Survival analysis of new intra-European scheduled air services

R. Klophaus[a,*], M. Viehmann[a]

[a] Competence Center Aviation Management (CCAM), Worms University of Applied Sciences, Germany
klophaus@hs-worms.de, viehmann@hs-worms.de
* Corresponding author

Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI)
June 19-25, 2022, Mauritius Island

March 31, 2022

Keywords: Airline network management, air services, survival analysis, Cox regression

1 INTRODUCTION

Year after year, airlines offer new air services, some of them in route competition with other carriers. Planning and launching scheduled air services requires time and money of the operating carrier but also at the served airports. There might be substantial support through tax payers' money as well. Nevertheless, new air services are often abandoned within the first years. Some are discontinued shortly after the start even without exceptional events such as the COVID pandemic, although the respective carrier as a whole continues to operate or competing services by other carriers sustain on a given route. Accordingly, the important question for airlines, airports and other stakeholders arises as to the key factors of service survival. The present paper addresses this question. We analyze the survival of direct services on intra-European routes opened from 2013 to 2018. Our aim is to examine how airline, airport, route and service attributes affect the probability of service survival over time. Covariates include the airline business model and route competition. Our empirical analysis ends in 2019, the last regular year for the aviation industry before the COVID crisis led to an exceptionally high number of suspended services (Suau-Sanchez *et al.*, 2020).

Research on airline competition and market entry following the deregulation of passenger airline markets (e.g., Borenstein, 1992) do not consider the survival probability of individual direct services. Same holds for applications of graph theory to airline networks (e.g., Roucolle *et al.*, 2020). With a case study approach, Lohmann & Vianna (2016) examined how aviation and non-aviation factors affected route suspension in Australia between 2008 and 2013. de Wit & Zuidberg (2016) estimated discrete choice models of route termination in the networks of the four largest European low-cost carriers. They found that distance, the number of seats offered, market share, seasonality and route age have a significant effect on the chance of route termination. Calzada & Fageda (2019) mainly investigated route expansion in the European air transport market in the period 2002 to 2013, but also considered a sample of active routes and the factors determining their discontinuity using a logit model. More recently, Manello *et al.* (2021) linked route closures to the economic development of European regions using socio-economic and airport characteristics, however, without specification of airline or service attributes.

Our aim is to fill a gap in the literature by estimating probabilities of service survival on intra-European routes with a set of airline, airport, route and service characteristics. We estimate variants of a Cox proportional hazard model with data from the OAG airline schedules database. Among other findings, we derive a significant dependence of the service survival rate on the airline business model. Our results should be of particular interest for airline and airport managers not to invest time and money in the launch of scheduled air services that are more likely to fail than others.

2 METHODOLOGY

2.1 Model Formulation

The Cox proportional hazards model (Cox, 1972) makes a parametric assumption concerning the effect of the covariates on the hazard function, but no assumption regarding a particular form of the hazard function itself. This approach suits us as we are primarily interested in the effects of the covariates, not in the shape of the hazard function itself. The hazard function for the Cox proportional hazards model can be specified as:

$$\lambda(t|X) = \lambda_0(t)\exp(X\beta) = \lambda_0(t)\exp(\beta_1 X_1) \cdots \exp(\beta_k X_k) \quad (1)$$

This expression provides the hazard function for air services at time t dependent on the covariate vector X of k explanatory variables. The hazard function denotes the rate or probability of an event (termination of a service) in the interval of infinitesimal length $[t, t + \Delta t]$ conditional on the covariates X. The model assumes multiplicative effects of the covariates on the baseline hazard function $\lambda_0(t)$, i.e., the baseline hazard function that is time-dependent does not depend on the covariates itself. The influence of the covariates is expressed by the hazard ratios $\exp(\beta_j)$ that are assumed to be constant over time. If β_j is estimated close to 0, the hazard ratio $\exp(\beta_j) \approx 1$, i.e., the covariate j is predicted to have no significant influence on the hazard rate. Similarly, $\beta_j > 0$ results in a hazard ratio $\exp(\beta_j) > 1$ and the predictor increases the risk of an event. For a covariate predicted to decrease the risk or increase survival probabilities, $\beta_j < 0$ or $\exp(\beta_j) < 1$.

