Available online at www.sciencedirect.com
SCIENCE DIRECT®

Transportation Research Part E 42 (2006) 1–19

TRANSPORTATION RESEARCH PART E

www.elsevier.com/locate/tre

The net benefit of airline overbooking

Yoshinori Suzuki *

Department of Logistics, Operations, and Management Information Systems, College of Business, Iowa State University, 2340 Gerdin Business Building, Ames, IA 50011-1350, USA

Received 11 May 2004; received in revised form 9 September 2004; accepted 14 September 2004

Abstract

Benefits of airline overbooking are often measured by considering only the revenue gains attained for overbooked (congested) flights, and ignoring the potential loss of revenues that may take place in other (un-congested) flights (i.e., the "gross" benefit). This study explores the "net" benefit of overbooking by considering revenue implications of overbooking on both congested and un-congested flights. A series of simulation experiments are conducted to investigate the nature of the relationship between the gross and net benefits. The results imply that the net benefit may be substantially lower than the gross benefit under many practical conditions.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Airlines; Revenue management; Overbooking; Simulation

1. Introduction

Today, most airlines practice overbooking to enhance their revenues. It is well known that about 10–15% of travelers with confirmed reservations do not show up for their flights without giving prior notice to airlines ("no-shows"). If no overbooking is used, therefore, airline flights must frequently depart with some vacant seats, even when the seats are completely sold out. Smith

* Tel.: +1 515 294 5577; fax: +1 515 294 2534.
E-mail address: ysuzuki@iastate.edu

1366-5545/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tre.2004.09.001

et al. (1992), for example, report that for closed flights (sold-out flights) airlines may still experience roughly 15% of seats empty during the actual flights, if airlines do not practice overbooking. Thus by selling their seats beyond the physical capacity, airlines can reduce the number of vacant seats during the actual flights, and improve their unit (flight) revenues.

The benefit of airline overbooking is well documented in the literature. [1] Alstrup et al. (1989) claim that, for a typical major airline, no-shows can cause a direct loss of about \$50 million per year, and that overbooking can reduce such losses substantially. Curry (1990) mentions that overbooking can generate an additional 3–10% of gross passenger revenues for airlines. Davis (1994) reports that the revenue management tools such as overbooking saved American Airlines an estimated \$1.4 billion over a three-year period, and that the airline expects to generate, by using these tools, at least \$500 million additional revenues annually in the future.

These figures, however, may not reflect the true benefit of overbooking to airlines. Typically, the benefit of overbooking is estimated by taking the difference of the loss of revenues caused by no-shows (and late cancellations) in closed flights when overbooking is *not* used, and that when overbooking *is* used (e.g., Alstrup et al., 1989; Nambisan, 2003). Thus, the benefit is typically estimated by simply calculating the amount of revenue gains attained for overbooked (congested) flights, and ignoring the revenue impact of overbooking on other (un-congested) flights. This condition implies that, if overbooking affects revenues of not only congested flights but also un-congested flights, the benefit calculated by the standard approach may be misleading.

In theory, overbooking can affect revenues of both congested and un-congested flights. When overbooking is used, reservations for congested flights will increase, as airlines will accept reservations by "additional" passengers who demand seats after the flight capacity is reached. Theoretically, these "additional" passengers can be classified into two types. The first is "new customers". Without overbooking, some travelers demanding seats on an airline's congested flights may have to give up flying or fly by other airlines. The use of overbooking allows an airline to accept reservations by these travelers whose demand would otherwise be lost to other airlines or to other modes of transportation. The second is "flight switchers". In the absence of overbooking, travelers who cannot use their preferred flights (congested flights) may reserve seats in other flights *of the same airline* (e.g., flights with inconvenient arrival times). [2] If overbooking is used, these "compromising" travelers may "come back" to their preferred flights.

Note that if all the "additional" passengers are new customers, overbooking has no revenue impact on un-congested flights. If, however, the "additional" passengers include flight switchers, overbooking affects revenues of not only congested flights but also un-congested flights, because in this case the revenue gains for the former flights are realized, at least partially, at the expense of reduced revenues for the latter flights. Since the existence of flight switchers is likely, these conditions imply that the true benefit of overbooking, which is obtained by taking the *net of* revenue gains realized for congested flights *and* revenue losses incurred for un-congested flights ("net" benefit), should always be *lower* than that calculated by the standard approach ("gross" benefit), which considers only the revenue gains attained for congested flights.

[1] For a detailed review of the airline overbooking research, see, for example, McGill and Ryzin (1999).

[2] This type of flight-switching behavior of air travelers is conceptually equivalent to what Davis (1994) referred to as the "passenger recapture".

From the practical standpoint, understanding the true, or "net", benefit of overbooking is important for airlines. Without knowing the net benefit, airlines may *overestimate* the revenue implications of overbooking and *underestimate* the revenue impacts of other factors, such as airline promotional activities. Of particular interest to airlines may be to understand the extent to which the net benefit is lower than the gross benefit under a variety of conditions, because such information may help airlines estimate the net benefits from the available gross benefit figures.

In this study, we perform a series of simulation experiments to investigates the nature of the relationship between the gross and net benefits of airline overbooking. The study objective is twofold. First, we examine whether the use of gross benefits by airlines *in lieu of* net benefits (i.e., the current approach to measure overbooking benefits) can be justified or not. If we find that the net benefit is similar to the gross benefit under all conditions, the current approach may be justified. If we find the opposite, however, the current approach may not be justified. Second, based on the simulation results, we derive a set of formulas that explain the functional relationship between the gross and net benefits. These formulas may be used conveniently by airlines to transform the (available) gross-benefit figures into the net overbooking benefit figures.

2. Measuring overbooking benefit

2.1. Individual-traveler revenue

Consider a congested flight j where the demand for seats is greater than the capacity, so that the number of reservations for the flight at the time of departure is either: (1) greater than the capacity (if j is overbooked), or (2) equal to the capacity (if j is *not* overbooked). Let R_i be the revenue that an airline receives from an individual traveler i who has a reservation on flight j at the time of departure. If this airline does not practice overbooking, R_i can be written as follows: [3]

$$R_i = SH_iP_i + (1 - SH_i)(RF_iNR_iP_i + (1 - RF_i)P_i), \tag{1}$$

where SH_i is a binary variable coded 1 if traveler i shows up for flight j (coded 0 otherwise), P_i is the airfare paid by traveler i, RF_i is a binary variable coded 1 if traveler i has a refundable ticket (coded 0 otherwise), and NR_i is a binary variable coded 1 if (given that i has a refundable ticket and does not show up for flight j) traveler i chooses *not* to refund the ticket (coded 0 otherwise).

Eq. (1) shows that the revenue from traveler i is realized if: (1) i shows up, (2) i does not show up but has a non-refundable ticket, or (3) i does not show up, has a refundable ticket, and does not refund the ticket (e.g., miss flight j and fly on a later flight). If, however, traveler i does not show up, has a refundable ticket, and refunds the ticket, the revenue from i becomes zero. [4]

[3] Since we consider only one flight (j), we do not use, for simplicity, the subscript representing flights in the equations that follow.

[4] Some tickets require service fees to receive refunds. In this case the ticket purchase that results in a refund may generate a small amount of revenue (service fee) to an airline. If, however, the charge is truly the "service charge" that reflects the cost of processing a ticket refund, the ticket purchase that eventually results in a refund may be considered as generating *no revenue* to an airline.

