



Airline demand distributions: passenger revenue management and spill

William M. Swan *

Boeing Commercial Airplanes Group, P.O. Box 3707, MC21-28, Seattle, WA 98124, USA

Abstract

Both revenue management and airline schedule optimization need to characterize the distribution of likely demand outcomes. Sources have proposed both Gamma and Normal shapes for these distributions. Data suggests that a model combining both distributions is appropriate. The model explains when the Gamma shape will dominate and when the Normal will determine the shape. One consequence of this understanding is that Gamma shapes are probably better for revenue management and Normal for spill modeling. However, it takes a compound process combining the two to generate all the observed characteristics of various cases. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Airline; Spill; Revenue management; Demand distribution; Scheduling

1. Introduction

Airlines need to know the shape of demand distributions in two common situations. The first is revenue management. Reservation systems are supposed to protect the right number of seats for late booking high-fare demand. To do this they need to know the distribution for the expected high-fare demand for each day. The second case is scheduling which flights get larger airplanes. For this, the distribution of demand for a flight leg over a month is important for calculating net loads for various seating capacities. The same underlying structure gives insights into both cases. The one-day problem is known as revenue management. The literature on revenue management is vast. An introduction can be found in Tretheway and Oum (1992). The one-month difference between demand and load is known as spill. The spill problem has been addressed largely within the airline industry. Li and Oum (2000) and Swan (1983, 1993) are among the few publicly available reviews.

* Tel.: +1-206-766-2515; fax: +1-206-766-1022.

E-mail address: bill.swan@boeing.com (W.M. Swan).

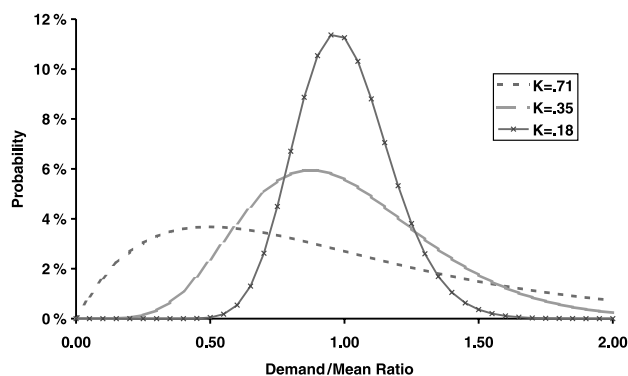


Fig. 1. Gammas change shape with width.

Distributions of demand can be described by the usual statistics: mean, standard deviation, and higher-order moments. These are referred to as the average demand, the width,¹ and the “shape” of the demand distribution. Typically, the average is measured from data at hand for a particular case, while the width is calculated from larger historical trends. The shape is often an assumption buried in the derivation of the optimization model. This paper focuses on the idea of shape—how it changes and how changes coordinate with changes in width.

Modeling often assumes one of two basic shapes for variations in demand—the Normal and the Gamma distribution. The Normal is familiar. It maintains its shape at any width. Higher-order moments are always zero. The Gamma is less common, and the shape changes depending on the width. This is illustrated in Fig. 1. Reasonable models of the underlying statistical phenomena indicate that the true shapes are combinations of these two.

2. Underlying statistical description

A single statistical story captures what drives demand variations for air travel demands. That model starts by portraying demand as a series of coin flips. Later, the probability for the coin will change from day to day.

First, consider the demand for a single departure of a flight on a particular day. For instance, the 9:00 flight from Seattle to Chicago on 28 April. This would be the situation for revenue management. Imagine that in Seattle each person has a likelihood of wanting to travel on the 28th to Chicago at 9:00. This population may be viewed as making a Bernoulli-trial “coin flip” to decide whether to seek the flight. If the probabilities for the coins were known, the expected demand would be known. However, a city’s worth of coin flips still means that there is uncertainty in the outcome for any one day. This uncertainty gives a distribution of demands, and a width. A minimum amount of underlying randomness becomes part of the demand phenomenon. Data supports this picture.

¹ This paper uses the short-hand term “width” for the coefficient of variation, i.e. the ratio of the standard deviation to the mean. The term “K-factor” or “K” is commonly used by airlines.

If travel decisions are correlated between individuals, the width increases. Some people travel or stay home together. Others travel to the same meeting, or cancel trips for the same external reason. Any grouping of decisions is like several people sharing the same coin flip. It increases the width of the overall distribution.

