Logistics Research (2019) 12:6
DOI_10.23773/2019_6

Airline business models and their network structures

Anne Lange, Tobias Bier

Received: 01 June 2018 / Accepted: 27 May 2019 / Published online: 24 June 2019

ABSTRACT

Network structure is one parameter that feeds into airline business models. We analyze the networks of 58 European airlines including full service carriers, low cost carriers, regional and charter airlines. Eight network metrics are calculated to describe the various aspects of network structure. A principal component analysis is conducted and indicates two components in the metrics. The components address network coverage and service network. The airline networks are clustered based on the identified components. Findings from the analysis include that the resulting clusters based on only network structure appear to be consistent groups of airline business models. It indicates that only judging from network structure allows to reason on airline business models. In more detail, full service carriers are structured in three subgroups differing in coverage as well as in the operated services. At the same time, low cost carriers, charter and regional airlines appear as clusters in the analysis. A few airlines are identified as outliers and investigating their business model confirms this network observation.

KEYWORDS: metrics · graph theory · principal component analysis · airline business model · network structure

"This article is part of a focus collection on "Understanding Future Logistics" based on the BVL's 9th International Scientific Symposium on Logistics in Magdeburg in 2018"

Anne Lange
University of Luxembourg, Luxembourg Centre for Logistics and Supply Chain Management,
162A, avenue de la Faïencerie,
L-1511 Luxembourg, Luxembourg,
anne.lange@uni.lu,
Tel: +352 46 66 44 6308

Tobias Bier
TU Darmstadt,
Department of Law and Economics,
Hochschulstr. 1,
64289 Darmstadt, Germany

1 INTRODUCTION

Airlines provide air transportation service to passengers and cargo. To do so, they operate a fleet of aircraft and make use of them to transport passengers and cargo from one place to another. The underlying route network is essential for the operations. The route network is the entirety of flights offered by an airline. Hence, it is the backbone of the operations. Given that aircraft and fuel expenses constitute major costs for airlines, route network decisions are tightly linked to the cost of operations for an airline. At the same time, the destinations, connecting services, flight frequencies and travel time offered by an airline define the essential product offered by the airline to its customers.

Airline business models explain the underlying rationale on how an airline provides service to its customers and airline network structure is an established element of airline business models. One common assumption is that Full Service Carriers (FSC) tend to operate hub and spoke networks (HS) whereas and Low Cost Carriers (LCC) operate point-to-point networks (PP).

The structure of transportation networks has been studied in recent years and advancements have been made on describing various aspects of transportation network structure quantitatively. For instance, Hu and Zhu [1] describe maritime networks, Wang and Cullinane [2] work on seaport centrality, Dobruszkes [3] illustrates LCC network structures, and Mishra et al. [4] suggest metrics for the connectivity of urban public transport networks. Mattsson and Jenelius [5] review research on transportation network vulnerability, one of the typical application areas where network structures need to be quantified.

Thus far, these two streams of research are not yet well connected. That is, on the one hand there are extensive metrics to account for transportation network structure. On the other hand, airline networks are described on a rather high level when it comes to studying airline business models. Our work is positioned at this interface. We aim to argue that a more detailed operationalization of airline network structure benefits the description of airline business models. Thus, we explore the research question in how far a quantitative assessment of airline network structures supports the clear description of airline business models. We focus solely on exploring the information contained in the airline network structure. In doing so, we contribute findings to airline business model research regarding the description of airline networks and wish to support the development of more clear-cut models of airline business models.

Our approach is novel for airline business model analysis: We calculate established metrics of airline network structure and apply a principal component analysis in order to reduce the dimensions of the dataset. This allows us to assess airline network structure from multiple perspectives and account for the correlation of the different metrics. When clustering the airlines according to the network structures alone, similarities between airlines already appear that relate well to similarities in airline business.

The structure of the remainder is as follows: Section 2 provides the necessary theoretical background on transportation network structure and airline business models. Section 3 summarizes the analyzed dataset of 58 airlines and presents the metrics from graph theory used. It further outlines the principal component analysis we use to reduce the dimensionality in the collected data and shows how we apply a cluster analysis to detect airlines of similar network structure. Section 4 presents the interpretation of the factor loadings as well as the findings from the cluster analysis. Section 5 discusses the findings and compares to prior research. Section 6 concludes.

2. THEORETICAL BACKGROUND

Choosing a suitable network structure is a complex yet important task for an airline and it is not surprising that it has received attention in the literature. Early contributions address the issue from the perspective of changes in network structure as related to the liberalization of aviation markets. The economic literature was interested in understanding the formation of HS in deregulated markets (see, for instance, Oum et al. [6]). Pros and cons of HS as compared to PP have been studied (see, e.g., [7], [8]). More recent work results from the observation that FSC and LCC do, among others, compete based on their network structure (e.g., [9]). Obviously, competition not only results from the network structure. Babić and Kalić [10] argue that it is the interplay of network structure and pricing policy that airlines compete on, whereas Brueckner [11] and Brueckner and Flores-Fillol [12] factor in average flight delay.

