

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/300299980>

Airport Operational Performance and Its Impact on Airline Cost

Chapter · July 2013

DOI: 10.1002/9781118535844.ch5

CITATIONS

10

READS

5,360

2 authors:



Mark Hansen

University of California, Berkeley

141 PUBLICATIONS 3,649 CITATIONS

SEE PROFILE



Bo Zou

University of Illinois at Chicago

60 PUBLICATIONS 1,672 CITATIONS

SEE PROFILE

Airport Operational Performance and its Impact on Airline Cost

Mark Hansen¹, Bo Zou²

¹ Department of Civil and Environmental Engineering, University of California at Berkeley

² Department of Civil and Materials Engineering, University of Illinois at Chicago

1 Introduction

Delays, many of them the result of limited or reduced airport capacity, are a fact of life in air transport. In 2007, nearly one in four US airline flights arrived at its destination over 15 minutes late (BTS, 2009). About a third of these late arrivals were a direct result of the inability of the aviation system to handle the traffic demands that were placed upon it, while another third resulted from airline internal problems. Most of the remainder was caused by an aircraft arriving late and thus having to depart late on its next flight (BTS, 2009). Between 2002 and 2007, as the air transport system recovered from the 9/11 attacks, scheduled airline flights increased about 22 per cent, but the number of late-arriving flights more than doubled. Since 2007, traffic and delays have declined somewhat because of the recession, but the FAA expects growth to resume, with air carrier flight traffic reaching 2007 levels by 2013, and growing an additional 32 per cent by 2025 (FAA, 2011). It is widely recognized that delay increases nonlinearly as demand approaches the capacity in the system (Figure 1-1).

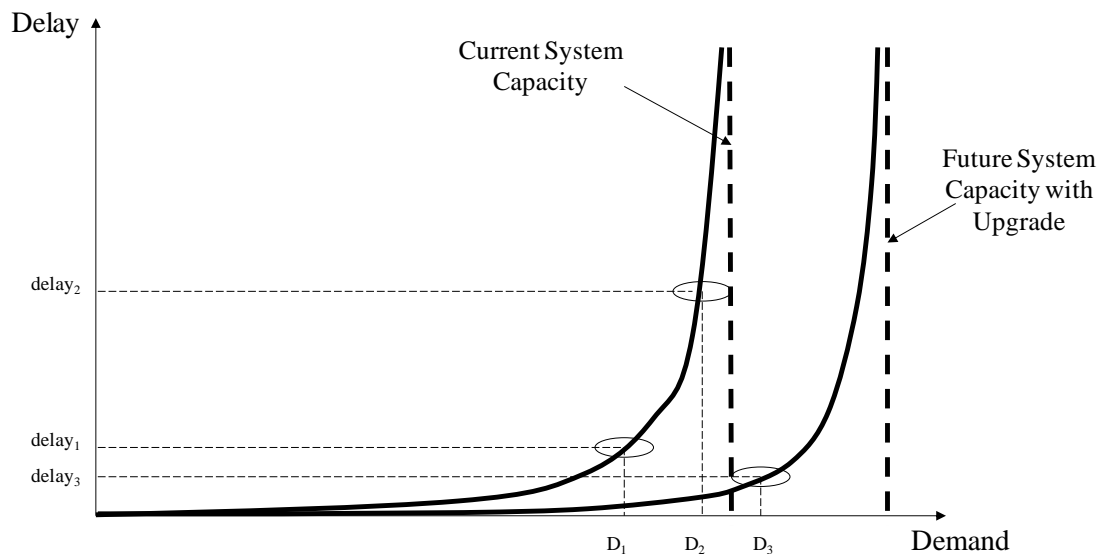


Figure 1-1: Illustration of the relationship between delay, demand and system capacity

Substantial investments are required in order to modernize and expand our aviation infrastructure so that it can accommodate anticipated growth without large increases in delay. In the US, the Next Generation Air Transportation System (NextGen) will deploy improved systems for

communications, surveillance, navigation, and air traffic management and also require flight operators to invest in new on-board equipment. Substantial improvements in air transportation capacity also require airport infrastructure enhancement. Estimates of these combined investments reach well into the tens of billions of dollars (GAO, 2008; ACI, 2009).

Much of the business case for these large expenditures rests on the value of reducing delay and its associated costs. The cost of delay includes many elements. Passengers are inconvenienced when their flight arrives late, or especially if they are forced to arrive on a different flight as a result of a cancellation or missed connection. Delays thus degrade the quality of air travel products, diminishing passenger willingness to pay for them and discouraging some from flying altogether. Adaptations to avoid or mitigate delay—such as leaving early for a business meeting to make sure it is not missed, or scheduling flights at less than ideal times to avoid congestion—also entail costs.

The delay cost element that receives the most attention, however, is that incurred directly by airlines through increased capital and operating expenses. This is the core element of most business cases for airport infrastructure investment to improve airport operational performance. Many in the community perceive that, since airlines must pay these costs with “real money”, they merit stronger consideration than the costs of passenger time loss. By the same token, there is a stronger basis for quantifying airline delay costs, and as a result a larger literature on methodologies for doing so.

This chapter reviews current knowledge and thinking about the costs of delay and related phenomena to airlines. We review—with a somewhat critical eye—various methods for determining the cost of delay for airlines. We also present results from the application of these methods. We focus on the US and Europe, where the flight delay problem gets the most attention and where the value of delay reduction is the most critical input to the benefit-cost analysis for airport and other aviation capital investments.

Section 2 reviews methods for measuring delay or—more broadly—operational performance in air transport. Section 3 discusses valuation methodologies. Section 4 identifies some unresolved issues regarding both measurement and monetization, while Section 5 summarizes and concludes this chapter.

2 Quantifying Operational Performance

2.1 Arrival delay against schedule and schedule buffer

To quantify delay cost impact we must first quantify delay. This is not as simple as it may seem. To highlight the issues involved, we start with a stylized view of how airlines would operate in the absence of congestion and delays. An airline might start the process of scheduling a flight by determining an *ideal departure time (IDT)*. The ideal departure time would take into account not only preferred passenger travel times, but also internal airline constraints, such as those necessary to create efficient crew schedules and fleet plans. As part of this process, the airline would choose the most appropriate aircraft type from its fleet for the flight. Based on the characteristics of that aircraft and assuming it could fly the optimal, unimpeded origin-to-destination trajectory, an *ideal arrival time (IAT)* could be computed as illustrated in Figure 2-1. The interval between the ideal departure and arrival times, called *unimpeded flight time*, is a key quantity and will be discussed throughout this chapter.

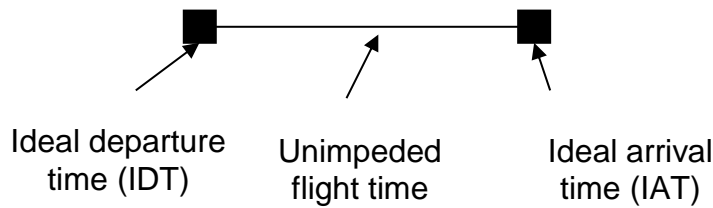


Figure 2-1: Unimpeded flight time

Now let us consider how congestion and delays alter this situation. As illustrated in Figure 2-2, the airlines will typically increase scheduled flight times over unimpeded ones in order to account for delays resulting from flight restrictions imposed to organize traffic, congestion, and a variety of other factors. We call this added time the *schedule buffer*, which is the difference between the *scheduled arrival time (SAT)* and *IAT*. Schedule buffers allow airlines to absorb some delays while preserving flight punctuality and predictability, and thus making the schedule more robust to flight time variations. Adding buffer also helps airlines keep good on-time performance, which has become an increasingly critical marketing tool. In principle, once the unimpeded flight time is determined, the schedule buffer can be obtained by comparing this with the flight schedule. This, of course, begs the question of how to define the unimpeded flight time, which will be discussed later.

The amount of buffer added depends in part upon the statistical knowledge of past delays (Cook et al., 2004), and is also subject to several constraints. First, as buffer is now part of the schedule, it entails costs, for example by increasing crew pay hours and the travel time advertised to potential customers. A large buffer also decreases aircraft utilization and thereby increases capital cost per block hour. In short, while buffer time is valuable, particularly in a delay-prone system, it is also expensive. Setting the right amount of buffer requires trading off these costs with the potential cost savings from the buffer when delay occurs.

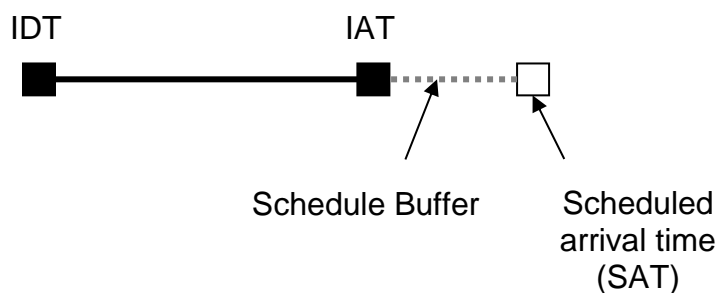


Figure 2-2: Schedule buffer

Of course, even after buffer is added, flight delays can still occur. Arrival delay, the type of delay most commonly considered, measures flight lateness against the *SAT* (Figure 2-3). Schedule buffer and arrival delay both represent excess travel times that would not exist in an operationally perfect air transportation system. While buffer is predetermined for a particular flight, arrival delay varies randomly from day to day and flight to flight, its magnitude shifted upward or downward according

to the amount of buffer. Arrival delay is negative when the buffer exceeds the amount of time between the actual and ideal flight arrival times. Early arrivals are widely believed to have little cost impact since they save little in the way of resources as compared to an on-time flight. Thus, arrival delay truncated at zero, or positive arrival delay, is often used for the delay metric.

