Models of the Spiral-Down Effect in Revenue Management

Anton J. Kleywegt* William L. Cooper† Tito Homem-de-Mello‡

*School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, USA
anton@isye.gatech.edu

†Department of Mechanical Engineering
University of Minnesota
Minneapolis, MN 55455, USA
billcoop@me.umn.edu

‡Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208, USA
tito@northwestern.edu

1 Introduction

Revenue management involves the application of quantitative techniques to improve profits by controlling the prices and availabilities of various products that are produced with scarce resources. The best known revenue management application occurs in the airline industry, where the products are tickets (for itineraries) and the resources are seats on flights. In almost every instance of published work, the starting point of the analysis is some set of assumptions regarding an underlying stochastic or deterministic demand process. With these assumptions in hand (and assumed to be correct), most papers proceed to "solve" the model and derive "optimal" operating policies. In the airline context, such a policy usually details which types of tickets are available at which times, and under which circumstances.

However, the situation faced by revenue managers in practice is different from published work in at least two key regards: assumptions may be incorrect, and model parameters are not known. It is widely understood that the forms of the demand models used in practice are not likely to be correct—one simply cannot construct practical models that take into account all the trade-offs considered by customers. Typically the parameters of the chosen models are estimated with available data, but even if the data are good and a good forecasting method is used, it is likely that parameters are being estimated for an incorrect model. In practice, there is an iterative process whereby a control (e.g., protection levels) is enacted, sales occur, a flight departs, and parameter estimates are updated based upon observed data. The updated estimates are then used to choose a new control for the next instance of flight, and so on. An important question is what can happen if the airline uses a good forecasting method, but the chosen controls are based on an incorrect model, specifically, a model with erroneous assumptions regarding customer behavior.

We study a stylized model of the situation described above, and show what can occur in certain cases when an airline makes incorrect modeling assumptions. In particular, if the airline uses the standard Littlewood/expected marginal seat revenue (Littlewood-EMSR) technique to set protection levels, and if it estimates "the probability distribution of high-fare demand" using past high-fare sales (truncated or untruncated) without correctly taking into account how customer behavior depends on the controls (the protection levels), then the protection levels, high-fare sales, and revenues may "spiral down" over time.

A quick description of the spiral down effect is as follows. Suppose that there are two classes of tickets and that customers are flexible, that is, they are willing to buy either low-fare or high-fare tickets, but they will buy the low-fare tickets if both are available. Suppose also that the airline decides how many seats to reserve for high-fare tickets based on past sales of high-fare tickets, while neglecting to account for the fact that availability of low-fare tickets will reduce sales for high-fare tickets. Then, if more low-fare tickets are made available, low-fare sales will increase and high-fare sales will decrease, resulting in lower future estimates of high-fare demand, and subsequently lower protection levels for high-fare tickets and greater availability of low-fare tickets. The pattern continues, resulting in a spiral down of high-fare sales, protection levels, and revenues. In this paper, we describe this effect by establishing limit theorems that give conditions under which the protection levels converge, in many cases downward. Boyd et al. (2001) have used simulation to demonstrate this spiral-down effect, which is known to some practitioners.

The fundamental issue is that the optimization model used by the decision maker is incorrect. In addition, not only does the incorrect optimization model produce suboptimal decisions (which is not surprising), but the decisions produced can become worse as the revenue manager attempts to refine the incorrect model with observed data. The combination of an incorrect optimization model with parameters that are adjusted based on data, and data that are affected by the controls produced by the incorrect optimization model, can cause the optimization model and the resulting controls to deteriorate as the forecasting-optimization process evolves.

2 A Single Problem Instance

Consider a single flight with capacity c. Suppose that there are class-1 and class-2 tickets for sale. The price of class-i tickets is denoted with f_i. Suppose that $f_1 > f_2 > 0$. If the airline sells s_1 tickets in class 1 and s_2 tickets in class 2, then it will receive revenue $r = f_1 s_1 + f_2 s_2$.

Next we address the demand for the tickets. The notion of customer demand for different products is central to revenue management. At the same time, the research in revenue management has not yet produced a widely accepted model of demand. We do not propose a particular model of demand. Rather, the purpose of our work is to illustrate some effects that can occur if the demand model used by the revenue manager does not accurately describe the customer behavior.

