Journal of the Eastern Asia Society for Transportation Studies, Vol. 6, pp. 303 - 317, 2005

PATTERN ANALYSIS ON THE BOOKING CURVE OF AN INTER-CITY RAILWAY

Chi-Kang LEE
Professor
Department of Transportation and
Communication Management Science
National Cheng Kung University
1, Da-Shiue Rd., Tainan, 70101 Taiwan
Fax: +886-6-275-3882
E-mail: cklee@mail.ncku.edu.tw

Tzuoo-Ding LIN
Associate Professor
Department of Transportation and
Communication Management Science
National Cheng Kung University
1, Da-Shiue Rd., Tainan, 70101 Taiwan
Fax: +886-6-275-3882
E-mail: tdlin@mail.ncku.edu.tw

Chao-Hui LIN
Master
Department of Transportation and
Communication Management Science
National Cheng Kung University
E-mail: chaohui@mail.tcm.ncku.edu.tw

Abstract: We present a pattern analysis on the booking curve of an intercity passenger railway. It consists of three steps: cluster analysis, discriminant analysis, and prediction analysis. Basic booking curve patterns are built on the base of risk preference and booking magnitude. Using four-week ticket sales data from Taiwan Railway Administration, a numerical case study of the rapid train on a long distance market is illustrated to demonstrate the characteristics of the proposed method. We obtain pretty good statistical results in cluster analysis and discriminant analysis. Although its overall prediction accuracy is not good, the proposed method provides a useful pattern analysis procedure. Many forecasting techniques of growth curve may be included in the process so as to improve overall fit of the method in the future.

Key Words: Booking curve, Multivariate analysis, Revenue management

1. INTRODUCTION

Booking curve is an important tool for booking control policy in revenue management. Revenue management is the art of maximizing profit generated from a limited capacity by selling each product to the right customer at the right time for the right price (Belobaba, 1987; McGill and van Ryzin, 1999). For railway problem, it encompasses practices of price-discrimination to the demand that occurs over time before the train's departure, over

space for various origin/destination pairs, and/or over service for different train or car classes. On the other hand, the heart of revenue management is to control the seat inventory, and its objective is to find the right combination of passengers of various price classes. Then, these optimal considerations have to be translated into a booking control policy, which determines whether or not to accept a booking request when it arrives. In order to support the strategy and tactic decisions, accurate demand forecasting is essential. For the railway with an advance reservation system, pattern and prediction analysis on booking data is one way to predict and control the seat, so as to reach the objective of revenue management. Booking curve is a graph of the expected cumulative demand for a train service at the time before the train's departure, for an origin/destination market. An example booking curve is the solid line illustrated in Figure 1. The booking control policy sets trigger points, which presented by the dash lines in Figure 1. A booking control action will be activated at the time, either increasing or decreasing the booking limit for demand, if the predicted booking curve is over or under the trigger points. The objective of this study is to analyze the patterns of booking curves for Taiwan Railway Administration (TRA), an intercity passenger railway.

Taiwan is an island with 36000 square-kilometers, and has population of 23 million. TRA has more than 2000 kilometers railway and runs more than 1000 intercity trains per day. Due to fierce competition in intercity transportation markets and some other reasons, highly regulated TRA cannot obtain enough operation revenue, and is currently in a process of business reengineering and regulation reform. Given limited supply of equipment and train, TRA is going to allocate capacity to the most profitable business and implement revenue management. TRA is highly regulated and currently has no price-discrimination policy. In order to implement revenue management and support demand modeling, collecting reservation data, analysis and refining that data, and communicating testing results with key decision makers is the first step.

Nowadays, passengers can obtain tickets by an advanced reservation system using Internet or telephone from 16 days before departure, or buy tickets at any TRA station from 6 days before departure. On average, about 30% passengers obtain tickets by the reservation system and 70% passengers buy tickets at station. TRA's ticket sales system is very old, and is undertaking an update project. There is now no database for advanced ticket reservation, and no database for ticket sales of some stations. For this study, we record detailed ticket sales data, both advanced reservation system and station sales system, for four weeks, from the March 24th to April 20th in 2003. Because there are a lot of testing results, in the paper, we focus our analysis on the most popular train and most important origin/destination pair. The market is from the capital city, Taipei, to the number two city, Kaoshuing, and its distance is 350 kilometers.

The organization of the paper is as follows. In the next section, we describe observed booking curve data, and basic patterns of booking curve. Two criteria are used to build the patterns of booking curve, and they are risk preference and demand level. In the third section, we develop a research procedure using multivariate analysis methods for pattern analysis on booking curve. It consists of three steps, and they are cluster analysis, discriminant analysis, and prediction analysis. After that, we present the testing results of the patterns of booking curve and discuss their implications in the fourth section. Finally, in the fifth section, we make the concluding remarks of the study, and discuss some directions of future research.