The Cox model is a standard method, e.g. in medical studies to analyze the effects of different covariates such as different treatments or physical characteristics. The subjects in our case are non-stop air services between two airports provided by a specific airline. The subjects are followed up regularly during the study with their status and their characteristics being recorded. We are interested in the event of termination of the air service by the respective carrier. After the event, the subjects are no longer followed-up. Hence, we do not consider services interrupted for more than one year. We check for the first week in October each year whether a service is continued or terminated. In order to exclude the special effects of the ongoing COVID crisis, we only record schedule data until 2019. We are therefore making censored observations, as we cannot observe the further survival of the services that were still active in October 2019. As we examine new air services opened from 2013 to 2018, truncation is of no relevance.

We calibrate different models to evaluate model quality and potential differences in termination risk stemming, for example, from business models or specifics of individual airlines. First, we fit a single model including all business models (mainline and low-cost) and analyze effects of the covariates commonly shared between different carriers (Model 1). We expect the airlines' particularities to influence the risk but also expect differences due to the airlines' business models. In a stratified model, a common extension to standard Cox models, we estimate separate baseline hazards $\lambda_{0,1}(t)$, $\lambda_{0,2}(t)$, ... for each level of the stratification variable:

$$\lambda_s(t|X) = \lambda_{0,s}(t)\exp(X\beta) = \lambda_{0,s}(t)\exp(\beta_1 X_1)\ \cdots\ \exp(\beta_k X_k), \tag{2}$$

where $\lambda_s(t|X)$ is the hazard function for subjects belonging to stratum s. While the baseline hazard depends on the stratum, the proportional hazards do not. Hence, we model specifics of the homogeneous group in a stratum in the baseline hazard and assume that effects of the covariates influence the hazard independently of the stratum. We build models that include the business model and the carriers as stratification variables (Model 2 and Model 3). In order to analyze potentially different effects of the predictors, we continue to build separate models that only contain subsets of a single business model (Model 4 / mainline and Model 5 / low-cost). To account for effects particular to individual airlines, we further stratify the airlines in each subset in Model 6 (mainline) and Model 7 (low-cost).

2.2 Data and Covariates

Our data contains 18,836 records of 5,820 distinct intra-European city pairs (non-directional), which are operated by 100 airlines in continuous operation during the entire study period up to the year 2019. 18 of these airlines are categorized in the OAG database as airlines with a low-cost business model (3,753 city pairs), while the other 82 airlines are classified as mainline carriers (2,587 city pairs). Table 1 provides an overview of the covariates analyzed.

Table 1 – *Description of Covariates*

Variable	Description
BusModel	Airline business model (mainline or low-cost)
AirlineSize	Airline size as available seat kilometers offered world-wide (ASKs, in millions)
Distance	Route distance in NM
Frequency	Service frequency offered by respective airline
CapacityShare	Market share of the respective carrier on city pair by seat capacity
ArrApFreq	Airline presence at the arrival airport (# frequencies)
DepApFreq	Airline presence at the departure airport (# frequencies)
ArrApFreqShare	Airline relative presence at the arrival airport (share of all frequencies)
DepApFreqShare	Airline relative presence at the departure airport (share of all frequencies)
LargerApFreqTotal	Airport size in terms of total frequencies (world-wide) for the larger airport
SmallerApFreqTotal	Airport size in terms of total frequencies (world-wide) for the smaller airport

3 RESULTS AND DISCUSSION

The results of the different models are detailed in Table 2. We state the proportional hazards estimated and additionally provide the significance levels. Model 1 shows a significantly increased likelihood of service termination when airlines follow a mainline instead of a low-cost business model. This is a rather surprising result, considering that low-cost carriers' footloose strategies have been emphasized in the academic literature (e.g., Graham, 2013). Mainline carriers may close unprofitable new services early before they become established with increasing demand. Airline size is not a differentiator, a somewhat counter-intuitive finding as well. However, this might largely depend on our assumption to only examine new air services by carriers with ongoing operations. Due to modal competition, increasing route distance has a positive impact on the survival probability, a finding in line with de Wit & Zuidberg (2016) and Manello *et al.* (2021). Frequency has a positive impact as well. Route competition between carriers measured by capacity shares, decreases termination risk only in Model 5, i.e., for low-cost carriers. Hubbing effects for mainline carriers strongly increase service survival as shown in airlines' relative presence at either endpoint. Effects of the absolute size of the respective arrival and destination airports are limited.