There is a wide spectrum of potential airline network structures, extending far beyond the archetypes of HS and PP [13]. Thus, describing them in more detail is relevant. Airline network structures have been quantified by graph theoretic metrics commonly used in social network analysis. The seminal work by Guimerá and Amaral [14] analyzes the worldwide air transportation network and showed that it belongs to the classes of scale-free and small world networks, indicating an efficient connectivity. To do so, the authors study (scaled) node degree and node betweenness centrality. Similar approaches have been applied to national networks, for instance by Guida and Maria [15] for the Italian network (node degree and node betweenness) and by Wang et al. [16] for the Chinese network (average path length, clustering coefficient, degree distribution, node degree, node closeness and node betweenness). Alderighi et al. [17] provide a set of metrics to study the structure of airline networks. This set includes the Gini index and network betweenness for spatial analysis. In their literature review, Lordan et al. [18] highlight that the analysis of airline network structure has received limited attention from airline management research. Instead, research has rather been conducted mainly from the complex network point of view. Some examples of studies in airline management research do exist, however. For instance, Reggiani et al. [19] analyze the network of Lufthansa as well as that of Star Alliance and convincingly highlight the network impact of strategic choices made by Lufthansa and Star Alliance. The authors make use of a large set of metrics, including node degree, node closeness, node betweenness, diameter, clustering coefficient, Gini index, network betweenness, and entropy function. Bowen [20] describes the cargo networks of UPS and FedEx. He includes an analysis of the network structures by studying the number of nodes, the number of edges, the beta and the gamma indices. Analyzing airline network structure is of interest to various disciplines and certain graph theory metrics have repeatedly been used to do so. However, no standard set of well-established metrics exists for this purpose. Further, the multitude of metrics used to account for the structure of airline networks oftentimes provides comparable yet not identical insights into the structure.

Since airline network structure is essential for the service and cost of airline operations, it is an element in airline business models. Passenger airline business models have traditionally been categorized into four groups: full service carriers, low cost carriers, regional carriers and charter airlines. It is not trivial to clearly distinguish between FSC and LCC since the business models have been evolving. Button and Ison [21] point out that LCC, in general, implement specific strategies that differentiate them from FSC. Notably, LCC offer a limited range of service in their basic fares as compared to FSC. That is, they strip the core service down to the transportation and offer additional services at an extra charge, creating additional revenues. Service is focused on a single booking class and bookings are often possible only online. Airports served tend to be second-tier airports. Further, LCC focus on short aircraft ground times and reduce cost by keeping the variation in aircraft types in the fleet low. Moreover, they make use of their market power towards suppliers to create efficiencies in sourcing [21]. It is common to associate LCC with PP and FSC with HS. As such, LCC and FSC differ in both their service offer and their cost positions and compete against each other. Regional carriers are active only in a restricted geographic area. Charter airlines do not provide their service directly to passengers, but rather to wholesale intermediaries such as travel and tourism companies. The operated routes are requested by the intermediaries. However, as competition increases in the challenging airline industry, the boundaries of airline business models are blurring and business models are evolving [22-25].

Table 1: Network aspects in airline business models

Reference	**Type**	**Structural network aspects touched upon in identifying airline business models**
Bachwich and Wittman [23]	N	Stage length
Casadesus-Masanell and Ricart [27]	N	Stage length
Daft and Albers [28]	FQ	Stage length, service frequency
Dobruszkes [3]	FN	Gamma connexity index, number of airports served, number of routes operated
Dobruszkes [29]	FN	Number of cities served, number of routes operated, centralization of the network
Gillen and Gados [25]	N	Stage length, geographic coverage, service frequency
Klophaus et al. [30]	FQ	Point-to-point service
Lohmann and Koo [31], Jean and Lohmann [32], Moir and Lohmann [33]	FN	Network density, number of destinations, service frequency, stage length
Mason and Morrison [34]	FN	Network density, number of routes, service frequency
O'Connell [35]	N	Hub-and-spoke operations
Pereira and Caetano [36]	FQ	Network architecture
Soyk et al. [24]	FQ	Network concentration, service frequency
Urban et al. [37]	FQ	Network system (Point-to-point, Hub-and-spoke, Multi-hub, Point-to-point & Hub-and-spoke), geographical coverage
Wensveen and Leick [26]	FQ	Stage length, number of destinations, frequency

Types: (FN) Framework mostly quantitative, (FQ) Framework mostly qualitative, (N) Narrative

Passenger airline business models describe and explain how passenger airlines create and capture value. Researchers do so by following various approaches. Table 1 highlights, how the aspect of *network* figures in the description of airline business models. It is striking that *network* is typically considered at a simplified level, taking into account one or very few features of an airline network. Stage length appears to be the most prominent feature included. Dobruszkes [3, 29] addresses the airline network at a more granular level. For one, he takes into account the gamma connexity index as well as the centralization of the network, both of which relate to graph theory. Moreover, his understanding of the operating network goes far beyond its structure. In addition to the features in Table 1, he includes parameters on how far the airline network relies on the 5th through 9th freedom of the air, in how far destinations tend to be 'warm water' destinations, the rivalry at the airports and the share of international flights. In doing so, he consciously adds additional elements that describe the nature of the considered airline networks but that are not covered by graph theory metrics. These elements are specific to airline networks and provide the analyst with more insights.

3. METHODOLOGY

3.1. Dataset

We collected flight data to be analyzed from flightradar24.com. Our analysis is based on the information of flights that were operated by the individual airlines during the observation period. The data we collected originates from Automatic Dependent Surveillance-Broadcast (ADS-B) technology. Its benefit for our purposes is that it retraces aircraft as they move in the air and on ground, providing us with the information of operated flights. It can be used to observe and analyze flight trajectories (e.g. [38, 39]), a feature that we do not even need to employ. We cross-checked a sample of the collected information with actual flight schedules and are confident that the information we collected accurately reflects flight operations. Budd [40] has compared the reliability of ADS-B based information with that of proprietary airline information and classifies this data source as an "innovative and welcome source of empirical data". Other studies have made use of the commercial OAG database before for this purpose (e.g. [13, 37, 3, 29]). The OAG database contains data on scheduled flights and is known to be reliable. However, since OAG is built around flight schedules, charter airline flights are not included in the data provided by OAG [29]. We opted to use the freely available ADS-B based flight information because it covers all operated flights, at a low cost of data collection and because we are confident that the data collected is accurate.