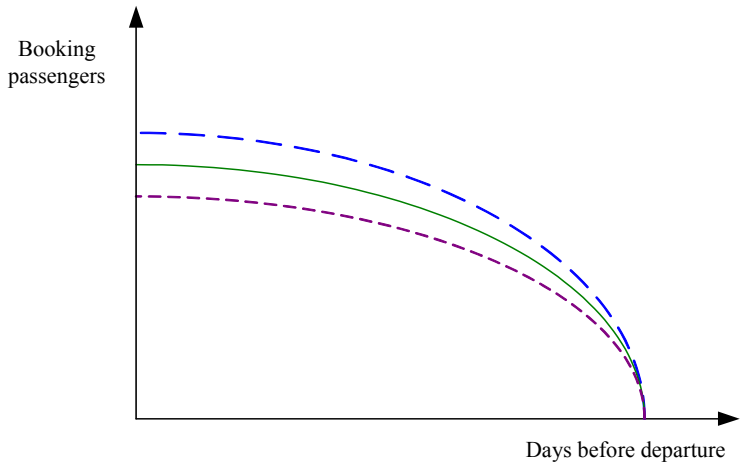
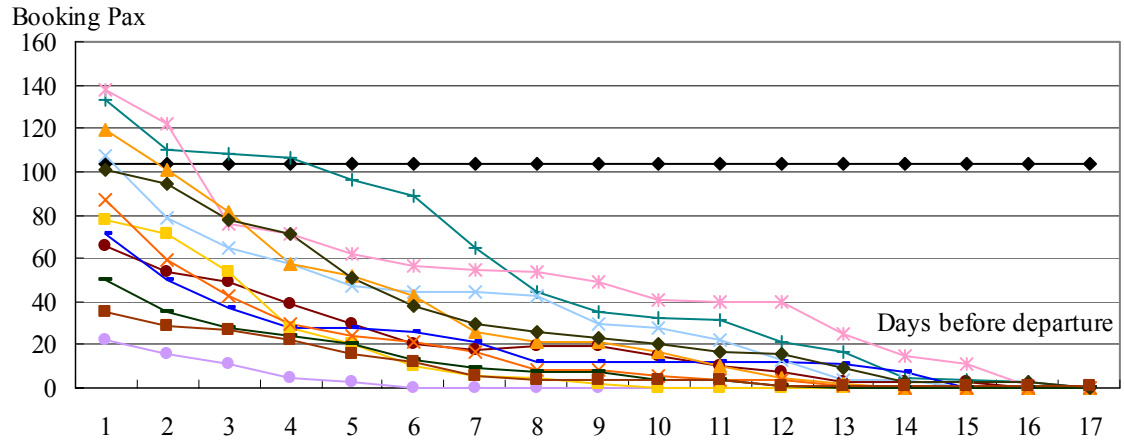


Figure 1. Expected Booking Curve and Booking Control Boundary

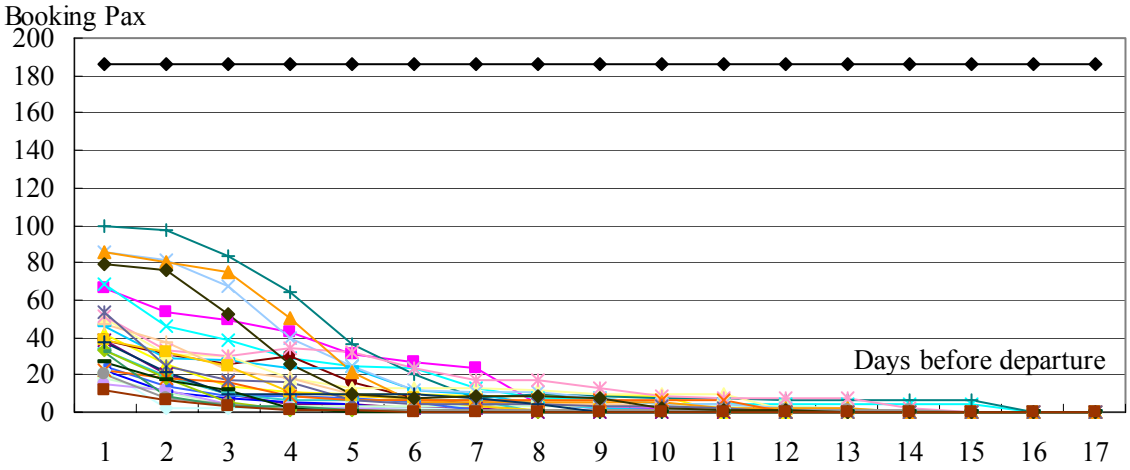
2. **BOOKING CURVE**

The most basic data used in the paper is $N_r^{w,d,t}$, the number of booked tickets for origin/destination pair *w*, departure date *d*, and train number *t*, at reservation date *r*. As the example shown in Figure 1, a booking curve is the cumulative booked tickets at every reservation date. For the specific market studied in the paper, booking curves of three trains are illustrated in Figure 2. Train 1005 only has service during weekend, that is Friday, Saturday, and Sunday; and its departure time is 7:30 A.M. and arrival time is 11:30 A.M. Train 1007 has service every day; and its departure and arrival time is 8:00 A.M. and arrival time is 12:41 A.M. Train 1009 has service every day; and its departure time is 8:36 and arrival time is 12:44. It is clear that these booking curves have different patterns, in terms of shape of the curve and the magnitude of booking level. In order to reduce the effect of magnitude, we standardized the data using booking rate, i.e. booking number/ booking limit. Figure 3 shows standardized booking curves for the three trains. The basic data or a point in a standardized booking curve is $X_r^{w,d,t}$, the cumulated booking rate for origin/destination pair *w*, departure date *d*, and train number *t*, at reservation date *r*. A booking curve is a vector of $X_r^{w,d,t}$ and is written as $X^{w,d,t}$. It is evident that some curves