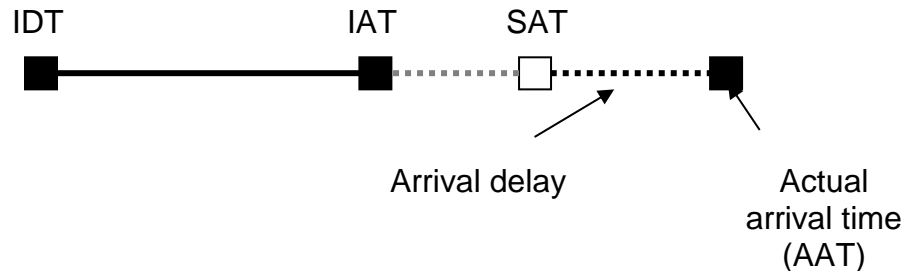


Figure 2-3: Flight delay against schedule

2.2 Alternative metrics

The above description only provides the most conventional way of characterizing flights' operational performance. Various alternative metrics exist, looking at flight activities relative to schedule from different perspectives. For example, one can focus on the departure end, and develop a set of corresponding delay metrics. Statistical evidence shows strong correlation between departure and arrival delays (Hansen et al., 2001; Zou and Hansen, 2011a). However, it is more convenient to consider arrival delay due to its direct connection with schedule buffer. Mayer and Sinai (2003a) used excess travel time relative to the minimum feasible travel time observed on each flight segment as a single measure of flight operational performance. Essentially, this measure equals the sum of schedule buffer and arrival delay, with the minimum feasible travel time defined as the unimpeded flight time. As we shall see in the following section, since the definition of unimpeded flight time is not unique, neither is the excess travel time.

Another potential metric is delay variance. It has been argued that schedule disruption, an important delay impact, results mainly from delay variability rather than delay *per se*. At a hub airport, for instance, half of the flights arriving late by 20 minutes whereas the other half being on time would be more disruptive than all flights being delayed by 20 minutes. Hansen et al. (2001) used the Principal Component Analysis and found that, at the airline-quarter level, the variance of delay and the average delay have different patterns of variation. They identified the “variability” factor that has much higher impact on the variance of delay than on the other delay metrics. Nonetheless, variance of delay has not been widely considered so far in delay cost analysis, in part because it is not clear how to assign it a cost factor.

It is also common to summarize delays in terms of the fraction of flights whose delay exceeds, or does not exceed, a certain value. We refer to these as on-time metrics. Airlines have developed a special nomenclature for this. The A-14 metric, for example, is the fraction of flights that arrive more than 14 minutes late. FAA and DOT have used A-14—or, more precisely, its complement—as a key metric for airline on-time performance. Airlines attach greater significance to D-14, the fraction of flights that depart more than 14 minutes late, because this is a metric over which their own personnel, as opposed to the air navigation service provider, have more control. Other measures of this type that airlines follow include A/D-0, 30, 60, and 120. By tracking these different

metrics, airlines develop a more complete picture of operational performance, and in particular the incidence of long, disruptive delays, than the average provides. As with delay variance, however, little has been done to relate on-time performance to airline costs.

In sum, while the alternative operational performance metrics discussed above have their uses, arrival delay and schedule buffer are the most widely used in delay cost analysis, because of their simplicity and mutual complementarity. For this reason the subsequent discussion of delay cost impact is based primarily upon these two metrics.

3 Estimating the Cost Impact of Imperfect Operational Performance

We consider two approaches to estimating the impact of imperfect operational performance on airline costs. The first, which we term the *cost factor approach*, is based on assigning unit costs to different categories of delay based on estimates of the resources consumed when a given category of delay occurs. The second, which we will call *the total cost approach*, is built on firm or industry level relationships between total operating cost and delay. The cost factor approach is more of a “bottom-up” approach since, implicitly at least, it assumes that aggregate delay cost is the sum of the costs of individual delay events. Conversely, the total cost approach takes a “top-down” perspective which allows for the possibility that delay can seep into every facet of air transport production cost.

3.1 Cost Factor Approach

The cost factor approach estimates delay cost as a linear function of one or more delay quantities, the coefficients of which are cost factors. In the simplest case with a single delay variable, the total cost impact of delay is:

$$C = P \cdot X \quad (3.1)$$

where C is the total delay cost, P denotes the cost per unit time (e.g. measured in \$/min), and X represents the total delay minutes. X can be obtained by summing up delays experienced by individual flights.

The above equation is overly simplistic in that delay is not homogenous. One source of heterogeneity that affects cost is where the delay is taken, which may be at the gate, on the taxiway, or in the air. Other sources of heterogeneity with potentially significant cost implications include the length of delay, aircraft size, and whether the delay is caused by the flight itself or propagated from other ones. This suggests that delays should be disaggregated into categories before applying cost factors. The cost formula then becomes:

$$C = \sum_i P_i \cdot X_i \quad (3.2)$$

where subscript i denotes the delay category. Equation (3.2) represents the general formula of the cost factor approach, which admits many possibilities for classifying delay. How to determine X and P will be addressed in the ensuing two sub-sections. Further issues in applying (3.2) will be discussed in 3.1.3.

3.1.1 Categorizing delay

Since delays in different categories can have different unit cost impact, how to appropriately categorize and quantify delays is a critical issue in applying the cost factor approach. There are two major bases for delay classification. The first is based on the phase of a flight in which the delay is taken. Gate, taxi, and airborne delays are the primary categories. Gate delay is measured as the difference between the actual and scheduled times a flight leaves the gate. Taxi and airborne delays are measured as differences between actual taxi and airborne times and unimpeded times. A flight delay can also be distinguished according to whether it is propagated from delay on an earlier flight, or is caused by some occurrence on the delayed flight itself. The latter are termed primary delays; they can occur at the gate, on the taxiway, or in the air. Propagated—sometimes termed reactionary—delays, on the other hand, are always manifested in late departure from the gate. Depending upon whether a propagated delay is caused by delay on previous legs flown by the same airframe, or delay on legs flown by a different aircraft, propagated delay can be further distinguished as rotational and non-rotational. Figure 3-1 illustrates these different concepts.

