The hidden cost of airline unpunctuality

A. Cook\textsuperscript{a*}, G. Tanner\textsuperscript{a} and A. Lawes\textsuperscript{b}

\textsuperscript{a} Department of Transport Studies, University of Westminster, 35 Marylebone Road, London, NW1 5LS, UK
\textsuperscript{b} Consumerdata Ltd, Dorset House, Regent Park, 297 Kingston Road, Leatherhead, KT22 7PL, UK

Abstract

The ‘soft’ cost of delay, such as loss of market share due to unpunctuality, is a dominant component of airline delay costs, yet it is very poorly understood. The relationships between market share, punctuality and customer satisfaction are examined. Saturation of delay inconvenience and crossovers in Kano satisfaction factors contribute towards the fit of a logit function as a new method for the distribution of passengers’ airline-switching propensity. ‘Soft’ Euro costs are estimated for unpunctuality for seven delay ranges. Future work is summarised, which is required for the achievement of truly comprehensive 4D trajectories in air traffic management.

Keywords: Airline, cost of delay, customer satisfaction, Kano model, punctuality, soft cost.

1. Introduction

Lack of airline passenger satisfaction typically results in reduced market share, even in the current economic climate. Punctuality is a key attribute of satisfaction for many passengers. Unpunctuality may thus cause a reduction in market share, whereby the airline incurs a ‘soft’ cost – i.e. a hidden cost which is not itemised in accounts, but impacts the bottom line.

Although the major cost components associated with airline delay (delayed passenger costs to the airline, marginal crew and maintenance costs, plus fuel burn) may be dominated by ‘soft’ costs (Cramer and Irrgang (2007), Cook et al. (2004)), almost nothing quantitative has been published on this. The primary objective of this paper is to formulate a viable distribution for soft costs as a function of delay. A previously-derived average cost is distributed over delay ranges and correspondingly scaled up to the airline network level. The proposed distribution is derived in the context of an extensive literature review and survey work carried out for this paper.

For airlines, in terms of decision making on when to use accelerated fuel burn for delay recovery, quantifying passenger soft costs is a vital component of the equation – an introduction to ‘dynamic cost indexing’ is furnished by Cook et al. (2009). Understanding these costs is not only important to the airlines, but also informs decision making in air traffic management, for example in the trade-off between strategic costs (such as airspace design) and tactical impacts (such as re-route costs).

Soft costs can only be properly understood through market research. However, punctuality can be a difficult concept to convey to respondents. The situation is not helped by the fact that a number of airline attributes may be conflated by respondents, such as service frequency and flexibility, or higher fares and flexibility – see Proussaloglou and Koppelman (1999). Analogous, is the ‘halo’ effect. This was originally proposed by Thorndike, as a “…tendency to think of the person in general as rather good or rather inferior and to color the judgments of the qualities by this general feeling” (Thorndike, 1920). Rooted in personnel appraisal, this principle may be applied equally to product, service, system or process evaluation, whenever the rating of one or more attributes is partially determined by the ratings of other attributes.

A comprehensive model of airline market share needs to embrace numerous choice factors, not only those of airline preference, but also of modal choice and even airport choice and access mode. Several authors have addressed a number of airline choice attributes, or the perspective of joint airport and airline choice. Few have simultaneously dealt with airline, airport and access mode attributes, an

\* Corresponding author. E-mail address: cookaj@westminster.ac.uk (A. Cook).
exception being Hess and Polak (2006). A primary objective of our paper is to estimate the
distribution of soft costs as a function of airline delay, although the results of such analyses should be
considered within the wider choice context. Since the focus is exclusively on airline costs, passenger
‘value of time’ is not included in the calculations. A separate literature exists on this concept, which
is important in cost-benefit analyses, particularly in the transport sector.

2.1 Modelling customer satisfaction in the airline context

One approach to customer satisfaction modelling, which has gained significant attention, is the Kano
model, Kano et al. (1984). This is summarised in Table 1, which defines a three-tier approach to
customer satisfaction ‘requirements’ (excluding a neutral ‘indifferent’ term in the original paper). We
have used an alternative terminology with ‘factors’ instead, with the original terms proposed by Kano
given in parenthesis under each one. Sauerwein et al. (1996) explain in some detail an application of
this model for assessing product requirements in terms of customer satisfaction. These authors point
out that such ‘requirements’ are likely to vary by customer segment and that merely satisfying the
‘must-be’ and ‘one-dimensional’ requirements will result in a product being perceived as “average”
and therefore “interchangeable”.

Table 1. Kano customer satisfaction requirement levels

<table>
<thead>
<tr>
<th>Factor (requirement) level</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic factor</td>
<td>Taken for granted. If unfulfilled generates dissatisfaction but if fulfilled, does not lead to increased satisfaction. If absent from product/service, customer will have no interest in it.</td>
</tr>
<tr>
<td>(“Must-be requirement”)</td>
<td></td>
</tr>
<tr>
<td>Linear factor</td>
<td>Customer satisfaction is here proportional to the degree of fulfilment. Usually demanded explicitly.</td>
</tr>
<tr>
<td>(“One-dimensional requirement”)</td>
<td></td>
</tr>
<tr>
<td>Premium factor</td>
<td>Large influence on satisfaction: fulfilment leads to more than proportional satisfaction. Neither explicitly expressed nor expected by the customer. If not fulfilled, does not generate dissatisfaction.</td>
</tr>
<tr>
<td>(“Attractive requirement”)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Sauerwein et al. (1996)

Where does airline delay fit into this picture? This is clearly dependent on the traveller’s expectation
as regards delay and may also be a function of other factors, such as journey purpose. Marketing can
also change expectation, even for products/services of which the consumer has no direct experience.