Now consider a situation where overbooking is used by the airline. The congested flight *j* is now overbooked. In this case, the revenue from traveler *i* must be modified as follows:

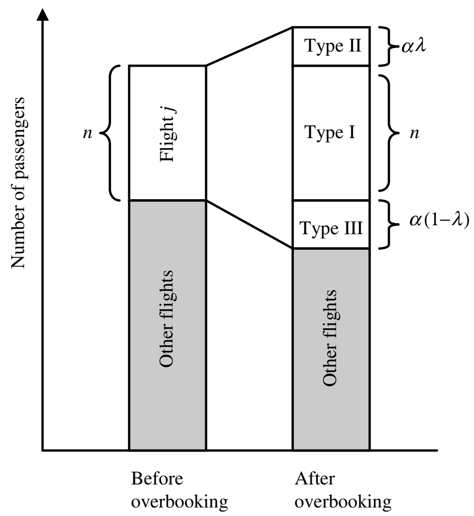
$$R(OB)_i = SH_i(P_i - BP_iCOMP_i) + (1 - SH_i)(RF_iNR_iP_i + (1 - RF_i)P_i), \tag{2}$$

where BP_i is a binary variable coded 1 if (given that *i* shows up for flight *j*) traveler *i* is denied a seat (i.e., bumped) on flight *j* (coded 0 otherwise), and $COMP_i$ is the amount of denied-boarding compensation which the airline pays to traveler *i*, if *i* is bumped from flight *j*.

Observe in Eq. (2) that if the flight is overbooked, the revenue from a traveler must be adjusted by the amount of compensation for possible bumping of this traveler. This condition indicates that the expected revenue of a traveler on flight *j* is affected by whether *j* is overbooked or not. Also observe that Eq. (2) implicitly assumes that if a traveler is bumped on flight *j*, he/she will not be bumped *again* on the substitute flight. This assumption may be considered reasonable, as airlines generally do not assign bumped travelers to other overbooked flights.

2.2. Revenue gains by type of traveler

Without the loss of generality, we assume that there are three types of travelers who have reservations on flight *j* at the time of departure when *j* is overbooked (see Fig. 1). The first is "early birds", who represent the first *n* reservations on flight *j*, where *n* is the seat capacity of *j*. These travelers either: (1) reserve their seats before the capacity is reached, or (2) reserve seats after the capacity is reached, but still quality as the first *n* reservations on flight *j*, because of cancellations of earlier reservations. We refer to this type of travelers as Type I customers. The second and third types represent the "new customers" and "flight switchers" that were discussed previously. The "new customers" are those who: (1) make reservations after the seat capacity is reached, (2) do not quality as the first *n* reservations, and (3) would fly on *other airlines*, or simply *give up flying*, if unable to fly on flight *j*. The "flight switchers", on the other hand, are those who: (1) make



Type I (*Early Birds*): Travelers representing the first *n* reservations on flight *j*, where *n* is the seat capacity of *j*. They either: (1) reserve their seats before the capacity is reached, or (2) reserve seats after the capacity is reached, but still quality as the first *n* reservations on flight *j* because of cancellations of earlier reservations.

Type II (*New Customers*): Travelers who: (1) make reservations for flight *j* after the seat capacity is reached, (2) do not quality as the first *n* reservations, and (3) would fly on other airlines, or simply give up flying, if unable to obtain seats on flight *j*.

Type III (*Flight Switchers*): Travelers who: (1) make reservations for flight *j* after the seat capacity is reached, (2) do not quality as the first *n* reservations, and (3) would fly on other flights *of the same airline* if unable to obtain seats on flight *j*.

Fig. 1. Travelers reserving seats in flight *j* (at the time of departure).

reservations after the seat capacity is reached, (2) do not quality as the first n reservations, and (3) would fly on other flights *of the same airline*, if unable to fly on flight j. We refer to these two types of travelers as Type II and Type III customers, respectively. Notice that Type II and Type III customers cannot fly on flight j, if j is *not* overbooked.

Given these definitions, we now consider the benefit of overbooking by the type of customer (type I, II, or III). Our approach is to calculate, for each traveler i on a overbooked flight j, the difference of: (1) the revenue that is actually realized from i on j, and (2) the revenue that would have been realized from i if j is not overbooked (e.g., zero for Type II customers, and airfares of substitute flights for Type III customers). We then add the resulting figures across all i, by customer type, to derive the total benefit attained for each customer type. By using this approach, we consider not only the revenue gains attained for flight j by overbooking the flight, but also the possible loss of revenues that can take place in other flights (i.e., net benefit).

We first consider Type I customers. Here, we make two assumptions. First, for each Type I customer i, the airfare of flight j stays the same regardless of whether j is overbooked or not (i.e., P_i of Eq. (1) and that of Eq. (2) are identical for all i). We need this assumption to calculate the "pure" effect of overbooking on airline revenues. Second, exactly the same travelers constitute Type I customers on flight j regardless of whether j is overbooked or not. We make this assumption to simplify our formulas. [5] Under these assumptions, the benefit derived from each Type I customer i on flight j can be obtained by subtracting Eq. (1) from Eq. (2). (This difference reflects the amount by which the revenue of each Type I traveler i increases by overbooking flight j.) Thus the total benefit derived from all Type I customers, $B(\mathrm{I})$, is given by

$$B(\mathrm{I}) = -\sum_{i=1}^{n}\{SH_i BP_i COMP_i\}. \tag{3}$$

Notice that Eq. (3) is always non-positive. This condition indicates that overbooking *reduces* the expected revenues of those travelers who make early reservations on congested flights, because when overbooking is used the revenue must be adjusted by the cost for possible bumping.

Next, we consider Type II customers. For each Type II customer, the overbooking benefit is simply given by Eq. (2), because the expected revenue of a Type II customer is zero when overbooking is not used, whereas it is the value of Eq. (2) when overbooking is used. Let α be the number of overbooked passengers on flight j (sum of Type II and Type III customers), and λ be the share of Type II customers within α ("new-customer proportion"). The total overbooking benefit derived from all Type II customers, $B(\mathrm{II})$, can then be written as follows:

$$B(\mathrm{II}) = \sum_{i=n+1}^{n+\alpha\lambda}\{SH_i(P_i - BP_i COMP_i) + (1 - SH_i)(RF_i NR_i P_i + (1 - RF_i)P_i)\}. \tag{4}$$

[5] Without this assumption, we must consider how the characteristics of travelers on flight j would change before and after the use of overbooking strategy (these characteristics can affect the amount of overbooking benefit). Since such changes are unpredictable, we use this assumption to "hold" the traveler characteristics constant before and after the use of overbooking strategy. Note that by using this assumption, SH_i, RF_i, and NR_i of Eqs. (1) and (2) become identical, so that taking the difference of the two Eqs (i.e., derivation of Eq. (3)) becomes simple.

In theory, Eq. (4) can be either positive or negative. In practice, however, it is unlikely to be negative. The necessary conditions for Eq. (4) to be negative are: (1) $P_i < COMP_i$ (for most i), and (2) the majority of Type II customers showing up for flight j must be bumped. The second condition does not hold in practice (unrealistic), as it generally requires that flight j be overbooked by more than 100% of its seat capacity. Practically, therefore, Eq. (4) should always be positive. This condition implies that, basically, any new customer brought to fight j (by overbooking the flight) will make a positive contribution to the unit (flight) revenue.

