**Journal of the Operational Research Society (2004) 55**, 535–541  0160-5682/04 $25.00

www.palgrave-journals.com/jors

# Maximizing revenue in the airline industry under one-way pricing

MF Anjos*, RCH Cheng and CSM Currie

*University of Southampton, Southampton, UK*

The paper describes a methodology that has been implemented in a major British airline to find the optimal price to charge for airline tickets under one-way pricing. An analytical model has been developed to describe the buying behaviour of customers for flights over the selling period. Using this model and a standard analytical method for constrained optimization, we can find an expression for the optimal price structure for a flight. The expected number of bookings made on each day of the selling period and in each fare class given these prices can then be easily calculated. A simulation model is used to find the confidence ranges on the numbers of bookings and these ranges can be used to regulate the sale of tickets. A procedure to update the price structure based on the remaining capacity has also been developed.

*Journal of the Operational Research Society* (2004) **55,** 535–541. doi:10.1057/palgrave.jors.2601721

**Keywords:** air transport; revenue management; simulation; yield management

## Introduction

The paper describes a methodology that has been developed for a major British airline to find the optimal pricing strategy for tickets under one-way pricing, that is, where single rather than return tickets are sold. The same rules and restrictions apply to all of the tickets sold; therefore, the tickets are identical and the price is assumed to vary only with the time left until departure. In general terms, the problem consists of maximizing the revenue made from a fixed stock of identical, perishable products. This arises in a number of different sectors of the travel and leisure industry (eg pricing of hotel rooms, cabins on cruise liners, cinema tickets) and in the sale of fashion or seasonal products in the retail industry. Other authors have concentrated on the sale of fashion goods.[1,2]

An analytical model has been developed to describe booking patterns for flights under one-way pricing. Using the analytical model and calculus of variations with Lagrangian multipliers, a standard analytical technique for constrained optimization, we can derive an expression for the optimal price structure.

Parameter values in the model are selected by finding the best fit of the model's predicted booking profile for a particular flight to historical booking profiles for that flight. Once the best-fit parameter values have been calculated, we can determine the optimal price structure and the expected booking profile for the selling period. For practical purposes, we also need confidence ranges for the numbers of bookings as a means of checking that the sales are sufficiently close to the behaviour predicted by the model. We obtain the confidence ranges using a simulation model that is based on the analytical model of buying behaviour. The optimal price structure is updated throughout the selling period using data arriving during the selling period on the numbers of bookings made.

The organization of the article is as follows. We describe the context of the work in the second section and in the following four sections describe the analytical model, how the model is fitted, the simulation model used to generate confidence ranges on the booking profiles, and the method used to update the optimal price structure. We have applied this methodology to real data from our industrial partner but due to its sensitivity, this cannot be presented here. We therefore illustrate our methodology with a fictional example and this is presented in the Results section. We conclude the paper with an assessment of the model's performance and further work to be carried out in this area.

## Context

The advent of low-cost carriers into the short-haul European airline market has led to a dramatic shake-up of the pricing strategies of many of the major airlines. An increasing number of tickets are now sold on a one-way basis rather than as a return and this has removed many of the rules and restrictions traditionally associated with lower priced airline tickets, such as the requirement that a stay is over a Saturday night or is of a sufficient duration. In this context, all tickets are now identical (with the exception of those in business

*Correspondence: MF Anjos, Operational Research Group, Faculty of Mathematical Studies, University of Southampton, Southampton, SO17 1BJ, UK.
E-mail: anjos@stanfordalumni.org

class), and the price must be determined solely by the number of days left until departure.

The methodology that we describe in the paper has been developed for a British airline for setting economy fares on domestic and European routes. We do not consider issues concerning the optimal mix of business and economy passengers nor the optimal prices for business class passengers. For confidentiality reasons, the airline's name will not be disclosed.

