.0000**
Cancel day11	**0.0000**	**0.0000**	**0.0000**
Cancel day12	**0.0000**	**0.0000**	**0.0000**
Cancel day13	**0.0000**	**0.0000**	**0.0016**
Cancel day14	**0.0000**	**0.0000**	**0.0047**

Cancel day15	**0.0000**	**0.0000**	**0.0000**

Alternative	**Observed**	**Predicted (Static)**	**Predicted (Dynamic)**
Cancel day16	**0.1487**	**0.1060**	**0.1392**
Keep day1	**0.9167**	**0.1293**	**0.9382**
Keep day2	**0.9984**	**0.1458**	**0.9984**
Keep day3	**0.9984**	**0.2355**	**0.9953**
Keep day4	**0.9984**	**0.2406**	**1.0000**
Keep day5	**0.9984**	**0.2425**	**0.9953**
Keep day6	**1.0000**	**0.2397**	**0.9984**
Keep day7	**0.9984**	**0.2445**	**0.9953**
Keep day8	**0.9984**	**0.2480**	**0.9937**
Keep day9	**1.0000**	**0.2421**	**1.0000**
Keep day10	**1.0000**	**0.2832**	**0.9984**
Keep day11	**1.0000**	**0.2848**	**1.0000**
Keep day12	**1.0000**	**0.2975**	**0.9968**
Keep day13	**1.0000**	**0.3244**	**0.9984**
Keep day14	**1.0000**	**0.3544**	**0.9953**
Keep day15	**1.0000**	**0.3892**	**1.0000**
Keep day16	**0.2342**	**0.3323**	**0.2722**
D		**22.2140**	**0.1760**

5.0 EXPERIMENT WITH REAL TICKET RESERVATION DATA

5.1 DATA CONSTRUCTION

Observations relative to ticket exchange and cancellation from 696 individuals were obtained from internet purchases of railway intercity tickets; the source of the data cannot be revealed here due to confidentiality issues. Based on the trip schedule of real data, ticket fare of the original departure time and other departure times within the same departure day are constructed for each day over the decision horizon based on historical data. In each time period, if the passenger decides to change or cancel the current ticket, then the same passenger will no longer be in the decision process in the next time period; all observations occurred in the period after ticket exchange are excluded from the dataset. This results into a total of 7,268 observed decisions valid for model estimation.

5.2 MODEL SPECIFICATION

The specification follows the same structure proposed in the simulated data experiment. The model specification considers 16 discrete time periods defined by $t \in \{0,1,\dots,15\}$ where t also represents the number of day from original ticket purchase. The first time period is the day when original ticket is purchased $(t = 0)$, (day1). The last time period is departure day $(t = 15)$, (day16) The utility specification is defined as follows:

$$\left.\begin{aligned}
&U_{i5t}=\beta_{cost}(f_{5t}-f_{b0})+\beta_{dfi_exc}t+\varepsilon_{i5t}+\varepsilon_i\\
&\quad\dots\dots\dots\dots\dots\dots\dots\dots\\
&U_{ijt}=\beta_{cost}\left(f_{jt}-f_{b0}\right)+\beta_{dfi_exc}t+\varepsilon_{ijt}+\varepsilon_i\\
&\quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&U_{i19t}=\beta_{cost}(f_{19t}-f_{b0})+\beta_{dfi_exc}t+\varepsilon_{i19t}+\varepsilon_i
\end{aligned}\right\}\text{15 Exchange alternatives (5:00 – 19:00).}$$

$$\begin{aligned}
U_{ict}=\ & ASC_{cnl}+\beta_{gp}gp+\beta_{mr}mr+\beta_{ev}ev+\beta_{Mon}Mon+\beta_{Fri}Fri+\beta_{STA1}STA1\\
&+\beta_{STA3}STA3+\beta_{ref}f_{b0}+\beta_{dfi_cnl}t+\varepsilon_{ict}+\varepsilon_i
\end{aligned}$$

$$U_{ikt}=\begin{cases} c+\varepsilon_{ikt}+\varepsilon_i & \text{if } t=15\\ \varepsilon_{ikt}+\varepsilon_i & \text{if } t<15\end{cases} \tag{19}$$