The dataset contains 134652 flights from 58 European airlines during the week of November 19-26, 2017. Airlines were selected based on their fleet size. In order to draw conclusions regarding network structure, the airlines need substantial operations. Azur Air and TAP Express operate the smallest fleet in the sample with 22 aircraft each. We include an airline if its headquarters is located in a country on the European continent. Hence, we include Russian and Turkish airlines in the sample. Table A.1 in the Appendix lists the airlines included as well as their fleet size and country of origin.

For each flight, data on the origin, destination and operating date were collected. This information is used to reconstruct the as-operated flight networks of the airlines. The networks are interpreted as graphs where the nodes are the individual airports and directed edges represent the connections of each airline between two airports in the timespan. The arcs are weighted by the number of flights on the connections per week.

Note that the aircraft are allocated to the individual airlines based on their registration. That is, wet leased aircraft appear as flying for the owning company. Further, code share flights are allocated to the operating airline. In addition, we do not aggregate airlines to their mother companies, but allocate aircraft to airlines solely based on their registration. This implies, for instance, that HOP! appears as an individual airline in our sample, independent from its mother company Air France. This seems reasonable as HOP! acts on the market independently from Air France at the time of the data collection. It is certainly true that this approach may be challenged as it also implies that TAP Express appears individually even though it operates as a capacity provider to TAP airlines. The key question to be addressed would be to identify which flights are actually planned under individual responsibility and which ones are jointly managed with another airline. As this question cannot be answered consistently without inside knowledge on all airlines, we refrained from this approach apart from one exception. We aggregated Germanwings and Eurowings flights since it is public knowledge that at the time of data collection they had already been merged and were operating jointly.

3.2. Graph theory metrics

A large set of graph theory metrics are available to study network structures. The networks are modeled as graphs with airport nodes and flight edges. The data needed for this representation of the network structure may come directly from airlines or aggregated flight schedules. It is thus possible to collect the information at a very limited cost as it is generally available publicly.

We use eight network metrics to describe the network structures. These metrics have been chosen as they are well known and have previously been used to analyze the structure of airline networks (e.g., [18, 41, 14]). Four of these metrics describe the network as a whole and the four remaining metrics describe individual nodes. The node-individual metrics are aggregated to a network metric by calculating their network-specific Hirschman-Herfindal index (HHI). Thus, the analysis focuses on the level differences between the node-individual metrics within a network. Furthermore, the HHI also normalizes the metric with regard to network size.

Average edge weight is the first network metric. It is defined as the total number of flights over the total number of routes in the network. Hence, it expresses how often on average the airline operates a flight on one of its routes.

Latora and Marchiori [42] introduced *network efficiency*, a metric applied for the airline network context by Lin and Ban [43]. It is defined as $E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}$, where N is the number of nodes in the network, and d_{ij} is the distance between nodes i and j, measured in the number of edges between them. If a directed network is fully connected, it will have $N(N-1)$ edges. In this case, the distance between all nodes is 1 and $E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} = \frac{1}{N(N-1)} * N(N-1) * \frac{1}{1} = 1$.

Edge density in a directed graph is defined as $D = \frac{M}{N(N-1)}$, where M is the number of edges in the graph. Edge density calculates the ratio of the existing edges in a network over the maximum number of potential edges in the network. A complete graph has $D = 1$. Edge density has been used in the airline context among others by Sun and Wandelt [44].

Finally, network *transitivity* is a metric to account for clustering in a network. Formally, it is defined as $T = \frac{3*(\text{number of triangles on the graph})}{\text{number of connected triples of nodes}}$ [45]. A triple of nodes refers to a node and its two directly connected neighbors. The triangles on the graph refer to the situation where three nodes are connected by a loop of three edges. Network transitivity accounts for how often neighbors of one node are also connected directly. It is a variant of a clustering coefficient.

The first node-specific metric is *closeness*. It is one of the well-known point centralities summarized by Freeman [46]. It is calculated as $C_C(j) = \frac{1}{\sum_{i=1}^{N} d_{ij}}$. That is, it is the inverse of the sum of all shortest distances from one node to all other nodes. This metric is easily confounded with network efficiency, yet the two metrics are far from identical. Important to note is that the metric decreases as networks grow, hence comparing node closeness centrality across networks needs to correct for network size. Freeman [46] suggests measuring graph centrality based on the existing point centrality as the normalized difference to the node with the highest point centrality. The graph centrality is then independent of the number of nodes in the network. We deviate from this suggestion and calculate the HHI of the closeness node centralities in the network. Thereby, we put an emphasis on how equally distributed closeness centrality is.

Node strength is the weighted sum of all incoming and outgoing edges in a node: $s_i = \sum_{j \in N} w_{ij}$ [43]. w_{ij} is the weight of the edge from i to j and N_i is the set of neighbor nodes to i. The airport-specific metric increases, as the airport is connected to more destinations and/or if existing routes are operated at higher frequencies. Very important airports such as FSC hubs will have a high node strength. Again, the variable analyzed per network is the HHI of node strength, hence the measure of how equally distributed the strength of the nodes in the network is.