Recognising the importance of journey purpose focuses attention on trip-end constraints and,
subsequently, the importance of arrival delay over departure delay. Furthermore, a delay on the first
leg of a journey of 45 minutes could mean a missed connection and a subsequent arrival delay of four
hours, after waiting for the next onward flight. Passengers (such as these) with high levels of
disutility associated with relatively small delays, will be more averse to, and aware of, delay.

Experience also informs expectation – the more frequently a route is travelled, the more accurate the
passenger’s estimation of delays is likely to be. Recent good or bad experience may sensitise, or even
desensitise, the passenger to delay. The third delay in three consecutive weeks of, say, 30 minutes,
could be the ‘last straw’ for one passenger, whereas another passenger could have already adapted
their plans for the third trip. The Kano model includes such a temporal dimension.
2.2 Exploring response to delay using a Kano model

Wittmer and Laesser (2008) undertook a survey of 2834 passengers at Zürich airport over a one week period in 2006. Respondents were categorised into five groups according to their number of business-related flights per year, to evaluate, using a Kano model, differences in attitudes to delay as a function of the number of such flights taken. If the flight were to be on time instead of having an expected delay\(^1\) of 15 minutes, then for around 25% of travellers this did not create any satisfaction, whereas a greater delay led to dissatisfaction. On-time performance may thus be said to be a basic factor\(^2\) in terms of generated satisfaction for these passengers. This was quite a flat effect across frequency of business-purpose travel, as shown in Figure 1. For around half the passengers travelling least often on business (left-hand side of figure), on-time performance was a linear factor in terms of satisfaction, this proportion rising to two-thirds of those travelling most frequently for this purpose. On-time performance was a premium factor for fewer than 10% of least frequent business travellers and fewer than 1% of most frequent. Responses classified as “indifferent, questionable or contrary” (labelled “indifferent”) ranged from 20% (least frequent travellers) to 7% (most frequent).

![Figure 1. Kano satisfaction ‘factors’ for bettering 15 minutes of expected delay](image)

Indifference and the premium effect both (broadly) fall with frequency of travel. The linear effect increases with frequency of travel, dominating throughout, whilst the basic effect holds at around 25% of respondents. The data for 30 minutes’ expected delay are extremely similar to those for 15 minutes, in both magnitude and trend.

Turning next to an expected delay of 45 minutes, if the flight were to be on time, this did not create any satisfaction, whereas a greater delay led to dissatisfaction, for around 15% of travellers. On-time performance may thus be said to be a basic factor for these passengers. This was again quite a flat effect across frequency of travel, but notably lower than the 25% for 15 minutes and 30 minutes, as would be expected. (Indeed, by 60 minutes, this had fallen again, to fluctuate around a value of only 5%). At 45 minutes, the dominating effects were (equally) the premium factors and ‘indifference’ for

---

1. Note that expected delays were apparently hypothesized to respondents (rather than allowing them to base responses on their actual expectation of delay).

2. This is an example of how the original term “must-be requirement” may sit somewhat uncomfortably in certain contexts, suggesting a ‘requirement’ that expectation be exceeded. “Factor” seems to be a more versatile terminology. Wittmer and Laesser (2008) also use this terminology, referring to the three factors in Table 1 as “basic”, “power” and “enthusiasm”, in that order.
least frequent travellers, with both of these falling as the linear factor increased, as a function of travel frequency. The picture was broadly similar for 60 minutes’ expected delay, except that the premium factor did not fall with travel frequency and the linear factor was much weaker.

This suggested that the data tabulated by Wittmer and Laesser (2008) could be restructured and plotted for each travel-frequency group as a function of the expected delay. This gave a series of graphs which demonstrated a smooth transition in pattern, moving successively from the least frequent (Figure 2) to the most frequent (Figure 3) business-purpose travellers. Since the transition was smooth, the intermediate graphs are not shown.

In each plot (all five; three not shown), for expected delays of 15 and 30 minutes, the linear factor dominated, followed by the basic factor. By 60 minutes, in each plot, these effects had fallen off, to be overtaken by the premium factor and ‘indifference’. The crossover point at which the premium factor became the dominant one, by simple interpolation, was later in each of the successive frequency groups (at approximately 42, 48, 49, 51 and 53 minutes, respectively). Whilst it is not appropriate to over-interpret the crossover points in an absolute sense, since we do not know the actual power of the premium factor (other than it exceeds unity), it is clear that as (business) travel frequency increases, the linear effect persists for longer delays.

The premium effect – deriving greater-than-linear satisfaction from delay reduction without additional dissatisfaction if the delay worsens – generally falls with increasing frequency of travel for expected delays of 15, 30 and 45 minutes. This could be attributable to relatively substantial delay being already factored in to the travel planning of more frequent travellers or more generalised disgruntlement at delay levels.
The strength of the premium factor at 60 minutes’ expected delay for each frequency group, implying extra delay does not create additional dissatisfaction, indicates that a common saturation of disutility might prevail at such higher levels of delay, although in each group ‘indifferent’ responses also increase as a function of the expected delay. It is possible to speculate that these higher delays were seen by some respondents as less likely to occur, or less likely to be recovered to an on-time performance, thus reducing their engagement with the question. Possible support for this in the tabulated data lies in the fact that for each level of expected delay, ‘indifference’ fell as a function of travel frequency.