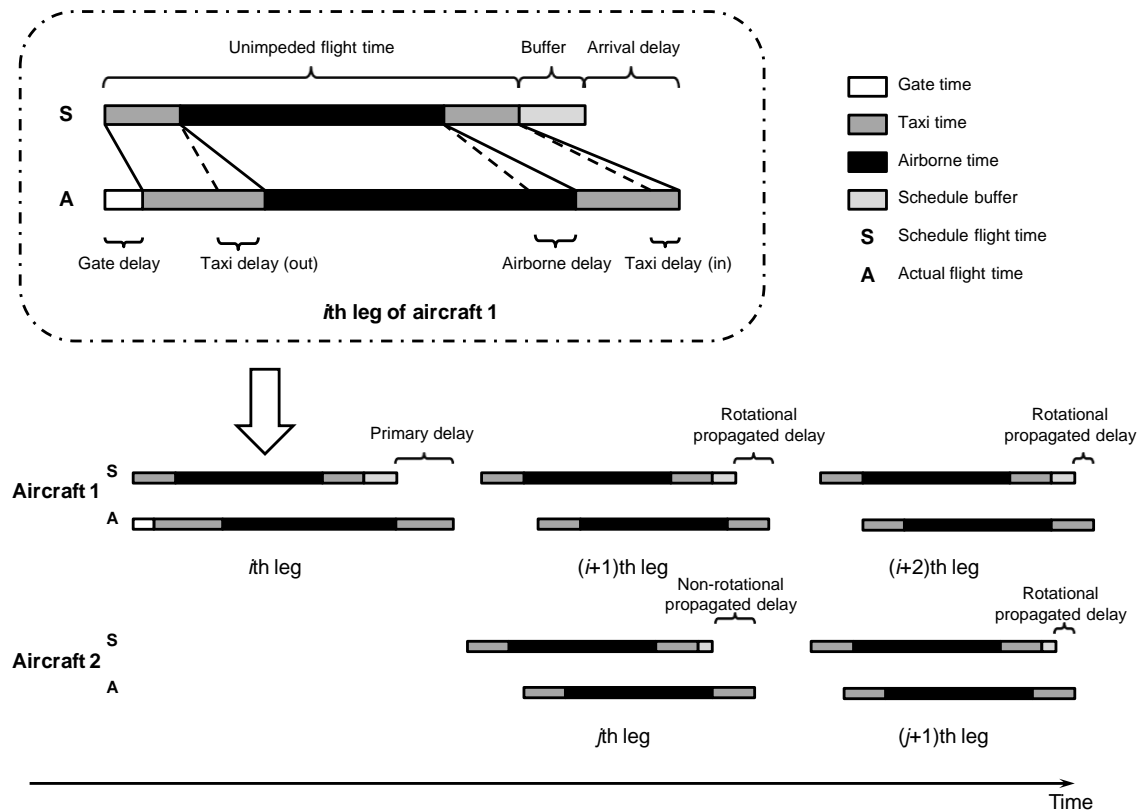


Figure 3-1: Illustration of different delay concepts

Also shown in the above figure is the role of schedule buffer. When buffer is built into the schedule of a flight leg, it absorbs all or some of the delays on previous legs. In the above figure, we assume that the i th leg of aircraft 1 experienced no propagated delay. At the arrival end, an arrival delay equal to the sum of gate, taxi, and airborne delays minus buffer is incurred. Suppose no primary delay occurs on the successive two legs of the flight, but the primary delay on the i th leg propagates to the next two legs. As a result rotational propagated gate delays occur on the following two legs ($i+1$ and $i+2$). Due to the schedule buffer, as well as additional schedule layover time above that required to turn around the aircraft, such propagated delays are partly offset and diminish on legs farther downstream from the i th leg. Similarly, suppose some passengers from the delayed i th leg of aircraft 1 need to connect to the j th leg of aircraft 2, and/or the crew of the i th leg of aircraft 1 is assigned to the j th leg of aircraft 2. Thus, as a result of the delay on the i th leg of aircraft 1, aircraft 2 has to depart late on its j th leg, an example of non-rotational propagated delay. Given no other delays during the j th leg of aircraft 2, this flight will end up arriving at the destination airport with a smaller delay due to the schedule buffer and excess layover time. The propagated delay will be similarly further reduced on the next leg. Cook et al. (2004) proposed a delay categorization based on the above characterization, as shown in Figure 3-2. Schedule buffer is included and termed “strategic delay.”

While conceptually appealing, applying this decomposition to calculate delay cost is difficult in practice, given existing data and models. Instead existing cost studies employ parts of this framework. For example, the location of delay is widely considered. If one believes that one minute of gate delay costs the same irrespective of whether it is primary or propagated, then it is sufficient to decompose delay into gate, taxi, and airborne components. Identifying taxi and airborne delays is difficult, however, because they require comparison between the actual and unimpeded (or nominal) times a flight spends in these two phases. The unimpeded taxi time depends on detailed information such as airfield geometry and gate location, which is often unavailable to researchers. Even the proper definition of nominal taxi time is unclear. For example, FAA calculates nominal taxi times that allow sufficient time for a plane to wait for one aircraft ahead in the take-off queue, while others assume no interference from other aircraft in the nominal scenario. Similar issues arise in determining the unimpeded airborne time. Conceptually, unimpeded airborne time represents the amount of time a flight spends in the air, flying an optimal trajectory and encountering no delays as the result of other flights in the system. Winds, aircraft type, and even the relative importance that airlines attach to fuel and time all affect the optimal trajectory and its associated time, but such detailed information is rarely available. To ensure the plausibility and cross validate results, it may be good practice to consider multiple delay measures. In JEC (2008) two nominal airborne times are considered (Table 3-1). Airborne delay is the actual elapsed time in the air minus the nominal airborne time.

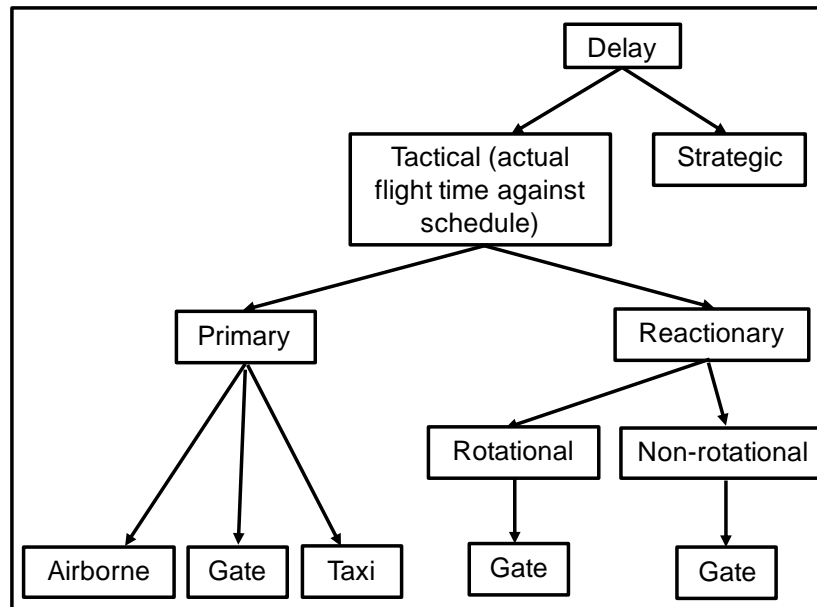


Figure 3-2: A comprehensive categorization of flight delay (Source: Cook et al., 2004)

Table 3-1: definition of nominal airborne time in JEC (2008)

Measure of nominal airborne time	Nominal airborne time
C	Min(planned flight time, scheduled block time minus the nominal taxi time)
D	5 th percentile of the observed airborne time for a give segment and month

Some differences between these two nominal airborne times measures are worth noticing. The estimated flight time included in a flight plan, a component of Measure C, may include anticipated delays. The other component, scheduled block time minus the nominal taxi time, includes schedule buffer. Measure D, in contrast, will count as airborne delay any additional flight time above the 5th percentile. Some of this difference, however, may arise as a result of aircraft type and winds that are not the result of delay. It should also be noted that delays calculated using measure D for the nominal flight time comingle schedule buffer and delay against schedule.

Another stream of cost studies place more emphasis on distinguishing between primary and propagated delay. This also imposes difficulties. As pointed by Cook (2009), if an aircraft arrives 30 minutes late at the gate, and then leaves 45 minutes late on the next outbound leg, the portion in the 45 minutes that should be counted as delay propagation from the last leg is generally not known. In Europe, the Central Office for Delay Analysis (CODA) data allows the distinction of delay in different categories as defined in the IATA standard codes, in which one category is “reactionary” (i.e. propagated) delay. Using the CODA database, ITA (2000) calculates cost associated with both primary and propagated delays. Since primary delays can take place at the gate, on the taxiway, or airborne, the cost factors the ITA applied are likely to be an average over those three possibilities.

On the other hand, if delays are measured against flight schedule, buffer needs to be identified separately. Again, because of the difficulty in measuring unimpeded flight times, there seems no widely acknowledged definition of buffer. ITA (2000) assumed an increased flight time of either 5 or 10 percent because of schedule padding. Inspired by the minimum feasible flight time introduced in Mayer and Sinai (2003a) and the nominal total flight time in JEC (2008), Zou and Hansen (2011a) proposed three buffer measures, defined as the difference between scheduled time and the 5th, 10th, and 20th percentiles of the observed gate-to-gate time over all flights for each directional flight segment, airline and quarter. Since a consensus on defining buffer has not yet been achieved, trying different measures may be worthwhile in order to help a further understanding of buffer and related scheduling behavior.

In sum, the availability of delay data is a critical challenge for the cost factor approach. It is widely acknowledged that delays in different categories exhibit different cost impacts, and the categorization can be by location of delay or by the primary/propagated dichotomy. However, some difficulties exist in both ways of categorization; which one to choose largely depends on the specific data available to researchers. But of equal significance is the need to determine appropriate cost factors, to which we now turn.

3.1.2 Determining cost factors

Cost factors are based on the assumption that delay causes additional consumption of largely the same inputs as the airlines' normal line production process. These include fuel, labor (e.g. cabin crew, flight attendants, and ground personnel), capital (e.g. depreciation, rental, and lease), and others (e.g. airport charges, maintenance). These delay cost factors can be inferred from operating cost information that is regularly tracked—and in the US, publicly reported—by airlines on a per block-hour basis.

Delays occur on individual flights involving specific aircraft and circumstances. Ideally, therefore, cost data at the flight level would be used to develop flight-specific cost factors. Airlines themselves are unlikely to possess data at such fine detail; even if they did, this information would be highly sensitive and not publicly accessible. As a result, researchers often resort to two surrogates: one is more aggregate cost data; the other is information gathered from interviews. This is reflected in Table 3-2, which documents the data sources used by four major cost-of-delay studies.

Aircraft operating cost data by carrier and aircraft type for US airlines are publicly available in the Form 41 database; US studies typically rely on this database as the primary resource for cost factors. Form 41 contains detailed airline financial and operating information, such as salary of pilots, direct expenses for maintenance of flight equipment, equipment depreciation cost, and the total number of flight hours for each combination of aircraft type, airline, and quarter. This allows construction of block hour cost for various cost categories—crew, maintenance, fuel, etc. Cost factors for different types of delay are obtained by assuming which categories of cost are incurred for delays of a given type, as discussed further below.

In contrast, perhaps because cost data are less readily available publicly in Europe, cost factors are based on interviews in many European studies. Interview subjects include airline, airport, Eurocontrol, and research personnel. Such personal inquiries provide researchers with a richer picture of how delays affect airline operations and business methods. On the other hand, cost estimates obtained from interviews are inherently subjective. Interviewees usually tend to incorporate cost impacts that are obvious to them while omitting those that are not directly visible. In addition, even for the impacts that are manifest, estimates (like those based on block hour cost data) may be unduly influenced by accounting conventions that often have little empirical basis.