## Analytical model of buying behaviour

For any given flight, we assume that the expected number of seats sold $x$ days before departure at a price $y$ is equal to the number of people who want to buy a ticket for the flight on day $x$ multiplied by the probability that someone who wants to buy a ticket on day $x$ is prepared to pay a price $y$ for that ticket.

We define $p(x, y(x))$ to be the probability of buying a seat on day $x$ at price $y$, given that you want to travel $x$ days later. The number of people who wish to buy a ticket for the flight $x$ days before departure is defined to be $f(x)$. Therefore, the number of bookings made $x$ days before departure if the price structure is $y(x)$ is

$$n(x,y) = \int_{x}^{x_{\text{start}}} f(x')p(x', y(x'))\,\mathrm{d}x' \tag{1}$$

the integral from the start of the selling period $x_{\text{start}}$ up to day $x$ of the potential demand multiplied by the probability of buying a ticket. Note that time is measured in days left until departure, therefore $x$ decreases as departure is approached. The model of buying behaviour is continuous and we are therefore assuming that customers can buy a ticket at any time during the selling period.

The choice of functions $p(x, y(x))$ and $f(x)$ is made by careful examination of the behaviour of the bookings as the price $y$ and the number of days until departure $x$ vary. For the fictional example, we choose

$$\begin{aligned} p(x, y(x)) &= \mathrm{e}^{-y(a+bx)} \\ f(x) &= (gx + d)\mathrm{e}^{-hx} \end{aligned} \tag{2}$$

This choice of functions seems reasonable for the sale of airline tickets. As the price increases the probability of someone buying a ticket on a given day $x$ decreases, that is, $p(x, y(x))$ decreases as $y$ increases for constant $x$. As we approach departure, and $x$ decreases, the probability of someone buying a ticket at price $y$ increases, that is, $p(x, y(x))$ increases as $x$ decreases for constant $y$. For zero price, the probability of purchasing a ticket is 1. The potential demand, on the other hand, rises to a peak before departure and then declines to a value $d$ at departure.

The same methodology can be used for different forms of $f(x)$ and $p(x, y(x))$ and is thus appropriate to a wide variety of other contexts of pricing of perishable goods.[3]

### *Finding the optimal price*

The expected revenue on day $x$ is the price of a ticket on day $x$ multiplied by the number of tickets sold on that day. To obtain the total revenue for the flight $R$, we integrate over the selling period.

$$R = \int_{0}^{x_{\text{start}}} y(x)f(x)p(x, y(x))\,\mathrm{d}x \tag{3}$$

We wish to maximize the total revenue $R$, subject to the constraint that the number of tickets sold does not exceed the capacity of the aircraft (assuming no overbooking is allowed). Where $C$ is the capacity of the aircraft, the capacity constraint can be written as

$$\int_{0}^{x_{\text{start}}} f(x)p(x, y(x))\mathrm{d}x \leqslant C \tag{4}$$

We use the method of calculus of variations with Lagrangian multipliers to maximize $R$ subject to the capacity constraint in Equation (4) and find that $y(x)$ should be of the form

$$y(x) = \frac{-p(x, y(x))}{\partial p/\partial y} + \lambda \tag{5}$$

where $\lambda$ is the non-negative Lagrangian multiplier associated with the capacity constraint.

Substituting the expression for $p(x, y(x))$ given in Equation (2) into the expression for the optimal price in Equation (5), we deduce that for our fictional example,

$$y(x) = \frac{1}{a + bx} + \lambda(a, b, d, g, h, C) \tag{6}$$

where $\lambda$ is dependent on the capacity of the plane $C$ and the model parameters $a$, $b$, $d$, $g$ and $h$. If the total demand is smaller than the capacity of the plane, the capacity constraint (Equation (4)) is said to be inactive and $\lambda$ is zero. As the demand increases beyond the capacity of the plane, the capacity constraint is said to be active and $\lambda$ becomes positive and then increases as the difference between the demand and the plane's capacity increases.