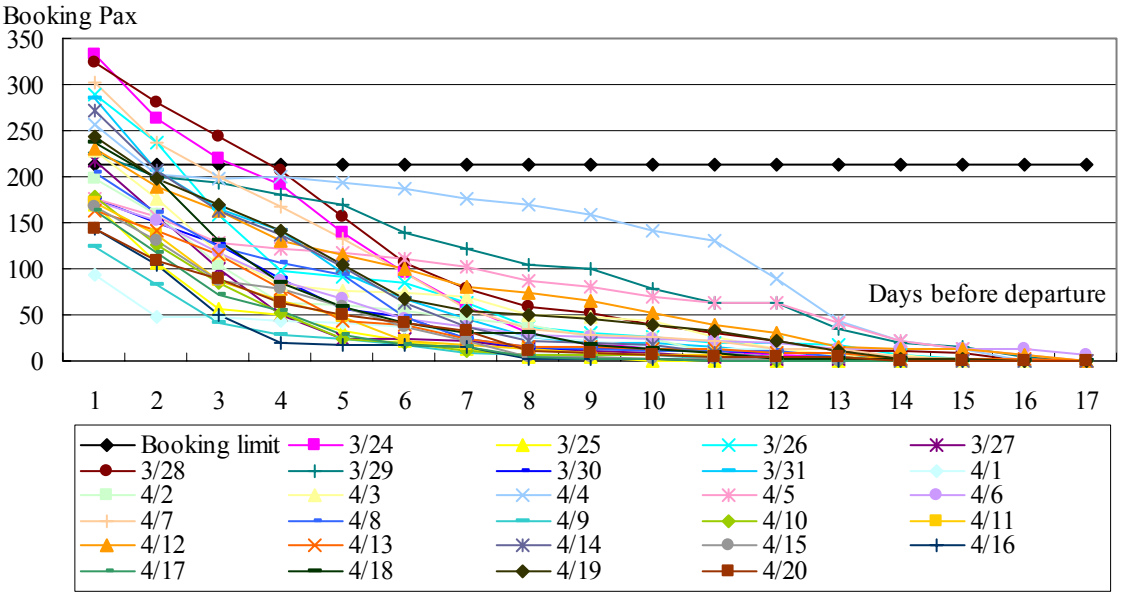
different in Figure 2, are similar in Figure 3



(a) Train Number 1005.

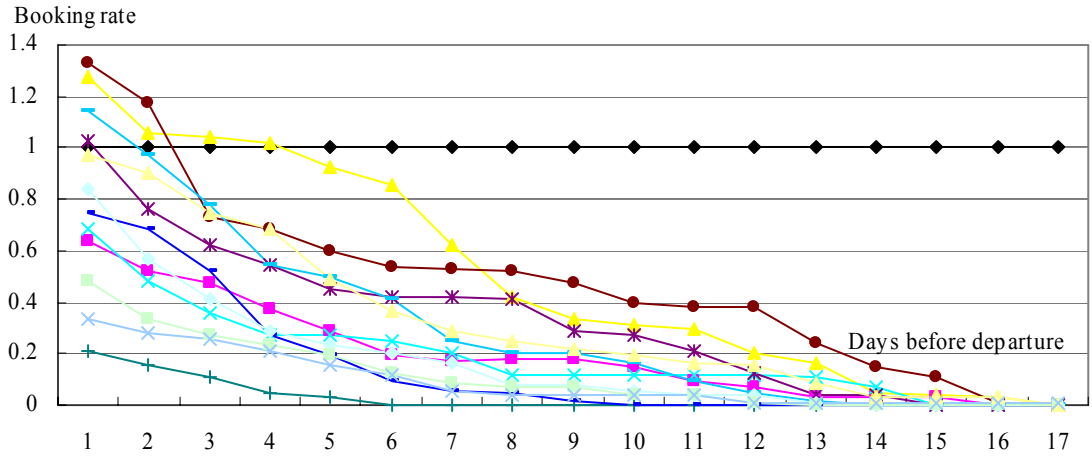


(b) Train Number 1007.

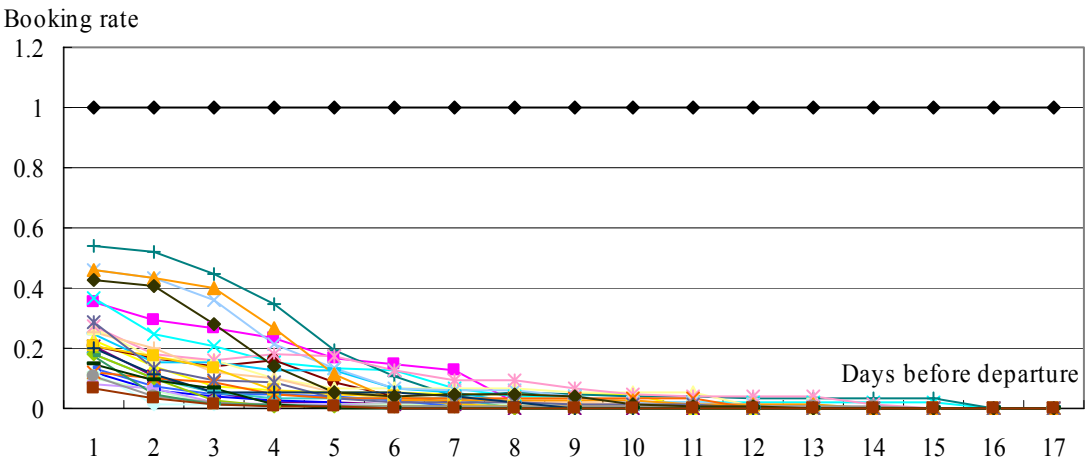


(c) Train Number 1009.