Table 3-2: Data sources for determining cost factors

List of research	Data source
Cook et al. (2004)	Primary data from interviews with airlines, handling agents, aircraft operating lessors and other parties (e.g. Eurocontrol, research institutions, offices that set airport charges, and IATA). Supplementary data from ICAO, the Airline Monitor, and simulation results using Lido (for calculating fuel burn rate).
ITA (2000)	Interviews with individual airlines.
JEC (2008)	US DOT Form 41 (Schedule P5.2, T-2); FAA Aviation Environmental Design Tool System for fuel burn data; FAA rules for flight attendant hourly wage information.
ATA (2010)	US DOT Form 41 data for U.S. scheduled passenger airlines with annual revenues \geq \$100 million.

In developing cost factors from detailed cost data, as available in the US, judgments must be made about which cost categories to include for which forms of delay. Aircraft operating costs are categorized into various cost components such as fuel, crew, maintenance, depreciation, rental, airport charges, and so on. Decisions about which components to include in a delay cost factor are based upon knowledge and understanding of the researchers and/or subject matter experts. As is the case with delay categorization, this exercise can be carried out at varying levels of detail, ranging from a single cost factor representing the “average” unit cost impact of all delays, to hundreds of different cost factors each describing the effect of one minute of delay for a given aircraft model, delay duration, delay location, cost scenario – even including consideration of network effects (e.g. Cook et al., 2004). Tables 3-3 to 3-6 reflect the cost components considered in several studies, and how they were used in constructing different cost factors. Table 3-3 provides a snapshot of the cost components that are included, as shown by checks, in different cost factors. In the first column, for example, the cost factor is defined by a combination of aircraft type (B733), delay duration (15-min), delay location (taxi), cost scenario (low), and network effect consideration (without network effect).

Table 3-3: Cost components and cost factors in Cook et al. (2004)

Cost components	B733 15-min taxi delay: low cost scenario, w/o network effect	A319 65-min gate delay: high cost scenario, w/ network effect	B744 15-min gate buffer: used, low cost scenario	A321 65-min taxi buffer: unused, high cost scenario
Fuel	✓	✓	✓	
Maintenance	✓	✓	✓	✓
Crew		✓	✓	✓
Aircraft depreciation, rental, and lease			✓	✓

Airport charges	✓	✓		
Ground and passenger handling		✓		
Aircraft operator passenger cost		✓		

Table 3-4: Cost components and cost factors in ITA (2000)

Cost components	Primary delay	Propagated delay	Buffer
Aircraft operating cost	✓		✓
Operating staff cost	✓	✓	✓
Structural cost	✓	✓	✓
Passenger driven cost	✓	✓	
Hub and connection cost	✓	✓	✓
Induced (long term) airline cost	✓	✓	

Table 3-5: Cost components and cost factors in CAA (2000)

Cost components	Long airborne delay	Short airborne delay	Long ground delay	Short ground delay
Flight crew	✓		✓	
Cabin crew	✓		✓	
Fuel and oil	✓	✓		
Flight equipment insurance	✓		✓	
Rental of flight equipment	✓		✓	
Maintenance and overhaul	✓	✓		
Flight equipment depreciation	✓	✓		

Handling charges & parking fees	✓		✓	✓
Passenger services (e.g. meals)	✓		✓	

Table 3-6: Cost components and cost factors in JEC (2008)

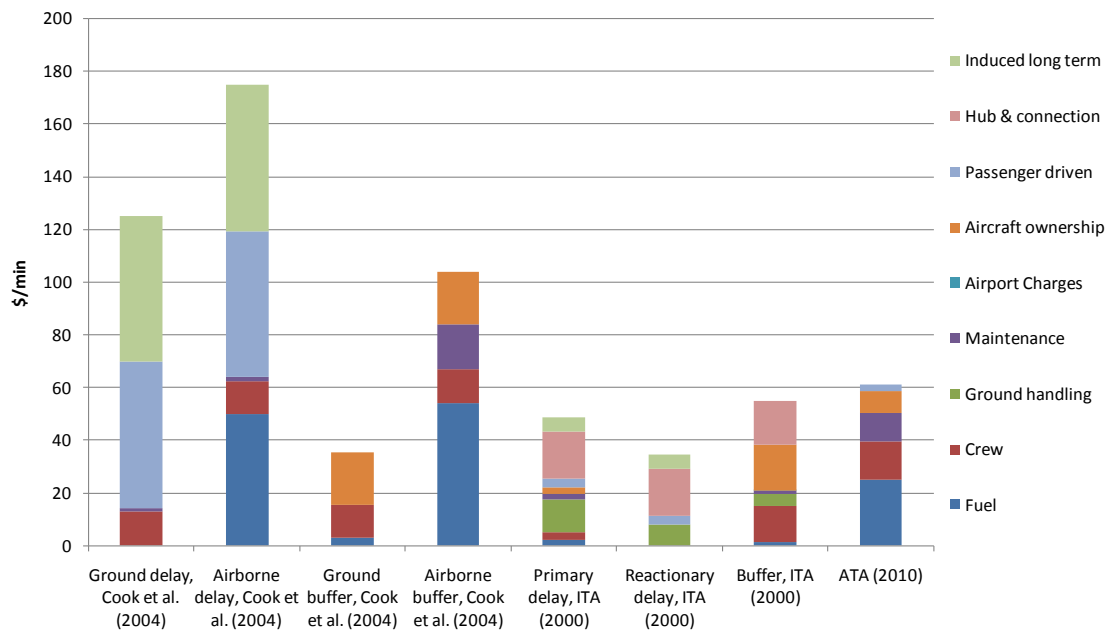
Cost components	B733 gate delay	B744 taxi delay	A321 airborne delay
Taxiing fuel cost		✓	
Airborne fuel cost			✓
Pilot salary	✓	✓	✓
Flight attendants expenses	✓	✓	✓
Maintenance and depreciation			✓

The common cost components appearing in the above tables are fuel, crew, and maintenance. Inclusion of other items varies from study to study. For example, airport charges can increase if there are long gate delays. Late arrivals induce additional ground and passenger handling costs. These are included in Cook et al. (2004) and ITA (2000). Probably due to more stringent delay compensation rules in Europe, all studies on the European side include passenger cost as an integral component. This could involve extra expenses on meals and hotel accommodation for passengers with long delays. ITA (2000) added a “hub and connection” term, to capture the cost impact resulting from schedule disturbance and loss of operational efficiency at hub airports, e.g. lengthening of connecting times, missed connections, and flight cancellations. Cook et al. (2004) further considered a potential drop of airlines’ market share and revenue loss in the future due to a lack of punctuality. Some of these losses may be recouped by competing airlines, but this possibility is not typically considered.

Figure 3-3 illustrates the estimates of delay cost factor values and their composition from several studies, all converted to US 2008 dollars.¹ In several delay factors, fuel, crew, and maintenance cost account for the bulk in the total. When a buffer cost factor is measured, aircraft ownership, including airframe depreciation, rental, and lease, becomes an important cost component. The hubbing and connection cost factor, not estimated in other studies, accounts for most of the propagated delay cost factor and is generally estimated to be of the order of \$15-20 per minute in ITA (2000). The passenger-related cost refers to compensation to passengers due to delay, and is substantial in Cook et al. (2004) compared to ITA (2000). The induced long term cost, referring to the loss of market share and revenue due to delay, accounts for a major portion in Cook et al. (2004),

¹ The conversion to 2008 prices was made using the June value of the harmonized index of consumer prices (HICP) provided by EUROSTAT. All values were converted to dollars using the foreign exchange reference rate for Dec 1, 2008 published by European Central Bank. The above information is available in Eurocontrol (2009).

as well. Overall, there is a significant variation of cost factors, with the high end reaching almost \$175 per minute of delay and the lower at less than \$40 per minute.



Notes: 1. All numbers updated to 2008 value in dollars, based on inflation and exchange rates provided by Eurocontrol (2010). 2. The numbers from Cook et al. (2004) are averaged across different aircraft models, as in Eurocontrol (2010). Passenger driven cost only includes “hard” cost (passenger compensation and rebooking expenses). “Soft” cost is counted as “induced long term cost”.

Figure 3-3: Cost factors and their composition

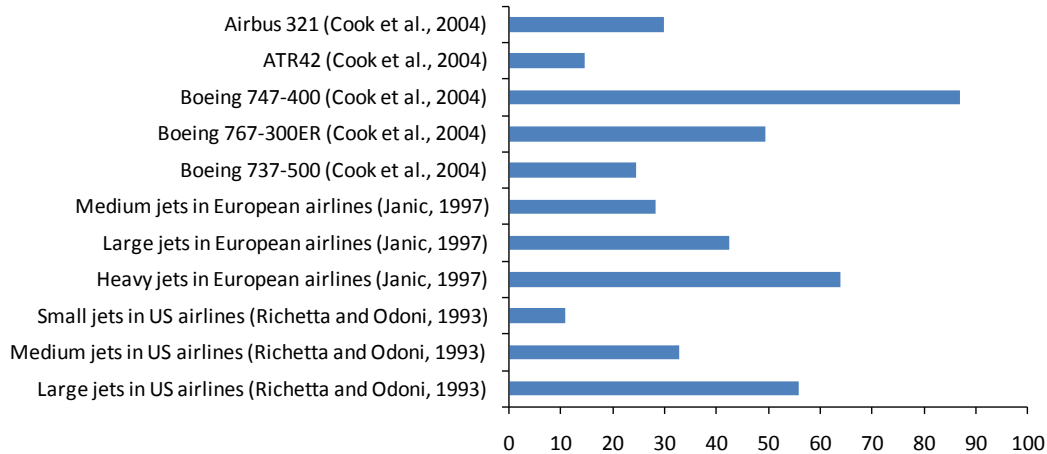
The different dimensions of delay characteristics specified in each study, as shown in Table 3-7, suggest the level of cost details the cost factors are able to provide if relevant delay information is available. The location of delay significantly affects the magnitude of delay impact; this is considered in most studies. It is also important to account for the impact of aircraft size. Unless carefully aggregated, using cost factors not differentiated by aircraft size would likely yield unreliable results. For example, aircraft size contributes to the large differences among various ground delay cost factor estimates reported in the literature, as shown in Figure 3-4. Delay cost factors are also differentiated between primary and propagated delays in two European studies,² where data permit such a distinction.