The utility of individual i on alternative j is denoted by U_{ijt}. For ticket exchange decision, the index j indicate 15 exchange departure time (5:00 AM to 7:00 PM). The utility of exchange (U_{ijt}) includes exchange cost which is defined as the difference between the original fare (f_{b0}) and new fare (f_{jt}) at time t, and day from issue (dfi) which is the number of day from original ticket purchase equal to t where $t = 0$ on the day of original purchase and $t = 15$ on departure day. The utility of cancel (U_{ict}) includes alternative specific constant (ASC), refund, dummy of group traveler, dummies of original departure in the morning (5:00-9:00 AM.) and evening (3:00-7:00 PM.), dummies of original departure on Monday and Friday, dummies of STA1 and STA3 destination. The utility of keep (U_{ikt}) is defined in two cases. In the last time period ($t = 15$) passenger deciding to keep the ticket obtain an utility that includes the constant term relative to the utility of traveling with the original ticket. In other time periods ($t < 15$) the systematic term of the keep utility is normalized to zero. ε_{ijt} is the random error term for each alternative at a given time period. ε_i is the individual error term which is assumed to be constant across all observations produced by the same respondent.

5.3 ESTIMATION RESULT

The results obtained from model estimation are shown in Table 5. Most of the variables are statistically significant at 5% confidence level. The results obtained from the dynamic model shows negative sign in a number of variables associated with cancel decision which are: group traveler (party size includes more than one passenger), evening departure (original departure time from 3:00-7:00 PM.), original departure on Friday, and STA1 destination. This indicates low tendency of passenger with these characteristics to cancel their ticket. On the other hand, passengers with morning departure (original departure time from 5:00-9:00 AM.), original departure on Monday, and STA3 destination have a positive sign for the corresponding structural coefficients, indicating that passengers with these characteristics have higher likelihood to cancel the ticket. In particular, passengers traveling early in the week and traveling alone (typically associated with business travelers) are more likely to cancel their ticket which is in line with the results of Iliescu (2008).

The exchange cost and refund have the expected sign indicating disutility associated with paying additional cost to exchange ticket and the utility of receiving refund when ticket is canceled respectively. The variable of keeping the ticket on departure day (day16) shows negative sign which could be explained by the fact that the fare of the original ticket possessed by the passenger is higher compared to a ticket hypothetically exchanged to other departure times. Another reason could be that passengers intentionally want to exchange/cancel the ticket but could not find an alternative departure time which economically matches their schedule.

The day from issues (number of days since the original ticket is purchased) has positive sign for the variable associated with exchange and cancel decision; this indicates that it is preferable for passengers to adjust their ticket later. This is line with expectations and consistent with results obtained by Iliescu (2008), who found that the odds of ticket change increase as the departure date approaches due to a strong effect of “last minute” change of plan. More specifically, the day from issue coefficient for the cancel decision has larger magnitude compared to the day from issue coefficient for the exchange decisions. This is intuitive based on this operator’s refund policy; passengers are fully refunded if the ticket is exchanged up to one hour before departure, while late tickets exchange are possible but limited by the uncertainty about seats availability.

The dummy variables of cancel on the original purchase date (day1) and exchange on the departure day (day16) show large magnitude indicating that a high number of cancellation and exchange occurs on the day they purchase ticket and on the departure day respectively. These results are in line with Iliescu (2008) and Graham et al. (2010) which found that ticket changes are more likely to happen in recently purchased ticket (especially within the first week) and are more likely to occur as the departure date approaches. Finally, the variable associated with early exchange (exchanging to departure time earlier than original ticket) shows negative sign which indicates that passengers gain less utility when making early exchange compared to later exchange (which is the base case).

Table 5: Estimation Result: Real Data.