PageRank is an index to account for the importance of nodes in networks. PageRank is well-known today as it is the basis for googles website scorings based on the work by Brin and Page [47]. It goes back to the observation in social network analysis that the importance of a group leader not only results from the number of ties she has to other actors, but is also strongly influenced by the importance of her connected actors. With this observation, Katz [48] suggested an index to take both the number of edges as well as the neighboring nodes' importance into account. In the airline context, PageRank is used to measure the likelihood that an arbitrary flight lands at a specific airport [49]. An airport is perceived important if this likelihood is high. It is intuitive that the likelihood that a flight reaches an airport increases with the number of connected airports as well as the importance of the connected airport. The individual nodes' PageRank values are aggregated to a network metric by their HHI.

Betweenness centrality is defined as the number of shortest paths in the network that a node is part of. Following Freeman [46], the index calculation starts by determining the number of existing shortest paths from node i to node j in the network, denoted as g_{ij}. The number of these shortest paths going through node k is defined as $g_{ij}(k)$. With these two parameters, the betweenness centrality of node k reads $C_B(k) = \sum_j \sum_{i<j} \frac{g_{ij}(k)}{g_{ij}}$. Betweenness centrality in airline networks is commonly associated with the capability of an airport to offer connecting services (e.g. [16]). Freeman [46] suggests a graph betweenness centrality to account for the deviations from the average node betweenness centralities. Again, we do not make use of this index but calculate the HHI of the node betweenness centrality to focus more on the inequality among nodes.

The metrics were selected such that they are independent of the network size. The application of the HHI eliminates the potential effect of the network size in the data for the respective metrics.

It is important to keep in mind that none of the individual metrics will assume negative values. Table 2 reports the descriptive statistics of the calculated metrics.

Table 2: Descriptive statistics

Variable	min	median	Mean	max
Average edge weight	1.687	7.859	8.216	16.765
Network efficiency	0.048	0.124	0.153	0.395
Edge density	0.011	0.048	0.052	0.158
Transitivity	0.000	0.041	0.085	0.319
Closeness (HHI)	0.004	0.017	0.018	0.051
Node strength (HHI)	0.014	0.102	0.109	0.257
PageRank (HHI)	0.019	0.113	0.116	0.223
Betweenness (HHI)	0.028	0.209	0.306	1.000

3.3. Principal component analysis

A principal component analysis (PCA) used as a factorial method serves to reduce the dimensions in a dataset to a few principal components (PC). The idea behind PCA is to find a few weighted linear combinations of the multiple variables in the dataset that nevertheless describe the dataset well; that is they maximize the explained variance in the data. This approach works particularly well if the original variables are sufficiently correlated. Handling and interpreting the few linear combinations is easier than dealing with the original variables. See Härdle and Simar [50] for an introduction into the mathematical background of PCA. We follow Backhaus et al. [51] in conducting this PCA.

Table 3: Correlation of network metrics

	Average edge weight	Network efficiency	Edge density	Transitivity	Closeness (HHI)	Node strength (HHI)	PageRank (HHI)	Betweenness (HHI)
Average edge weight	1	-0.781	-0.502	-0.398	-0.220	0.490	0.485	0.423
Network efficiency		1	0.624	0.555	0.127	-0.638	-0.655	-0.568
Edge density			1	0.655	0.690	-0.393	-0.412	-0.323
Transitivity				1	0.129	-0.729	-0.746	-0.636
Closeness (HHI)					1	0.180	0.165	0.185
Node strength (HHI)						1	0.973	0.953
PageRank (HHI)							1	0.895
Betweenness (HHI)								1

An essential question before conducting a PCA is whether the multiple variables are sufficiently correlated. Backhaus et al. [51] suggest using the Kaiser-Mayer-Olkin (KMO) criterion to test if the matrix of correlation of the variables allows for a PCA. Generally, a PCA should not be conducted if the matrix displays a KMO below 0.5.

The weights assigned to the original variables to obtain the PC are referred to as factor loadings. The factor loadings support the interpretation of the principal components. Factor loadings below a certain absolute value do not support the interpretation of the PC. Specifically, we shall not include absolute factor loadings below 0.25. Our data suggests this value as most variables load strongly to one of the PCs. Dobruszkes [29] chooses an alternative approach by calculating a significance threshold based on the number of observations.

It is not surprising from the description above that some of the information aggregated in the eight metrics points towards the same aspects of network structure. Table 3 presents the correlation between the different metrics and finds very strong effects for some metrics such as node strength (HHI), PageRank (HHI) and betweenness (HHI). It is worthwhile to note that closeness (HHI) and network efficiency hardly correlate despite similarities in their definition at first sight. Our data shows a KMO of 0.69. Thus, we may thus safely assume that the collected variables are sufficiently correlated to conduct a PCA. We turn to the Eigenvalue criterion to define the number of factors present in the data. Two PC have an Eigenvalue greater than 1, which indicates that they are suitable to explain the data. The following then proceeds with two PC. Table 4 lists the PC loadings. Together, the two PC explain 82% of the variance in the data (PC1 0.57 and PC2 0.24).

3.4. Cluster analysis

Given that we find two PC, we may conduct a cluster analysis in two dimensions, allowing for an intuitive representation in a coordinate system. Cluster analysis is "one of the most used multivariate statistical techniques for segmentation" [52].