Table 3-7: Some characteristics of cost factors determined in existing studies

Source	Location where delay occurs	Aircraft size	Primary/propagated
Cook et al. (2004)	✓	✓	✓

² In Cook et al. (2004) multipliers were introduced in the cost factors to account for the “knock-on” effect of primary delays.

ITA (2000)	✓	×	✓
CAA (2000)	✓	×	×
JEC (2008)	✓	✓	×
ATA (2010)	×	×	×



Note: the values from Cook et al. (2004) are based on the high cost scenario for a typical delay duration of 15 minutes, without considering the network effect.

Figure 3-4: Unit ground delay cost in literature (\$/min, updated to 2008 values)

3.1.3 Shortcomings of the cost factor approach

This section discusses several shortcomings of the cost factor approach. First, choosing the appropriate delay categorization scheme is difficult. Second, the cost factor approach may overlook potentially important aspects of the relationship between delay and airline cost.

3.1.3.1 Choosing the categorization scheme

Although multiple options for categorizing delays and associated cost factors may be available, choosing an appropriate one depends on the availability of necessary data from both the cost and delay sides. The cost side may be the less problematic, as long as all cost components are identified. In principle, one only needs to determine what cost components should be included for a specific type of delay. While this process involves a certain degree of judgment, the more significant questions concern the level of detail at which delay data are available in databases available to researchers. Sometimes applying a single cost factor to all delays may be necessary because delay cannot be further disaggregated with the data available.

If delay information by both location and aircraft type is available, having corresponding cost factors is clearly desirable. However, two related issues may arise. On the delay side, when making distinctions among delays according to the location where they occur, it is necessary to define the unimpeded taxi and airborne time, which may impose considerable uncertainty about measure accuracy. On the cost factor side, data specific to an aircraft type may be difficult to obtain. If data availability by aircraft type is incomplete, linear interpolation/extrapolation with respect to the number of seats may provide a good approximation, as shown in Figure 3-5.

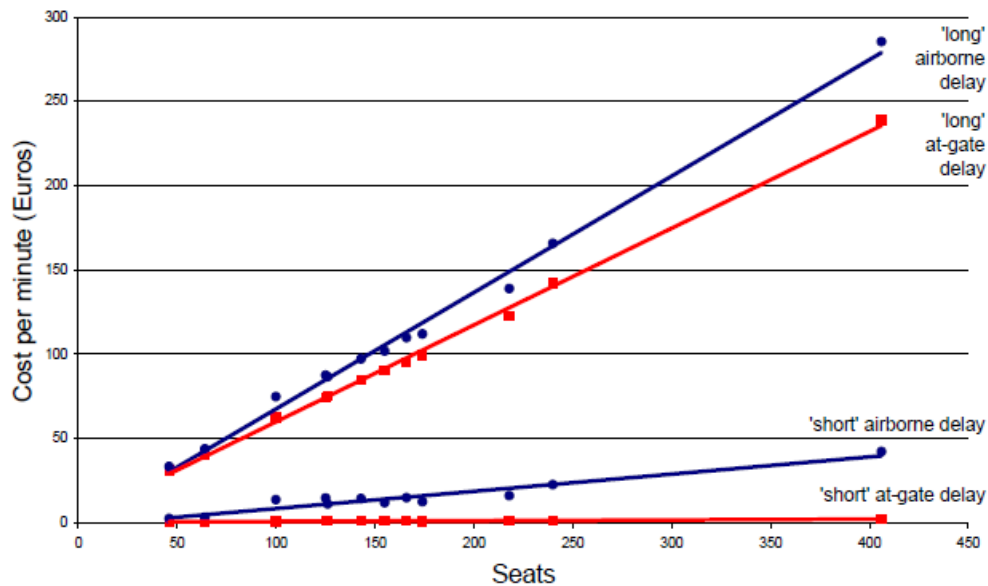


Figure 3-5: Cost of at-gate and airborne delays as a function of the number of aircraft seats (source: Cook et al., 2004)

Several issues should be considered when making distinctions between primary and propagated delays. First, although primary and propagated delays may entail quite different cost factors, it is difficult to decompose delays into these two categories. Second, even if primary and propagated delays can be distinguished in the data, determining corresponding delay cost factors, especially by aircraft type, can be challenging. As a final caveat, when cost factors for primary delay already account for the cascade effects of primary delay on other flights, only primary delay needs to be included in the cost calculation. Otherwise double-counting could occur.

3.1.3.2 Effects not captured

Recall that the cost factor approach implicitly assumes that cost increases linearly with delay, across all flights. In reality, however, a single four-hour delay will probably be more costly than 24 ten-minute delays, for two reasons. First, as delay becomes longer, additional cost items may be incurred. Figure 3-6 shows that handling surcharges, expenses for providing passengers with meals and accommodations, and other costs appear when delay exceeds a certain time threshold. In addition, the delay propagation effect increases non-linearly with the size of the initial delay, as demonstrated by Beatty et al. (1998). Longer delays propagate to more flights and are more likely to disrupt ground operations, gate assignments, crew schedules, and passenger itineraries (Figure 3-7).

A few cost factor studies, notably Cook et al. (2004) and CAA (2000), have attempted to take this effect into account. Cook et al. (2004) chose two specific delay durations (15 and 65 minutes) to typify “short” and “long” delays (Figure 3-5). Similarly, in CAA (2000), 20 minutes was selected as the threshold separating “short” from “long” delays. Both studies conclude that longer delays have higher cost factors. While such categorizations help address the issue of non-linearity, they are at best first steps toward representing the true relationship between length of delay and cost. Further efforts in this direction are certainly warranted.

Taking this one step further, the relationship between delay duration and cost can even be non-monotonic. Airlines sometimes add delays to flights in order, for example, to avoid having a flight arrive at a hub in the middle of a departure “bank” (or “wave”). In a hub-and-spoke network, the interaction of delays for many flights scheduled in a connecting bank plays an especially important role in maintaining the integrity of airport and airline operations. If all the flights in an inbound bank are delayed by the same amount, then the effect may be far less severe than if half the flights are delayed by a smaller amount of time than the others (Hansen et al., 2001). The cost factor approach cannot capture such effects, because of its underlying assumption that the cost of delay is an additively separable function of delay variables for individual flights.

Finally, the presence of delays may generate sizable indirect effects that may be difficult to capture with the cost factor approach. Carriers may take a variety of measures in flight scheduling to make their operations more robust to delay. One routine practice is building more padding into scheduled block times, which explains the necessity of including the buffer side when evaluating the whole delay cost impact. In addition, airlines may require extra aircraft, flight crew, and ground personnel, and load flights with more fuel. These adaptations entail overhead and capital costs that are not accounted for in determining cost factors. The cost of delay may thus permeate the entire cost structure of the airline in ways that are not tied to individual delay events (Hansen et al, 2001). It would be extremely difficult to capture such effects using cost factors. The majority of studies based on cost factors simply ignore these issues. (Only ITA (2000) tried to quantify such a “structural cost” item, albeit in a vague manner). As a consequence, it is likely that the cost factor approach may underestimate the true cost impact of flight delays.

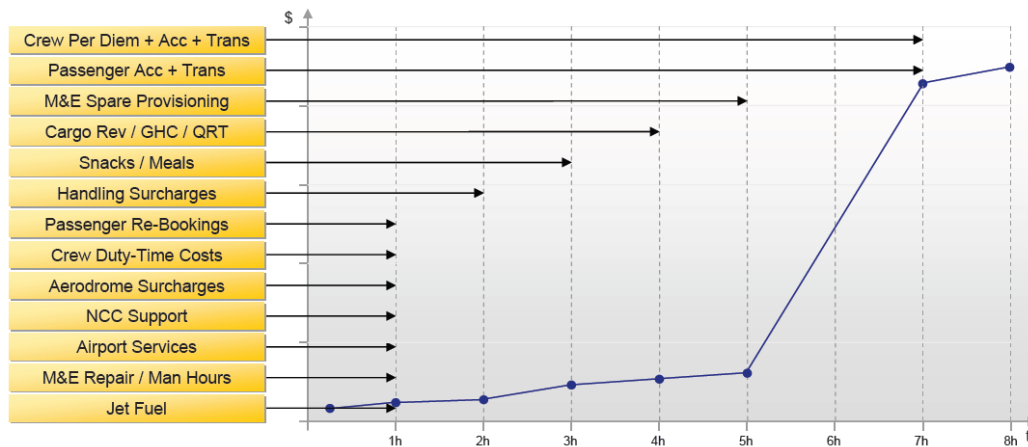


Figure 3-6: Delay cost increase as a function of delay duration (source: m2p consulting, 2006)