	Exchange	Cancel	Keep	MNL			Dynamic (2-SL)		
				Est	T-stat		Est	T-stat	
ASC cancel		x		**-6.297**	**12.9**	*	**-3.652**	**57.1**	*
>1 psg		x		**-0.869**	**2.1**	*	**-1.090**	**1.5**	
Orig Deptt 5-9 am		x		**0.143**	**0.8**		**0.639**	**1.2**	
Orig Deptt 3-7 pm		x		**-0.327**	**1.9**		**-0.760**	**1.4**	
Depart Monday		x		**0.556**	**1.8**		**2.740**	**3.0**	*
Depart Friday		x		**-0.286**	**1.8**		**-0.451**	**1.0**	
STA1 destination		x		**-0.435**	**2.3**	*	**-0.306**	**0.6**	
STA3 destination		x		**0.557**	**2.5**	*	**1.648**	**2.6**	*
Exchange cost	x			**-0.011**	**19.3**	*	**-0.026**	**3.7**	*
Refund		x		**0.014**	**6.0**	*	**0.042**	**9.6**	*
Keep (day 16)			x	**1.885**	**11.0**	*	**-3.547**	**12.8**	*
Day from issue	x			**-1.217**	**35.3**	*	**0.189**	**5.8**	*
Day from issue		x		**0.163**	**5.9**	*	**0.266**	**35.4**	*
Cancel (day 1)		x		**5.629**	**18.3**	*	**3.169**	**42.7**	*
Exchange (day 16)	x			**17.050**	**30.5**		**1.578**	**10.2**	*
Early exchange	x			**-3.299**	**24.5**	*	**-1.751**	**12.1**	*
Log-likelihood (0)					**-20,592**			**-4,324**	
Log-likelihood (final)					**-7,629**			**-3,117**	
Likelihood ratio index					**0.63**			**0.28**	
R-square wrt 0					**0.6295**				
No. individual					**696**				
No. observations					**7,268**				

* Statistically significant at 5% significance level.

5.4 MODEL VALIDATION

To test the prediction capabilities of the model proposed, the resulting coefficients of the model have been used to replicate the choice observed in the sample. Results are reported in Table 6.

Table 6: Validation Result: Real Data.

Choice		Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	No.	Observed	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Exchange	Pred	MNL	5	3	1	0	0	0	0	0	0	0	0	0	0	0	0	1
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
2	No.	Observed	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Exchange	Pred	MNL	6	3	1	1	0	0	0	0	0	0	0	0	0	0	0	1
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	No.	Observed	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
Exchange	Pred	MNL	6	4	2	1	0	0	0	0	0	0	0	0	0	0	0	1
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	No.	Observed	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2
Exchange	Pred	MNL	13	7	3	1	0	0	0	0	0	0	0	0	0	0	0	2
	Pred	DDCM	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
5	No.	Observed	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0
Exchange	Pred	MNL	20	12	5	2	1	0	0	0	0	0	0	0	0	0	0	3
	Pred	DDCM	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	No.	Observed	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Exchange	Pred	MNL	24	14	6	2	1	0	0	0	0	0	0	0	0	0	0	3
	Pred	DDCM	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
7	No.	Observed	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1
Exchange	Pred	MNL	29	19	7	2	1	0	0	0	0	0	0	0	0	0	0	4
	Pred	DDCM	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
8	No.	Observed	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	9
Exchange	Pred	MNL	29	19	7	3	1	0	0	0	0	0	0	0	0	0	0	4
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	No.	Observed	3	1	0	0	0	0	0	2	0	0	0	0	0	3	0	9
Exchange	Pred	MNL	29	18	8	3	1	0	0	0	0	0	0	0	0	0	0	4
	Pred	DDCM	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	No.	Observed	1	0	0	0	1	0	0	0	3	0	0	0	0	0	0	6
Exchange	Pred	MNL	22	14	7	2	1	0	0	0	0	0	0	0	0	0	0	4
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	No.	Observed	2	2	1	1	0	0	0	1	0	0	0	0	1	0	0	12
Exchange	Pred	MNL	26	16	7	3	1	0	0	0	0	0	0	0	0	0	0	0
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
12	No.	Observed	2	0	1	0	0	0	1	0	0	0	0	0	0	0	0	8
Exchange	Pred	MNL	32	21	9	3	1	0	0	0	0	0	0	0	0	0	0	5
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
13	No.	Observed	4	1	0	0	0	0	0	1	1	0	1	0	1	0	0	5
Exchange	Pred	MNL	40	23	10	4	1	0	0	0	0	0	0	0	0	0	0	6
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Choice		Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
14	No.	Observed	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	3
Exchange	Pred	MNL	62	34	13	4	1	0	0	0	0	0	0	0	0	0	0	8
	Pred	DDCM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	No.	Observed	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
Exchange	Pred	MNL	95	154	194	213	221	218	219	214	213	200	203	190	183	175	167	36
	Pred	DDCM	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
16	No.	Observed	181	6	8	1	6	2	8	4	6	0	5	4	5	6	20	25
Cancel	Pred	MNL	181	1	2	3	4	4	5	6	7	8	9	11	12	14	16	4
	Pred	DDCM	85	0	0	0	0	0	0	0	0	0	0	0	0	2	0	20
17	No.	Observed	495	483	469	467	460	457	447	438	428	427	420	414	405	393	369	285
Keep	Pred	MNL	77	135	201	224	233	235	233	227	219	221	215	220	219	216	210	283
	Pred	DDCM	606	494	483	468	467	460	457	447	438	428	427	419	414	403	393	344
Total	No.	Observed	20	6	6	1	1	1	2	5	4	1	2	2	4	6	4	59
Exchange	Pred	MNL	438	359	280	243	230	221	219	215	213	200	203	190	183	175	167	80
(1-15)	Pred	DDCM	5	1	0	1	0	0	0	0	0	0	0	1	0	0	0	5
		Total	696	495	483	469	467	460	457	447	438	428	427	420	414	405	393	369

