Journal of Air Transport Management 26 (2013) 40–43

Contents lists available at SciVerse ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

An application of survival and frailty analysis to the study of taxi-out time: A case of New York Kennedy Airport

Tony Diana*

Federal Aviation Administration, 800 Independence Avenue, SW Washington, DC 20591, USA

Keywords:
Survival analysis
Flight delay
Airport performance
Airport taxing

ABSTRACT

This study uses survival models to evaluate how selected operational factors affect the duration of aircraft taxi-out times at John F. Kennedy Airport, New York. Frailty models help assess whether fixed or random effects are likely to explain differences between two summers, 2006 and 2007. The hourly departure records for summer are censored when operations occurred below the airport's ceiling and visibility minima, that is, in instrument meteorological conditions. Cox regression models showed that block delay and the percent of airport utilized capacity are most likely to increase the risk of longer taxi-out times in instrument meteorological conditions compared with other factors such as departure delays, arrival delays and the volume of departures. Frailty analysis reveals that taxi-out times are not significantly affected by either fixed or random effects.

1. Introduction

This study compares taxi-out time between June and August 2006 and same period for 2007 at New York John F. Kennedy International Airport (JFK). Based on the Airline Service Quality Performance data (ASQP),[1] the percent of on-time gate departures declined from 73.1% in the summer 2006 to 64.6% in 2007, while taxi-out time increased from 36.2 to 43.46 min over the same period. Based on ASQP, 31,275 departures were reported by the US carriers in summer 2007, up 12.8% from summer 2006. Delta Air Lines and JetBlue increased their scheduled operations respectively 20.1% and 19.0% in summer 2007 compared with previous summer. On the other hand, departure cancellations went up from 535 in summer 2006 to 1218 in summer 2007. Overall, 10.65% of takeoff and landing operations in summer 2006 occurred in instrument meteorological conditions (IMC) compared with 11.81% in summer 2007. Instrument meteorological conditions at JFK are determined by a cloud ceiling lower than 2000 feet and visibility below 4 nautical miles.

While there was not a significant difference in the percentage of operations in IMC and despite more cancellations between the two summers, taxi-out time still went up. To understand the reasons for longer taxi-out times between the two time periods, we focus on the relationship between selected operational variables and taxi-out times in semi and parametric models as well as to determine whether the duration of taxi-out time can be attributed to fixed or random effects. Frailty models are extensions of the proportional hazards model, the most common model in survival analysis. Frailty models are used to determine whether differences in taxi-out times between summer 2006 and 2007 could be explained by unobserved individual departure-level factors otherwise unaccounted for by other survival models' estimates.

2. Survival and frailty analysis

Survival analysis, also called reliability analysis in the engineering field, represents a tool to compute time-to-event probabilities given specific circumstances such as the treatment for a disease in the medical field or component failure in engineering. It has not, however, been widely applied to the analysis of airport efficiency and airline delay.

Proportional hazard models assume that observations have hazard functions proportional to one another and they can accommodate a variety of shapes for the common hazard function

* Tel.: +1 410 564 0286.
E-mail addresses: Tony.Diana@faa.gov, tonydiana1@verizon.net.

[1] The data compiled by the US Bureau of Transportation Statistics are available at http://www.bts.gov as well as at https://aspm.faa.gov/asqp. ASQP statistics are based on the flights reported by the US carriers. Under 14 C.F.R Part 234 (Airline Service Quality Performance Reports), the reporting carriers are those certificated under 49 U.S.C. 41102 that accounted for at least 1% of domestic scheduled-passenger revenues in the 12 months ending March 31. A reportable flight is any nonstop flight, including a mechanically delayed flight, to or from any airport within the contiguous 48 states that accounts for at least 1% of domestic scheduled-passenger enplanements in the previous calendar year, as reported to the Department pursuant to part 241 of this title.

0969-6997/$ – see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.jairtraman.2012.10.002

across observations. Presently, Cox regression models determine whether the risk of longer taxi-out times is likely to increase with block, arrival and departure delays, the volume of departures, and the percentage of airport capacity utilized. By contrast, accelerated failure time model are parametric and their probability distribution is specified. The Weibull distribution is selected here for the following reasons: (1) it can be used with smaller samples, (2) all 'failure' modes fit in Weibull and (3) it can be used to derive important information such as the shape parameter, the characteristic life, the failure percentage, and the failure-free time. Accelerated failure time models are appropriate because they focus on the relationship between taxi-out time and other covariates likely to influence the duration of taxi-out operations. Parametric accelerated time to failure models examine the multiplicative, proportional, effects of block delay, departure delay, the percent of airport capacity utilized, arrival delay and the volume of departures with respect to the time it takes for an aircraft to takeoff after leaving the gate. In frailty models, fixed or random effects have a multiplicative effect on the hazard function.