We now consider Type III customers. Since they are "flight switchers", the benefit from each Type III customer i can be obtained by taking the difference of Eq. (2) (expected revenue of i if j is overbooked) and the airfare that i would have paid for a seat on a substitute flight (of j) if i could not have obtained a seat on flight j (expected revenue of i if j is *not* overbooked). Thus, the overall benefit derived from all Type III customers, $B(\mathrm{III})$, is given by the following formula:

$$\begin{aligned} B(\mathrm{III}) = \sum_{i=n+\alpha\lambda+1}^{n+\alpha} & \{[SH_i(P_i - BP_iCOMP_i) + (1 - SH_i)(RF_iNR_iP_i + (1 - RF_i)P_i)] \\ & - [SH_iP(-OB)_i + (1 - SH_i)(RF_iNR_iP(-OB)_i + (1 - RF_i)P(-OB)_i)]\}, \end{aligned} \tag{5}$$

where $P(-OB)_i$ is the airfare that traveler i would have paid for a seat on a substitute flight. Note that the expected revenue of i on a substitute flight is given by Eq. (1) (not Eq. (2) because by definition the substitute flight is a non-overbooked flight), except that P_i is replaced by $P(-OB)_i$.

Eq. (5) indicates that the necessary condition for $B(\mathrm{III})$ to be positive is that $P_i > P(-OB)_i$ (for most i), which implies that, for an airline to attain positive benefits from Type III customers, the airline must charge higher fares for overbooked flights than for other flights. This condition *alone*, however, does not guarantee positive values of $B(\mathrm{III})$. Notice that even when $P_i > P(-OB)_i$ for all i, $B(\mathrm{III})$ can still be negative if one or more Type III customers are bumped on flight j. The value of Eq. (5), therefore, is highly unpredictable, even under practical conditions.

2.3. Gross and net benefits

Given $B(\mathrm{I})$, $B(\mathrm{II})$, and $B(\mathrm{III})$, we can now calculate the overall (global) benefit that an airline can realize by overbooking flight j; i.e., the gross and net benefits. Both the gross and net benefits can be obtained by calculating the extra revenue generated for each traveler i by overbooking the flight, and adding the resulting figure across all i. The gross and net benefits, however, require different assumptions regarding the composition (mix) of travelers on flight j.

To obtain the gross benefit, we assume that all the "additional" passengers (α) are Type II (new) customers (i.e., $\lambda = 1$). Since the gross benefit considers *only* the revenue gains attained for overbooked flights, we ignore the possible loss of revenues in other flights by excluding flight switchers from the formula. This condition indicates that the gross benefit of overbooking flight j, $B(Gross)$, can be obtained by adding $B(\mathrm{I})$ and $B(\mathrm{II})$, but not $B(\mathrm{III})$, and fixing λ to 1:

$$\begin{aligned} B(Gross) = & -\sum_{i=1}^{n}\{SH_iBP_iCOMP_i\} \\ & + \sum_{i=n+1}^{n+\alpha}\{SH_i(P_i - BP_iCOMP_i) + (1 - SH_i)(RF_iNR_iP_i + (1 - RF_i)P_i)\}. \end{aligned} \tag{6}$$

To obtain the net benefit, on the other hand, we allow "additional" passengers to include both Type II and Type III customers, so that the possible loss of revenues in other flights (the opportunity cost of overbooking flight j) is considered. Thus, the net benefit of overbooking flight j, $B(Net)$, is obtained by adding $B(\mathrm{I})$, $B(\mathrm{II})$, and $B(\mathrm{III})$, without fixing λ to a certain value:

$$\begin{aligned} B(Net) = &-\sum_{i=1}^{n}\{SH_iBP_iCOMP_i\} \\ &+\sum_{i=n+1}^{n+\alpha\lambda}\{SH_i(P_i - BP_iCOMP_i) + (1-SH_i)(RF_iNR_iP_i + (1-RF_i)P_i)\} \\ &+\sum_{i=n+\alpha\lambda+1}^{n+\alpha}\{SH_i(P(\Delta)_i - BP_iCOMP_i) + (1-SH_i)(RF_iNR_iP(\Delta)_i + (1-RF_i)P(\Delta)_i)\}, \end{aligned} \tag{7}$$

where $P(\Delta)_i = P_i - P(-OB)_i$. Notice that if $\lambda = 1$, Eq. (7) reduces to Eq. (6). Eq. (7), therefore, is a *generic formula* that calculates both the gross and net benefits depending on the value of λ.

In the sections that follow, we evaluate, for a variety of λ values (0–1), Eqs. (6) and (7) and analyze their relationships. A particular emphasis is placed on investigating the ratio of Eq. (7) to Eq. (6). This ratio represents the extent to which the net benefit is lower than the gross benefit, and therefore is a useful statistic to measure the relationship between the gross and net benefits. For clarity, we will refer to this statistic succinctly as the "benefit ratio" from now on.

3. Simulation

One approach to calculate the values of Eqs. (6) and (7) (and their ratios), for a variety of λ values, is to derive the expected values of the equations. We do not, however, use this approach as it gives only the *means*. Notice that the right-hand-side (RHS) variables of Eqs. (6) and (7) are *random variables*, whose values change randomly from one traveler to another. This condition indicates that $B(Gross)$ and $B(Net)$ are also random variables. Given the stochastic nature of $B(Gross)$ and $B(Net)$, we should investigate, in addition to the means, their variances and distribution shapes, because these statistics give important insights into the characteristics of random variables. For this reason, we perform a series of simulation experiments and "empirically" derive the means, variances, and distributions of $B(Gross)$, $B(Net)$, and their ratios.

3.1. Experiment

We simulate the unit revenue of a hypothetical US domestic flight j, which is assumed to be always congested and overbooked. The number of available seats (n) is assumed to be 200. Each experiment represents a flight departure, and is performed according to the following steps.

First, $n + \alpha$ reservations are generated for flight j, where α is a non-negative random variable representing the number of overbooked travelers. For each $n + \alpha$ travelers making reservation for flight j, traveler (i) characteristics such as the customer type (Type I, II, or III), SH_i, P_i,

RF_i, NR_i, and $P(-OB)_i$ are determined by using a PC random number generator (RNG). Second, the number of travelers showing up for flight j (denoted as S) is calculated by counting the number of travelers whose SH_i value is 1 (recall that $SH_i = 1$ if i shows up for j). Third, the value of S is compared with the seat capacity n to determine the number of travelers who must be bumped from flight j. (If $S \leqslant n$, no traveler needs to be bumped. If $S > n$, $S-n$ travelers must be bumped.) Fourth, if one or more travelers need to be bumped from j, the RNG selects (randomly) the specific traveler(s) to bump. (This process assigns the value of BP_i to all $n + \alpha$ travelers, where $BP_i = 1$ if i is bumped.) Fifth, for each bumped traveler, the amount of compensation ($COMP_i$) is determined by the RNG, subject to certain upper and lower limits.

Once the above simulation procedure is complete (i.e., at the end of each experiment), $B(Gross)$ and $B(Net)$, as well as their ratios, are calculated by using Eqs. (6) and (7). We perform this experiment 1000 times for each value of λ ranging from 0 to 1 in the increment of 0.01 (101 values). This task involves simulating 101,000 flight departures and creating approximately 23.3 million seat reservation data. We perform this entire simulation process twice to test two different scenarios (details of the scenarios are discussed later). Thus, in total, we simulate 202,000 flight departures and generate roughly 46.5 million seat reservation data.