A continuous pricing strategy is optimal under the assumption that price changes have no associated costs. In reality, this is not the case. For example, it costs money to advertise fares and to inform travel agents or the general public of those fares. If fares change frequently, the company must spend more money informing its customers of the changes and risks upsetting customers who may expect prices to be reasonably static over most of the selling period. Without knowing the exact cost of price changes, it is hard to factor these into the analysis. We have therefore assumed that a prescribed structure of discrete fares is provided in advance by the airline, and set the price charged to be the fare closest to the optimal continuous price. The resulting methodology is easy to explain and works reasonably well in practice.

### *Cancellations*

In the airline industry, the cancellation rate on some flights can be considerable, therefore it is important that the methodology includes a means of coping with cancellations. We outline here an extension to the basic model of buying behaviour that accounts for cancellations. Unfortunately, we have been unable to test this due to a lack of reliable data.

Suppose that the probability of a booking being cancelled on any day after it was made is constant and equal to $q$. The expected number of cancellations $x$ days before departure is then equal to $q$ multiplied by the sum over all the uncancelled bookings on day $x$ of the number of days since each booking was made. In terms of the one-way pricing model, the expected number of cancellations by day $x$ is therefore

$$X(t) = q \int_{x}^{x_{\text{start}}} (x' - x) f(x') p(x', y(x'))\,\mathrm{d}x' \tag{7}$$

where $f(x)$ and $p(x, y(x))$ are the functions used in the basic one-way pricing model described in the previous section. The cancellations model therefore uses all of the parameters of the one-way pricing model plus the additional parameter $q$. After a cancellation, the ticket becomes available for resale. The model of cancellations could therefore be used to decide on the level of overbooking that should be allowed.

## Fitting the model

Fitting the basic analytical model of buying behaviour involves first deciding what functions to use for $f(x)$ and $p(x, y(x))$ and then estimating the associated parameters from the available historical data. Fitting the cancellations model would involve estimating $q$ in addition to the one-way pricing parameters. As discussed above, we were unable to fit the cancellations model due to a lack of reliable data on cancellations. The data available to us were in the form of booking profiles. These include records of the price charged on a day $x$, $y(x)$ and the net number of seats booked by day $x$, $n(x, y)$. Data were available for a prescribed set of days during the selling period, which we call 'data collection days'. No record is kept of cancellations and these could only be detected if the net number of bookings decreased between two data collection days.

In recent years, the Internet has played an increasing role in the sale of airline tickets, especially among the low-cost carriers. The methodology we present is particularly well suited for modelling sales in this context, as we do not directly model demand as a function of price, but only as a function of time to departure. The potential demand function $f(x)$ could be well represented by the number of hits on a company's website for a particular flight or the number of phone calls made to a call centre inquiring about that particular flight. This could enable $f(x)$ to be fitted independently of $p(x, y(x))$.

For our fictional example, we assume that we have no data available on the potential demand, and we fit the parameters $a$, $b$, $d$, $g$ and $h$ by minimizing the sum of squared errors between the model estimates for the number of bookings between two data collection days and the actual number of bookings between those days. The model uses the same price structure as that used in the data. As cancellations are not properly accounted for in the basic model, we ignore data collection days between which there is a decrease in the net number of seats booked. This allows us to remove some of the cancellations from the data.

The main challenge in the fitting process is coping with the inherent messiness of the data. This can lead to uncertainty over the actual numbers of bookings received. Group bookings and cancellations probably constitute the major complicating factors. No strategy for coping with these factors has yet been incorporated.

We use the Nelder Mead optimisation routine[4] to find the minimum of the sum of the squared differences between the model and the actual data. The best-fit parameters corresponding to this minimum can then be input into Equation (6) to estimate the optimal price structure. The expected number of bookings under this (or any other) price structure can be calculated using Equation (1). A simulation model, described in the next section, is then used to find the confidence ranges on the expected number of bookings.