As a special case of survival models, frailty models can prove instrumental in introducing variability among the sampled flights unaccounted for in other survival models. First, in time-to-event or duration models, "failure to account for unobserved heterogeneity causes the estimated hazard rate to decrease more with the duration than the hazard rate of a randomly selected member of the population" (Woutersen, 2007). Second, departing flights cannot be construed as homogeneous due to differentiating factors such as traffic mix at specific times of the day, airline policies that regulate taxi operations speeds and gate location that affects the time from gate departure to take-off, among others. These attributes are difficult to measure and integrate into parametric models. Finally, frailty models can account for the two sources of variation in time-to-event models: variability originating from observable risk factors and heterogeneity caused by unknown covariates (Hougaard, 1991).

Frailty models include random components and variability in unobserved factors. They are thus appropriate for time-to-event or duration model data such as taxi-out time where the random effect has a multiplicative weight on the baseline hazard function. The models assume that the population is not homogeneous: Not all the departures in the samples are subject to the same risk of longer duration. Here, risks can be related to airport operations (i.e. availability of gates) or conditions (i.e. ramp congestion, available airport capacity, peak hourly traffic, poor weather conditions, delays) that 'weaken' on-time departure performance and capacity utilization, hence the term 'frailty' coined by Vaupel et al. (1979).

The survival and frailty models may be of interest to aviation analysts and practitioners for the following reasons. First, they can help them evaluate how airport congestion is likely to arise as slower moving aircraft in the ramp area and taxiways are likely to slow down the departure flow. Second, they are designed to provide some indications as to what variables may influence the duration of taxi-out times. Finally, improved taxi operations represent a key component in the NextGen Implementation Plan.[2] The 'Improved Surface Operations' portfolio includes operational increments such as 'initial surface traffic management' and 'enhanced surface traffic operations' to improve arrival and departure throughput, as well as to accelerate departure sequencing to help schedule adherence and minimum wait time in departure and takeoff queues.

[2] The 2012 NextGen Implementation Plan is available at the following website: http://www.faa.gov/nextgen/implementation/media/NextGen_Implementation_Plan_2012.pdf.

3. Methodology

The summers of 2006 and 2007 embrace June, July, and August. The summer season is selected to measure changes in taxi-out times because traffic is at peak due to the vacation season, and surface movements are likely to be disrupted by thunderstorms. The observations pertained to the operating hours of 07:00 to 21:59 (local time). To determine the impact of heterogeneous factors, taxi-out times are censored by the existence of instrument meteorological conditions. In summer 2006, 43.12% of the 1250 observations are censored (IMC $= 1$) compared with 43.44% in summer 2007.

The data originated from ARINC's Out-Off-On-In[3] times compiled by the US Bureau of Transportation Statistics (BTS) and available in the Aviation System Performance Metrics (ASPM) data warehouse. The variables used in the parametric and non-parametric models are defined as follows:

- *Taxi-Out Time* is measured as the average minutes elapsed between gate departure and takeoff.
- *Block Delay* represents the difference in minutes between actual and scheduled gate-out to gate-in times.
- *Departure Delay* is the difference in minutes between actual and scheduled gate-out times.
- The *Percentage of Capacity Utilized* is computed as the number of operations (arrivals and departures) divided by the sum of the airport arrival rates (AAR) and airport departure rates (ADR). Each facility provides the called rates on a daily basis to the Air Traffic Control System Command Center. The rates reflect airport configurations, meteorological conditions, and estimated volume of traffic.
- *Arrival Delay* is the difference in minutes between actual and scheduled gate-in times.
- *Departures* are the number of operations (arrivals and departures) reported by the major US carriers in ASQP.

Three SAS procedures are used to analyze the data. First, the LIFETEST procedure provided non-parametric estimates of the survival distribution function for summer 2006 and 2007. It is used to obtain the product-limit or the life-table estimates of the distribution. Second, the LIFEREG procedure enabled to fit the parametric models to failure time, the parameters by maximum likelihood and the standard errors of the accelerated failure time models. The parameters of the accelerated failure time model are estimated using regression techniques in which the response variable is the logarithm or a known monotone transformation of a failure time (Kalbfleisch and Prentice, 1980). Third, the NLMIXED procedure is used to fit the accelerated time model and to generate the cumulative distribution function of taxi-out times. It is also designed to generate nonlinear mixed effects models.