The 58 airline network observations are positioned in a coordinate system expanded by PC 1 and PC 2. We are interested in identifying groups of airlines with similar airline structure. We thus conduct a 2-dimensional cluster analysis based on a k-means algorithm. The scree plot suggests four to eight clusters. Based on the interpretability of the data, a solution with seven clusters is presented below. Since all eight original metrics are defined in the positive range, a positioning of an airline in the negative range of PC 1 or PC 2 is the result of negative factor loadings.

3.5. Data analysis software

The analysis was conducted fully in the R software, version 3.5.1. The igraph package (version 1.2.1) for R provides an extensive library of metrics for network structure. PCA and cluster analysis are among the statistical tools implemented in R.

4. FINDINGS

4.1. Interpretation of factor loadings

Table 4: Factor loadings

	Principal component	
	1	2
Average edge weight	0.60	-0.40
Network efficiency	-0.76	0.33
Edge density	-0.47	0.79
Transitivity	-0.79	(0.20)
Closeness (HHI)	(0.15)	0.93
Node strength (HHI)	0.98	(0.18)
PageRank (HHI)	0.97	(0.15)
Betweenness (HHI)	0.93	(0.22)

In approaching an interpretation of the two factors presented in Table 4, note that most metrics correlate heavily with one factor. Network efficiency and transitivity load negatively, HHI of node strength, HHI of PageRank and HHI of betweenness centrality load positively to PC 1. Edge density and HHI of closeness influence PC 2 positively and PC 1 negatively. Average edge weight loads to both factors but with varying signs. It is the only metric negatively correlated with PC 2, rendering it important for the discussion below.

PC 1 can be interpreted as the *coverage of the network*. An airline network will score high in this dimension when node strength, PageRank and betweenness of the nodes in the network are unevenly distributed as this will increase the respective HHI. Node strength increases with many flights at an airport. If this metric is unevenly distributed, this suggests that many airports have few flights and few airports operate many flights. This is achieved in a centrally oriented star-type structure where few central airports serve a very large number of smaller airports. A betweenness HHI value equally suggests a star-type structure as the central airports lie on many shortest paths between other airports while spoke airports do not have this property. A structure with some important airports would create a high value on the PageRank HHI. The PageRank of an airport increases in particular when it is connected to few other important airports in the network. This is achieved for instance, when several hub airports are interconnected. Furthermore, in order to score high on the PC 1 dimension, network efficiency and transitivity should be low. Recall that network efficiency is the average of the inverse of the distances in the network. The longer the paths in the network, the lower the network efficiency values. Hence, airline networks where passengers can keep connecting from airport to airport (even though it may not be very smart to do so for passengers) tend to have low network efficiency values. Transitivity is low when there are no clusters of airports, i.e., neighboring nodes are not connected.

This is common for hub networks again. In a hub network, several spoke cities are connected to a hub, but there typically is no direct connection between the spokes. Summing up these points suggests that high values on PC 1 relate to centrally oriented multi-hub structures where spoke airports are served distinctly from one of the hubs. A network that scores high in this dimension maximizes its coverage while limiting the number of routes.

In contrast, PC 2 focuses on *frequency of service*. Edge density and closeness HHI will contribute positively to PC 2. Edge density increases as the network moves from sparse to complete. Hence more complete networks c.p. have a higher position on the PC 2 axis. Recall that node closeness is the inverse of the distance of a node to all others. To achieve a high HHI few nodes need to acquire a major share of the total closeness centralities in the network. Highest node closeness centrality is achieved by the center of a star and it decreases for multi-star networks. Average edge weight furthermore negatively influences the positioning on the PC 2 axis. Average edge weight refers to the average number of flights per operated route. Thus, airlines with a high frequency will score lower on PC 2 than ones with fewer flights on the routes. Bringing these three aspects together, a network scoring positively on the PC 2 axis offers many routes from one central base or hub but the routes are not operated at high frequencies.

4.2. Findings from the cluster analysis

Figure 1 shows the positioning of the airlines in the sample along the two axis. The seven resulting clusters are identified by colors and labelled by cluster number. Cluster I contains a subset of the traditional full service carriers, including Air France, Iberia, British Airways and Lufthansa. The airlines share strong hub operations and operate large networks. Cluster II features further full service carriers. This cluster subsumes the slightly smaller counterparts to the airlines in Cluster I. These airlines are positioned higher on PC 2 particularly as their average edge weight is lower than their counterparts in Cluster I. This is worth noting since the selected metrics are independent of network size. Hence, what is observed here is a specific network structure that differs between Clusters I and II and not the effect of the network size per se. Given that the airlines spread further right on PC 1, this implies that the node strength, PageRank and betweenness are less evenly distributed than in the networks of Cluster I. It indicates that the spread between the operations at the hubs in comparison to the served airports is more pronounced than for Cluster I. It is noteworthy that Pegasus Airlines figures in Cluster I: It can rather be attributed to the LCC context, yet it operates a star-shaped network from a few Turkish airports.

Closely related to these two is Cluster III that also includes FSCs. This cluster contains large airlines from smaller countries that act as their flag carrier, such as Ukraine International Airlines or Aer Lingus. The networks are operated at lower frequencies, that is, lower edge weight, than the airlines in Cluster I, moving Cluster III upwards on the PC 2 axis in comparison. Thus, based on the analysis of the network metrics, the business model of FSC is separated into three subgroups.