3.2. Simulation inputs

This section describes how the RHS variables of Eqs. (6) and (7) (simulation parameters) are specified in this study. To obtain realistic results, we specified the parameters carefully by referring to a variety of data sources. These sources include: (1) *DataBank 1B* (DB1B) of the US Department of Transportation (a 10% sample of all the US domestic airline tickets), (2) research publications, (3) telephone interviews with major US airlines, [6] (4) magazines such as *US News & World Report*, (5) newspapers such as *USA Today*, and (6) internet home-pages such as www.airline-fares.com. A summary of parameter specifications is shown in Table 1.

Overbooked travelers (α): In general, US airlines overbook their congested flights by about 10–20% of seat capacities (see, e.g., US News and World Report, 1990, Smith et al., 1992, USA Today, 1999). Given this condition, we determine the value of α in each experiment by using a normally-distributed random variable whose mean is 30 (15% of 200). The standard deviation is set to 3.5, so that roughly 99% of the time the α will lie between 20 and 40 (between 10% and 20% of 200). Since α must be integer and non-negative, we: (1) round the value of α to the nearest integer, and (2) truncate the distribution so that it produces only non-negative values.

Customer type: In each experiment, there are $n + \alpha$ travelers having reservations on flight j at the time of departure, each of whom must be classified as a Type I, Type II, or Type III customer. By definition, travelers making the first n reservations always constitute Type I customers. Each of the remaining α travelers is classified as either a Type II or Type III customer, by a random Bernoulli variable in which the probability of being Type II is given by λ.

Refundable ticket (RF_i): According to Fan (2002), the proportion of travelers with refundable tickets on US domestic flights ranges roughly from 6% to 16%. Since the proportion of refund-

[6] We contacted nine US major airlines, of which six were willing to answer our questions, at least partially. The interviewees were mostly the senior personnel in customer service or public relations departments.

Table 1
Simulation parameter specifications

Parameter	Distribution	Mean	Std. dev.	Probability	Information sources
α[a]	Normal[b]	30	3.5	–	*US News & World*, Smith et al. (1992), *USA Today*
λ[c]	N.A.	–	–	–	N.A.
RF_i (Type I)	Bernoulli	–	–	0.060	Airline interviews, Fan (2002)
RF_i (Type II & III)	Bernoulli	–	–	0.800	
SH_i (refundable)[d]	Bernoulli (*A*)	–	–	0.600	*USA Today*, Alstrup et al. (1989), Airline interviews
	Bernoulli (*B*)	–	–	0.750	
SH_i (non-refundable)[d]	Bernoulli (*A*)	–	–	0.926	
	Bernoulli (*B*)	–	–	0.898	
P_i (refundable)	Normal[b]	311	46	–	*DataBank 1B* (2003), Suzuki (2002)
P_i (non-refundable)	Normal[b]	165	25	–	
δ_i (refundable)	Normal[b]	0.90	0.04	–	Internet search of selected airlines and dates
δ_i (non-refundable)	Normal[b]	0.75	0.10	–	
NR_i	Bernoulli	–	–	0.050	Airline interviews
BP_i[e]	N.A.	–	–	–	N.A.
$COMP_i$	Normal[b]	300	40	–	Airline interviews, *USA Today*

[a] Rounded to the nearest integer.
[b] These distributions are truncated. See the manuscript for detailed descriptions.
[c] The range between 0 and 1 is tested by the increment of 0.01. In the simulation, each overbooked traveler is classified as either a Type II or Type III passenger by a random Bernoulli variable in which the probability of Type II is given by λ.
[d] Two scenarios (*A* and *B*) are tested for this parameter.
[e] Determined by the random number generator (0 or 1) when the number of travelers showing up for flight j is larger than n. The expected number of bumped passengers per flight is approximately 2.69, which is about 1.3% of all passengers who show up for flight j.

able-ticket holders is usually higher in congested (overbooked) flights than in un-congested flights, [7] we assume that: (1) the lower-bound figure (6%) reflects the proportion of refundable-ticket holders in un-congested flights, and (2) the upper-bound figure (16%) reflects the proportion of refundable-ticket holders in congested or overbooked flights (flight j).

Given these assumptions, we determine the ticket type (refundable or non-refundable) of each traveler i on flight j by a random Bernoulli variable in which the probability of a refundable ticket is assumed to be 0.06 for Type I customers, and 0.80 for Type II and III customers. The probability of 0.06 is assumed for Type I customers, as the proportion of travelers with refundable tickets is roughly 6% in un-congested flights (which contain only Type I customers) (see above). The probability of 0.80 is assumed for Type II and III customers, because this figure makes the overall proportion of travelers with refundable tickets on flight j to be roughly 16% (given the

[7] Our airline interviews indicate that non-refundable tickets are mostly held by travelers who make early reservations (most non-refundable tickets require 14-day advance purchases), while refundable tickets are mostly held by travelers who make late reservations. This pattern implies that: (1) the probability of having refundable tickets is higher for Type II and III customers than for Type I customers, and (2) therefore the proportion of refundable-ticket holders is higher in congested (overbooked) flights than in un-congested flights, as the latter flights do not contain Type II and III customers who are more likely to have refundable tickets than Type I customers.

refundable-ticket probability of 0.06 for Type I customers), which reflects the average proportion of travelers with refundable tickets on most overbooked flights (see above). [8]

Show-up probability (SH_i): According to our airlines interviews, almost all no-shows are caused by double bookings by travelers with refundable tickets; i.e., they reserve seats in two flights, show up for only one, and refund the unused tickets (also see American Airlines' internet web-site for similar discussions). This condition implies that no-show probabilities should be higher for travelers with refundable tickets than for travelers with non-refundable tickets.

We may reasonably assume that, for travelers with refundable tickets, the upper-bound no-show probability is 0.5 (since their no-show probability becomes 0.5 if they all use double-bookings), while the lower-bound is 0.125 (the average no-show probability of all travelers [9]). We test two scenarios within these upper and lower bounds. Specifically, we test (1) *Scenario A* in which the no-show probability of refundable-ticket holders is 0.4, and (2) *Scenario B* in which the no-show probability of refundable-ticket holders is 0.25. The no-show probabilities of non-refundable ticket holders in scenarios *A* and *B* are set to 0.074 and 0.102, respectively, which are reverse calculated from the overall (flight j total) no-show probability of 0.125 (see footnote 9).

Note that the expected number of bumped passengers per flight is given by the formula:

$$\sum_{S=0}^{n+\alpha} \left\{ \max(0, S-n) \cdot \frac{(n+\alpha)!}{S!(n+\alpha-S)!} (1-p)^S p^{n+\alpha-S} \right\}, \tag{8}$$

where p is the overall no-show probability of passengers on flight j (0.125). Calculation of Eq. (8) indicates that, in our simulation, the expected number of bumped passengers per fight is roughly 2.7 (1.3% of travelers who show up for flight j). This figure is similar to that of Alstrup et al., 1986), who report that in overbooked flights about 1–2% of passengers are bumped. [10]

Airfares (P_i): We use the average one-way US domestic airfare calculated from the DB1B (3rd quarter, 2003, all routes included) as the mean of P_i distribution ($P_i \sim$ normal). [11] Since we found strange airfares in the data (e.g., airfare < 0, and >60,000), we sort the data by airfares, and use only the middle 80% of the data (eliminate the lowest 10% and highest 10%). We compute average airfares separately for travelers with refundable tickets and for those with non-refundable tickets. Only the data for economy-class tickets are used (because overbooking is most frequently used in this class). We do not use the standard deviation (STD) obtained from the DB1B as the STD of P_i distribution. Suzuki (2002) argues that the STD calculated from the DB1B tends to overestimate the true value, and that the true value may be approximated by multiplying the average airfare by

[8] The refundable-ticket probability of 80% for Type II and Type III customers may not be unrealistic, because in most situations these customers may purchase tickets only after all the non-refundable tickets are sold out.