4. Results

4.1. Cox models

The Cox proportional hazard models using the SAS procedure (PHREG) do not make any assumption about the shape of the underlying hazards. Hazards for summer 2006 and 2007 are assumed to be proportional over time. The results are seen in Table 1.

Taking the example of the percent of airport capacity utilized in summer 2006 in IMC, an holding other variables constant, the

[3] ARINC's website is http://www.arinc.com.

Table 1
Proportional hazard model for summers 2006 (top) and 2007 (bottom).

Parameter	IMC = 1			IMC = 0		
	Parameter estimates	Pr > ChiSq	Hazard ratio	Parameter estimates	Pr > ChiSq	Hazard ratio
Summer 2006						
Block delay	1.6759	<0.0001	5.344	0.0050	0.9282	1.005
Departure delay	−0.8298	<0.0001	0.436	−0.0401	0.3698	0.961
Capacity utilized	0.8212	<0.0001	2.273	0.0309	0.4838	1.031
Arrival delay	−0.8183	<0.0001	0.441	−0.0837	0.0773	0.920
Departures	−0.7934	<0.0001	0.452	−0.0723	0.1428	0.930
Summer 2007						
Block Delay	1.7457	<0.0001	5.730	0.0399	0.4640	1.041
Departure delay	−0.9039	<0.0001	0.405	−0.0538	0.2239	0.948
Capacity utilized	0.8326	<0.0001	2.299	−0.0556	0.2340	0.946
Arrival delay	−0.8414	<0.0001	0.431	−0.0804	0.0902	0.923
Departures	−0.9151	<0.0001	0.400	−0.0516	0.2960	0.950

average minutes of taxi-out time is 82.1% higher for each 1% increase in the airport capacity used. The percentage slightly increased to 83.26% in summer 2007. In the case of departure delay, the average minutes of taxi-out time in IMC are 83.0% lower for each 1% increase in the average minutes of departure delay in summer 2006 compared with 90.4% in summer 2007. Block delay in IMC had the greatest impact on the hazard function as measured by the magnitude of the parameter estimates. While all the parameter estimates are significant in IMC, none of them is in VMC for both summers at a 95% confidence level.

The hazard ratios provide an indication of the risk of longer taxi-out times: In summer 2006 and 2007, the estimated taxi-out times are likely to increase fivefold in IMC if block delay increased by 1%, holding the effects of other dependent variables constant. The estimated taxi-out times are also likely to increase, by 227% in IMC if the percent of capacity utilized increased by 1%, holding the effects of other dependent variables constant.

4.2. Accelerated failure time models

The LIFEREG procedure is used for the parametric accelerated failure time models. The assumption is that survival times accelerates or decelerates by a constant factor. In this study, the Weibull distribution is used to model survival data for the following reasons:

- If the accelerated failure time assumption holds, then the proportional hazard assumption also holds.
- The log−log of the survival function is linear with the log of time.

With SAS, accelerated failure time models use Weibull and exponential distributions. When based on a Weibull distribution, the hazard function can be described as follows:

$$h(t,\beta) = \gamma\alpha(\alpha t)^{\gamma-1}\text{and } \alpha = \exp\{-x'\beta\} \tag{1}$$

Based on Cox and Oakes (1984), the distribution of survival past time can be computed as;

$$g(t,\beta) = \exp\{-(\alpha t)^{\gamma}\} \tag{2}$$

The outputs from the LIFEREG models are displayed in Table 2.

An examination of the outputs suggests three things. First, all the independent variables had a significant impact in IMC, but not in VMC—whether in summer 2006 or 2007. Second, if the Weibull shape or slope is less than one—as it is the case in summer 2006 and 2007, then the 'failure' rate decreases with time. In summer 2006 and 2007, the estimated median taxi-out time is three times as high in IMC as in VMC as derived from the exponential value of the scale parameter. Moreover, it does not differ much in IMC when comparing both summers. Third, for every 1% increase in capacity utilized in IMC, there is a corresponding change of 62.7% in the expected duration of taxi-out time in summer 2006 compared with 61.1% in summer 2007. Similarly, for every 1% increase in block delay in IMC, there is a corresponding change of 86.6% in expected duration of taxi-out time in summer 2006 compared with 86.4% in 2007.

4.3. Frailty models

The NLMIXED procedure is used to derive frailty models with and without random effects; outputs are displayed in Table 3.