Figure 6 to Figure 8 briefly summarize the predictions over different time periods (days) where exchange decisions are aggregated for all exchange departure times. The validation results show that the DDCM slightly under-predicts cancellation and although it is not able to predict the cancellation on the first time period (day1) as well as MNL, it is capable of predicting cancellation on the last time period (day16) reasonably well. In term of exchange, DDCM slightly under-predicts the total number of exchange except for the first (day1) and the last time period (day16) which are characterized by a relatively high exchange rate; however, the MNL drastically over predicts exchange decisions throughout all time periods. The prediction of keep obtained from DDCM is reasonably close to the observed value while the MNL significantly under predicts the keep decision as a consequence of over prediction in exchange.

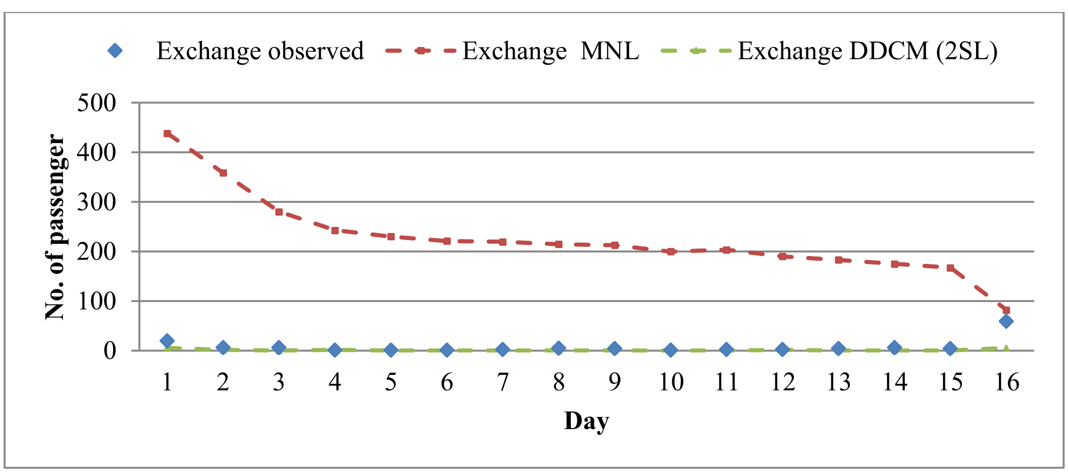

Figure 6: Validation of Exchange Decision: Real Data.