Parameter values are not updated during the life of the flight and we rely on a regular recalibration of the model to provide up-to-date estimates of the best-fit parameters. Most airlines will have some general categorization for flights and this could be used initially to pool flights for fitting. This increases the data available for fitting and so allows the fitting to be focused on more recent flights while still having enough data available to account for natural variability. Further work could consider how fits of the model to data might inform pool definitions.

## Confidence ranges

We estimate the confidence ranges around the numbers of bookings by day and fare class using a stochastic simulation model. This is based on the basic one-way pricing model described earlier and generates sample booking profiles for a given price structure and set of parameter values. Estimates of the expected numbers booked by day and fare class and the expected variation in these can be calculated using the output from the simulation model. We describe the variation using confidence ranges and these are expressed as ranges in the numbers of bookings that you would expect to observe in a set percentage of flights. The user can set the width of the ranges by changing the confidence level. For example, if the user requests a 50% confidence range, and the model is run 1000 times, the lower limit will be given by the 250th value in a non-decreasing list of booking numbers and the upper limit by the 750th value. The confidence ranges are intended to provide alerts of any anomalies in buying behaviour to analysts monitoring the flights. They therefore act as a safety net preventing either all the seats on a plane being booked at a very low price or few seats being sold on a plane due to the prices being set too high.

We simulate the demand using a non-homogeneous Poisson process with rate parameter given by $f(x)$ where $x$ is the midpoint of the day (eg 1.5 if we are looking at bookings made between days 2 and 1) and $f(x)$ is the potential demand in the one-way pricing model and is given in Equation (2).

Of those who are interested in flying, not all will actually buy a ticket and within the simulation model, each potential customer undergoes a test to determine whether they will make a purchase. A random number between 0 and 1 is generated and if this is below $p(x, y(x))$ (the probability that an interested customer buys a ticket) they buy, otherwise no sale is recorded.

The model is run 1000 times to obtain the booking limits with a total run time of about 20 s on a Pentium III, 933 MHz. If this process is to be run several times for many different flights, the time taken to run the program could become prohibitively long. It was hoped that booking limits could be obtained analytically but solutions have proved to be too complex for practical use.

### *Validation*

We validate the mean behaviour of the simulation model by comparing its results with analytical results for the expected number of tickets sold given in Equation (1). This is useful for determining the optimal time step to use and the minimal number of runs. As the basic one-way pricing model is fitted to the data, we assume that this has the correct mean behaviour.

A time step of 1 day is used in the model, which means that bookings are generated for each day of the selling period. The choice of time step must take into account contradictory requirements. On the one hand, if the time step is too small, the probability of there being any potential customers is very small and the model's prediction of the number of units sold is too low. On the other hand, if the time step is too long, the rate parameter used in the Poisson process and the probability of purchasing are not a sufficiently accurate reflection of customer behaviour over the whole of the time step.

Kincaid and Darling[5] suggest using a time scale that can be compressed and expanded locally to take account of the time-variant arrival rate in an inhomogeneous Poisson process. This method has been applied by Zhao and Zheng[6] who describe it as the conversion of an inhomogeneous Poisson process to a homogeneous Poisson process. We have not implemented a model with a time-dependent time step, but this is something that could be considered in the future.

Matching the variability of the simulation model with the variability of the data has proved to be very difficult in this application of the methodology as the expressions for the variability in the numbers of bookings vary with day and fare. If the price structures used for a particular flight vary significantly from day to day, it is hard to find sufficient datapoints for sales at particular prices on particular days to enable estimation of the variance in the number of sales. In order to obtain some measure of the variability of the booking numbers, albeit a fairly inaccurate one, we sampled from the data set to obtain the variability in the numbers booked irrespective of the fare being charged. The variability in these numbers was greater than that observed for the simulation model. This is to be expected, as the variations in price structures in the data set will have led to a higher degree of variability in the numbers booked.