In the fixed-effect model, the estimate for gamma and b0 are significant at 95% confidence level as opposed to b1 These parameters allow the computation of the distribution of survival past time t represented by $G(t,\beta) = \exp\{-(\alpha t)^{\gamma}\}$. Therefore, if

$$\alpha(\text{summer 2006}) = \exp\{-3.4978 + 0.0664\} = 0.0323 \tag{3}$$

Table 2
Accelerated failure time model for summers 2006 (top) and 2007 (bottom).

Parameter	IMC = 1		IMC = 0	
	Estimates	Pr > ChiSq	Estimates	Pr > ChiSq
Summer 2006				
Intercept	−6.0446	<0.0001	−5.5230	<0.0001
Block delay	−2.0140	<0.0001	−0.0082	0.8802
Departure delay	1.0012	<0.0001	0.3610	0.4092
Capacity utilized	−0.9856	<0.0001	−0.0272	0.5286
Arrival delay	0.9801	<0.0001	0.0827	0.0738
Departures	0.9486	<0.0001	0.0659	0.1689
Scale	1.1892		0.9937	
Weibull shape	0.8409		1.0166	
Summer 2007				
Intercept	−6.0642	<0.0001	−5.5733	<0.0001
Block delay	−1.9959	<0.0001	−0.0301	0.5671
Departure delay	1.0388	<0.0001	0.0491	0.2515
Capacity utilized	−0.9433	<0.0001	0.0558	0.2220
Arrival delay	0.9688	<0.0001	0.0771	0.0926
Departures	1.0437	<0.0001	0.0521	0.2763
Scale	1.1434		0.9768	
Weibull shape	0.8746		1.0237	

Table 3
The fixed and random-effects models.

Fixed effects			Random effects	
Parameter	Estimates	Pr > ChiSq	Estimates	Pr > ChiSq
Gamma	3.4978	<0.0001	12.3582	–
b0	3.9305	<0.0001	3.8292	<0.0001
b1	−0.0696	0.2525	−0.06537	0.2192
Logsig			−1.2527	<0.0001

$$\alpha(\text{summer } 2007) = \exp\{-3.4978\} = 0.0302 \tag{4}$$

then the probability that taxi-out time will reach at least 30 min is computed as:

$$1 - G\left(t, \text{summer } 2006\right) = 1 - \exp\left\{-(0.0323*30)^{3.4978}\right\} = 0.5917 \tag{5}$$

$$1 - G\left(t, \text{summer } 2007\right) = 1 - \exp\left\{-(0.0302*30)^{3.4978}\right\} = 0.5074 \tag{6}$$

In the fixed-effects failure time model, the '−2 log likelihood' output is 1000.68 compared with 963.1 in the random frailty model. Logsig represents the log of the standard deviation of random departure effects. The variance of the taxi-out random effect is 0.08, that is significant at a 95% confidence level. Since the *p*-value is greater than 0.05 in the fixed-effects and random-effects models, we cannot infer that IMC leads to significantly longer taxi-out times whether effects are fixed or random.

5. Final remarks

Surface movement efficiency is attracting more attention from airport and airline analysts because it impacts on-time performance as well as fuel burned in ground surface movements. We assume that the duration of taxi-out time is a function of several factors that may not be expressed in parametric and semi-parametric models. As a result, survival and frailty analytical models help us to understand the factors that may affect the hazard of longer taxi-out times as well as those that may not be expressly included in the parametric and in the Cox models. The analysis shows that block delay and a high percent of airport capacity utilized are likely to impact the duration of taxi-out time in instrument meteorological conditions. The study also suggests that there is no significant impact of fixed and random effects on taxi-out time in summer 2006 and 2007.

Note

This article does not reflect the opinion of the Federal Aviation Administration.

References

Cox, D.R., Oakes, D., 1984. Analysis of Survival Data. Chapman and Hall/CRC Press, Boca Raton.

Hougaard, P., 1991. Modelling heterogeneity in survival data. Journal of Applied Probability 28, 695–701.

Kalbfleisch, J., Prentice, R., 1980. The Statistical Analysis of Failure Time Data. John Wiley & Sons, New York.

Vaupel, J.W., Manton, K.G., Stallard, E., 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16, 439–454.

Woutersen, T., 2007. Testing for Heterogeneity in Duration Models: a Semi-parametric Test. Centre for Microdata Methods and Practice (CEMMAP). April. http://www.cemmap.ac.uk/lancaster/woutersen.pdf.