Cluster IV closely relates to Cluster V. Both clusters score similar values in PC 1 yet they differ in PC 2. Cluster IV is positioned below Cluster V on the PC 2 axis. The networks of the airlines in Cluster IV are more star-type than those of Cluster V. That is, the Cluster IV networks rely more heavily on their hubs or bases. The average flight frequency is greater. In sum, this suggests star-shaped networks with many airports directly connected to a few important bases. Judging from the business side, Cluster IV includes what is commonly perceived as large LCC on the European market: Ryanair, EasyJet, Lufthansa's low cost offspring Eurowings, Air France's offspring Hop! In comparison, Cluster V contains mostly charter airlines, some of them with a regional focus. Their flights are offered at a lower frequency. Keep in mind that the collected data is from November in Europe, which is an off-peak season for holiday flights. During this season, charter airlines will not operate at high frequencies on few routes as it would be expected during the summer months.

Cluster VI is comparable to Cluster II on the PC 1 axis. It contains airlines with a rather regional focus of destinations. For instance, TAP Express and Iberia Express both are regional feeder partners. Icelandair focuses on flights from Iceland. However, Cluster VI airlines show a higher position on the PC 2 axis than the airlines in Cluster II. The networks in this cluster have fewer destinations and flights are operated at lower frequencies.

Finally, Cluster VII contains three airlines whose networks have hardly any characteristics in common with the other clusters. Jet2.com is a low cost carrier operating with eight bases from within Great Britain. Loganair is a Scottish regional carrier with very particular services such as short connections to and between Channel Islands. Rusline is a Russian regional carrier serving mostly Russian destinations.

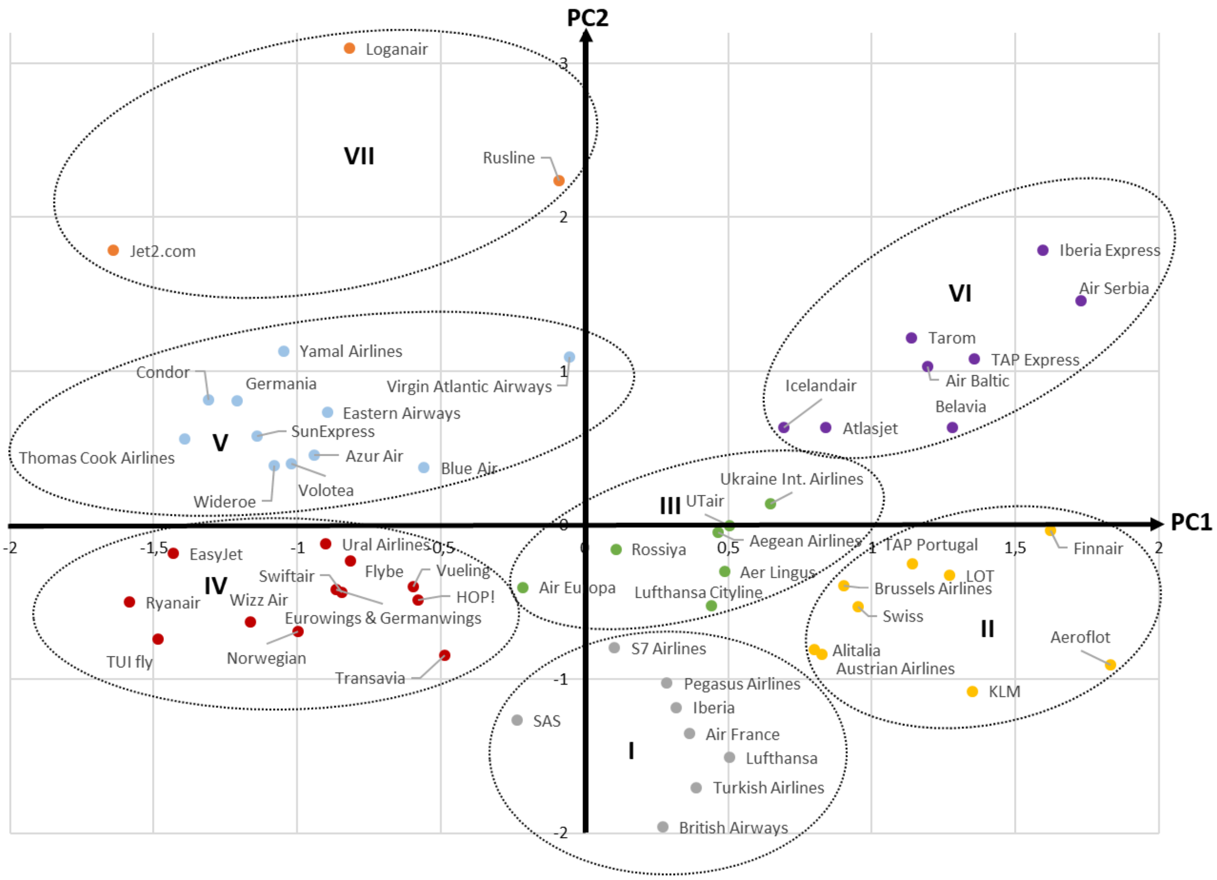

Figure 1: Positioning of airlines (clusters are identified by color and number)

Observe that the far ends of the coordinate system remain mainly void. Among the analyzed airline networks, only the airlines in Cluster VII score extreme values on both components. The other airlines appear to emphasize only one component.

The airline industry is dynamic and the strategic position of airlines in our sample has changed since the data collection. For instance, HOP! has been moving its operations significantly closer to Air France since October 2018. Most notably, Germania Fluggesellschaft mbH declared bankruptcy in February 2019 and ceased operations the following day. Moreover, Flybe was sold to the Connect Airways consortium in February 2019 with the intention that Flybe will operate for Virgin Atlantic. Strategic changes such as these will affect the airlines' networks and revisiting the networks in the future will most likely allow identifying the operational differences.