## Updating the price structure

As a flight progresses through the selling period, information becomes available about the customer buying behaviour for that flight and the number of bookings that have been made. This information is used to update the optimal price structure and subsequently similar methods to those used at the start of the selling period can be used to update the expected number of bookings and the limits on these by day and fare class. The updating method that we propose here does not update the parameter values, relying on regular recalibration using all available relevant data, to update these for all open flights.

As bookings are made, the remaining capacity decreases, and the capacity constraint on day $x$ is

$$\int_0^x f(x')p(x', y(x'))\,\mathrm{d}x' \leqslant C - n_x \tag{8}$$

where $n_x$ is the number of seats that have already been sold $x$ days before departure. Unless bookings are following the model exactly or the original and updated capacity constraints are not active, the updated constraint will give a new Lagrangian multiplier $\lambda_x$, resulting in an updated optimal price of

$$y(x) = \frac{1}{a + bx} + \lambda_x \tag{9}$$

Therefore, the updating process has the effect of shifting the optimal prices up or down, with the optimal prices to charge on each of the remaining days of the selling period being shifted by the same amount. If more bookings than normal have been made on a busy flight, the optimal price is likely to increase and if less bookings than normal have been made it should decrease.

This method can have little or no effect on weak flights for which capacity is never reached and for which $\lambda_x$ remains at zero for the whole of the selling period. In this case, the confidence ranges become more important, as they will alert the airline that the flight is not performing as expected and further action is required.

If the prices are changed, the expected number of bookings will also change. We estimate the expected numbers of bookings by day and by fare class and the confidence ranges around those numbers in the same way during the selling period as is done at the start of the selling period.

## Results

Analysis has been carried out for a number of different flights at the airline we have worked with. We are unable to reproduce results of this work here and instead present results for a fictional example to demonstrate how the methodology works. Bookings for the example are generated using the simulation model described in the Confidence ranges section.

We began by fitting the bookings data from ten sample flights to the model. The actual parameter values used and the best-fit parameter values are given in Table 1. The optimal continuous price structure for the flight using the best-fit parameters was deduced using Equation (6). We then converted the continuous price structure to a discrete price structure with fare bands given in Table 2.

The price was updated on each of the data collection days (days 75, 50, 40, 30, 20, 15, 10 and each of the remaining days until departure) taking into account the number of bookings made. Updated estimates of the expected number of bookings and confidence ranges around these were then found. Bookings were generated using the simulation model with the updated price structure and the model parameters used to generate the sample data set. Figure 1 shows how the optimal price varied through the selling period. The optimal prices and the expected numbers of bookings with confidence ranges were found using the best-fit parameters.

The expected numbers of bookings made between the data collection days with confidence ranges on those bookings are given in Figure 2 alongside the actual numbers of bookings. The confidence level is set to 90%, which means that approximately 90% of the time the numbers booked on the flight will be inside the confidence range. Changing the

**Table 1** Original and best-fit parameter values

| | *Original parameters* | *Best-fit parameters* |
|---|---|---|
| *a* | 0.00667 | 0.00667 |
| *b* | 0.00216 | 0.00291 |
| *d* | 7.49 | 13.0 |
| *g* | 15.4 | 4.95 |
| *h* | 0.125 | 0.117 |

**Table 2** Discrete fare structure

| *Fare bands (£)* |
|---|
| 15 |
| 25 |
| 40 |
| 50 |
| 60 |
| 75 |
| 85 |
| 95 |
| 100 |
| 150 |