2.3 Punctuality and market share

Flights which are delayed beyond a passenger’s tolerance limit are likely to shift the perception of that service into the “interchangeable” category identified by Sauerwein et al. (1996). ‘Switching’ behaviour, from one airline to another, or from a flight to a different mode or action, is a determinant of airline market share, and hence profitability. Such tendencies will not only vary by market segment, but by person, by trip, and even within trip (a delayed flight with good customer service recovery could prevent the passenger from choosing a different airline next time). Airline priorities for passenger treatment will be strongly influenced by the yield from that passenger. They may also vary as a function of route and time of day, and, more rarely, time of year.\(^3\)

There is, however, little evidence in the literature on how punctuality drives airline markets. Sultan and Simpson (2000), examining US and European airlines competing on trans-Atlantic routes, show that Americans and Europeans concur on some aspects of service delivery priorities, and differ on others. They observe that “reliability” is the most important aspect of airline service quality on both sides of the Atlantic. Bieger et al. (2008), in another Zürich case study, compared service priorities for passengers flying with ‘traditional’ carriers and low-cost carriers (LCCs), with “punctuality” being ranked similarly by both, although the LCC segment was not unexpectedly price dominated.

---

\(^3\) Abdi and Sharma (2008) comment, in an Emirates case study, that during the Eid holidays or the festive season, Emirates’ priority is to get passengers home as punctually as possible, thus giving a relatively higher weighting to customer value over other cost-based decisions.
In a stated preference survey, Teichert et al. (2008) interviewed 5829 frequent-flyer programme (FFP) members who had travelled on at least one of eleven selected European short-haul routes. For the punctuality attribute, two choice levels of “punctual almost every time” and “delayed for about 30 min many times” were used. In an initial (a priori – by class flown) segmentation, the multinomial logit parameters showed punctuality to dominate both class segments, being almost identical for each, with a higher coefficient only for total fare for business-class passengers. In a more refined (latent class) five-segment solution, punctuality dominated one segment and was second only to fare in two others.

Suzuki et al. (2001) modelled US domestic airline market share as a function of service quality (and other variables). Previous studies, it is asserted, all assume no “sudden change” of gradient for functions modelling passenger demand as a function of service quality. The ‘asymmetric-response’ model developed is based on loss aversion theory, which suggests that consumers react more strongly to performance below a reference point (such as expected service) than to a comparable performance exceeding it. This means that the utility function should be steeper for losses.

Compared to conventional demand models linking service quality and airline market share (linear regression, Cobb-Douglas and linearised logit), only in the authors’ asymmetric-response model were service quality effects significant at the 5% level. The conclusion presented is that if an airline’s service quality falls below the market reference point (median service quality), market share will decrease significantly, whereas a corresponding service increase above the reference point may not increase market share. This clearly echoes Wittmer and Laesser (2008).

Morash and Ozment (1996) investigated the connection between airlines’ “consumer-perceived quality” with each of two groups of recorded performance measures: “external time advantages” and “internal time advantages”. The ‘consumer-perceived quality’ is split into three dimensions: ‘total’, ‘flight’ and ‘baggage handling’ – each wholly derived from complaint rates.

Membership of ‘good’ or ‘mediocre’ benchmark groups for “percent of on-time flight arrivals” was significantly (p ≤ 0.05) differentiated using ‘total quality perception’. Benchmark group membership for other external time advantages, such as service frequency, were more strongly differentiated than ‘on-time’ arrivals, however. Similar results obtained using ‘flight quality perception’. Regarding internal time advantages, it was hypothesised that although most of the network qualities remain hidden from the passenger, if the network appears to be extensive and offers fast and easy access, such ‘value delivery’ to the passenger may in itself increase customer satisfaction.

Dresner and Xu (1995) examined, using two-stage least-squares regression for 13 major US carriers, the two links (“→”) in the relationship:

\[ \text{customer service} \rightarrow \text{customer satisfaction} \rightarrow \text{corporate performance}. \]

Customer service was assessed based on published statistics for ‘on-time performance’ (percentage of arrivals within 15 minutes of the scheduled time), mishandled baggage reports and denied boarding rates. Customer satisfaction was based on total complaints data, i.e. not differentiated by category. Corporate performance was measured as a profitability ratio, thus off-setting differences in accounting practices and scale effects.

An increase in customer service levels will only increase profits if the (indirect) revenue effect through increased satisfaction outweighs any (direct) cost effect associated with actually improving or

---

4 For other analyses of FFPs as choice determinants, see Proussaloglou and Koppelman (1999) and Nako (1992).
5 A clustering method that generates statistically derived segments/classes simultaneously, rather than sequentially through heuristics.
6 These include: on-time reliability, frequency of service, cancellation rates, denied boardings, customer services.
7 Ten network measures, grouped across connectivity, size and density.
8 The best-in-class percentage of 92.8% suggests this had the usual US tolerance window of 15 minutes.
maintaining the service delivery. Ceteris paribus, airlines that achieved higher on-time performance had significantly fewer customer complaints ($p < 0.05$). The greater the number of customer complaints, the lower the airline’s profitability ratio ($p < 0.01$). However, increasing on-time performance significantly (and directly) contributed to reduced profitability ratios ($p < 0.01$), which was “likely due” to higher costs. Further examination of the data for Continental Airlines for 1988 suggested that these increased costs were not off-set by the positive effect of the associated reduction in customer complaints.