5. DISCUSSION

Our analysis intentionally rests solely on metrics of network structure. Unlike, e.g. Dobruszkes [3, 29], we do not include elements to describe the nature of the airline network such as if the network serves 'warm water' destinations or the share of international flights. This difference in approach leads to different insights. Consider Ryanair and EasyJet in Figure 1. Previously, Graham [53] has compared the network structures of LCC, including these two airlines. Equally, Dobruszkes [29] includes these two airlines in his sample on LCC. Both find relevant differences between Ryanair and EasyJet and it appears reasonable that differences between the two business models appear more pronounced than for our study, which considers a broader sample of airlines to start with. Graham [53] points out that at the time, Ryanair was serving more destinations than EasyJet, and at a higher average frequency per route. At the same time, Ryanair's network relied strongly on its base in London Stansted while EasyJet distributed its flights almost evenly between London Stansted, Gatwick and Luton. Dobruszkes [29] observes that Ryanair and EasyJet differ in particular regarding his first two PC: the volume of supply and the density of the network. The difference in the other two PC, namely centralization and charter-like operations, is less significant. Our analysis takes into account not only LCC but reveals insights on the positions of more diverse airlines. Hence, it is natural that the differences between Ryanair's and EasyJet's networks appear less pronounced. In addition, we intended to correct our analysis for network size. Hence, the differences in network structure resulting from the number of destinations served between the two competitors should be reduced in our analysis. We do observe that EasyJet's network scores slightly higher on the PC 1 dimension than Ryanair's, which is partially driven by EasyJet operating not from one but from three key airports in the London area.

The analysis of the data reveals that the network structures of FSC can be separated in three clusters. This is an indication that they do operate different types of networks. When comparing solutions with fewer clusters, we see that the next two clusters to be merged are Cluster II and III and subsequently Cluster I will be added to this large cluster. Thus, it can easily be argued that the networks among the FSC are somewhat similar, yet there are three types of networks to be distinguished. The analysis identified large flag carriers with more or less global networks with many spokes and potentially several hubs, hence positioned in lower ranges of PC 1. Furthermore, we find smaller counterparts to these used as feeders as well as smaller alliance partners. Network size is corrected for in the analyzed metrics, hence the finding here is that the smaller carriers share a similar structure.

Urban et al. [37] conduct a cluster analysis only of FSC and LCC on a global scale. Their focus is on the convergence of these two business models. Parameters for the entire breadth of airline business models feed into their study. The route network is taken into account as a categorical variable (HS, PP, HS and PP, multi-hub). Similar to our findings, they observe subgroups for both FSC and LCC. The identified groups for FSC are *medium-size network carrier, global niche market network carrier, high quality network carrier* and *large-size network carrier*. There is an overlap of 13 airlines with our analysis. EasyJet, Ryanair, and Vueling are consistently clustered by our approach jointly in Cluster IV such as in the group of PP LCC by Urban et al. [37]. For the FSC, the allocation to clusters is not fully identical. Urban et al. [37] group Aeroflot, Finnair, and TAP and among the global niche market network carriers. They all show in Cluster II above with KLM joining them in our analysis. In contrast, Cluster I of large FSC includes Air France, British Airways, Iberia, Lufthansa and Turkish Airlines while Urban et al. [37] allocate them to three different subgroups of FSC, differentiating between medium- and large-size and high quality. It is a relevant finding that both approaches – covering the full breadth of airline business parameters and studying only the network structure – produce mostly consistent findings. It would be interesting to extend our analysis similarly to a global sample to verify if subgroups among the LCC can equally be identified. Omitting the regional and charter airlines may also bring in additional focus on the LCC networks.

As is true for all empirical studies, our findings only apply to the investigated dataset of European airlines. However, the airline business is mostly considered a global market and airline competition is not limited to Europe. Further, we identified similarities to the results of a global sample by Urban et al. [37]. It appears reasonable to assume that similar findings may be generated for other geographic markets.

6. CONCLUSIONS

We have collected and analyzed information on European airline network structures. We did so with the intention to discuss the benefit of detailed metrics for airline network structure for describing airline business networks. We found that airlines of related business models do share significant similarities in their network structures. In fact, clustering the airlines based on the information collected about their network structures by itself suggests a strikingly consistent landscape of airline business models. Full service carriers are found in Clusters I, II and partially III. Low cost carriers form Cluster IV. Many charter airlines figure in Cluster V and most regional carriers in Cluster VI. This observation supports the idea that studying airline business networks can greatly benefit from the objective assessment of airline network structure based on metrics from graph theory.

The airline network is commonly considered as one element of an airline business model. Jointly, all of the elements define an airline business model. Describing the element of network more granularly is a step towards exploring the distinctions between business models.

It is furthermore relevant to observe that the PCA identifies two components that characterize airline networks. They hint at differences between HS und PP networks. PC 1 aims at identifying the network coverage whereas PC 2 points at the service network. Obviously, both aspects contribute to the understanding that a HS network increases its coverage and operates the connections between hubs at high frequency.

From a methodological perspective, it is worth mentioning that PCA proves to be a powerful tool to condense the network information obtained from multiple metrics originating in graph theory. In general, the metrics are strongly correlated but not identical, as is also the case with our data. This may be one reason why research has applied a wide variety of metrics to account for HS and PP networks. Trying to identify the one best metric among them is virtually impossible. Our analysis includes eight network metrics and the PCA reduces them to two components. It is a valid path for further research to validate that other network metrics lead to consistent components.