One type of demand model specifies, for each given set of alternatives, the probability that a customer chooses each alternative in the given set. Most revenue management models published in the literature do not use such models of individual customer behavior, but rather more aggregate models of demand. (For an example of a model that *does* incorporate individual customer behavior, and for additional references, see Talluri and van Ryzin 2001.) Models of individual customer behavior, such as the one referred to above, can be used to obtain aggregate models of demand. For example, given a stochastic process describing how customers arrive over time to request products, and a model of individual customer behavior, and a rule or policy describing how the set of available alternatives changes as the process unfolds, one can derive the probability distribution of the quantity of each product demanded during the life of the process. We emphasize that under such a scheme, the probability distribution of the quantity of each product demanded depends crucially on the policy that determines the set of available alternatives. Many aggregate demand models that have been used in the literature do not incorporate this dependence.

In our study, the model of actual demand is an aggregate model that depends on the policy that determines the set of available alternatives, and the model used by the revenue manager does not depend on the policy. The revenue manager uses a good forecasting method, but attempts to estimate a demand model that turns out to be an incorrect model.

Of course, the policy used by the revenue manager should be compatible with the demand model used by the revenue manager. Next we describe the demand model and the policy used by the revenue manager in our study, as well as the observed data. Recall the case with a single flight and two fare classes described above. We assume that the revenue manager uses the well-known Littlewood-EMSR rule (see, e.g., Littlewood 1972, Belobaba 1989, Wollmer 1992, Brumelle and McGill 1993, or van Ryzin and McGill 2000) to control the availability of the two alternatives, namely class-1 and class-2 tickets. That is, the policy used by the revenue manager specifies a protection level ℓ that is chosen by the Littlewood rule as follows. Suppose that the revenue manager's demand model has a cumulative probability distribution H for the demand for class-1 tickets. Then ℓ is chosen to satisfy

$$\ell \in H^{-1}(\gamma) \tag{1}$$

where $\gamma := 1 - f_2/f_1$ and $H^{-1}(\gamma)$ denotes the set of γ-*quantiles* of H.

Here, we must emphasize that in the actual demand model, "demand for class-1 tickets" is not a well-defined quantity because, in general, customers decide what to purchase based on their own preferences as well as the available set of alternatives. Nevertheless, the revenue manager bases his decisions on a supposed (and estimated) probability distribution H for the demand for class-1 tickets and on the well-known Littlewood-EMSR rule. Once the revenue manager has decided to use the Littlewood-EMSR method, he needs to estimate H based upon some data. In practice, this data would typically include historical values of class-1 tickets sales, possibly after some so-called unconstraining to remove effects caused by censored and/or missing data (see Section 4.2 of Boyd and Bilegan 2003). The data used by the revenue manager to estimate H consist of values of an "observed quantity" X, that the revenue manager may call the aggregate demand for class-1 tickets. For the reasons explained above, we assume that X depends on the chosen parameter ℓ. Let $G(\ell, \cdot)$ denote the cumulative distribution function of the observed quantity X if the booking control process uses protection level ℓ.

We illustrate these ideas with some examples.

Example 1. Suppose there are three types of customers, namely type a, type b, and type ab customers. Type a customers buy class-1 tickets only, type b customers buy class-2 tickets only, and type ab customers buy either class-1 or class-2 tickets, but prefer the cheaper class-2 tickets if they are available. Unlike some models in the revenue management literature, we impose no conditions on the order of arrival, so that type a, type b, and type ab customer arrivals may be interspersed. Let D_a, D_b, and D_{ab} denote the number of type a, type b, and type ab customers respectively. Let $D_a(\ell)$, $D_b(\ell)$, and $D_{ab}(\ell)$ denote the number of type a, type b, and type ab customers respectively who arrive until $\ell \leq c$ tickets remain unsold, and class-2 tickets become unavailable. Suppose that the observed quantity X is equal to the number of class-1 tickets sold (which is "demand for class-1 tickets" truncated by the limited capacity). Then, $X = D_a(\ell) + \min\{\ell, D_a - D_a(\ell) + D_{ab} - D_{ab}(\ell)\}$, and $G(\ell, x) = \text{Prob}\,[D_a(\ell) + \min\{\ell, D_a - D_a(\ell) + D_{ab} - D_{ab}(\ell)\} \leq x]$, which depends on ℓ.

Example 2. In this example the demand is the same as in Example 1. However, the observed quantity X is equal to the number of type a customers who arrive during the time horizon plus the number of type ab customers who arrive during the time horizon when class-2 tickets are not available anymore, that is, the number of type ab customers who either purchase class-1 tickets or who arrive when no more tickets are available. Thus, in this example the revenue manager gets the benefit of the doubt and continues to observe customers even after c tickets have been sold. Here, $X = D_a + D_{ab} - D_{ab}(\ell)$, and $G(\ell, x) = \text{Prob}\,[D_a + D_{ab} - D_{ab}(\ell) \leq x]$, which also depends on ℓ.