Estimation errors add to this width. For instance, revenue management problems have to be solved with imperfect knowledge of the coins being flipped. Usually, average demand has to be estimated from earlier similar days. Any estimate is imperfect. Errors in estimating this average add to the width of the distribution of demand. Revenue management must operate with demand distributions including estimation errors.

Additional width is added when the problem is expanded beyond the one-day-at-a-time view of revenue management. For assignment of an airplane capacity to a schedule for a month or more, the probabilities of the coins being flipped change day-by-day for the various days of the week and weeks of the month being considered. These predictable cyclic variations add a third component to the width of the demand distribution.

The statistical model of the underlying phenomena now has four parts:

1. People flip coins to determine whether to take a flight.
2. People travel in groups.
3. The expected value of the coins is estimated with estimation errors.
4. The value changes daily when a group of flights is being considered.

We now discuss these parts in detail referring to data and issues from revenue management and spill modeling applications. In the end, our story will explain the shapes we see for demand distributions.

3. Gamma shapes for randomness—people flip coins

The statistical model is of a city full of people who “flip a coin” to determine whether to travel on a particular day, where to go, and which flight to take. The simplest version of this is a city with N people each with a probability P of choosing the flight in question. The result of N such Bernoulli trials are the well-worn conditions:

The expected demand $\mu = NP$.

The standard deviation of demand $\sigma = \sqrt{NP(1 - P)}$.

For the case where P is small, $(1 - P) \cong 1$, the familiar answer becomes $\sigma = \sqrt{\mu}$.

A number of simplifications have been made without doing damage to the basic concept that there is randomness in people’s behavior. First, each person’s coin may have a different value for P , but the model uses a (root mean square) average. Second, a person makes several choices between deciding to travel, picking the time, and selecting a particular flight. If each of these choices is a sequential coin flip, then the P used here represents a product of several component probabilities. Finally, we are eventually going to approximate the compound Bernoulli distribution with a Gamma with the appropriate mean and standard deviation. When μ is small, this

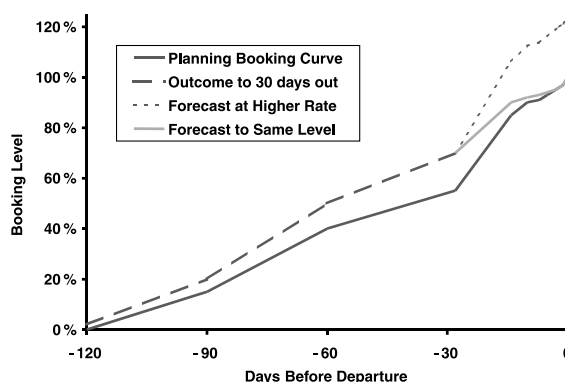


Fig. 2. Booking curves planned, actual, and forecasts.

Gamma is loaded up near zero and has a long positive tail. When μ is above 50, the shape is approximately Normal and the higher-order moments are small.

The Bernoulli-trial motivation is comfortable for spill modeling. An alternative derivation will seem more familiar for revenue management. Revenue management conventionally thinks in terms of “booking curves.” Visually, a booking curve is the plot of total accumulated bookings against time. This is illustrated in Fig. 2. Total bookings at departure are the integral of a time-varying Poisson arrival rate of reservations requests, $\lambda(t)$. The resulting compound Poisson distribution can also be approximated by the Gamma distribution with the mean the expected demand $\mu = \int \lambda(t) dt$ and standard deviation $\sigma = \sqrt{\mu}$. As before, the varying probabilities are usually approximated by their average.

4. Information in booking curves

There is a difference in the booking-curve thinking and the coin-flip model that is significant. The coin-flip model does not suggest that parameters can be reestimated half-way through. The booking rate model makes it intuitive to consider updating the estimate of total arrivals and of resulting variances based on early bookings. Forecasts of $\mu(T)$ for a particular departure at time T use the partial booking levels $\mu(T)$ at earlier times τ : $\mu(\tau) = \int_{-\infty}^{\tau} \lambda(t) dt$. Some have found that the early and late bookings are not correlated. Others have found that errors in estimates of early booking rates are correlated with errors in estimates of late booking rates. Yet others have either found or assumed that errors in estimates of early bookings are negatively correlated with errors in estimates in late bookings.