2.4 European legislation and passenger complaints

On 17 February 2005, the European Union’s air passenger compensation and assistance scheme (Regulation (EC) No 261/2004) was introduced. In addition to affording passengers with additional rights in cases of flight disruption (denied boarding, cancellation and delay), the Regulation also requires airlines to inform passengers of their rights when a flight is disrupted. This includes giving the contact details of a body designated by the member state to receive complaints. In the UK, this is the Air Transport Users’ Council (AUC).

Steer Davies Gleave (2007a) made the general observation in 2007 that in many member states the number of complaints was increasing. In specific reference to the Spanish Civil Aviation Authority (Dirección General de Aviación Civil, DGAC), it was commented that: “Since the Regulation 261/2004 came into force in February 2005, the number of complaints received by DGAC has increased considerably. According to their figures, there was an increase of 71.14% between 2004 and 2005 […] A further increase of 206.75% is forecast with respect to 2005 […]”.

![Figure 4. Total complaints and enquiries by category, received by AUC](source: Compiled from AUC Annual Report 2007 – 2008)

In 2005-2006, the UK’s AUC received 6094 written complaints, nearly three times as many as the year before. Complaints about cancellations increased six-fold, and those regarding delays five-fold. A general upward complaints trend continued in 2006-2007: over half of these related to the Regulation. Most of the increase was accounted for by telephone enquiries, the scale of which was “most likely” (Air Transport Users Council, 2007) explained by the increase in staff time allocated to the telephone advice line. However, total complaints and enquiries about delays fell because such written complaints decreased markedly: “possibly because there appears now to be less confusion for passengers about their rights under the Regulation following delays” (ibid.). Figure 4 shows the total number of written complaints and telephone enquiries/complaints (combined), as received by the AUC, for the 2006 – 2007 and 2007 – 2008 reporting periods, where it may be seen that overall delay complaints fell again.

---

Such data on delay complaints is not trivial to interpret. It is a function of the capacity of the receiving organisation (such as the AUC) to receive such complaints, the way in which the airlines deal with the complaints themselves, actual delay levels, and passenger acceptance of delay. Based on the AUC data, we may cautiously postulate that there has at least not been a marked increase in delay complaints in more recent years in the UK. Even if the airlines are absorbing increases in such complaints, these are presumably being dealt with to the reasonable satisfaction of passengers, with no marked increase in onward referral.

3 In-house survey results

This section reports on analyses of the UK Customer Care Alliance (CCA) survey and on a dedicated follow-up survey undertaken for this paper. The on-line annual CCA survey, sponsored by Consumerdata (UK) took place July – September 2008. A convenience sample, invitations were e-mailed to a proprietary database of consumers who had previously completed customer satisfaction surveys. The objective was to determine how customers were treated when a problem arose, across all industry sectors. 671 respondents declared that their worst problem was with air travel (see Table 2 for distribution).

Table 2. Travel causing most serious problem

<table>
<thead>
<tr>
<th>Air travel sector causing most serious problem</th>
<th>Number of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional carriers</td>
<td>219</td>
</tr>
<tr>
<td>LCCs</td>
<td>144</td>
</tr>
<tr>
<td>Package/charter carriers</td>
<td>196</td>
</tr>
<tr>
<td><strong>Sub-total</strong></td>
<td><strong>559</strong></td>
</tr>
<tr>
<td>Other (e.g. airport) / missing</td>
<td>112</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>671</strong></td>
</tr>
</tbody>
</table>

Respondents were asked how long prior to the interview they had experienced the problem. Removing the ‘unsure’ and missing categories, and plotting these against which statements best described “how upset” the respondents were as a result of the problem (from 1 to 5: ‘not at all’ to ‘extremely’), furnishes Figure 5. Comparing problems which occurred more than six months ago, to those which were more recent, significantly higher recalled ‘distress’ (Mann-Whitney U test, p < 0.001) is demonstrated for the problems which occurred longer ago. (This is revisited later in this section, in the context of frequency of travel). The degree of ‘distress’ declared across the three airline segments of Table 2 was almost identical, however.

The relationship between how much “out of pocket” the respondent was, and the extent to which they declared themselves to be upset, was almost flat. Further analysis suggested that this was probably not because the loss was subsequently recovered or compensated for through a claim.
Consider the following key variables:

1. satisfaction with the way the problem was resolved
2. satisfaction with the airline finally, overall
3. likelihood of recommending the airline to others
4. likelihood of flying with the same airline again

Analysing these for correlations, problem resolution and overall satisfaction – (1) and (2) – had only a medium correlation \((r_s = 0.66, p < 0.001)\), whilst (3) and (4), often used as mutual proxies – were indeed strongly correlated \((r_s = 0.83, p < 0.001)\). Notwithstanding these collinearities, both likelihood of recommendation and likelihood of flying again with the same airline were fairly weakly correlated with problem resolution (respectively: \(r_s = 0.57, p < 0.001; r_s = 0.49, p < 0.001\)) and rather strongly correlated with overall satisfaction (respectively: \(r_s = 0.83, p < 0.001; r_s = 0.72, p < 0.001\)).