Example 3. This example is the model that usually is associated with the Littlewood rule. There are only two types of customers, namely type a and type b customers, that is, $D_{ab} = 0$. In addition, all type b customers arrive before any type a customers arrive. The observed quantity X is the number D_a of type a customers. Thus, in this example the revenue manager observes all the type a customers, whether the number of type a customers exceeds the capacity c or not. Hence $X = D_a$, and $G(\ell, x) := \text{Prob}\,[X \leq x] = \text{Prob}\,[D_a \leq x]$, which is independent of ℓ. In the situation described here, using (1) to choose the protection level is the right thing to do.

Example 4. In this example the demand is the same as in Example 3. However, in this example the observed quantity X is the actual sales of class-1 tickets. Thus, in this example, the actual sales of class-1 tickets is equal to $X = \min\{D_a, c - \min\{D_b, (c-\ell)^+\}\}$ Hence, $G(\ell, x) = \text{Prob}\left[\min\{D_a, c - \min\{D_b, (c-\ell)^+\}\} \leq x\right]$, which again depends on ℓ.

3 A Sequence of Problem Instances

In order to describe the spiral-down phenomenon, we consider a sequence of problem instances; that is, a sequence of booking processes indexed $k = 1, 2, 3, \ldots$ of a particular flight, say an 8am Monday flight from New York to Los Angeles. Initially, suppose that the revenue manager selects a protection level L^0 for flight 1, and subsequently observes quantity X^1. The distribution of the observed quantity X^1 for flight instance 1 is $G(L^0, \cdot)$. Based upon what is observed, the revenue manager selects a new booking limit L^1 for flight 2. The distribution of the observed quantity X^2 for flight 2 is $G(L^1, \cdot)$. The revenue manager continues in this fashion, getting sequences $\{L^k\}$ and $\{X^k\}$.

As mentioned in the previous section, a typical method for selecting $\{L^k\}$ is for the airline to determine an estimate of the distribution of "the demand for class-1 tickets", and then to use this estimate as if it were correct (i.e., not an estimate) as input into an optimization scheme, which results in the protection level L^k being chosen according to Littlewood's rule (1). That is, to obtain L^k, the revenue manager obtains an estimate $\hat{H}^k$ based upon observed data and then chooses $L^k \in (\hat{H}^k)^{-1}(\gamma)$. To provide a formal description of the iterative forecasting and booking control procedure, we introduce some more notation. Let $\mathcal{D}$ denote the space of distribution functions on $\mathbb{R}$. For problem instance k, we consider a generic update function $\phi^k : \mathcal{D} \times \mathbb{R}^k \mapsto \mathcal{D}$ that maps the initial forecast $\hat{H}^0 \in \mathcal{D}$ of class-1 demand, and the sequence $\{X^i\}_{i=1}^k \in \mathbb{R}^k$ of observed quantities, to a new forecast $\hat{H}^k$.

The initial estimate $\hat{H}^0$ of the class-1 demand distribution is specified and a protection level $L^0 \in (\hat{H}^0)^{-1}(\gamma)$ is chosen. For each $k \in \mathbb{Z}_+$, we assume that

$$\mathbb{P}[X^k \leq x | \mathcal{F}^{k-1}] \quad = \quad G(L^{k-1}, x) \quad \text{for all } x \in \mathbb{R}. \tag{2}$$

Forecasts and protection levels are updated according to

$$\begin{aligned} \hat{H}^k &:= \phi^k(\hat{H}^0, X^1, \ldots, X^k) & (3) \\ L^k &\in (\hat{H}^k)^{-1}(\gamma). & (4) \end{aligned}$$

4 Types of Results Obtained

Due to space constraints, we cannot thoroughly cover our results here. We describe the spiral-down effect by establishing limit theorems that give conditions under which the protection levels converge, in many cases downward.

Example 5. The time horizon is 100, and the capacity is $c = 100$. Class-1 tickets have price 1, and class-2 tickets have price 0.7. The revenue manager's demand model is the same as in Examples 3 and 4. More specifically, suppose that in the revenue manager's model, first type b customers arrive according to a Poisson process with rate 1.5 over the interval $(0, 50]$, and next type a customers arrive according to a Poisson process with rate 1.5 over the interval $(50, 100]$. Suppose that the revenue manager's model is not quite correct, and in fact type b customers arrive according to a Poisson process with rate 1.5 over the interval $(0, 100/3]$, type ab customers arrive according to a Poisson process with rate 1.5 over the interval $(100/3, 200/3]$, and then type a customers arrive according to a Poisson process with rate 1.5 over the interval $(200/3, 100]$.