[9] In the US domestic market, the proportion of travelers who do not show up for their flights is usually in the range of 10–15% (see, e.g., USA Today, 1999; Alstrup et al., 1989).

[10] US DOT data, such as *Air Travel Consumer Report*, indicate that the proportion of bumped passengers (voluntary and involuntary bumped passengers) is far less than 1%. The DOT figures, however, are obtained by dividing the number of bumped passengers by the *total enplaned passengers*. The DOT figures do not reflect the proportion of bumped passengers *on overbooked flights*.

[11] When airfares are for round-trip tickets (which is true for most data found in DB1B), the one-way fare is obtained by dividing the round-trip fare by two.

0.15. For this reason, the STD of P_i distribution, for both the refundable and non-refundable tickets, is set to 15% of the average P_i in our experiments.

Substitute airfare ($P(-OB)_i$): Using the internet, we investigated the airfares of selected US airlines, and analyzed the extent to which the fares differ by flight within the same route, booking class, and trip date. [12] Our results indicate that the ratio of the lowest fare to the highest fare (within a route, class, and date) typically ranges from 0.8 to 1 for refundable tickets and from 0.5 to 1 for non-refundable tickets. We assume that these ratios reflect the values of $P(-OB)_i/P_i$. [13] Given these assumptions, we model $P(-OB)_i$ as $P_i \times \delta_i$, where δ_i is a normally-distributed random variable whose mean is 0.9 for refundable tickets and 0.75 for non-refundable tickets, subject to $0 \leqslant \delta_i \leqslant 1$ ($\forall i$). The standard deviations of δ_i are set to 0.04 and 0.10 for refundable and non-refundable tickets, respectively, so that roughly 99% of the time the δ_i will lie within the observed range (0.8–1 for refundable tickets; 0.5–1 for non-refundable tickets).

No-refund probability of no-show travelers with refundable-tickets (NR_i): As mentioned previously, almost all no-shows may be caused by double bookings by travelers with refundable tickets (those who reserve seats in two flights, show up for only one, and refund the unused tickets). Thus, no-shows by travelers with refundable tickets may almost always result in ticket refunds. There may be, however, a limited number of refundable-ticket holders who: (1) miss their flights and wish to fly on later flights, or (2) simply forget to refund the un-used tickets. In our simulation, each no-show traveler with a refundable ticket is classified as either a no-refund or a refund traveler, by a random Bernoulli variable in which the probability of no-refund is 0.05.

Compensation for bumped passengers ($COMP_i$): During our telephone interviews, we asked airlines if they knew the average compensation paid per bumped passenger. While none knew the exact amount, many gave us the typical range, which were usually between \$200 and \$400. [14] In our simulation, we use this typical range. Specifically, once a traveler is bumped, the amount of compensation for this traveler is determined by a normally-distributed random variable whose mean and standard deviation are \$300 and \$40 respectively. Notice that by using these figures the $COMP_i$ will lie, roughly 99% of the time, within the typical range of \$200 to \$400. The distribution is truncated such that the lower limit is \$0 and the upper limit is \$800. [15]

---

[12] We investigated the airfares of three major US carriers (American, Delta, United) on randomly selected 20 domestic origin-destination routes (economy-class fares only).

[13] In general, the highest fares are found in congested flights, while the lowest fares are found in un-congested flights (within a route, class, and date). This pattern implies the followings. First, since flight *j* is a congested flight, its fare (P_i) may be the highest within a route, class, and date. Second, since the substitutes of flight *j* for Type III customers (if they cannot obtain seats on flight *j*) are un-congested flights (by definition), fares of the substitute flights ($P(-OB)_i$) may be the lowest within a route, class, and date. Third, consequently, the ratio of the lowest fare to the highest fare (within a route, class, and date) may reflect the ratio $P(-OB)_i/P_i$.

[14] When an airline indicated that (a part of) the compensation is given in the form of a free future ticket, we used the monetary value of the ticket (average airfare of all the routes for the airline) to calculate the amount of compensation, because the ticket value may reflect the "lost future revenue" of the airline (opportunity cost).

[15] The maximum amount of compensation a US domestic airline pays for a bumped passenger is \$800 (see, e.g., USA Today, 1999).

4. Results

Simulation results are shown in Figs. 2–6. In each figure, the results are reported separately for scenario *A* (no-show probability for refundable-ticket holders = 0.4) and for scenario *B* (no-show probability for refundable-ticket holders = 0.25).

Fig. 2a and b show the mean values of *B*(*Gross*) and *B*(*Net*) for the entire range of λ tested in our experiments (from 0 to 1). From the figures we see that, for both scenarios *A* and *B*: (1) *B*(*Gross*) seems to be *unaffected* by λ, (2) *B*(*Net*) appears to be affected by λ, such that the lower the value of λ the lower the *B*(*Net*), and (3) the *B*(*Net*) curve always lies beneath the *B*(*Gross*) curve regardless of the value of λ. These conditions imply that, as expected, the value of *B*(*Net*) is always *lower* than that of *B*(*Gross*), unless $\lambda = 1$ (where *B*(*Gross*) = *B*(*Net*)).

Fig. 3a and b show the variability of *B*(*Gross*), and Fig. 4a and b show that of *B*(*Net*). In this study, we use the 10th and 90th percentile data points to establish the range of variability for

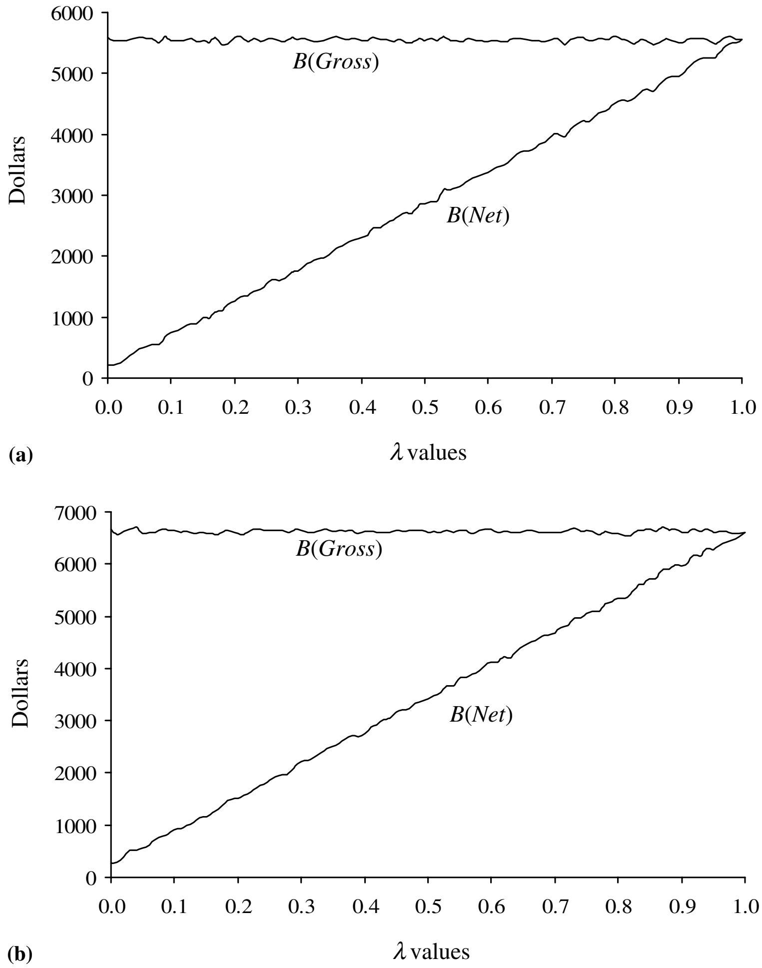


Fig. 2. Net vs. gross benefit: (a) Scenario *A* and (b) Scenario *B*.