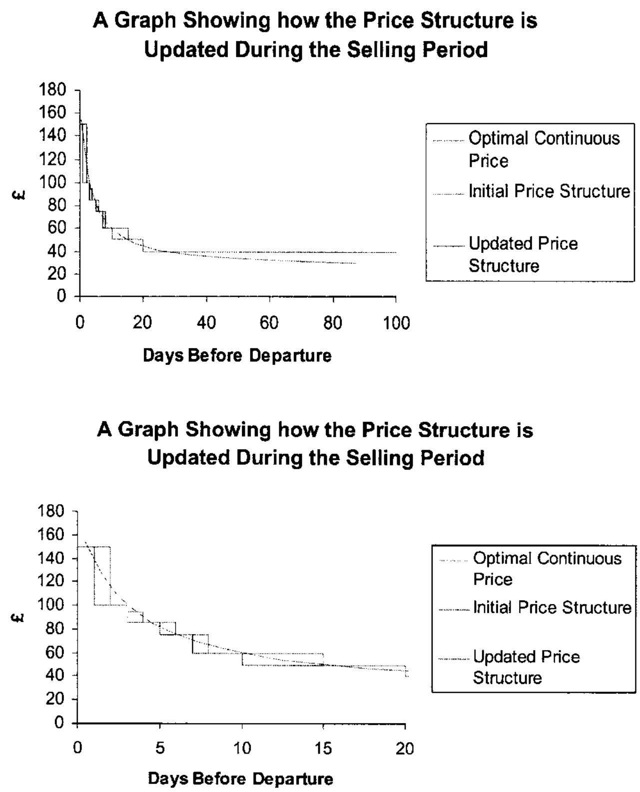


**Figure 1** Graphs showing the optimal continuous price structure, the initial discrete price structure and the updated discrete price structure for a fictional example for (a) the whole selling period and (b) the final 20 days of the selling period.

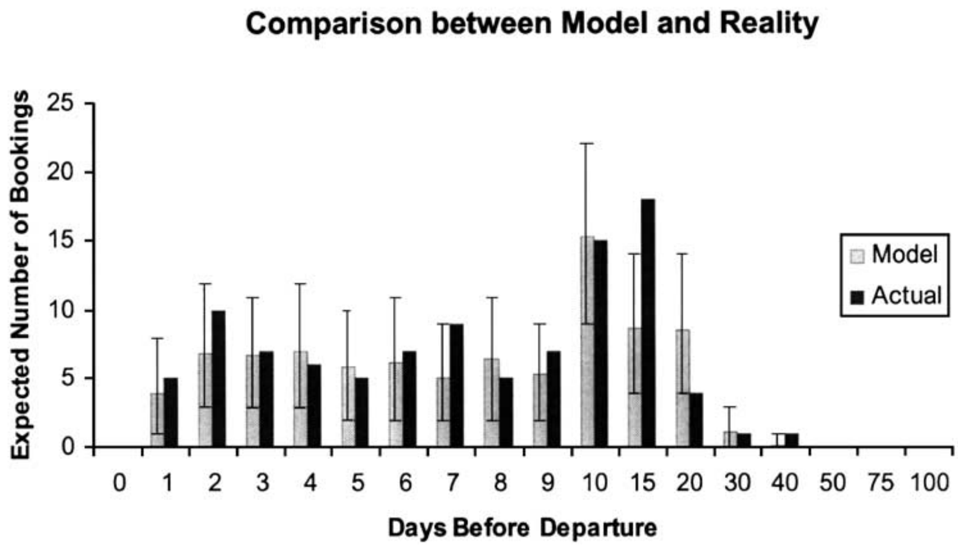


**Figure 2** A graph comparing the expected number of bookings calculated using the one-way pricing model and the best-fit parameters and the actual number of bookings generated using the original parameter set.

confidence level to 50% would result in the width of the confidence ranges approximately halving and these would be halved again if the confidence level were changed to about 30%.

In this example, the model gives reasonable estimates of the number of bookings. The model's estimates are generally lower than the actual number of bookings towards the end of the selling period; therefore, the updated price is slightly higher than the original optimal price at the end of the period.