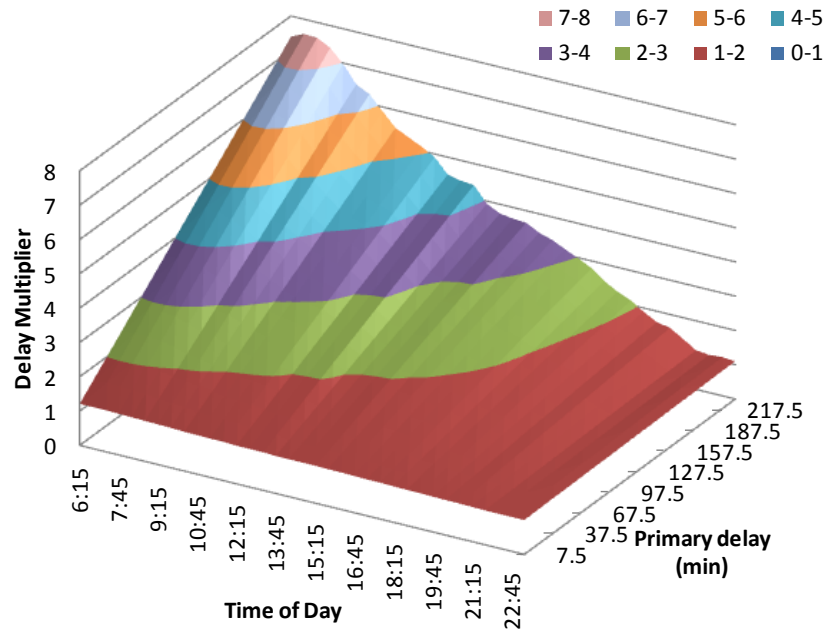


Figure 3-7: Delay propagation as a function of time of day and the amount of primary delays
(source: Beatty et al., 1998)

3.1.3.3 Cost estimation results

Table 3-8 lists the estimated annual cost of delay obtained by several studies in Europe and the US, all using the cost factor approach. The ITA (2000) and Cook et al. (2004) are representative on the European side and widely cited by delay cost researchers. Since the publication of Cook et al. (2004), Eurocontrol has periodically reported cost estimates for ATFM delays using updated cost factors. The estimates for the years 2007 and 2009 from Eurocontrol are presented here, in order to compare with the estimates from two recent US studies, the JEC (2008) and ATA (2010).

In addition to geographical and demand level differences, the seemingly large cost estimates for 2007 and 2009 in the US, compared to their European counterparts, may be explained by several reasons. First, the Eurocontrol estimates only considered delays due to ATFM, while the US estimates included all types of delays. Second, as discussed before, the delay measured in JEC (2008) is against some unimpeded, optimal flight time; schedule buffer is at least partly incorporated. This results in a greater amount of delay than one would obtain when only delay against schedule is considered. Block hour cost was also used for the ATA (2010) estimate as a proxy for a single delay cost factor. As has been discussed, many delays occur at the gate, where fuel consumption is marginal. Thus, using block hour cost, of which fuel expense constitutes a significant portion, will yield an overestimate of the cost of delays at the gate. On the other hand, neither the US nor the European estimates fully account for the non-linear, combinatorial, or indirect cost impacts described above.

Table 3-8: Existing estimates of delay (and buffer) cost to air carriers using the cost factor approach

Scope	Source	Evaluation year	Cost Estimate	(Averaged) cost factor	Note
Europe	ITA (2000)	1999	€3.0-5.1 billion (delay: €1.6-2.3 billion; buffer: €1.3-2.7 billion)	Delay: €35.5/min (using marginal cost), €50.9/min (using average cost); buffer: €45.2/min	Buffer is calculated as 5% to 10% of scheduled flight time.
	Cook (2004)	2002	€840-1200 million	€72/min	Focus on departures. Assume all delays are primary; classified into two types (< 15 min and >15 min). Buffer not considered.
	Eurocontrol (2010)	2007	€1.40 billion	N/A	Use updates of cost factors in Cook et al. (2004). Consider total ATFM delays. Buffer not considered.
		2009	€1.00 billion	N/A	Similar updates as for the 2007 estimate.
United States	Geisinger (1988)	1986	\$1.2 billion		
	Odoni (1995)	1993	\$2-4 billion		
	Citrenbaum and Juliano (1998)	1996	\$1.8 billion		
	JEC (2008)	2007	\$3.6/6.1 billion (corresponding to delay measures C and D in Table 3-1)		Delay is measured against “unimpeded flight time..
	ATA (2008)	2007	\$7.8 billion	\$60.46/min	Single cost factor; use block hour cost as a proxy. Buffer not considered.

	ATA (2010)	2009	\$6.1 billion	\$60.99/min	Same as for ATA (2008) estimate.
--	------------	------	---------------	-------------	----------------------------------

3.2 Aggregate Cost Approach

The weaknesses in the cost factor approach suggest that alternative methods for estimating the cost of delay to airlines should be considered. To a large extent, the weaknesses stem from allocating cost components to each cost factor, and assuming that cost bears a linear relationship with the delay duration. In order to address these two concerns, an “aggregate cost” (or top-down) approach is proposed in this section. Rather than investigating explicitly the cost impact of each minute of delay, the top-down approach assesses the impact of delay impact in a more holistic manner. One way of doing this, the total time approach, is simple and intuitive, and uses readily available aggregate airline accounting information. Unfortunately, this approach is also almost certainly wrong. The second approach, an econometric one, is more rigorous, data intensive, and methodologically sophisticated, and hence probably more accurate.

3.2.1 Total time approach

This approach assumes that airline operating costs are proportional to total aircraft operating time, measured, for example, in plane-hours. If this were true, then delay cost could be estimated as the fraction of total aircraft operating time due to delay, multiplied by total airline operating costs.

The total time approach is based on the premise that if an airline operated zero aircraft hours, its costs would also be zero. Thus, the relationship between operating cost and aircraft hours includes the origin, and the point corresponding to the actual cost and hours. A straight-line interpolation between these two points can be used to predict how costs would change if delay hours were eliminated. This approach thus avoids the numerous details of determining cost factors, and only requires calculating the aggregated delay time. Since total operating cost, aircraft time, and delay time for an airline can be readily obtained, the total time approach is straightforward to implement. This was done in JEC (2008), and yielded much larger cost estimates than the cost factor approach. Using the same delay measures as employed under the cost factor approach (see Table 3-1), they conclude with considerably higher estimates using the total time approach (\$19.1 vs. \$3.6 billion based on measure C; \$23.4 vs. \$6.1 billion based on measure D).

While it is appealingly simple, the validity of this approach is very questionable. First of all, like the cost factor approach, it assumes a linear relationship between delay cost and time. Second, the approach ignores differences where delay and planned aircraft hours occur; for example, a much larger proportion of delay time occurs on the ground. Moreover, the total operating cost includes not only fuel, crew salaries, maintenance, and depreciation, but also advertising, ticket agents, landing fees, legal fees, and other items that may be relatively insensitive to delays. As a result, cost numbers generated through this approach will almost certainly be too high.

3.2.2 Econometric approach

3.2.2.1 Basic concepts and cost model functional form

A second version of the aggregate cost approach rests on the idea that the provision of air transportation service to passengers constitutes a production process. Economic theory suggests that in the production process, each firm minimizes the cost C at which it can produce a given amount of output Y , given the prices it pays for inputs, \bar{w}_t . The airline cost function can be conceptually expressed as: $C = f(Y_i, \bar{w}_t)$, where subscript i denotes a particular firm (airline), and t identifies the time period. A typical output measure is revenue ton-miles. Inputs include labor, fuel, capital, and materials. The function represents the cost of acquiring the optimal set of inputs, given the output and input prices (Hansen et al, 2001). In reality, however, capital inputs cannot be adjusted to the optimal level instantaneously (Caves et al., 1984; Gillen et al., 1990). We therefore relax the assumption of optimal capital stock by treating capital input, denoted by S , as quasi-fixed

and employing a variable cost function to reflect the short-run cost minimization process. The airline variable cost function can then be written as a function of its output Y_{it} , the price of the three variable inputs (fuel, labor, and materials) \bar{W}_{it} , and capital input S_{it} , i.e. $VC_{it} = f(Y_{it}, \bar{W}_{it}, S_{it})$.

In the airline cost literature, it has long been recognized that costs depend on the nature and quality of the airlines' output, as well as the quantity. Because the nature and quality of output also vary over time and across carriers, the specification of the airline cost function above needs to take these into account. A set of additional variables \bar{Z}_{it} describing the nature of the output are introduced. Variables of this kind that often appear in the literature can be the size of the airline's network (often measured by the number of points served) and the average flight distance (stage length). We hypothesize that flight delays, or in a broader sense, imperfect operational performance also affects airline cost. Therefore we add a new variable or vector of variables \bar{N}_{it} to the cost function. The cost function then becomes $VC_{it} = f(Y_{it}, \bar{W}_{it}, \bar{Z}_{it}, S_{it}, \bar{N}_{it})$.