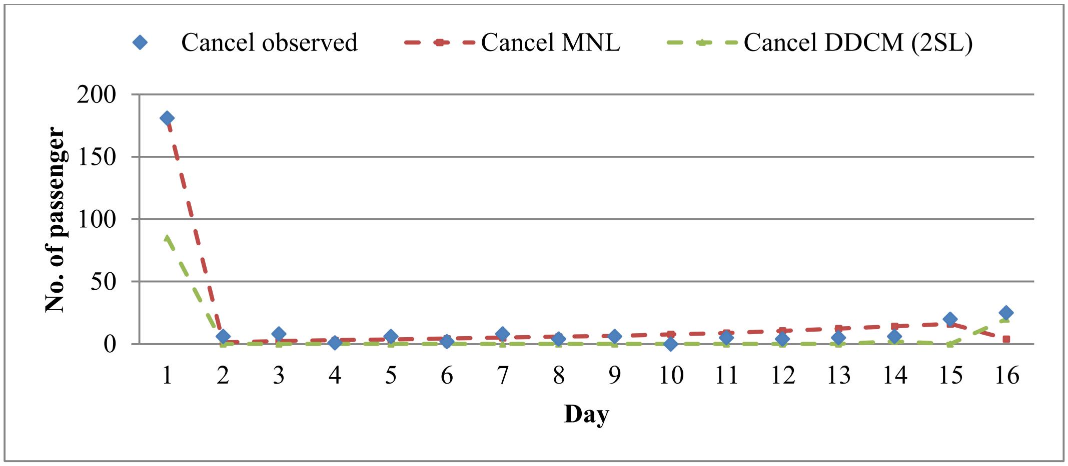

Figure 7: Validation of Cancel Decision: Real Data.

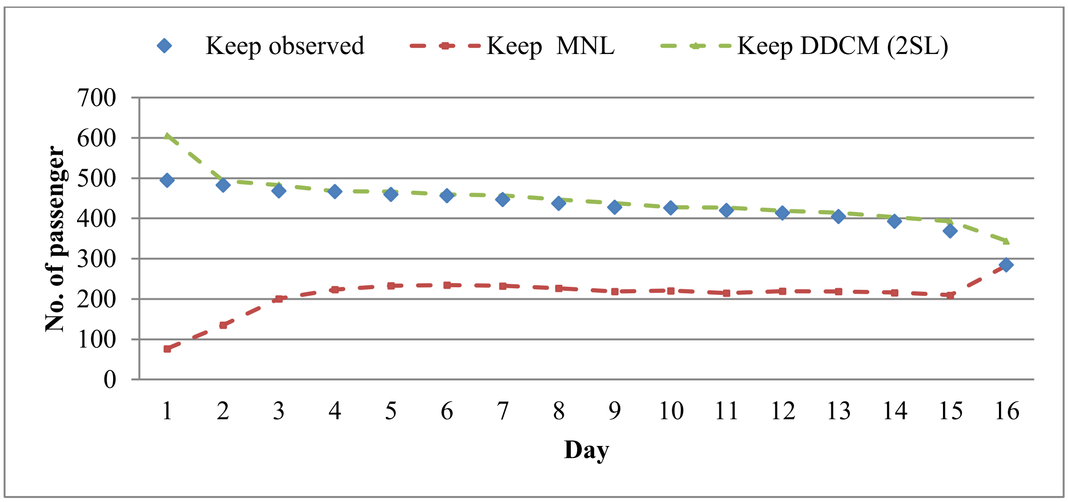

Figure 8: Validation of Keep Decision: Real Data.

The choice probability for each alternative observed and predicted together with measure of errors for the real data experiment is reported in Table 7. It shows that the *D* value of the dynamic model is significantly smaller than the correspondent value obtained with the MNL model (1.194 compared to 15.179) indicating a much better prediction capability of the dynamic model over the MNL model.

Table 7: Model Validation: Choice Probability of Real Data Experiment.

Alternative	Observed	Predicted (Static)	Predicted (Dynamic)
Exchange day1	0.0287	0.6297	0.0072
Exchange day2	0.0121	0.7244	0.0020
Exchange day3	0.0124	0.5797	0.0000
Exchange day4	0.0021	0.5171	0.0021
Exchange day5	0.0021	0.4929	0.0000
Exchange day6	0.0022	0.4804	0.0000
Exchange day7	0.0044	0.4799	0.0000
Exchange day8	0.0112	0.4799	0.0000
Exchange day9	0.0091	0.4856	0.0000
Exchange day10	0.0024	0.4662	0.0001
Exchange day11	0.0047	0.4763	0.0000
Exchange day12	0.0048	0.4521	0.0024
Exchange day13	0.0097	0.4420	0.0000
Exchange day14	0.0148	0.4314	0.0000
Exchange day15	0.0102	0.4254	0.0000
Exchange day16	0.1599	0.2220	0.0136
Cancel day1	0.2601	0.2601	0.1221
Cancel day2	0.0121	0.0026	0.0000
Cancel day3	0.0166	0.0048	0.0000
Cancel day4	0.0021	0.0064	0.0000
Cancel day5	0.0128	0.0079	0.0000
Cancel day6	0.0043	0.0093	0.0000
Cancel day7	0.0175	0.0112	0.0000
Cancel day8	0.0089	0.0130	0.0000
Cancel day9	0.0137	0.0148	0.0000
Cancel day10	0.0000	0.0182	0.0000
Cancel day11	0.0117	0.0206	0.0000
Cancel day12	0.0095	0.0252	0.0000
Cancel day13	0.0121	0.0300	0.0000
Cancel day14	0.0148	0.0351	0.0049
Cancel day15	0.0509	0.0410	0.0000
Cancel day16	0.0677	0.0111	0.0542
Keep day1	0.7112	0.1102	0.8707
Keep day2	0.9758	0.2729	0.9980
Keep day3	0.9710	0.4155	1.0000
Keep day4	0.9957	0.4765	0.9979