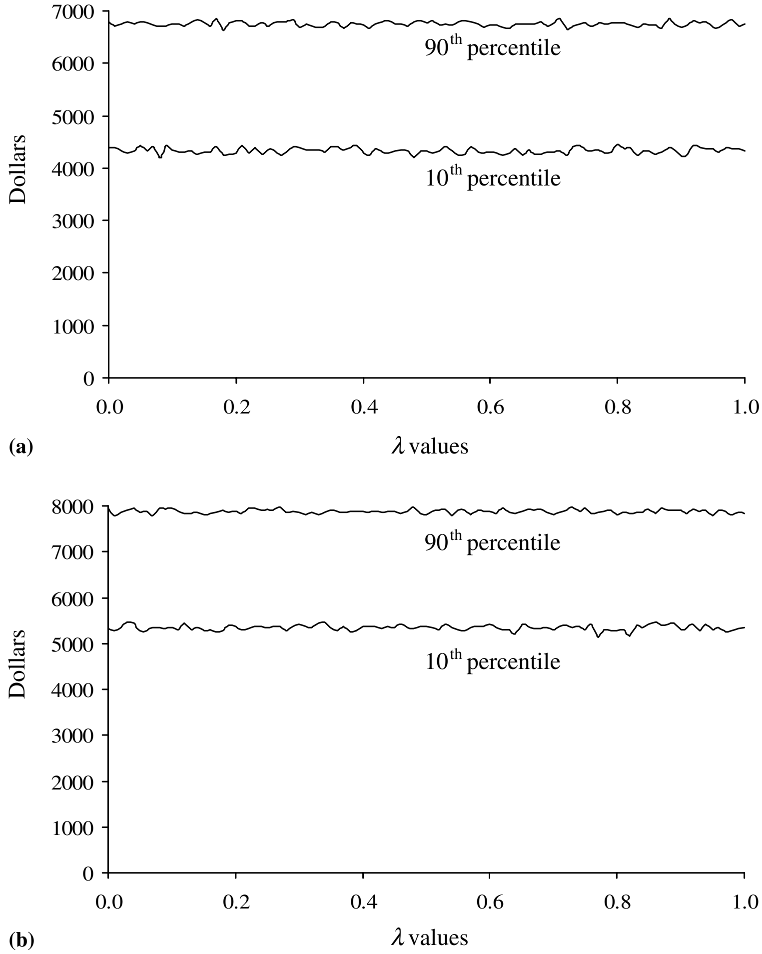


Fig. 3. Gross benefit variability: (a) Scenario *A* and (b) Scenario *B*.

B(*Gross*) and *B*(*Net*) (i.e., by using the middle 80% of "observed" data). While we could have used the standard deviations and established the confidence intervals instead, we do not use this approach. The confidence intervals indicate the likely locations of the *population means*, and do not reflect the actual range of variability (of *B*(*Gross*) and *B*(*Net*)) which an airline may experience in practice. The 10th and 90th percentiles, on the other hand, reflect the range in which *B*(*Gross*) and *B*(*Net*) are actually observed in our experiments (roughly 80% of the time), and therefore may be viewed as the practical range of variability for airlines.

We see from Fig. 3a and b that *B*(*Gross*) is most likely to be positive under all conditions, because its lower bound (10th percentile) is always positive in the entire range of λ. Fig. 4a and b, on the other hand, indicate that, while *B*(*Net*) is likely to be positive in most situations, it can be negative if the value of λ is less than 0.2 (observe that for both scenarios *A* and *B* the 10th percentile of *B*(*Net*) becomes negative if $\lambda < 0.2$). This condition implies that in order for an airline to obtain a truly positive overbooking benefit, the new-customer proportion (λ) must be at least 20%. Thus,

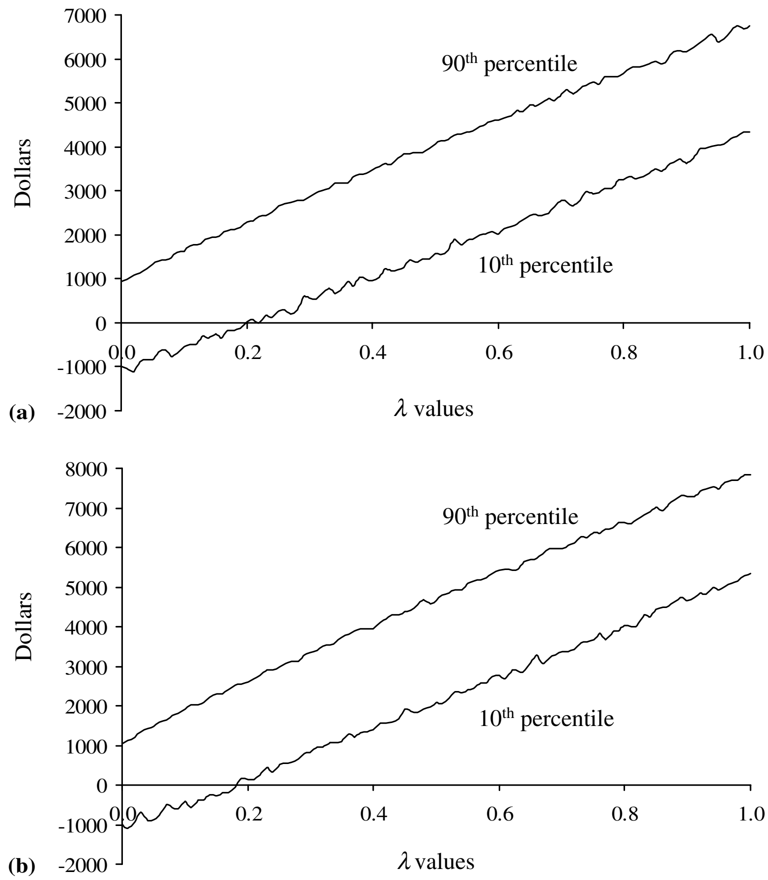


Fig. 4. Net benefit variability: (a) Scenario *A* and (b) Scenario *B*.

Figs. 3a, b and 4a, b jointly suggest that: (1) while the gross benefit may be positive under all conditions the net benefit can be negative if λ is less than 0.2, and (2) so that the use of overbooking can result in *revenue losses*, if the "additional" passengers generated by overbooking do not contain sufficient number of "new" customers.