## Conclusion

We have presented a methodology that has been developed for a major British airline to maximize revenue under one-way pricing of flight tickets. The methodology consists of a model of booking behaviour for flights under a one-way pricing system that can be used to find an expression for the optimal price structure. The optimal price structure can be updated throughout the selling period based on the number of bookings already made for the flight. The model can be used to calculate the expected number of bookings between different days in the selling period and within fare classes for a given price structure a given number of days before departure. A simulation model has also been developed to determine confidence ranges on the number of bookings made either between days in the selling period or within fare classes.

In the example, model parameters are found by fitting the model's estimates of the numbers booked between data collection days to data from booking profiles, which detail the number of bookings on each data collection day before departure and the price structure used on the flight. The best-fit parameters for the model are assumed to be those that minimize the sum of the squared difference between the model's estimates of the number of seats booked between data collection days and the actual number of bookings between those days. Satisfactory results have been obtained for the company concerned using this method although messiness in the data, caused mainly by cancellations and group bookings, makes fitting more difficult. A model of cancellations has been proposed but has not been used due to a lack of data. No method of coping with group bookings has been put forward.

If data on hit rates on airlines' websites for particular flights or other measures of levels of interest in flights are available, it may be possible to improve the fitting process by fitting the potential demand function separately from the function describing the probability of purchasing a ticket.

The fitting process currently makes no estimate of seasonality, relying on traditional methods to determine which flights can be pooled together and from which seasons of the year. In the future, it would be interesting to explore how the model might help define the flight pools.

Expected booking numbers for a given price structure are easy to calculate and acceptable ranges on these numbers can be estimated using a stochastic simulation model. The simulation model outputs confidence ranges for the number of bookings by day and fare class. In principle, analytical expressions could be obtained for the confidence ranges to avoid the need for the simulation model and so save processing time, but investigations conducted so far suggest that these would be too complex for practical use.

The variability in the booking profiles output by the simulation model has not been checked thoroughly, therefore we cannot be sure that the model is an accurate representation of the actual output. This is an area in which further investigation is definitely required and should it be found that the variability of the simulation model output does not match the variability of the actual data, the simulation model will have to be adapted to take account of this. For the moment, the user can adjust the tightness of the upper and lower booking limits by changing the confidence level to find the width that best suits the data they are observing.

Price structures can be updated based on the number of bookings that have been received by a given day. The expected numbers of bookings and limits on these numbers can then be updated given the new price structure. This updating routine does not include an update of the parameters used within the model and it is assumed that parameter values will be updated by means of a regular recalibration in which all available data for a pool of flights is used to refit the parameter values. Data from flights that have not yet departed can be used in this recalibration exercise allowing the parameter values to reflect current trends in booking behaviour. Updating parameters during the life of the flight using exponential smoothing, Kalman filters or some similar technique was considered, but it was decided that the huge variability in the number of bookings on a flight would result in very unstable parameter values. Future work could investigate more elaborate strategies to update parameters during the life of the flight.

Finally, probably one of the main influences on a customer's buying behaviour of economy class airline tickets is the level of competitors' prices. Incorporating information about competitors' pricing strategies would substantially increase the complexity of the methodology but is likely to improve the modelling of buying behaviour.

## References

1 Bitran GR and Mondschein SV (1997). Periodic pricing of seasonal products in retailing. *Mngt Sci* **43**: 64–79.
2 Gallego G and van Ryzin G (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Mngt Sci* **40**: 999–1020.
3 Anjos MF, Cheng RCH and Currie CSM (2003). Optimal pricing policies for perishable products. Research Report OR106, Faculty of Mathematical Studies, University of Southampton, UK.
4 Nelder JA and Mead R (1965). A simplex method for function minimization. *Comput J* **7**: 308–313.
5 Kincaid WM and Darling DA (1963). An inventory pricing problem. *J Math Anal Appl* **7**: 183–208.
6 Zhao W and Zheng YS (2000). Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Mngt Sci* **46**: 375–388.

*Received June 2003;*
*accepted December 2003 after one revision*