Cook (1997), analysing post-flight questionnaire data of over 4000 passengers on the North Atlantic, also found ‘overall’ score ratings to be better predictors of likelihood of airline recommendation, than any individual service attribute.

---

10 All such results quoted are Spearman rank correlation coefficients and use two-tailed tests.
11 It is anticipated that data reduction techniques, such as principal components analysis, will serve to extract more robust relationships in future analyses.
These relationships are represented in Figure 6. Examining these further, some more subtle (and statistically significant) effects were found, suggesting that traditional carriers are more expected to resolve problems, i.e. overall satisfaction with other carriers is not as contingent on problem resolution. Problem resolution plus elapsed time, may synergistically result in repeat business. Plotting likelihood of flying with the same airline again, against how long before the interview the problem was encountered, resulted in almost flat lines for each category of final, overall satisfaction.

To explore such findings further, a sub-sample of those invited to complete the CCA survey, known to have flown at least once in the past few years, was e-mailed an invitation to participate in a dedicated, on-line flight delay survey, in October 2008. Questions were asked based on the most recent delay recalled on a direct flight, regardless of length, since January 2007. Inviting respondents to complete a survey about flight delays clearly introduced a ‘bias’ towards longer delays, the average delay at the destination airport being 5 hours 40 minutes, across the 507 respondents\textsuperscript{12}, with a range from 10 minutes, to 48 hours 30 minutes. The purpose of the survey was to gain insights into the disutility of airline delay: having delays which were representative of actual delays encountered in Europe (peaking below 15 minutes; see Table 4), would not have been useful. Approximately 20% of the surveyed cases related to delays of one hour or less. Dominant trip purposes over the preceding twelve months were 8% business, 88% leisure/non-work, 4% equally frequent for these two.

<table>
<thead>
<tr>
<th>Table 3. Effects of delay on subsequent travel behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct effect of delay since delay event</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Had not affected travel plans</td>
</tr>
<tr>
<td>Had not travelled since anyway</td>
</tr>
<tr>
<td>Made some flights with another airline</td>
</tr>
<tr>
<td>Made some journeys by train/other</td>
</tr>
<tr>
<td>Made fewer journeys</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Missing response</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Table 3 shows the responses to the question asking whether the experience of the delay had directly led to any changes in travel behaviour since the delay event. For the majority of respondents (82%), this had either had no effect on their plans or they had not travelled since anyway. Few had substituted mode or made fewer journeys. Of the ‘other’ category, most free-text responses indicated that the respondent would not use the same airline or airport again. Since this sample was strongly dominated by the leisure market, it is likely that more representative trip frequencies would be higher on many European routes, for both traditional carriers and LCCs, such that the ‘not travelled since anyway’ responses would be fewer.

\textsuperscript{12} After aggressive data cleaning, 14% of cases were dropped.
11

Figure 7. Inconvenience as a function of delay experienced

Figure 7 is a plot of the delay minute deciles (average of 46 respondents per group) experienced, against the stated inconvenience of the delay on a 5-point scale, with standard errors of each group mean. The first decile is for 10-50 minutes, each point being plotted for the range mid-point. Although only 9 respondents were in the 10-20 minutes range, such that the low end of the curve cannot be further detailed, the logarithmic fit shown ($r^2 = 0.86$) clearly indicates a rapidly rising disutility which flattens out at higher delay.

As with the CCA data, the relationship between how much out of pocket the respondent was, and the degree to which they declared themselves to be inconvenienced, was almost flat. In a further striking similarity to the CCA data, further analysis suggested that this was probably not because the loss was subsequently recovered or compensated for through a claim.

We next turn our attention to a comparison of respondents who had made some flights with another airline as a result of the delay (‘switchers’, 41), compared with those whose subsequent plans had not been affected by the delay but had travelled since (‘non-switchers’, 298). In this sample, the lowest delay for a respondent who consequently switched to another airline was 60 minutes. The 41 switchers had, at the time of interview, made an average of 3.7 flights with other airlines, at an average rate of 0.6 flights per month. For the switchers, the mean number of months the delay had been experienced prior to the interview was 10.0, whereas for the non-switchers it was 6.6 (Mann-Whitney U test, $p < 0.001$). Any effects of delay events changing in importance as a function of time were probably outweighed by frequency of travel: for delays experienced longer ago, there would be a greater number of subsequent opportunities for travel.

The ratio of switchers to non-switchers did not change as a function of compensation received (in cash or kind), but increased reasonably uniformly with each delay decile, maximising at around 1 in 3. This value seems, prima facie, rather low, especially in view of the rate at which inconvenience saturated (Figure 7). Indeed, Cramer and Irrgang (2007) suggest that after experiencing a five-hour delay, all passengers will switch. The lower value from our survey may be explained by the strongly dominating leisure-purpose passengers, quite probably in combination with fare and/or schedule constraints. On certain routes, e.g. operated by just one carrier, it might not actually be possible for the passenger to change operator on their next journey.