Depending on the characterization of operational performance, different versions of airline cost models can be developed and estimated. Two most popular ones are Cobb-Douglas and Translog cost functions, both assuming non-linear relationship between the cost variable and the functional arguments. In a Cobb-Douglas model, the non-linear relationship is characterized by a log-log functional form, in which the coefficient associated with each explanatory variable is interpreted as the corresponding cost elasticity. Because of the log-log relationship, these elasticities are assumed constant. This constant elasticity assumption can be relaxed if a more flexible Translog functional form is adopted, which can be regarded as a second-order Taylor expansion of any general cost function. The Translog model imposes fewer a priori assumptions about the relationships between airline cost and the explanatory variables, and the cost elasticities vary with the variable values. In particular, because relationships are derived from observed co-variation between operational performance variables and cost, the results entail a minimum of assumptions about the delay-cost interaction mechanism involved. Interested readers can refer to Hansen et al. (2000; 2001) which demonstrated the employment of Cobb-Douglas and Translog cost functions to characterize the relationship between airline cost and operational performance.

3.2.2.2 Data, model estimation and applications

Reliable airline financial and operational databases are critical to the estimation of airline cost models. In the US, the Bureau of Transportation Statistics (BTS) releases airline-level financial and operational information—which are reported as a mandate by major domestic airlines—on a quarterly basis. The BTS On-time Performance database, also publicly available, provides airline operational performance at the individual flight level. Aggregate airline financial and operational information outside the US can be accessed, for example, through the Digest of Statistics published by the International Civil Aviation Organization (ICAO) as well as other data sources including Monitor, IATA publications, and annual reports of airlines. In Europe, individual flight operation records that are comparable to the BTS On-time Performance data are collected and maintained by CODA. For the remaining parts of the world, however, access to detailed operational performance information may present a challenge.

The choice of estimation methods depends upon the nature of the dataset (e.g. cross section vs. panel) and the specification of the cost model. To improve estimation efficiency, the Translog cost function is often jointly estimated with additional equations derived from production theory. Standard estimation techniques include Zellner's seemingly unrelated regression and the maximum likelihood method. Further details about model estimation can be found in Caves et al. (1984), Gillen et al. (1985; 1990), Oum and Yu (1998), and Hansen et al. (2000; 2001).

The estimated airline cost model can then be used to gauge the cost impact on airlines of delay as well as other dimensions of imperfect operational performance. One can, in principle, change operational performance variables to an improved level of interest and perform cost prediction. The cost savings then equal the difference between the new and original predicted cost. For example, if delay and buffer are included as two operational performance variables, one can reduce delay, or both and buffer, to zero, or some other value deemed to be the lowest attainable. A recent study on the total delay impact in the US (Ball et al., 2010), where multiple versions of airline Translog cost models have been estimated with different sets of operational performance variables, found the system-wide delay cost for US airlines in 2007 to range from \$5 to \$9 billion. Together with buffer, the total impact can reach \$8 to \$13 billion. These estimates lie between the results from the cost factor approach and total time approach for the same year estimated by JEC (2008). As discussed before, the other two approaches are likely to provide lower and upper bounds of the true values. It is therefore plausible that the econometric approach yields a closer estimate to the true cost impact.

3.3 Summary

This section has reviewed possible approaches for estimating the cost impact of delay, or more generally, of imperfect operational performance. The cost factor approach has gained popularity due to its intuitiveness and simplicity. It has significant shortcomings, however, as a result of which the estimates it generates may be inaccurate and biased toward underestimating true costs. In contrast, by treating total operating cost as proportional to hours of flight activity and ignoring important differences in distribution of time among phases of flight for delays and regular flight operations, the total time approach probably yields overestimates of delay cost. Building on production theory, the econometric approach investigates the statistical relationship between the operational performance of airlines and their costs. By employing flexible functional forms, the econometric approach avoids the strong assumptions, such as linearity and additive separability, that are built into the cost factor approach. A recent system-wide delay cost study in the United States using airline cost models appears to yield plausible estimates (Ball et al., 2010). However, these econometric models also have limitations. They are data hungry and therefore can be applied only at a relatively aggregate level. If the focus is narrower, for example a specific airport, the cost factor approach may be the most appropriate way of gauging the cost impact of delays. Nevertheless, the existing estimates of cost impact of delay provide some assurance that, despite the intrinsic methodological differences, cost factor, total time, and econometric approaches all yield estimates of a similar magnitude when applied to the US system.

4 Further issues

4.1 Cancellations

While the focus of the studies discussed previously is primarily on delay (and occasionally on the schedule buffer, as well), it is important to recognize that flight cancellations can also be an important cost driver. The decision to cancel a flight is intertwined with flight delays, and reflects the trade-off between avoiding excessive delays of subsequent flight legs and maintaining flight schedule integrity. Canceling a flight involves very complicated considerations in practice. Airlines have to weigh the cost incurred as a result of excessively long delays against the cost of cancelling a flight. The latter may include providing passengers on the canceled flight with meals and hotel accommodations, as well as transporting those passengers to their scheduled destinations. The second component often requires spare capacity, which incurs opportunity cost of capital and loss of potential revenue, and/or the amount one airline has to pay to another if such capacity is not available. Cancellation decisions are further compounded by the heterogeneity of airline behavior, which are subject to the operational preference and even corporate culture of individual airlines.

Under similar circumstances, some airlines may choose to persevere through long delays while others may decide to cancel certain flights. This airline heterogeneity is also reflected at the schedule planning stage. As illustrated in Mayer and Sinai (2003b), risk-averse airlines may accept a low aircraft utilization rate to avoid cancellations. This involves another type of opportunity cost of capital. All these considerations make the quantification of airline cancellation cost a difficult task.

Partly as a result of those, but partly also because of the lack of data on airline cancellation decision-making, the body of literature on quantifying flight cancellation cost is rather limited. Among the few attempts, Eurocontrol (2009) suggested different cancellation cost factors by aircraft size, for which it considered service recovery cost (e.g. meals and hotel), interline cost, loss of future value (individual passenger delay expressed in value), and operational savings. Using post-operation ground delay program data, Xiong and Hansen (2009) modeled airline cancellation decision making in a discrete choice context. Their results revealed an upper bound on cancellation cost, at around \$5000 per cancelled flight. An aggregate approach to quantify the cancellation effect on airline cost is seen in Hansen et al. (2000; 2001), where the authors employed the Principal Component Analysis and found the cancellation factor an important cost driver in airline operations. Since cancellation decisions are intertwined with flight delays (and possibly schedule buffer as well), it is difficult to precisely estimate the cost of cancellation purged of delay under an aggregate approach. Further research in this area is clearly warranted. A better understanding of the relative magnitude of the overall cancellation cost vis-à-vis that of delay and schedule buffer will enrich our knowledge base on airlines' cancellation decisions, and on the contribution of cancellations to the overall cost of imperfect operational performance.

4.2 Optimal level of operational performance and system response

While most of the focus so far has been given on the cost impact of imperfect operational performance relative to cases that delay (and buffer) were entirely eliminated, it is important to recognize that a delay-free world is a limiting—and unreachable—case. As long as winds and storms exist, aircraft parts fail, and people make mistakes, delay will remain part of operations in any aviation system. Similarly, to make schedules more robust to wind variability and other exigencies, airlines will continue padding extra time to make their schedules more robust to wind variability and other exigencies. As a consequence, the preceding estimates, which focus on the cost savings from reducing delay/buffer to zero, provide an upper bound of the achievable cost savings, and a high one at that.

In addition, a level of investment in aviation infrastructure and technologies that could achieve complete elimination of delay is neither realistic nor advisable from the efficiency standpoint. The inherent peaking nature of flight demand suggests that, if delay were reduced to zero at peak periods, the capacity of the aviation system would be significantly underutilized during off-peak hours. However, defining the attainable level of operational performance is both difficult and controversial. In effect, the question of the right level essentially reflects the trade-offs one has to make between throughput and operational performance.

Perhaps more importantly, the cost assessment methods outlined above assume that, except for operational performance improvement, everything else would remain unaffected in the system. This is certainly a simplification of the real world situations. Reduced delay—often through enhanced infrastructure supply—improves the quality of air service, attracting more people to use the air transportation system. This has been empirically observed by Hansen and Wei (2006) at Dallas-Fort Worth airport after the completion of a major capacity expansion project in 1996. The authors found that a large portion of the direct delay reduction benefits may have been offset by

changes in airlines schedules. From a broader, longer-term perspective, both demand and supply in the air transportation system will respond to improved operational performance, leading to a shift of the system equilibrium. Adequately capturing this equilibrium shift should also account for the economies of density and the Mohring effect³ that are inherent in an air transportation system when delay is absent. At the airport level, the impact of capacity change and flight delays on equilibrium shift has been examined in Morrison and Winston (1983; 1989; 2007), Jorge and de Rus (2004), Miller and Clarke (2008). Recently Zou and Hansen (2011b, 2011c) have investigated this issue on a system-wide scale. Modeling the air transportation system respectively from the airline competitive and behavioral equilibrium perspectives, the latter two studies showed that realistic airport capacity increase will continuously reduce delay—at a diminishing rate. Zero delays can never be achieved. Facing reduced delay and increased passenger demand, airlines also adjust fare, frequency, aircraft size, and the number of passengers on each flight. Zou and Hansen (2011b) further showed that the net benefits of delay savings to airlines should be reflected by profit change rather than the operating cost alone. These studies present a first step to incorporate the equilibrium concept into airline delay cost analysis. Future research is warranted to advance the relevant methodologies, and make the equilibrium concept more practical, and easy to implement.