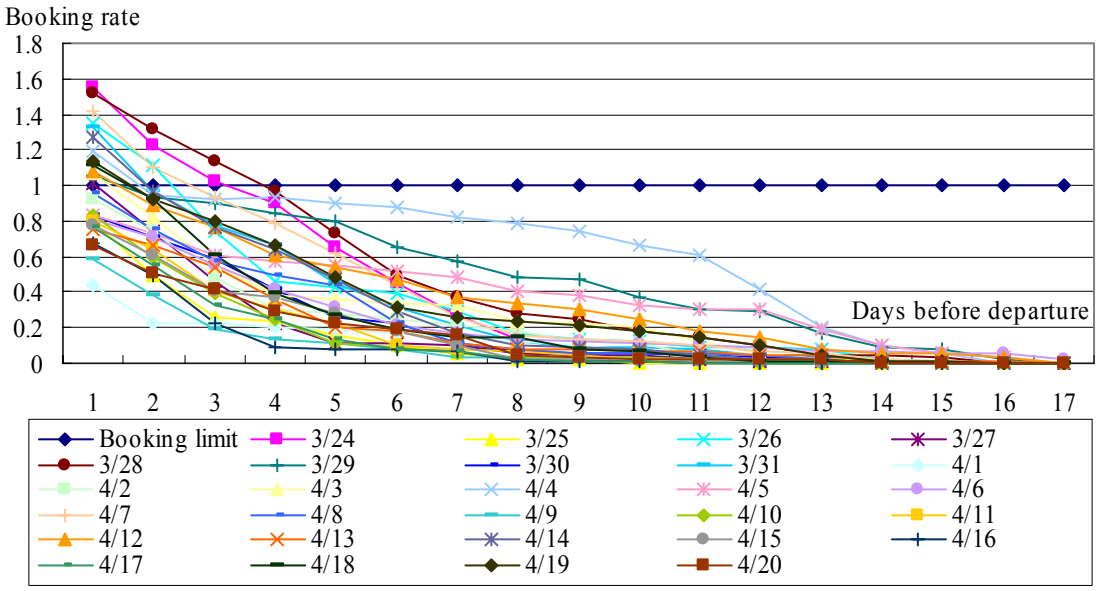
Figure 2. TRA's Booking Curves



(a) Train Number 1005.



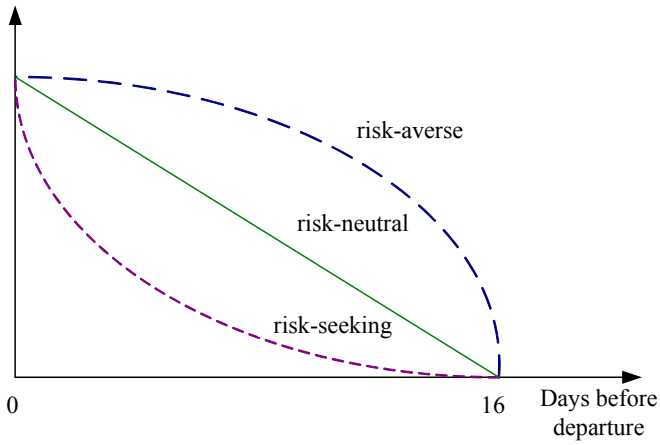
(b) Train Number 1007.



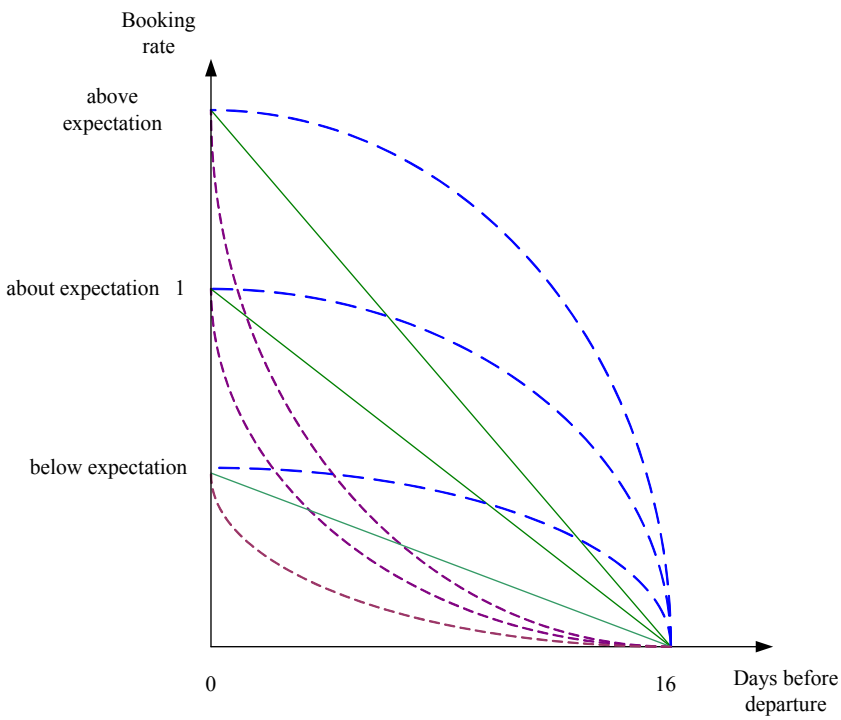
(c) Train Number 1009.

Figure 3. TRA' s Standardized Booking Curves.

After examining the observed booking curves, we generate basic patterns of booking curve using two criteria: risk level and booking level. As illustrated in Figure 4(a), there are three types of booking curve due to risk level: risk-averse, risk-neutral, and risk-seeking. For the risk-averse booking curve, many passengers book tickets at early dates in the reservation period. On the contrary, for the risk-seeking booking curve, most passengers obtain tickets at dates close to the departure time. Moreover, comparing to booking limit, there are several magnitude levels for the booking curves: e.g. above expectation, about expectation, below expectation, and far below expectation. Therefore, as shown in Figure 4(b), there are many types of booking curve based on the combinations of risk preference and magnitude level, and some of them will be shown in our empirical study.



(a) The Shape of Booking Curve



(b) The Shape and Magnitude of Booking Curve.

Figure 4. Booking Curve Pattern.

3. THE METHOD

The proposed method of pattern analysis consists of three steps. They are cluster analysis, discriminant analysis, and prediction analysis. The first and second steps are very popular in multivariate data analysis in all fields for pattern analysis (Hair, Anderson, Tatham, and Black, 1998). Prediction analysis is in the proposed process, because the purpose of pattern analysis on booking curve is to estimate the expected booking curve, so as to implement booking control policy. Hence, the effect of prediction accuracy should be at least one criteria considered in cluster analysis and discriminant analysis.