Therefore, rather than quantifying direct switching ‘probabilities’ per se, it is more appropriate as a first step to express this as a switching ‘propensity’ (a concept to be developed in the final section of this paper). As inconvenience and dissatisfaction maximise, so too will switching propensity. In the next section, saturation of inconvenience and crossovers in Kano satisfaction factors contribute towards the fit of a logit function as a distribution of switching ‘propensity’.
4.1 Distributing the soft cost as a function of delay duration

A form of the logit function, such as Equation 1, may be used to express the propensity, ‘$\Pi$’, of a passenger switching from a given airline, to some other choice, after a trip with a delay experience of duration time, ‘$t$’. The familiar ‘S’-curve produced (the black curve in Figure 8) has the desirable characteristics of maintaining a low switching propensity for some time, then rapidly increasing through a zone of ‘intolerance’, before levelling off after a duration of delay, beyond which, the passenger is already very likely to switch airlines (for example) for the next trip. Whilst other theoretical expressions may be employed, the form of Equation 1 as presented allows close manipulation of the distribution.

Equation 1

$$\Pi = \frac{1}{1 + e^{a-bt^c} - k}$$

If the plot in Figure 8 were of disutility per se, it is arguable that the curve would have a more complex shape on the right, for example oscillating as different onward connections were made and lost, or the delay involved an overnight stay at an airport, etc. However, by adjusting the constants (‘$a$’, ‘$b$’ and ‘$c$’) it is possible to produce a fit which accords in a semi-quantitative manner with several key findings of this paper.

![Figure 8. Hypothesised switching propensity by delay duration](image)

The grey markers are the inconvenience scores from Figure 7, normalised to the eighth decile (midpoint: 5½ hours’ delay), which was, it should be noted, 0.94 of the value of the tenth decile (midpoint: approximately 33 hours’ delay). This re-scaling adopts the position that the likelihood of switching to another choice after any delay exceeding five hours is the same. Cramer and Irrgang (2007), using a 1988 study from American Airlines, also suggest saturation at five hours\(^{13}\).

As discussed in Section 3, it seems likely that a more representative saturation would actually occur at lower delay than the Figure 7 markers: the ‘S’-curve runs higher than a log-fit (grey curve) through these markers. The general form of the ‘S’-curve has also been adjusted to fall off far more sharply

\(^{13}\) A simple (exponential) cost-versus-delay curve is presented, flatter at lower delay than the curve in Figure 8.
towards the origin than the log-fit of the grey curve suggests, as the latter would result in a high switching rate at 0 minutes of delay. A small constant, k, has been used to offset the residual \((1+e^a)^{-1}\), forcing the ‘S’-curve through the origin and to (naturally) saturate at \(p = 1 – k\). The curve in Figure 8 beyond 90 minutes is almost flat, and here unit propensity has been assigned. Arguments for \(k = 0\) could also be made.

Although the Wittmer and Laesser (2008) data were dominated by high-frequency business-purpose passengers, many of the distinctions between business- and leisure-purpose passengers have been eroded over recent years, particularly during the current economic downturn. This is, to some extent, supported by similarities across the business-purpose frequency groups in the data. (i) The Kano basic factor was quite flat across frequency of travel. (ii) Figures 2 and 3 suggested crossover points, where the Kano premium factor became dominant, which spanned only an 11 minute interval between the highest- and lowest-frequency groups. This occurred in the range of around 40 – 50 minutes. The dashed vertical lines in Figure 8 show this delay range and the ‘S’-curve passes the 50% point in this region. (iii) By 60 minutes, for each frequency group, the premium factor dominated, indicating common saturation of disutility at such higher levels of delay. The ‘S’-curve reaches 70% at just over 60 minutes – where it nearby intersects the grey inconvenience curve.

In their analysis of schedule delay\(^{14}\) and total carrier share, Koppelman et al. (2008) found that adding an ‘S’-shaped schedule delay penalty improved the overall goodness of fit with empirical data and suggested this makes sense behaviourally. Suzuki et al. (2001) adapt a simple binary logit equation to form a function similar to Equation 1, but with a detailed treatment for dealing with the asymmetries of loss aversion, as discussed earlier. Such forms may be developed for soft cost calculations in the future, allowing for differential segment sensitivities to delay, for example by carrier type, length of haul, journey purpose and frequency of travel. Here, however, a distribution of the soft cost as a function of delay duration will be made using the current form of Equation 1.

Using Airclaims and Association of European Airlines data to harmonise findings from two extensive European airline case studies, Cook et al. (2004) derive a soft cost of passenger delay of €0.18 per average passenger, per average delay minute, per average delayed flight, for 2003. In the current paper, the earlier approach (ibid.) is refined in four significant ways: (i) seven, instead of two, length of delay categories are used; (ii) total delay estimates are used to weight the values (not simply flow-management delay minutes); (iii) more realistic, non-linear switching propensities are estimated; and (iv) reactionary multipliers, which increase as a function of primary delay, are used to estimate network effects.

The starting premise is that the 2003 average value of the soft cost (€0.18/pax-min) has not increased. (This is in contrast to the airline hard costs of passenger delay, which generally have increased). European airline markets have become increasingly price driven, with many ‘traditional’ airlines no longer providing free catering on shorter hauls, and LCCs (have continued to) enjoy a considerable share of the business-purpose market. Increased distribution through the internet has also helped to keep fares down and competition up. Passengers may even be prepared to pay a higher price relative to a similar service offering from a competing airline that generates less satisfaction, or is perceived as likely to do so – see Proussaloglou and Koppelman (1999). The foregoing discussion of the AUC complaints data also supports the view that there has been no recent marked increase in delay sensitivity, despite a worsening of actual delay experienced.