The case of no correlation between late and early bookings implies that individuals are making independent decisions to travel, each at his own time. This is the simplest model and data supports the hypothesis in many cases.

Positive correlation between early and late errors between bookings and forecast has been observed in some cases. Indeed, positive correlation is implicit in the forecasting structure of several revenue management systems. There are two motivations for such cases. Positive correlation may mean that different people decide at different times to book the trip, but the reason for

travel may be the same, so there is some grouping effect between early and late booking people. A different explanation of positive early-to-late correlations is that there are coordinated errors in the estimates of booking rates. The early-to-late data tested for correlation are the errors in the booked demand compared against expectations for each period. That is, if early bookings are ahead of forecasts, then late bookings will also be ahead of forecasts. What if the forecasts missed the same underlying demand driver in both periods? (“You mean the Olympics are in Australia this year!”) Then the error is in the forecast, not the correlation of people’s behavior. It seems like the same thing, but it is not. If there is a true correlation of early and late bookings, it is part of the irreducible random errors and is caused by people being “grouped.” In this case, the grouping is between different bookings at different times. This grouping increases the unavoidable width of the demand distribution, at least as it is known at the outset. On the other hand, if there are merely correlated forecast errors early and late, as forecasts get better correlations will be reduced and total variance will go down. There is some evidence that this second situation is common.

The third possibility for correlation between early and late bookings is a negative correlation. This implies that demand is best known as a total, and what is not well known is when people will book. Two real-life situations illustrate such a case. One is when a fare promotion moves people to commit travel plans at an earlier moment. This is simply a rearrangement of $\lambda(t)$. Early bookings increase at the expense of later ones. The other is when demand is limited by something like hotel capacity. Once all the hotel rooms in Hawaii are full, there will be no more airplane bookings to Hawaii. So negative correlations in booking curves make sense, and do occur in practice.

5. Bigger random variations—people travel in groups

If absolutely everybody traveled in a group of four, then the estimates of the standard deviations derived above would all be off by $2 = \sqrt{4}$. For group size G , the number of Bernoulli trials becomes N/G , the probability P remains unchanged. The mean remains the same, but the standard deviation goes up:

The expected demand $\mu = G(N/G)P = NP$.

The standard deviation of demand $\sigma \cong \sqrt{GNP} = \sqrt{G\mu}$.

For the arrival rate $\lambda(t)$ model of the booking process, group size may be represented as bookings of more than one seat together. That means a lower arrival rate coupled with a distribution for the number of people represented by each reservation call. The distribution of group sizes is unimportant. What is needed is an explanation for randomness beyond the levels in the ungrouped case.

Data in the early 1980s suggested that demand acted as if average group size were as high as 5 or 6. Contrary to this, the average group size of reservations as they are booked is near 2. A value of 5 implies high correlation in the behavior among total strangers. Fortunately, there are alternate explanations for the early data.² Recent evidence suggests group sizes near 2 are correct

² Distributions for smaller average demands had greater width in monthly data, but these markets were more nearly one unmixed origin–destinations and therefore varying cyclic variances across on-board markets had less chance to cancel out. See later discussion of this point.

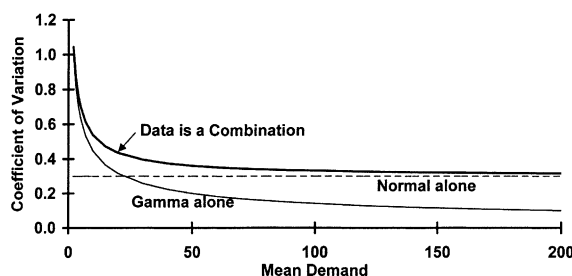


Fig. 3. Coefficient of variation declines to an asymptote.

for both revenue management and spill model cases.³ Evidence on this point is indirect and must wait until after the discussion of the cyclic variations in demand. The coefficient of variation data in Fig. 3 is the evidence for the size of grouping of travel choices. The group size (G) can be deduced by the rate of decline of the data line as average demand increases.

Clearly not all groups are the same size. However, the effect of groupings can be captured by using the (root-mean-squared) average group size as representative. Adding the group-size effect to the model does not change the shape. The shape of the random variances still follows the Gamma outlines.