3.1 Cluster Analysis

Cluster analysis is to partition the set of booking curves into several groups based on the pattern of curve. The objective of cluster analysis is booking curve, $X^{w,d,t}$. The similarity of the booking curves is focused on the proximity of objects, and it is measured by squared Euclidean distance. That is, the similarity of two booking curves is $\sum_r (X_r^{w,d,t} - X_r^{w,d',t'})^2$. In other words, we measure the correspondence of booking rate across the reservation date. Cluster algorithm is used to maximize the difference between clusters relative to the variation within the clusters. The nonhierarchical or K-means clustering method contained in SAS is used in the study (Sarle, 1983). Using the algorithm, the best cluster solution is to be found once the number of cluster is specified. In other words, the six-cluster solution is not a combination of two clusters from the seven-cluster solution. Selecting seed points, $[Z^1, Z^2, ..., Z^g]$, is the first step in the clustering algorithm; where Z^j is the cluster centroid of the booking curves in the *j*th cluster. In the study, the initial seed points are the average points in the data set. Secondly, every booking curve is assigned to its nearest seed. That is, a booking curve $X^{w,d,t}$ is assign to group *j* if $\sum_r (X_r^{w,d,t} - Z_r^j)^2$ is the minimum of { $\sum_r (X_r^{w,d,t} - Z_r^j)^2$, for *j* =*1~g*}. After that, the seed points are updated by $Z_{new}^j = Z_{old}^j + \eta(\overline{X^{w,d,t}} - Z_{old}^j)$ until the algorithm is converged, where η is a parameter of learning rate. In order to decide the number of clusters should be formed, we consider several stopping rules. First, we prefer to have low value of the average within cluster distance. Secondly, two statistics showing notable success in the literature, Pseudo F statistic and Cubic Clustering Criteria (CCC), are used to check the similarity improvement at each successive step (Sarle,, 1993). Pseudo F statistic is the ratio of the between-cluster variation to the average within-cluster variation, and it equals to $[R^2/(g-1)]/[(1-R^2)/(n-g)]$, where R^2 is the coefficient of determination, g is the number of clusters, and n is the sample size. CCC is written as $\ln[(1-E(R^2))/(1-R^2)]*K$, where *R* is the coefficient of determination and *K* is the variance stabilizing transformation.

3.2 Discriminant Analysis

In cluster analysis, we classify booking curves into *g* groups. Then, discriminant analysis is used to explain how the groups differ on relevant dimensions (Huberty, 1994). First, we have to identify dimensions of discrimination between groups and select independent variables. For a booking curve in origin/destination market *w*, $X^{w,d,t}$, characteristics associated with departure date *d* and train number *t* are the dimensions of discrimination between groups. Please refer to Table 1 for a list of the variables. With regard to departure date, there are 7 week variables, $W_1 \sim W_7$, and 5 festival holiday variables, $H_0 \sim H_4$. For example, H_1 is an index variable for the departure date just one day before a long holiday. With regard to departure time or arrival time, there are 4 time of day variables, $G_1 \sim G_4$ or $T_1 \sim T_4$. For example, D_3 is an index variable for the departure time at afternoon period 15:00 to 19:00. Consequently, considering the joint effect of two independent variables, there are several hundreds of interaction variables. For example, H_1T_3 is one, if the departure time is the afternoon just before a long holiday; otherwise, it is zero. With all single and interaction variables, the stepwise estimation method contained in SPSS is used to calibrate the discriminant functions for the *g* groups (Weinberg, 2002). After that, hit ratio and some statistics are used to assess overall fit of the discriminant model.

Table 1. Definition of Clustering and Discriminating Variables

W_i=0 or 1, i=1~7. W_i represents the ith day of week.

It is 1 if the train is in that day of week; otherwise, it is 0.

H_j=0 or 1, j=0~4. H_j represents the characteristics of festival holiday. H_0 represents normal time, H_1 represents one day before festival, H_2 represents the festival period, H_3 represents one day after festival, and H_4 represents the last day of festival.

It is 1 if the train is in that day; otherwise, it is 0.

G_k=0 or 1, k=1~4. G_k represents the kth period of day. Time periods are four: morning, noon, afternoon, and night.

It is 1 if the departure time of the train is that period of day; otherwise, it is 0.

T_l=0 or 1, l=1~4. T_l represents the lth period of day. Time periods are four: morning, noon, afternoon, and night.

It is 1 if the arrival time of the train is that period of day; otherwise, it is 0.

There are about four-hundred interaction variables, and they are W_iH_j, W_iG_k, W_iT_l, H_jG_k, H_jT_l, and G_kT_l. They represent "and" relationship between two variables.

It is 1 if both variables are 1; otherwise, it is 0.

3.3 Prediction Analysis

First, predication analysis is to check the capability of the above mentioned process, cluster analysis and discriminant analysis, for predicting the booking curve of a train service. For a future and unknown booking curve $X^{w,d,t}$, given the characteristics of origin/destination pair, a departure date, and a train number; we estimate its cluster number, *j*, using the discriminant functions. Then, with the cluster number, we have the cluster booking curve Z_r^j, which is the centroid of the historical booking curves in the cluster. We use the booking curve Z_r^j as the expected booking curve of $X^{w,d,t}$. In the study, the corresponding values in the real booking curve $X^{w,d,t}$ and in the expected booking curve Z_r^j, are compared by mean absolute percentage error (*MAPE*). *MAPE* is written as follows, where $X_r^{w,d,t}$ is the real booking rate, Z_r^j is the estimated booking rate, and n_r is the number of reservation dates. In general, the prediction capability is good, if MAPE is less than 20%. Moreover, we test the prediction oriented pattern analysis. That is, we choose the number of clusters and its associated discriminant functions, using the criteria of MAPE.