The calculation of the passenger soft costs proceeds as follows (see Table 4). For each primary delay range (row 1), the proportion of such flights is shown in row 3. It is important to note that values in other rows in Table 4 correspond to the mid-points of row 1 (apart from the ‘90+’ column).

---

14 This is a term used to describe the difference between a desired departure (or arrival) time and the closest option (realistically) available to the passenger. In the airline context, it thus reflects the inconvenience (or disutility) of flying at a non-preferred time, usually focusing on departure times and with reference to a ‘convenience or preferred window’ for the passenger. Such models have not been extensively reported in the literature.
The flight proportions are combined with the $\Pi$(switch) values taken from Equation 1\textsuperscript{15} to produce a joint-weighting factor (row 4). This factor has two properties. Firstly, the sum of the products of (i) the Euro costs (without network effect; row 5) and (ii) the proportion of flights (row 3), gives the original average value of €0.18/pax-min. Secondly, the ratios of the weighting factors between delay categories are the same as the ratios of the corresponding $\Pi$(switch) values. The joint-weighting factors were thus obtained by solving the seven simultaneous equations required for the delay proportions, subject to the $\Pi$(switch) constraints.

Notwithstanding the nomenclature ‘soft’, the resulting primary costs (row 5) represent significant bottom-line impacts for airlines. Since these are per-passenger, per-minute costs, they accumulate quickly and may even dominate the other costs associated with airline delay. Furthermore, the primary costs of row 5 still need to be scaled up to the network level. This is addressed in the next section.

Table 4. Calculation of per-passenger soft costs of delay, summarised by ranges of primary delay

<table>
<thead>
<tr>
<th>Primary delay range (mins)</th>
<th>1-15</th>
<th>16-30</th>
<th>31-45</th>
<th>46-60</th>
<th>61-75</th>
<th>76-90</th>
<th>90+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switching propensity\textsuperscript{a}</td>
<td>0.044</td>
<td>0.170</td>
<td>0.364</td>
<td>0.581</td>
<td>0.753</td>
<td>0.856</td>
<td>1.000</td>
</tr>
<tr>
<td>Proportion of flights\textsuperscript{b,c}</td>
<td>0.61</td>
<td>0.19</td>
<td>0.08</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Joint-weighting factor</td>
<td>0.24</td>
<td>0.95</td>
<td>2.04</td>
<td>3.25</td>
<td>4.21</td>
<td>4.78</td>
<td>5.31</td>
</tr>
<tr>
<td>Cost (without network effect; €/min)</td>
<td>0.04</td>
<td>0.17</td>
<td>0.37</td>
<td>0.58</td>
<td>0.76</td>
<td>0.86</td>
<td>0.96</td>
</tr>
<tr>
<td>Reactionary multiplier</td>
<td>1.48</td>
<td>1.74</td>
<td>2.00</td>
<td>2.25</td>
<td>2.51</td>
<td>2.77</td>
<td>3.14 \textsuperscript{d}</td>
</tr>
<tr>
<td>Total cost (with network effect; €/min \textsuperscript{e})</td>
<td>0.05</td>
<td>0.22</td>
<td>0.56</td>
<td>1.13</td>
<td>1.76</td>
<td>2.13</td>
<td>2.80</td>
</tr>
</tbody>
</table>

\(\textsuperscript{a} \Pi\text{(switch)}, \text{see Equation 1 / Figure 8.}\)
\(\textsuperscript{b} \text{Source: EUROCONTROL (2007) - CODA STATFOR series.}\)
\(\textsuperscript{c} \text{Source: EUROCONTROL (2008).}\)
\(\textsuperscript{d} \text{For 91-119 mins. Full tables available on request.}\)
\(\textsuperscript{e} \text{Averaged back over the primary delay; based on a B738.}\)

4.2 Scaling up the soft cost to the network level

The value of the basic reactionary delay multiplier for the latest year available is 1.8, EUROCONTROL (2008). This means that for each minute of primary delay, there are another 0.8 minutes of reactionary delay in the network, on average. Rather than using a common value (of 1.8) in order to obtain total network costs, disaggregate data were used from an American Airlines’ schedule model. Beatty et al. (1998) studied such delay propagation by building delay trees with schedule buffers included in the delay-tree scenarios. After sampling from the distributions modelled, linear fits were produced relating length of primary delay and corresponding reactionary delay multipliers. In the absence of European data for such a calculation, this fit was re-scaled to produce the average value of 1.8.

Whilst simple in concept, the way in which reactionary multipliers work in practice is less straightforward. Take for example a B738 with a 20 minute delay. This primary delay has an associated soft cost of €351 (€0.145/pax-min, interpolated from row 5; typical loading of 121 passengers assumed).

For 20 minutes, the reactionary multiplier is 1.69 (interpolated from row 6), i.e. it causes approximately 0.69 x 20 = 14 minutes of extra delay. In our European network model, 10 of these additional minutes are assigned to a subsequent rotation of the same B738, and 4 minutes to a rotation

\[a = 3.00, b = 0.10, c = 0.90, k = 0.05. \quad \Pi: 1.00 \text{ for } t > 90.\]
of a European ‘network-average’ aircraft. This produces\textsuperscript{16} a total soft cost of €434. In practice, these single impacts of 10 and 4 minutes might be distributed over a number of aircraft rotations, thus reducing their cost, but there are currently no data available to model this distribution. Since per-minute costs increase with length of delay, this means that higher primary delays in this model may be associated with overestimates of the network costs (assigning, for example, one rotational delay of 60 minutes, instead of two of 30 minutes). This particularly applies to the steeper part of the curve in Figure 8. On the other hand, for delays above 90 minutes, these primary costs are assumed to have saturated, whereas in practice this may not be the case, especially where additional step increases (e.g. due to cancellations) are considered. To quantify the extent to which effects such as these off-set each other, better network models would be required.