Fig. 5a and b report the means and variabilities of the "benefit ratio" ($B(Net) \div B(Gross)$) for scenarios *A* and *B*, respectively. As before, we measure the range of variability by using the 10th and 90th percentile data points. The most important finding is that when λ is close to 1 ($\lambda \geqslant 0.92$) the 90th percentile of the benefit ratio is equal to 1 (for both scenarios *A* and *B*). This condition indicates that when $\lambda \geqslant 0.92$ the net benefit is equal to the gross benefit in at least 10% of our experiments. Our results, therefore, imply that: (1) the net benefit may not be substantially different from the gross benefit when the new-customer proportion (λ) is 92% or higher, but (2) the net benefit may be substantially different from (lower than) the gross benefit otherwise. Given these results, the use of gross benefits by airlines in lieu of net benefits may be justified only when the new-customer proportion (λ) is greater than or equal to 92%.

To better understand the nature of the relationship between the benefit ratio and λ, we regressed the benefit ratio on λ by using the ordinary least squares (results are reported in

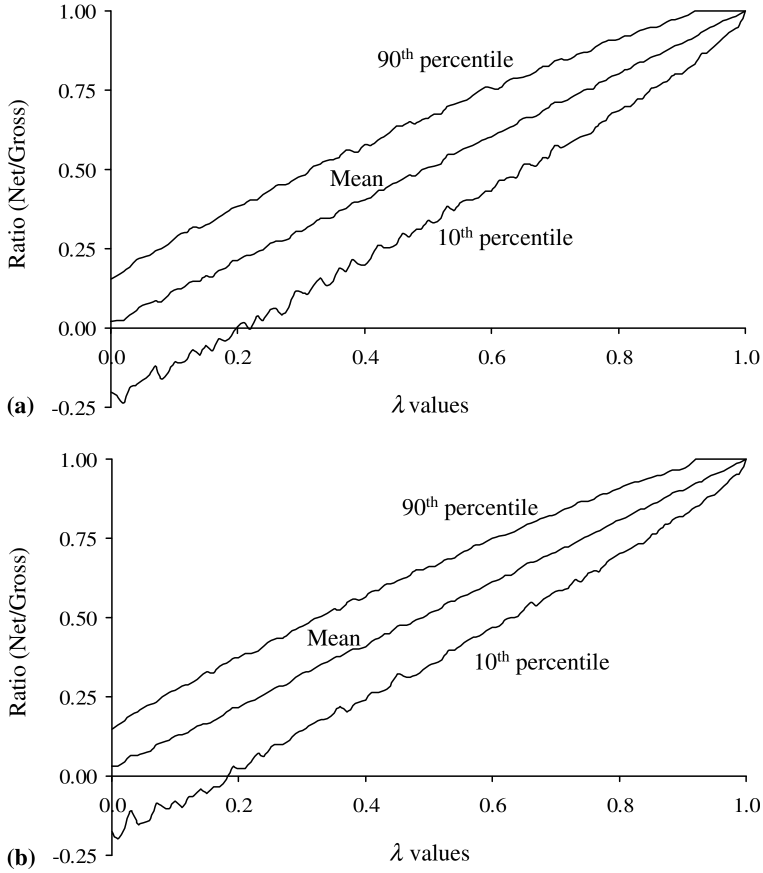


Fig. 5. Benefit ratio: (a) Scenario A and (b) Scenario B.

Table 2). We employed a linear-regression technique because Fig. 5a and b indicate that, for both scenarios A and B, the benefit ratio appears to be a linear function of λ. The regression results indicate that, for both scenarios A and B: (1) the slope of the fitted line (λ coefficient) is roughly one, (2) the model intercept is close to zero, and (3) the model fit is nearly perfect ($R^2 > 0.999$). These conditions imply three things. First, the expected value of the benefit ratio may be roughly equal to λ in most situations. For example, if λ is 0.7 (70%), the net benefit may be 70% of the gross benefit (i.e., 30% lower than the gross benefit). Second, a one-percent increase (decrease) of λ may result in a one-percent increase (decrease) of the benefit ratio. Thus, if λ declines by one percent, an airline's "extra" revenue generated by practicing overbooking may also decrease by one percent. Third, since the regression results are similar between scenarios A and B, the benefit ratio may be minimally affected by the no-show probabilities of individual travelers, as long as the overall proportion of no-show travelers within a flight is held fixed. [16]

[16] Recall that while the no-show probability of *individual travelers* with refundable tickets is different between scenarios A and B, the *overall proportion* of no-show travelers on flight j is identical for both scenarios.

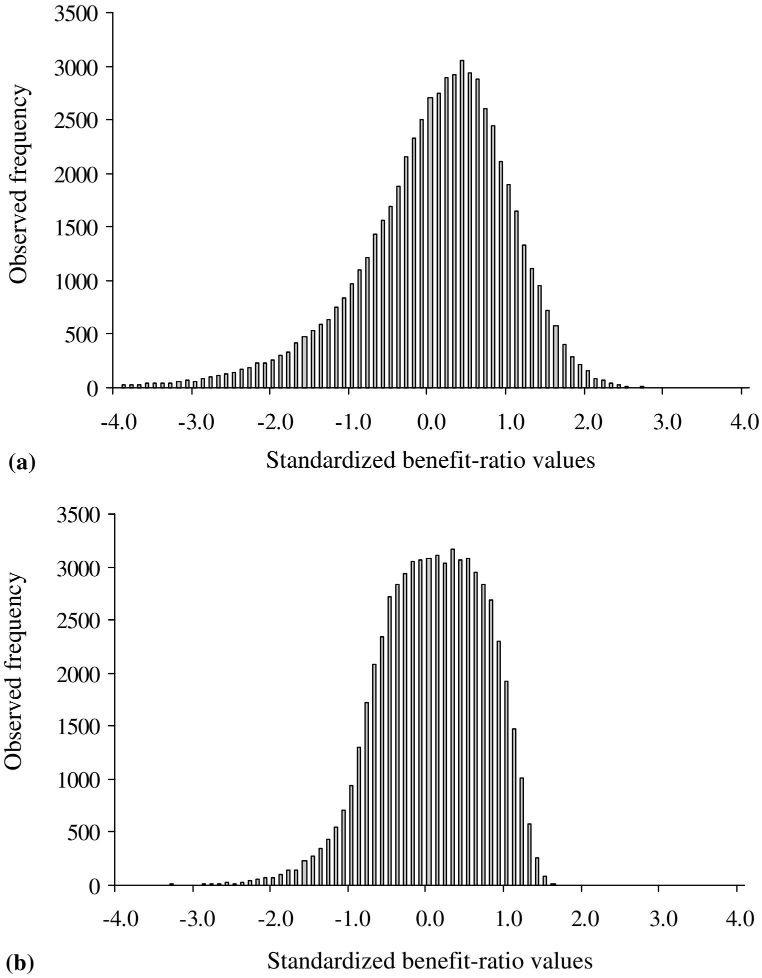


Fig. 6. Benefit ratio distribution: (a) Scenario A and (b) Scenario B.

Table 2
Regression results

	Scenario A[a]			Scenario B[b]		
	Coefficient	t-Stat.	p-Value	Coefficient	t-Stat.	p-Value
Model parameters						
Intercept	0.0142	16.7	<0.001	0.0254	37.9	<0.001
λ	0.9849	670.8	<0.001	0.9754	842.9	<0.001
Sample size	101			101		
Model fit (R^2)	>0.999			>0.999		
F-statistic	450,028			710,479		

[a] No-show probability of refundable-ticket holders = 0.40.
[b] No-show probability of refundable-ticket holders = 0.25.