$$MAPE = [\frac{1}{n_r}\sum_r \frac{|X_r^{w,d,t} - Z_r^j|}{X_r^{w,d,t}}]100\% \quad (1)$$

4. THE RESULT

4.1 Cluster Analysis

For the market of Taipei to Kaohsiung and the rapid train, we get 647 booking curves. Following the method discussed previously, we run the K-means method of cluster analysis by SAS, with standardized booking curves. In order to decide the number of clusters, we check several criteria. As the pseudo F statistic and Cubic Clustering Criteria shown in Figure 5, there is a peak at 7 clusters. With 7 groups, the average booking curves, or cluster centroids, are illustrated in Figure 6. It is evident that there are one booking curve of risk-averse, two booking curves of risk-neutral, and four booking curves of risk-seeking. Based on the basic patterns of booking curve described in Figure 4 and discussed previously, we interpret and name the clusters in Table 2. Some people in Taiwan are not used to have activity plan, and prefer to deal with trouble when it happens. It is also the case for passengers' booking behavior, even for some hot trains with high load factors, e.g. cluster 4 and 5. However, with the development of advanced reservation system and dynamic pricing strategy, it is in a transition process.

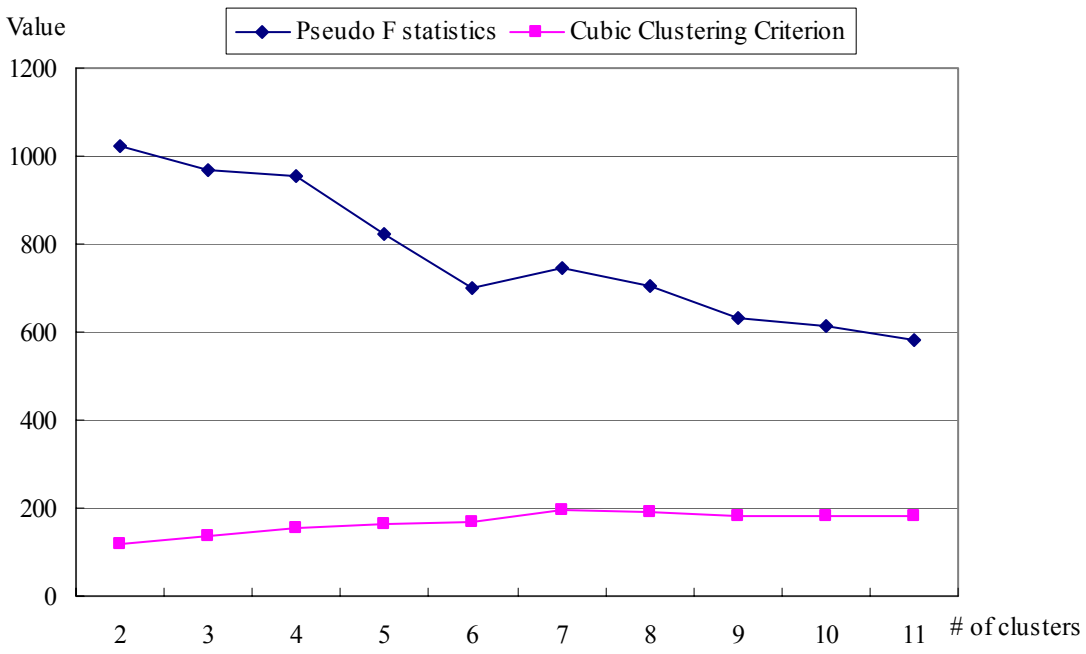


Figure 5. Pesudo F Statistic and Cubic Clustering Criteria in Cluster Analysis

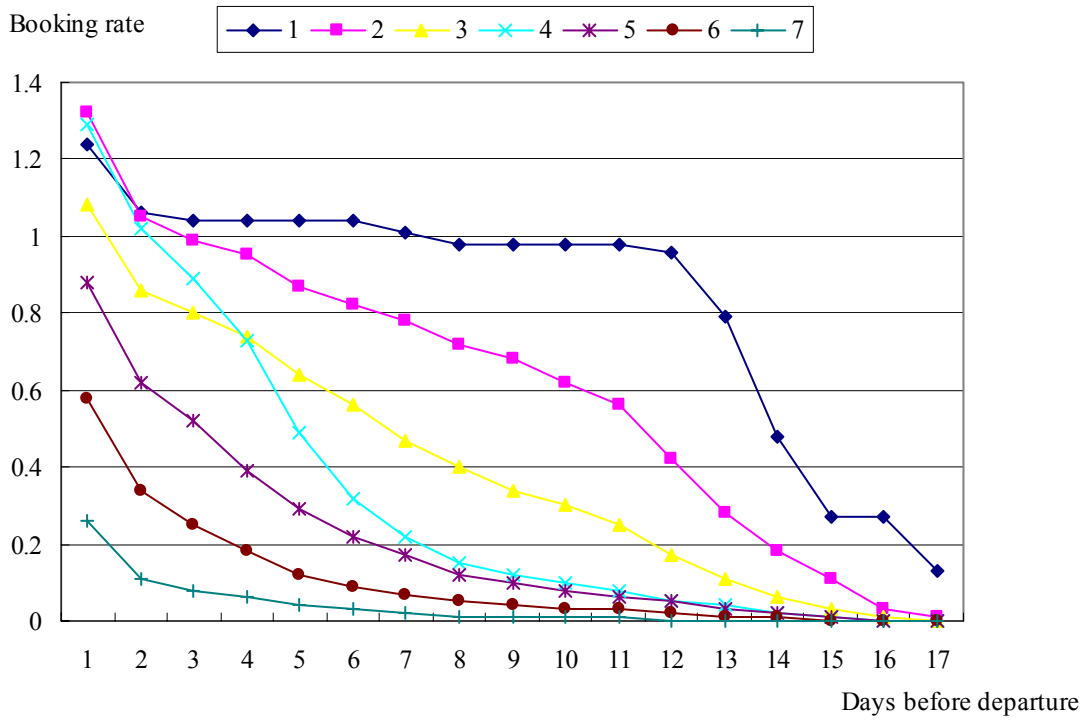


Figure 6. TRA's Booking Curve Pattern

Table 2. Seven-Cluster Solution in Cluster Analysis

Cluster	Name of Cluster	# of booking curves	Percentage (%)
1	Risk-averse/ Above expectation (1.2)	3	0.46
2	Risk-neutral/ Above expectation (1.3)	18	2.78
3	Risk-neutral/ About expectation (1.1)	42	6.49
4	Risk-seeking/ About expectation (1.3)	42	6.49
5	Risk-seeking/ About expectation (0.9)	105	16.23
6	Risk-seeking/ Below expectation (0.6)	158	24.42
7	Risk-seeking/ Far below expectation (0.3)	279	43.12

(Final booking rate = passenger volume/ booking limit)

4.2 Discrimination Analysis

For the 7-cluster solution, we first identify variables of discrimination. Using the variables, described previously and listed in Table 1, we describe the chacterisitcs of each cluster in Table 3. For example, passengers of cluster 1 have risk-averse booking behavior for the trains, running in Sunday or the end of festival holiday, departing Taipei at noon and/or arriving Kaohsiung in the evening. In order to have a good explanation, we