From this discussion, it is apparent that the total cost per minute (row 7) cannot be obtained by simply multiplying rows 5 and 6 (although the result would often be fairly similar). The values given in row 7 have been calculated for a B738 as the primary aircraft, with a European ‘network-average’ aircraft for non-rotational delays in the network. The grand total cost for each column was then averaged back over the primary delay minutes (mid-point of row 1) and passengers (121), to give per-minute, per-passenger costs, now with the full network effect included (row 7).

Only row 7 is aircraft-specific. For the B738 example, these final values lag behind the simple product of rows 5 and 6, although only by €0.2/pax-min by the final column (‘90+’). Delays with other primary aircraft display similar effects. For the higher delays with the smallest aircraft modelled, the (‘aircraft-weighted’) values of row 7 actually exceed the simple product of rows 5 and 6, as would be expected.

5 Conclusions and future research

This contribution to the field of delay cost management supports a rational approach towards soft costs, whereby they systematically increase as a function of delay duration. In this first step, a number of empirical findings and estimated effects have been consolidated. A valuable, but limited literature in this field has been reviewed. There remains an important opportunity to develop such modelling using primary, airline passenger satisfaction measurements with comprehensive attribute sets, since most existing work has used either service quality statistics or complaints data. The latter are actually using a proxy for dissatisfaction, which cannot be assumed to be a linear (negative) extrapolation of satisfaction, as has been discussed.

The definitions used to specify ‘punctuality’ from recorded data, or for presentation to respondents, will clearly affect the fit of the model in question. Furthermore, whilst a delay of 15 minutes relative to schedule may be recorded as ‘on time’ in the statistics, a connecting passenger with checked-through bags who misses a connecting flight with a tight connection time, will doubtless have a different view. Presenting ‘punctuality’ concepts to the interviewee is a problematic issue which requires considerable attention, as discussed by Bates et al. (2001).

In this paper, in consideration of potential fare and schedule constraints, switching ‘propensities’ have been proposed instead of ‘probabilities’ per se, with such propensities retaining a distribution between 0 and 1 (or 0 and 1 – k, etc). As a next step, to allow for different market conditions, it is possible to vary the actual distributed average of €0.18/pax-min. For example, on a LCC route, it might be argued that the average soft cost is far lower than €0.18/pax-min. Whilst we are currently preparing a further paper on such alternative average costs, we have not yet considered specific forms of Equation 1 for different markets. We consider that further market research should first be undertaken, to better understand the appropriate parameterisations, before this would be justified.

\textsuperscript{16} Calculations not shown due to restrictions on space. Full calculations in working papers are available at: www.eurocontrol.int/ee/c/public/standard_page/proj_CARE_INO_III_Dynamic_Cost.html
We propose a joint RP-SP model as one way forward, development of which is underway. Such behavioural models need to quantify not only the triggers for passenger defection, but how long it lasts and what the intervening behaviour may be (e.g., travel substitution) before the passenger returns to the original airline, if at all. They also need to take into account the collinearities of service attributes, notably frequency and punctuality. Other frequency effects remarked upon in sections 2 and 3, but not yet incorporated into the model presented in section 4, should also be integrated.

Returning to the Kano logic, Morash and Ozment (ibid.) suggest that some features of service delivery, including on-time reliability, may be ‘minimum customer expectations’, which will not ‘delight’ customers if delivered. This may be compounded by the fact, as also remarked upon by Wittmer and Laesser (ibid.), that airlines perceived as being punctual may be hostage to their own reputation: delays of less-than-expected levels may only prevent dissatisfaction, without generating increased satisfaction per se. However, expectations change, and as delays generally worsen, the threshold for satisfaction may be lowered, particularly in times of economic downturn, where price dominates more strongly.

Finally, at the strategic level of managing unpunctuality, it is vital to take into account the principle of which we are reminded by Dresner and Xu (ibid.): investment in customer service levels will only increase profits if the revenue effect outweighs such investment. Importantly for air traffic management and airlines alike, this principle also applies tactically for aircraft delay recovery.

As the gap between demand and capacity widens, improved network models will be required to make better use of the airspace. Tail-specific, reactionary models of the current system, with connecting passenger (and even crew) data, would be a significant step forward. Even then, the soft (and hard) passenger costs to the airlines of unpunctuality need to be better quantified and managed dynamically, before the ambitions of truly comprehensive 4D trajectories may be achieved.

Acknowledgement

The authors gratefully acknowledge the generous support of the European Organisation for the Safety of Air Navigation (EUROCONTROL) for enabling this research to be carried out. This work has been cofinanced by EUROCONTROL under its Research Grant scheme. The content of the work does not necessarily reflect the official position of EUROCONTROL on the matter. © 2009, EUROCONTROL and the named authors. All Rights reserved.

References


Nako, S. M. (1992). Frequent flyer programs and business travellers: an empirical investigation, Logistics and Transportation Review (28), No. 4