6. Normal shapes for estimation errors

The model of a coin flip means someone has to estimate the probability for the coin, P . Errors in estimates are conventionally taken as Normally distributed. The central limit theorem says if variations are the result of several not-so-Normal distributions combining, the combined distribution tends to be Normal anyway. For revenue management, the distribution is purely the errors of estimates. Errors are almost always assumed to be Normal. Frequently repeating an assumption does not constitute a proof. Nevertheless, the distribution for revenue management cases must be viewed as Normal until there is proof otherwise.

Negative values for demand cannot exist, and this provides a counter to a strict Normal distribution since Normal distributions imply a chance of negative values. Further, airlines do not schedule flights on days when demand is expected to be quite small. This removes observations just above zero from the distribution. Some have used a truncated normal or a lognormal shape for the demand distribution. Either shape could approximate the distribution of coin-flip probabilities, and there are interesting arguments for both. This discussion will continue to characterize the estimation errors using the complete Normal distribution. The simplification does no violence to the basic issues. Only the shape of the upper half of these distributions matters in applications.

Combining the estimation errors with the coin flip means drawing a coin value (P) from a Normal distribution and then creating a Gamma distribution from the result. This gives a distribution that has no accepted name. Simulations show the shape falls between the two

³ The average group size in premium (first, business, and full-fare class) demands may be smaller than 2. Not surprisingly, families travel discount.

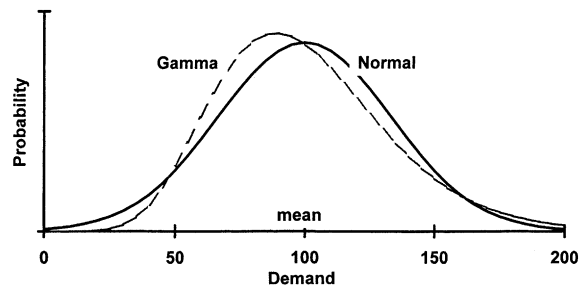


Fig. 4. Normal and Gamma shapes.

shapes—Normal and Gamma (Fig. 4). When the major source of variance is estimation error, the shape is nearly Normal. When the average demand is high enough, the Gamma is nearly Normal anyway, so Normal is the consensus when demands are the size of airplane seat counts. However, when the average demand is small, the width of the Gamma distribution is large and its shape is far from Normal. In these cases the Gamma variance dominates and the shape is predominantly Gamma.

This model suggests that the Gamma shape is right for most revenue management work. Revenue management deals with estimating how much space to protect for the demands for high-fare tickets. These demands have small average values, and for small values the event randomness dominates the shape.

7. Increased normal variations—different coins flipped for different flights

One addition to variation applies only for spill modeling. Airplane choices for schedules require understanding the variation of demand over the many individual flight legs the airplane is assigned. The usual case is a single airplane type or capacity assigned to the same flight every day of a month. In this case, even without random errors or estimation errors, the demand is not expected to be the same each day. For airlines, the day-of-week cycle is the largest source of variation. Demand on Wednesday is less than on Friday, and so on. Adding to this, for a month the first week may be lower than the last week, or the second week could contain a holiday. Altogether, these cyclic variations can be described by a Normal distribution. The more sources for the variation there are, the more Normal the shape. Assignments for a whole day of flying (several flight legs) for a month have more variation than for one leg, and assignments for a season, a year, or a whole fleet of airplanes of the same capacity has more variation still. The variances of these several cycles add, so they are independent distributions. Combinations of several independent sources of variation make the observed results more and more clearly Normally distributed, as the central limit theorem would suggest.

8. Combining demands reduces variation

Demand does not come from a single uniform source, as the model so far would suggest. Modern hubbed airline flights carry the demand from 30 or more different city pairs, and most

airlines would be sensitive to the differences between the business and pleasure component of each pair. Data supports consideration of both these points.

In the early days of spill modeling, observed cyclic variations were greater than those being deduced from recent data. That was because in the early days, much of the demand on a flight was from the single local city-pair market. If each city-pair market has somewhat different time-of-day or day-of-week variations, then some of these variations cancel out as several markets are combined. This is what happened as airlines moved from linear route systems to hubbed route systems. The counter to this is that for markets such as shuttle markets or hub by-pass flights that serve trips over a single city pair, demand distributions are wider than those measured from flights that serve connecting banks in hubs with multiple beyond destinations. This means, for instance, that carriers serving small market niches are likely to have a more difficult time achieving high load factors than carriers with major connected route systems.