Figure 1 shows the objective function for this example, namely the expected revenue as a function of the protection level. The optimal protection level is 95. The optimal protection level in the revenue manager's model is $G_1^{-1}(1 - f_2/f_1) = 70$, where G_1 denotes the Poisson distribution with mean 75. Thus, as expected, the revenue manager's model would give a suboptimal solution. However, next we show that the situation can be much worse if the revenue manager uses observed data, that depend on the chosen protection levels (where the dependence is not captured correctly in the revenue manager's model), to update the protection levels.

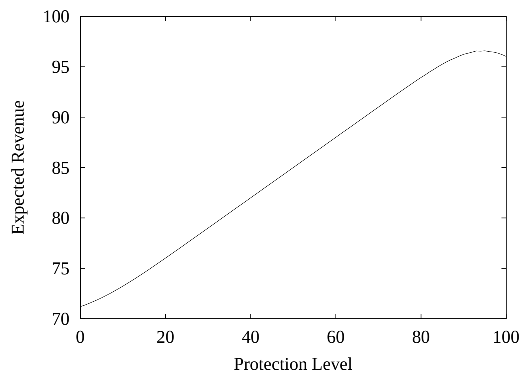

Figure 1: Example 5: Expected revenue as a function of protection level.

Suppose that the observed quantity X is equal to the number of class-1 tickets sold as in Example 4. Figure 2 shows how the protection levels spiral down along 10 sample paths from an initial level of $L^0 = 100$ if the revenue manager uses a stochastic approximation algorithm to choose L^k. Thus, due to the error in the revenue manager's model, the protection levels converge to their worst possible value instead of to the optimal value or some reasonable suboptimal value. Such disastrous spiral down behavior would be of great concern to any revenue manager. Figure 3 shows how the expected revenue obtained by the revenue manager spirals down.

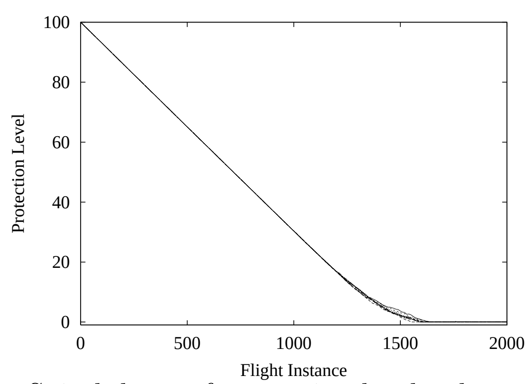

Figure 2: Example 5: Spiral down of protection levels, shown for 10 sample paths.

References

P. P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37:183–197, 1989.

E. A. Boyd and I. C. Bilegan. Revenue management and E-commerce. *Management Science*, 49:1363–1386, 2003.

E. A. Boyd, E. Kambour, and J. Tama. The impact of buy down on sell up, unconstraining, and spiral down, 2001. Presented at *1st Annual INFORMS Revenue Management Section*

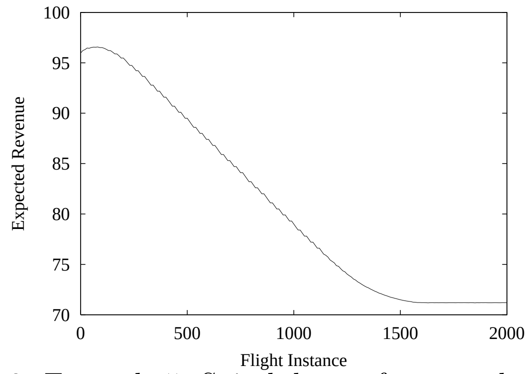

Figure 3: Example 5: Spiral down of expected revenue.

Conference, New York. Slides available at `http://209.238.188.14/html_files/roi/roi_agenda.htm`.

S. L. Brumelle and J. I. McGill. Airline seat allocation with multiple nested fare classes. *Operations Research*, 41:127–137, 1993.

K. Littlewood. Forecasting and control of passenger bookings. In *AGIFORS Symposium Proceedings*, volume 12, pages 95–117, 1972.

K. Talluri and G. van Ryzin. Revenue management under a general discrete choice model of consumer behavior. Working paper, Universitat Pompeu Fabra, 2001.

G. van Ryzin and J. McGill. Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels. *Management Science*, 46: 760–775, 2000.

R. D. Wollmer. An airline seat management model for a single leg route when lower fare classes book first. *Operations Research*, 40:26–37, 1992.