5 Concluding Remarks

This chapter presents a comprehensive review of existing knowledge about the impacts of operational performance on airline costs. No single methodology exists that can serve as a panacea for translating changes in operational performance into changes in airline costs under all circumstances. The econometric approach makes the fewest assumptions about the mechanisms through which flight delays and related phenomena affect airline costs. This makes the approach a useful one for assessing the economic cost of air transport congestion and delay, when the necessary data are available. The much more commonly used cost factor approach is probably more appropriate for the day-to-day business of investment analysis. It cannot be overemphasized, however, that this approach rests on strong and untested assumptions about the cost-generating mechanisms involved. The cost factor method can be applied with varying levels of refinement, as time and data permit, but it rests on a shaky foundation regardless. It is somewhat re-assuring, however, that estimates of the total cost of delay in the US based on the econometric and cost factor approaches yield results of a similar magnitude.

References

1. Air Transport Association (ATA), 2008. Cost of Delays: Direct Cost. <http://www.airlines.org/economics/specialtopics/ATC+Delay+Cost.htm>.
2. Air Transport Association (ATA), 2010, Annual and per-minute cost of delays to U.S. airlines. <http://www.airlines.org/Economics/DataAnalysis/Pages/CostofDelays.aspx>.
3. Airport Council International (ACI, North America) (2009) Airport capital development costs: 2009-2013. Available at: <http://aci-na.org/static/entransit/CapitalNeedsSurveyReport2009.pdf>.

³ The Mohring effect is an increasing return property that exists in many scheduled transport systems. It basically says that, as transport service frequency increases, passenger waiting time (or schedule delay) decreases, demand increases, which in turn results in further increase in service frequency.

4. Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B. (2010) Total delay impact study: a comprehensive assessment of the costs and impacts of flight delay in the United States. NEXTOR Report prepared for the Federal Aviation Administration.
5. Beatty, R., Hsu, R., Berry, L. Rome, J. Preliminary evaluation of flight delay propagation through an airline schedule. Presentation at the 2nd USA/Europe air traffic management R&D seminar, http://www.atmseminar.org/past-seminars/2nd-seminar-orlando-fl-usa-december-1998/papers/paper_038.
6. Bureau of Transportation Statistics (BTS) (2009) Airline on-time statistics and delay causes: on-time arrival performance. Available at: http://www.transtats.bts.gov/OT_Delay/ot_delaycause1.asp?display=data&pn=1.
7. Civil Aviation Authority (CAA), 2000. Delay cost based on UK airline actual P&L figures. SS4 working paper on delay cost methodology.
8. Cameron, A. C., Trivedi, P. K., 2003. Microeconometrics: methods and applications. Cambridge University Press, Cambridge, United Kingdom.
9. Caves, D. W., Christensen, L. R., Swanson, J. A., 1981. Productivity growth, scale economies, and capacity utilization in U.S. railroads, 1955-1974. American Economic Review, Vol. 71, pp. 994-1002.
10. Caves, D. W., Christensen, L. R., Tretheway, M. W., 1980. Flexible cost functions for multiproduct firms. The Review of Economics and Statistics, Vol. 62, No. 3, pp. 477-481.
11. Citrenbaum, D. and Juliano, R., 1998. A simplified approach to baselining delays and delay costs for the national airspace system. Federal Aviation Administration, Operations Research and Analysis Branch, Preliminary Report 12.
12. Cook, A., Tanner, G., Anderson, S., 2004. Evaluating the true cost to airlines of one minute of airborne or ground delay. Report prepared for the Eurocontrol Performance Review Unit, http://www.eurocontrol.int/prc/gallery/content/public/.../cost_of_delay.pdf.
13. Cook, A., Tanner, G., 2009. The challenges of managing airline delay costs. Presentation at the Conference on air traffic management (ATM) economics, Belgrade. http://www.garsonline.de/Downloads/090910/Papers/Paper_Cook%20&%20Tanner.pdf.
14. Eurocontrol (2009) Standard Inputs for EUROCONTROL cost-benefit analyses. 4th edition, http://www.eurocontrol.int/ecosoc/gallery/content/public/documents/CBA%20examples/Standard_inputs_for_CBA_2009.pdf.
15. Eurocontrol, 2010. Performance review report: an assessment of air traffic management in Europe during the calendar year 2009. Performance Review Commission, http://www.eurocontrol.int/prc/gallery/content/public/Docs/PRR_2009.pdf.
16. Federal Aviation Administration, 2011. FAA aerospace forecast: fiscal years 2011-2031. http://www.faa.gov/about/office_org/headquarters_offices/apl/aviation_forecasts/aerospace_forecasts/2011-2031/media/2011%20Forecast%20Doc.pdf.
17. Geisinger, K., 1988. Airline delay: 1976-1996. Federal Aviation Administration, Office of Aviation Policy and Plans.
18. Gillen, D. W., Oum, T. H., Tretheway, M. W., 1985. Airline cost and performance: implications for public and industry policies. Center for Transportation Studies, University of British Columbia, Vancouver, Canada.

19. Gillen, D. W., Oum, T. H., Tretheway, M. W., 1990. Airline cost structure and policy implications: a multi-product approach for Canadian airlines. *Journal of Transport Economics and Policy*, Vol. 24, No. 1, pp. 9-34.
20. Government Accountability Office (GAO) (2008) Next generation air transportation system: status of systems acquisition and the transition to the next generation air transportation system. Report prepared for the Congressional Requesters, available at: http://aci-na.org/static/entransit/gao_nextgen_9-10-08.pdf, pp. 7.
21. Hansen, M. M., Gillen, D. W., Djafarian-Tehrani, R., 2000. Assessing the impact of aviation system performance using airline cost functions. *Transportation Research Record*, 1073, pp. 16-23.
22. Hansen, M. M., Gillen, D. W., Djafarian-Tehrani, R., 2001. Aviation infrastructure performance and airline cost: a statistical cost estimation approach. *Journal of Transportation Research*, Part E, Vol. 37, pp. 1-23.
23. Institute du Transport Aérien, 2000. Costs of air transport delay in Europe. Report prepared for the Eurocontrol Performance Review Unit. <http://www.eurocontrol.int/prc/gallery/content/public/Docs/stu2.pdf>.
24. Joint Economic Committee (JEC), 2008. Your flight has been delayed again: flight delays cost passengers, airlines, and the U.S. economy billions. http://jec.senate.gov/public/?a=Files.Serve&File_id=47e8d8a7-661d-4e6b-ae72-0f1831dd1207.
25. Jorge, J. D., De Rus, G., 2004. Cost-benefit analysis of investments in airport infrastructure: a practical approach. *Journal of Air Transport Management*, Vol. 10, No. 5, pp. 311-326.
26. Mayer, C., Sinai, T., 2003a. Network effects, congestion externalities, and air traffic delays: or why not all delays are evil. *American Economic Review*, Vol. 93, pp. 1194-1215.
27. Mayer, C., Sinai, T., 2003b. Why do airlines systematically schedule their flights to arrive late? Working paper, the Wharton School of Business, University of Pennsylvania. Available at: <http://real.wharton.upenn.edu/~sinai/papers/Schedule-Mayer-Sinai-10-30-03-2.pdf>.
28. Miller, B., Clarke, J-P., 2007. The hidden value of air transportation infrastructure. *Journal of Technological Forecasting and Social Change*, Vol. 44, No. 1, pp. 18-35.
29. Morrison, S., Winston, C., 1983. Estimation of long-run prices and investment levels for airport runways. *Research in Transportation Economics*, Vol. 1, pp. 103-130.
30. Morrison, S., Winston, C., 1989. Enhancing the performance of the deregulated air transportation systems. *Brookings Papers on Economic Activity: Microeconomics*, pp. 61-112.
31. Morrison, S., Winston, C., 2007. Another look at airport congestion pricing. *American Economic Review*, Vol. 97, No. 5, pp. 1970-1977.
32. Odoni, A., 1995. Research directions for improving air traffic management efficiency. Argo Research Corporation.
33. Oum, T. H., Yu, C., 1998. Cost competitiveness of major airlines: an international comparison. *Transportation Research Part A*, Vol. 32, No. 6, pp. 407-422.
34. Oum, T. H., Zhang, Y. 1991. Utilization of quasi-fixed inputs and estimation of cost functions. *Journal of Transport Economics and Policy*, Vol. 25, pp. 121-134.

35. Varian, H. R., 1992. Microeconomic analysis. 3rd Edition. W. W. Norton & Company, New York.
36. Xiong, J., Hansen, M., 2009. Value of flight cancellation and cancellation decision modeling: ground delay program postoperation study. Journal of the Transportation Research Record, No. 2016, Transportation Research Board of the National Academies, Washington, D.C., pp. 83-89.
37. Zou, B., Hansen, M., 2011a. Impact of operational performance on air carrier cost structure: evidence from U.S. airlines. Transportation Research Part E: Logistics and Transportation Review, forthcoming.
38. Zou, B., Hansen, M., 2011b. Flight delays, capacity investment, and social welfare under air transport supply-demand equilibrium. Transportation Research Part A: Policy and Practice, forthcoming.
39. Zou, B., Hansen, M., 2011c. Assessing Benefits from Aviation Capacity Investment: An Equilibrium Approach. Presented at the Institute for Operations Research and the Management Sciences (INFORMS) Transportation Science and Logistics Society Workshop, Asilomar, California, Jun. 27, 2011.