The combination or isolation of business and pleasure is still an issue today. A typical hubbed flight leg with an average demand of 100 will have a coefficient of variation of 0.33. However, the loads are a combination of business and pleasure demands that have distinctly different variations through the days of the week and weeks of a month. A pure business market will have variation of perhaps 0.40, and a pure pleasure market's demand will have even more day-of-week peaking. (Values for pleasure demands are difficult to observe because pleasure markets tend to have full flights. Loads on full days are truncated values and not the underlying demand.) Only the combination of the two independent sources of demand bring flight leg values down to 0.33. Purely business or purely pleasure destinations tend to have higher values. This is illustrated in Fig. 5.

There is even evidence of some small negative correlation between the business and pleasure markets. The variation of the combined distributions is less than would be suggested for combining independent demands. This makes sense, since it would often be the same people who travel on both purposes. Work on Monday, pleasure on Saturday.

These cases illustrate that combining different sources of demand requires thinking through whether the sources are independent or motivated by the same cycles. Independent sources of variation tend to smooth out each other's cycles. This is the case for business and pleasure demand. Other sources are motivated by the same cycles and tend to add. For example, feeder flights for hubs may serve many different ultimate markets, but Friday is the peak day for most of them.

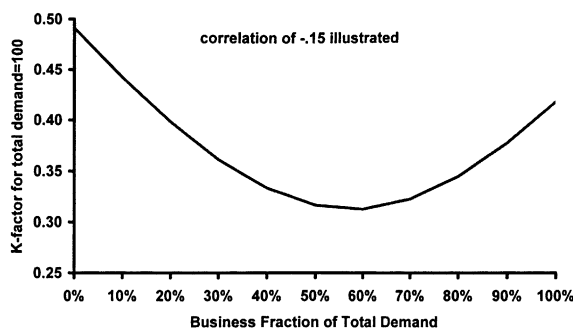


Fig. 5. Business and pleasure variations combined.

9. Evidence for combining Normal and Gamma models

The model of coin flips with coins of varying likelihoods does not allow direct calibration. However the model correctly forecasts the changes in width with average load for monthly flights whose combined distribution can be observed. The characteristic of the Normal part of story is that the coefficient of variation (σ/μ) is constant, no matter what the mean (μ). On the other hand, the Gamma model says the coefficient of variation should decline toward zero as the mean demand gets larger. Load distributions for the 30 days of a flight in a month show lower coefficients of variation for larger average loads, but the values asymptote to a high level. Data draws out information like that illustrated by the line in Fig. 3.⁴

The decline in coefficient of variation as loads grow from 2 to 50 suggests the random effects captured by the Gamma model are important. At the same time the fact that the value asymptotes at a value above zero means the Normal model is appropriate. This presentation of the data supports a model that combines both distributions. The Gamma dominates for smaller values of demand, and the Normal for larger values.

The entire model—a Gamma distribution driven by a draw from a Normal distribution—explains the observed data and the differences in data that occur in different cases. An important consequence is that forecasting can at best remove only part of the uncertainty in the distribution. Much of the variance is unavoidable.

10. Evidence from shapes

Which shape is right depends on where you look. The shape for small loads is different from the shape for larger loads; and the shape for all Tuesdays can be different from the shape for all flights or all months.

When observing demand, booked reservations gives cleaner data on than loads. Bookings are not obscured by delayed departures, nearby canceled flights, or by the randomness of the no-show rate. Bookings are also somewhat closer to demand in the operating sense, because the spilling of demand happens through refusing further reservations rather than from customers queuing at the airport gates and finding insufficient seats available. So when looking for data on the shape of demand distributions, booking data is slightly superior to load data, even though the mean demand value that matters is the mean value of those bookings that actually show up.

Reservation counts are particularly useful when observing demand distributions for premium first- or business-class cabins. Most airlines aggressively overbook coach seats whenever extra space in premium cabins is available. They do this knowing that customers can be upgraded at the gate if overbooking turns out to be too high. Modern pricing adds to these upgrades a host of last-minute upgrades purchased with frequent flyer miles or ready money. None of these passengers seated in first- or business-class represent real premium class demand. Their presence should not be counted in the distribution of demand for premium seats.