The data on benefit-ratio variability (Fig. 5a and b) indicate that, within the range of λ between 0.1 and 0.9, the 10th percentile data points are located roughly 0.15 below the means and the 90th percentile data points are located roughly 0.15 above the means. [17] This condition, in conjunction with the regression results reported in the previous paragraph, imply that the benefit ratio of an airline may be bounded, 80% of the time, between $\lambda - 0.15$ and $\lambda + 0.15$. Given these results, we can derive the following formulas, which may be used conveniently by airlines to approximate the net overbooking benefits from the available gross benefit figures:

$$B(Net) \geqslant B(Gross)(\lambda - 0.15), \tag{9a}$$

$$B(Net) \leqslant B(Gross)(\lambda + 0.15). \tag{9b}$$

Eq. (9) indicates that, for example, if an airline's gross benefit is \$100,000 and $\lambda = 0.7$, this airline's net benefit may be found, 80% of the time, in the range between \$55,000 and \$85,000. Airlines may wish to empirically estimate the value of λ (e.g., by using passenger surveys), and apply the estimated figure to Eq. (9) to find the likely values of their net overbooking benefits.

Fig. 6a and b illustrate the distributions of benefit ratios. These distributions are obtained by using: (1) the standardized benefit-ratios, and (2) the range of λ between 0.1 and 0.9 (see footnote 17 for the rationale of using this range). The sample size used to construct each distribution is 81,000. The figures indicate that in both scenarios the benefit ratio seems to be normally distributed. The formal statistical test, however, rejected the normal-distribution hypotheses for both scenarios with significant *p*-values ($p < 0.001$). The reason is that in both scenarios the distribution is skewed to the left (negatively skewed). This condition indicates that in both distributions the mean is *lower* than the median and the mode; i.e., the benefit ratios that are most frequently observed in practice may be *higher* than the mean ratios reported in this paper. Airlines may need to take this point into account when interpreting our simulation results.

5. Conclusions and limitations

Often, airlines estimate the benefit of overbooking by considering only the amount of revenue gains attained for overbooked flights, and paying little attention to the potential loss of revenues that may take place in other (non-overbooked) flights due to travelers' flight-switching behaviors. This study has calculated the "net" benefit of overbooking that considers both the positive and negative aspects of overbooking, and investigated the amount by which the net benefit may be lower than the gross benefit under a variety of conditions.

The most important airline implications are as follows. First, while the gross overbooking benefit may be positive under all conditions, the net benefit can be negative if the share of "new" customers within the "additional" passengers (new-customer proportion) is small. According to our

[17] We use the range of λ between 0.1 and 0.9 because: (1) this range may represent the realistic range of λ that an airline may experience in practice, and (2) the benefit-ratio variability is forced to be small when $\lambda > 0.9$ (because the ratio cannot exceed 1 theoretically), so that the variability may be underestimated if this range ($\lambda > 0.9$) is included.

results, in order for an airline to obtain a truly positive net overbooking benefit, the new-customer proportion must be at least 20%. Second, the net benefit may be substantially different from the gross benefit unless the new-customer proportion is greater than or equal to 92%. This condition suggests that the use of gross benefits by airlines in lieu of the true (net) benefits may be justified only when the new-customer proportion is at or higher than 92%. Third, the net benefit may be found, in most practical situations (80% of the time), in the range between $B(Gross)(\lambda - 0.15)$ and $B(Gross)(\lambda + 0.15)$. Airlines may conveniently use these formulas to estimate the net overbooking benefits from the available gross benefit figures.

Our study has its limitations, which may need to be addressed by future research. First, as mentioned previously, we did not consider the service charges that are required to refund certain types of tickets (refund processing fee). If the cost of processing a ticket refund (for airlines) is minimal, the ticket purchase that eventually results in a refund may generate a positive airline revenue (service charge). Future research may wish to incorporate service charges as potential sources of airline revenues when calculating the net benefit.

Second, in our simulation, the travelers to be bumped in each flight (when $S > n$) were determined *randomly* without regard to the type of travelers. In practice, certain types of travelers may be more likely to be bumped than other types. For example, the low-fare passengers with lower time-values may be more likely to be bumped than other types of travelers, as they may volunteer to be bumped (airlines first ask for volunteers to be bumped before they force travelers to give up seats). Future research may test the robustness of our study results by examining a variety of scenarios regarding the mix of passengers who are bumped.

Third, our study did not consider "ticket sell-ups". When overbooking is *not* used, travelers demanding economy-class tickets may purchase upper-class tickets *of the same flight* if they cannot obtain economy-class tickets. This phenomenon is referred to as "ticket sell-ups" (Curry, 1990). The ticket sell-ups imply that once overbooking is used, airlines may be selling economy-class tickets to those who may otherwise have purchased upper-class tickets (i.e., reduced revenues). If the "sell-up" effect is considered, the overbooking benefit may become even lower than the net benefit reported in this paper. Future research may wish to estimate (e.g., by using discrete-choice models), how overbooking may affect the booking-class choices of travelers, and take this effect into account when calculating the net benefit.

Fourth, we tested only two scenarios with regard to passenger show-up probabilities during our experiments (scenarios A and B). Although we believe, given the similarities of simulation results between scenarios A and B, that testing other scenarios does not affect or change the general implications reported in this paper, future research may wish to consider testing other scenarios for completeness.

Acknowledgments

This research was funded by the Iowa State University College of Business Mini Scholarship Grant. The author would like to thank the airlines that kindly provided the useful information to this study. Thanks also go to Christina Alverey and Ravi Vasikarla for their research assistance.

References

Alstrup, J., Boas, S., Madsen, O.B.G., Vidal, R., Victor, V., 1986. Booking policy for flights with two types of passengers. European Journal of Operational Research 27, 274–288.

Alstrup, J., Anderson, S., Boas, S., Vidal, R.V., 1989. Booking control increases profit at Scandinavian airlines. Interfaces 19, 10–19.

Curry, B.E., 1990. Optimal airline seat allocation with fare classes nested by origins and destinations. Transportation Science 24, 193–204.

Davis, P., 1994. Airline ties profitability yield to management. SIAM News 27 (5).

Fan, T., 2002. Smaller aircraft for more profits? A preliminary examination of airlines' fleet size decision with fare and demand distributions. Journal of the Transportation Research Forum 56, 77–93.

McGill, J.I., Ryzin, G.J.V., 1999. Revenue management: research overview and prospects. Transportation Science 33, 233–256.

Nambisan, S.S., 2003. A decision-support tool for airline yield management using genetic algorithms. Computer-Aided Civil and Infrastructure Engineering 18, 214–223.

Smith, C., Leimkuhler, J.F., Darrow, R.M., 1992. Yield management at American Airlines. Interfaces 22, 8–31.

Suzuki, Y., 2002. A heuristic method for effectively using frequent flyer miles for stochastic business trips. Transportation Research Part E (Logistics and Transportation Review) 38, 335–349.

US Department of Transportation, Air Travel Consumer Report, various editions.

US News and World Report, 1990. Bumped off at the airport. April 16, 108 (15), pp. 54–57.

USA Today, 1999. More fliers getting left at gate. April 19, p. 38.