Alternative	Observed	Predicted (Static)	Predicted (Dynamic)
Keep day5	**0.9850**	**0.4991**	**1.0000**
Keep day6	**0.9935**	**0.5102**	**1.0000**
Keep day7	**0.9781**	**0.5090**	**1.0000**
Keep day8	**0.9799**	**0.5072**	**1.0000**
Keep day9	**0.9772**	**0.4995**	**1.0000**
Keep day10	**0.9976**	**0.5156**	**0.9999**
Keep day11	**0.9836**	**0.5030**	**1.0000**
Keep day12	**0.9857**	**0.5226**	**0.9976**
Keep day13	**0.9783**	**0.5280**	**1.0000**
Keep day14	**0.9704**	**0.5336**	**0.9951**
Keep day15	**0.9389**	**0.5336**	**1.0000**
Keep day16	**0.7723**	**0.7669**	**0.9322**
D		**15.1790**	**1.1940**

6.0 CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

This paper has proposed a dynamic discrete choice model for ticket cancellation and exchange in the context of railway ticket purchase for intercity trips. The methodological framework proposed considers forward looking agents that maximizes their inter-temporal payoffs when deciding about exchanging or cancelling their ticket. The classical formulation based on the optimal stopping problem derived from dynamic programming is preserved here, while an innovative and elegant scenario tree formulation is proposed to solve the issue of calculating the passengers' expected utility over time. The model is estimated using maximum likelihood estimation, which seems particularly appropriated in this finite horizon problem. The analysis makes an important contribution in the context of discrete choice models for revenue management as it allows to account for temporal effects on individual decisions that are usually treated in a static context. The model has been successfully estimated using both simulated and real data; results shows that DDCM outperforms MNL in reproducing the initial values assumed in the simulated dataset and in reproducing the actual choices in both synthetic and real data.

Several extensions of this work warrant attention. It would be desirable to incorporate the unobservable (or latent class) segments within the population using a discrete segmentation approach; latent class (LC) models in which classes are based on trip characteristics (i.e. as group size, departure time) appears to be well suited for this kind of analysis. The model can account for taste heterogeneity by incorporating mixed logit choice model (ML) thus allowing for different passenger choice preferences in a continuous segmentation framework. Finally the modeling approach applied here to railway revenue management could be applied to test other refund and exchange policies and in general to other problems for which it is relevant to model passenger decision over time.

7.0 REFERENCES

ALOGIT, 2007. User Manual ALOGIT version 4.2. ALOGIT Software & Analysis Ltd.

Aviv, Y., A. Pazgal. 2008. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing Service Operation Management* 10(3) 339–359.

Bastin, F. 2011. Another Mixed Logit Estimation Tool (AMLET) Reference Manual. http://amlet.slashbin.net/

Besanko, D., W.L. Winston. 1990. Optimal price skimming by a monopolist facing rational consumers. *Management Sci.* 36(5) 555–567.

Dana, J. 1998. Advance purchase discounts and price discrimination in competitive markets. *Journal of Political Economics.* 106(2) 395–422.

Gale, I. L., T. Holmes. 1992. The efficiency of advance-purchase discounts in the presence of aggregate demand uncertainty. *International Journal of Industrial Organization* 10(3) 413–437.