have to use several variables simultaneously. That is, we need to consider not only independent explanatory variables, but also interaction variables. As the list in Table 1, there are 20 independent variables, and about 4-hundred interaction variables. With all these independent and interaction variables, we run the stepwise estimation method of discriminant analysist by SPSS. As the result illustrated in Table 4, only interaction variables are selected, with significant statistical effect. The discriminant functions and classification functions are listed in Table 5. The importance of the variables can be identified by the absolute values of their coefficients. It is clear that H_4G_3 (the end of festival holiday and departing in the afternoon), H_4G_2 (the end of festival holiday and departing at noon), H_1G_2 (one day before festival holiday and departing in the afternoon), have high values of coefficients, and there are especially the situations for passengers to book tickets in risk-averse manner. Moreover, by several statistics, such as Wilks' Lambda value and Chi-square test for discriminant functions, the difference between clusters is in general significant. At last, the overall hit ratio is 58.9%; and the hit ratio for each cluster is respectively 100%, 77.28%, 45.24%, 54.76%, 24.26%, 35.44%, and 86.02%. Therefore, we have pretty good overall fit, although we need to understand more on cluster 3, 4, 5, and 6.

Table 3. Clustering Variable Profiles for Seven-Cluster Solution

Cluster	Day of week	Holiday	Departure period	Arrival period
1	Sun	End of holiday	Noon	Evening
2	Fri	The day before holiday	Noon	Afternoon
3	Fri	The day before holiday	Afternoon	Evening
4	Fri ~ Sun	Not holiday	Noon	Afternoon and evening
5	Mon ~ Sun	Not holiday	Afternoon	Evening
6	Mon ~ Sun	Not holiday	Noon	Afternoon and evening
7	Mon ~ Sun	Not holiday	Morning and noon	Noon and evening

Table 4. Discriminant Analysis Results Using the Stepwise Estimation Method

Variables	Tolerance value	F statistic	Wilks' Lambda value
W_5H_1	0.56	11.23	0.08
H_4T_3	0.71	33.24	0.10
W_7T_4	0.67	32.38	0.10
W_5G_3	0.79	20.50	0.09
G_2T_3	0.71	11.52	0.08
W_6G_1	0.87	15.41	0.08
H_1T_3	0.54	4.95	0.08
H_2T_1	0.95	9.49	0.08
W_5G_4	0.89	10.08	0.08
W_7T_3	0.50	14.19	0.08
G_1T_3	0.65	4.66	0.08
H_4T_4	0.21	155.67	0.18
H_4G_3	0.29	85.00	0.13
H_4G_4	0.48	52.18	0.11
W_5G_2	0.53	5.59	0.08
H_2G_2	0.94	5.41	0.08
W_1G_1	0.95	4.39	0.08
H_1G_2	0.31	3.96	0.08