Table 2 – *Estimated Proportional Hazards*

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
BusModel	1.9634 (***)	NA	NA	NA	NA 0	NA	NA
AirlineSize	1.0000 (*)	1.0000 (**)	NA	1.0000	0.9999 (***)	NA	NA
Distance	0.9995 (***)	0.9995 (***)	0.9995 (***)	0.9995 (***)	0.9996 (***)	0.9995 (***)	0.9995 (***)
Frequency	0.9405 (***)	0.9406 (***)	0.9339 (***)	0.9465 (***)	0.9304 (***)	0.9358 (***)	0.9330 (***)
CapacityShare	0.9503	0.9542	0.9611	1.0536	0.8154 (**)	0.9673	0.9448
ArrApFreq	0.9992 (***)	0.9992 (***)	0.9993 (***)	0.9994 (***)	0.9990 (***)	0.9996 (***)	0.9989 (***)
DepApFreq	0.9992 (***)	0.9992 (***)	0.9992 (***)	0.9996 (***)	0.9984 (***)	0.9996 (***)	0.9985 (***)
ArrApFreqShare	0.8186 (**)	0.8145 (**)	0.9264	0.5249 (***)	1.4594 (**)	0.6924 (**)	1.2159
DepApFreqShare	0.9174	0.9146	1.1882	0.4248 (***)	2.6031 (***)	0.6818 (**)	1.9685 (***)
LargerApFreqTotal	1.0001 (***)	1.0001 (***)	1.0001 (**)	0.9999 (***)	1.0002 (***)	1.0000	1.0001 (***)
SmallerApFreqTotal	1.0000	1.0000	1.0000	0.9998 (**)	1.0001	0.9998 (***)	1.0001
Observations	18,836	18,836	18,836	7,178	11,658	7,178	11,658
Events	2,697	2,697	2,697	1,393	1,304	1,393	1,304

Significance levels shown in parentheses, $p<0.01$ (***), $p<0.05$ (**), $p<0.1$ (*)

REFERENCES

Borenstein, S., 1992. The evolution of US airline competition. *Journal of Economic Perspectives*, **6**(2) pp.45-73.

Calzada, J. & Fageda, X., 2019. Route expansion in the European air transport market. *Regional Studies*, **53**(8) pp. 1149-1160.

Cox, D. R. 1972. Regression Models and Life-Tables (with Discussion). J*ournal of the Royal Statistical Society*, Series B: Methodological **34** pp. 187-220.

de Wit, J.G. & Zuidberg, J., 2016. Route churn: an analysis of low-cost carrier route continuity in Europe. *Journal of Transport Geography*, **50** pp. 57-67.

Graham, A. 2013. Understanding the low cost carrier and airport relationship: a critical analysis of the salient issues. *Tourism Management* **36** pp. 66-76.

Lohmann, G. & Vianna, C., 2016. Air route suspension: The role of stakeholder engagement and aviation and non-aviation factors. *Journal of Air Transport Management*, **53** pp. 199-210.

Manello, A., Scotti, D. & Volta, N., 2021. Air connection dropouts and isolation risks across European regions. *Regional Studies*, **56** pp. 1-12.

Roucolle, C., Seregina, T. & Urdanoz, M., 2020. Measuring the development of airline networks: Comprehensive indicators. *Transportation Research Part A: Policy and Practice*, **133** pp. 303-324.

Suau-Sanchez, P., Voltes-Dorta, A. & Cugueró-Escofet, N., 2020. An early assessment of the impact of COVID-19 on air transport: Just another crisis or the end of aviation as we know it? J*ournal of Transport Geography*, **86** p.102749.