Gallego, G., O. Sahin. 2010. Revenue Management with Partially Refundable Fares OPERATIONS RESEARCH, Volume: 58 Issue: 4 Pages: 817-833, 2010

Gallego, G., G. Iyengar, R. Phillips, A. Dubey. 2004. Revenue management of flexible product. CORC Technical Report TR-2004-01, Department of Industrial Engineering and Operations Research, Columbia University, New York.

Gallego, G., R. Phillips. 2004. Revenue management of flexible products. *Manufacturing Service Operation Management* 6(4) 321–337.

Garrow, L., F. Koppelman. 2004a. Predicting air travelers' no-show and standby behavior using passenger and directional itinerary information. Journal of Air Transport Management 10, 401–411.

Garrow, L., F. Koppelman. 2004b. Multinomial and nested logit models of airline passengers' no-show and standby behavior. Journal of Revenue and Pricing Management 3 (3), 237–253.

Graham, R., L. Garrow, J. Leonard. 2010. Business travelers' ticketing, refund, and exchange behavior. Journal of Air Transport Management 16 (2010) 196–201

Iliescu, D. 2008. Customer Based Time-to-Event Models for Cancellation Behavior: A Revenue Management Integrated Approach, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Georgia Institute of Technology, 2008.

Iliescu, D., L. Garrow, R. Parker. 2008. Hazard models of US airline passengers' refund and exchange behavior. Transportation Research B 42, 229–242.

Koenigsberg, O., E. Muller, N. J. Vicassim. 2006. Should Easy Jet offer last minute deals? Working paper, Columbia University, New York

Landsberger, M., I. Meilijson. 1985. Intertemporal price discrimination and sales strategy under incomplete information. RAND J.Econom. 16(3) 424–430.

Levin, Y., J. McGill, M. Nediak. 2009. Optimal dynamic pricing of perishable items by a monopolist facing strategic consumers. *Production Oper. Management* 19(1) 40–60.

Liu, Q., G.J. van Ryzin. 2005. Strategic capacity rationing to induce early purchases. Working paper, Graduate School of Business, Columbia University, New York.

Miller, R. 1984. Job Matching and Occupational Choice *The Journal of Political Economy*, Vol. 92, No. 6. (Dec., 1984), pp. 1086-1120.

Neuling, S., S. Riedel, et al. 2004. New approaches to origin and destination and no-show forecasting: excavating the passenger name records treasure. Journal of Revenue and Pricing Management **3**(1): 62-72.

Pakes, A. 1986. Patents as options: Some estimates of the value of holding European patent stocks. Econometrica, 54:755-785, 1986.

Png, I. P. L. 1989. Reservations: Customer insurance in the marketing of capacity. *Marketing Sci.* 8(3) 248–264.

Ringbom, S., O. Shy. 2004. Advance booking, cancellations, and partial refunds. *Economic Bulletin.* 13(1) 1–7.

Rust, J. 1987. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. Econometrica, 55, 999–1033.

Rust, J. 1994. Numerical Dynamic Programming in Economics. Revised November1994 draft for Handbook of Computational Economics, 1994.

Shugan, S., J. Xie. 2000. Advance pricing of services and other implications of separating purchase and consumption. *J. Service Res.* 2(3) 227–239.

Spinler, S., A. Huchzermeier, P. Kleindorfer. 2002. An options approach to enhance economic efficiency in a dyadic supply chain. S. Seuring, M. Goldbach, eds. *Cost Management in Supply Chains*. Physica, Heidelberg, 350–360.

Spinler, S., A. Huchzermeier, P. Kleindorfer. 2003. Risk hedging via options contracts for physical delivery. *OR Spectrum* 25(3) 379–395.

Stokey, N. 1979. Intertemporal price discrimination. *Quart.J.Econom.* 93(3) 355–371.

Su, X. 2007. Inter-temporal pricing with strategic customer behavior. *Management Sci.* 53(5) 726–741.

Wolpin, K. 1984. An estimable dynamic stochastic model of fertility and child mortality. Journal of Political Economy, 92:852-874, 1984.

Xie, J., S. Shugan. 2001. Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Sci.* 20(3) 219–243.

Yu, M., R. Kapuscinski, H.S. Ahn. 2008. Advance selling with limited capacity. Working paper, Ross School of Business, University of Michigan, Ann Arbor.