Table 5. Discriminant and Classification Functions

	Discriminant Functions						Fisher's Classification Functions						
	1	2	3	4	5	6	1	2	3	4	5	6	7
W_5H_1	0.36	0.62	1.02	1.45	-0.19	2.04	3.61	6.39	3.48	3.40	3.83	2.67	1.37
H_4T_3	0.52	0.86	1.34	1.77	-0.61	1.25	3.44	7.17	2.52	3.20	3.68	2.37	0.45
W_7T_4	1.80	1.99	1.66	-2.86	0.76	-0.44	11.71	4.50	2.19	14.22	9.20	6.23	0.37
W_5G_3	2.45	2.53	1.11	-1.77	-0.86	-0.46	15.78	7.59	2.12	15.68	15.89	6.00	0.42
G_2T_3	1.20	1.50	1.66	-0.51	1.27	1.80	8.55	6.25	3.96	9.63	7.47	5.40	0.86
W_6G_1	1.58	1.50	0.15	-0.53	-0.85	1.50	10.86	5.16	2.21	9.53	12.20	3.08	0.72
H_1T_3	1.34	1.72	1.79	2.05	-4.62	-0.30	7.30	11.77	0.77	7.17	7.42	1.78	-0.21
H_2T_1	1.42	2.03	2.60	1.27	0.22	-1.89	9.00	11.64	3.42	9.31	8.69	7.16	0.81
W_5G_4	-1.26	-1.22	-0.09	-1.21	6.29	1.37	-5.79	-6.70	2.87	-4.31	-7.27	2.23	1.36
W_7T_3	0.34	0.57	0.89	0.56	1.90	0.98	3.30	3.74	3.20	3.14	3.46	3.67	1.06
G_1T_3	1.08	0.61	-2.39	-0.05	3.79	-4.48	6.75	-3.20	-1.33	-0.35	13.85	2.35	-0.42
H_4T_4	0.76	1.10	1.43	0.88	-0.27	-1.04	4.77	6.75	1.76	4.89	4.65	3.67	0.45
H_4G_3	-13.34	12.63	0.10	-4.57	-0.77	-0.27	-339.02	-2.15	-0.42	-3.28	-34.77	1.29	-0.06
H_4G_4	-14.91	11.43	0.02	-0.86	5.37	-7.05	-349.57	-3.95	-0.91	-17.80	-41.34	3.55	-0.12
W_5G_2	1.41	0.56	-3.00	-0.92	-4.57	2.63	8.53	-2.16	-2.84	3.98	13.08	-5.08	-1.01
H_2G_2	4.46	4.05	-2.14	5.26	-2.07	-0.97	28.98	18.06	3.61	13.31	44.65	6.33	0.88
W_1G_1	4.63	2.78	-7.81	-0.20	3.60	-1.21	31.11	-3.06	-0.48	10.10	52.12	2.51	-0.15
H_1G_2	14.59	-11.78	-0.41	-0.75	0.21	2.11	347.48	-2.15	0.05	15.17	39.79	-0.25	-0.01
Const	-0.81	-0.72	-0.57	-0.23	-0.36	-0.28	-183.61	-7.96	-2.43	-10.56	-25.23	-4.00	-0.94

4.3 Prediction Analysis

Prediction of booking curve is important in revenue management, for implementing booking control policy. We use mean absolute percent error (MAPE) to test, if the cluster centroid is a good estimate for the expected booking curve or not. As testing results of MAPE illustrated in Table 6, with 7-cluster solution, the error on predicting booking rate by cluster centroid is 56.7%. If the discriminant analysis in the process of pattern analysis has 100% hit ratio, every booking curve will be assign to the correct cluster, then MAPE is 31.5%. That is, the variation between clusters, or the possible improvement in discriminant analysis, is 25.2%; and 31.5% error is due to the variation in the cluster. Although there is a specific objective function in cluster analysis or discriminant analysis, the overall objective of the pattern analysis is to predict booking curve. Table 6 shows the result of a prediction oriented pattern analysis, that is, we choose the number of clusters and

do discriminant analysis accordingly based on MAPE. As the number of clusters increases, the hit ratio of discriminant analysis decreases, MAPE due to the variation between clusters increases, and MAPE due to the variation in the cluster decreases. It is clear that the hit ratio of discriminant analysis is not a good overall fit index for the prediction purpose. Overall MAPE is dependent on both cluster analysis and discriminant analysis. Six-cluster solution gives the best value of MAPE, 55%. However, the overall MAPE is in general not good for most cases; in order to fulfill the prediction requirement, some forecasting techniques of growth curves have to be included in the pattern analysis in the future.

Table 6. Prediction Oriented Cluster and Discriminant Analysis

The number of clusters	2	3	4	5	6	7	8	9
The Hit ratio in discriminant analysis	87%	80%	72%	64%	63%	59%	57%	56%
Mean Absolute Percent Error	66%	65%	58%	58%	55%	57%	57%	57%
Variation in the cluster	59%	50%	42%	41%	36%	32%	30%	29%
Variation between clusters	7%	16%	17%	17%	20%	25%	27%	27%

5. CONCLUDING REMARKS

We present a pattern analysis on the booking curve of an intercity passenger railway. It consists of three steps: cluster analysis, discriminant analysis, and prediction analysis. Basic booing curve patterns are built on the base of risk preference and booking magnitude. Using four-week ticket sales data from Taiwan Railway Administration, a numerical case study of the rapid train on a long distance market is used to demonstrate the characteristics of the proposed method. We obtain pretty good results in cluster analysis and discriminant analysis. The prediction error with the booking curve generated by cluster analysis and discriminant analysis is not acceptable. The prediction oriented process of three-step pattern analysis can make some improvement, but not much. Although its overall prediction accuracy is not acceptable, the proposed method provides a useful pattern analysis procedure. Many forecasting techniques of growth curve may be included in the process so as to improve overall fit of the method.

ACKNOWLEDGEMENTS

The author is grateful to his research assistant, Mr. Lin, T-H, for his computation works, and to National Science Council for the financial support under NSC93-2211-E-006-052.

REFERENCES

Belobaba, P.P. (1987) Airline Yield Management: An Overview of Seat Inventory Control, **Transportation Science, Vol. 21**, 63-73.

Hair, J., Anderson, R., Tatham, R. and Black, W. (1998) **Multivariate Data Analysis.** Fifth edition. Prentice-Hall Inc, USA.

Huberty, C. (1994) **Applied Discriminant Analysis.** John Wiley & Sons, Network York.

McGill, J.I. and van Ryzin, G.J. (1999) Revenue Management: Research Overview and Prospects, **Transportation Science, Vol. 33,** 233-256.

Sarle, W. (1983) **Cubic Clustering Criterion.** SAS Technical Report A-108. SAS Institute Inc, Cary, NC.

Sarle, W. et al. **The MODECLUS Procedure.** SAS Technical Report P-256. SAS Institute Inc, Cary, NC.

Weinberg, S.L. (2002) **Data Analysis for the Behavior Sciences Using SPSS**. Cambridge University Press, Cambridge, UK.