⁴ Data and observations mentioned in this paper are the result of proprietary airline information from multiple sources.

Looking exclusively at smaller first-class cabins, the distribution of demand has a shape that clearly favors the lower-order Gamma distribution. In fact, for those unfortunate flights where the average demand is only one or two, the distribution of first-class bookings for a flight for a month is likely to follow the exponential shape of a Gamma with a shape factor of 1.

These observations about first-class cabins apply for the distribution of demand for a particular flight for a month's worth of departures. A different shape might be seen selecting the distribution of bookings in first-class for a group of flight departures formed by searching a reservations system for cases where the forecast demand was very near a particular value, like 4.

On the other hand, bookings for total airplane loads or for loads on a fleet for month or a year are either Normal or of such a high-order Gamma that Normal is close enough.

These various shapes are all predicted by the underlying statistical model. For the first-class bookings of a single flight, the randomness of coin flips is a larger influence than the cyclic changes of the coin's values day-by-day. So a Gamma shape dominates. For a grouping of many days for flights forecast with a demand averaging above 8, the shape is more normal because the error in forecasting the average demand value is large enough to compete with the underlying randomness of the coin flip outcomes. For the shape for larger means, the randomness is a minor contribution to the overall variance, and more Normal cyclic influences dominate.

11. Practical application: Gamma or Normal?

Decisions based on the shape of demand distributions do not hinge on the size or shape of the entire positive tail of the distribution. Both revenue management and spill trade-offs depend on the likelihood of using one more seat. This is the value of the cumulative distribution taken at a point near the total capacity. With that standard in mind, differences across various proposed shapes are not severe. This is particularly the case if the techniques for calibrating the mean and standard deviation are tailored to the shape that will be assumed.

However, there is an ironic reversal in the most common selection of shapes. The Gamma shape has not been proposed for use in revenue management optimization, as far as the author has observed. This is surprising because revenue management depends on the demands for first-class and full-fare seats. These have mean values much smaller than airplane capacities and clearly in the range where the Gamma shape might be relevant. The only academically satisfactory explanation for this is that most of the variance in real applications is forecasting error.

On the other hand, there has been some advocacy for the Gamma for spill analyses in scheduling and fleet planning. This is surprising in that the decisions are based on demands of the magnitude of total airplane capacities. Most decisions involve monthly or even annual samples where the cyclic variations are known to be large. So the Normal shape should be the dominant influence. The academic explanation for this is that researchers have focused data collection on low-demand flights and demand for premium cabins. These cases have the least truncation of demand by capacity and would seem to give the cleanest data. So the cases with the cleanest data have been the cases with the stronger Gamma influences. However, this is not entirely a fair description. The author has seen exceptionally clean data with a distinctly Gamma shape for the demand for a group of flights. The markets in question happened to have demand distributed in that way. There is some suspicion that the case could be characterized as the combination of two

more random (Normal) groupings of markets with different averages, combined. Nonetheless, Gamma shapes do occur in practice, particularly for non-hubbed situations.

12. Summary and conclusions

The process that creates distributions of demand for revenue management or spill can be characterized as four rules: The distribution is driven by a coin-flip or analogous Poisson arrival process. The flip determines the outcome not for one passenger, but for groups averaging perhaps two passengers. The odds for the coin flip are known imperfectly and subject to estimation errors. And the odds for the coin flips for different flights in a group of flights change in cyclical patterns that is a distribution itself.

It takes all four of these rules to explain everything that is known about the means, width, and shape of various demand distributions and combinations of demand distributions. And most known patterns are explainable within the framework of these four rules.

The message here is that one straightforward statistical characterization of the process fits the data and explains the changing characteristics in different cases. More complicated models may not be necessary. Simpler ones are not sufficient.

References

- Li, M.Z.F., Oum, T.H., 2000. Airline passenger spill analysis-beyond the normal demand. *European Journal of Operations Research* 125 (1), 206–217.
- Swan, W.M., 1983. Traffic losses at high load factors. *AGIFORS Symposium Proceedings* 23, 155–178.
- Swan, W.M., 1993. Forecasting for Revenue Management Systems, unpublished working paper.
- Tretheway, M.W., Oum, T.H., 1992. Airline economics: foundations for strategy and policy. *Centre for Transportation Studies, University of British Columbia, Vancouver, Canada*, pp. 32–46 (Chapter 4).