The correlation of airline network structure and business model is intuitive at first, given that it is common to assume that LCC operate PP networks and FSC HS networks. However, it has been observed previously that these two types of networks cannot meaningfully describe the existing spectrum of network structures. We have illustrated that incorporating more granular and objective metrics of network structure in the description of airline business models overcomes this obstacle and allows to describe the network perspective better.

REFERENCES

1. Hu Y, Zhu D (2009) Empirical analysis of the worldwide maritime transportation network. Physica A: Statistical Mechanics and its Applications 388 (10):2061-2071. doi:https://doi.org/10.1016/j.physa.2008.12.016
2. Wang Y, Cullinane K (2016) Determinants of port centrality in maritime container transportation. Transportation Research Part E: Logistics and Transportation Review 95:326-340. doi:https://doi.org/10.1016/j.tre.2016.04.002
3. Dobruszkes F (2006) An analysis of European low-cost airlines and their networks. Journal of Transport Geography 14 (4):249-264. doi:https://doi.org/10.1016/j.jtrangeo.2005.08.005
4. Mishra S, Welch TF, Jha MK (2012) Performance indicators for public transit connectivity in multimodal transportation networks. Transportation Research Part A: Policy and Practice 46 (7):1066-1085. doi:https://doi.org/10.1016/j.tra.2012.04.006
5. Mattsson L-G, Jenelius E (2015) Vulnerability and resilience of transport systems – A discussion of recent research. Transportation Research Part A: Policy and Practice 81:16-34. doi:https://doi.org/10.1016/j.tra.2015.06.002
6. Oum TH, Zhang A, Zhang Y (1995) Airline Network Rivalry. The Canadian Journal of Economics / Revue canadienne d'Economique 28 (4a):836-857. doi:10.2307/135934
7. Brueckner JK (2004) Network Structure and Airline Scheduling. 52 (2):291-312. doi:doi:10.1111/j.0022-1821.2004.00227.x
8. Alderighi M, Cento A, Nijkamp P, Rietveld P (2005) Network competition – the coexistence of hub-and-spoke and point-to-point systems. Journal of Air Transport Management 11 (5):328-334. doi:10.1016/j.jairtraman.2005.07.006
9. Pels E (2008) Airline network competition: Full-service airlines, low-cost airlines and long-haul markets. Research in Transportation Economics 24 (1):68-74. doi:https://doi.org/10.1016/j.retrec.2009.01.009
10. Babić D, Kalić M (2018) Modeling the selection of airline network structure in a competitive environment. Journal of Air Transport Management 66:42-52. doi:https://doi.org/10.1016/j.jairtraman.2017.10.004
11. Brueckner JK (2010) Schedule Competition Revisited. Journal of Transport Economics and Policy (JTEP) 44 (3):261-285
12. Brueckner JK, Flores-Fillol R (2007) Airline Schedule Competition. Rev Ind Organ 30 (3):161-177. doi:10.1007/s11151-007-9140-1
13. Burghouwt G, Hakfoort J, Ritsema van Eck J (2003) The spatial configuration of airline networks in Europe. Journal of Air Transport Management 9 (5):309-323. doi:10.1016/S0969-6997(03)00039-5
14. Guimerá R, Amaral LAN (2004) Modeling the world-wide airport network. The European Physical Journal B 38 (2):381-385. doi:10.1140/epjb/e2004-00131-0
15. Guida M, Maria F (2007) Topology of the Italian airport network: A scale-free small-world network with a fractal structure? Chaos, Solitons & Fractals 31 (3):527-536. doi:10.1016/j.chaos.2006.02.007
16. Wang J, Mo H, Wang F, Jin F (2011) Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. Journal of Transport Geography 19 (4):712-721. doi:10.1016/j.jtrangeo.2010.08.012
17. Alderighi M, Cento A, Nijkamp P, Rietveld P (2007) Assessment of New Hub-and-Spoke and Point-to-Point Airline Network Configurations. Transport Reviews 27 (5):529-549. doi:10.1080/01441640701322552
18. Lordan O, Sallan JM, Simo P (2014) Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda. Journal of Transport Geography 37 (Supplement C):112-120. doi:10.1016/j.jtrangeo.2014.04.015
19. Reggiani A, Signoretti S, Nijkamp P, Cento A (2009) Network Measures in Civil Air Transport: A Case Study of Lufthansa. In: Naimzada AK, Stefani S, Torriero A (eds) Networks, Topology and Dynamics: Theory and Applications to Economics and Social Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 257-282. doi:10.1007/978-3-540-68409-1_14
20. Bowen JT (2012) A spatial analysis of FedEx and UPS: hubs, spokes, and network structure. Journal of Transport Geography 24 (Supplement C):419-431. doi:10.1016/j.jtrangeo.2012.04.017
21. Button K, Ison S (2008) The economics of low-cost airlines: Introduction. Research in Transportation Economics 24 (1):1-4. doi:https://doi.org/10.1016/j.retrec.2009.01.008
22. Daft J, Albers S (2015) An empirical analysis of airline business model convergence. Journal of Air Transport Management 46:3-11. doi:10.1016/j.jairtraman.2015.03.008
23. Bachwich AR, Wittman MD (2017) The emergence and effects of the ultra-low cost carrier (ULCC) business model in the U.S. airline industry. Journal of Air Transport Management 62:155-164. doi:10.1016/j.jairtraman.2017.03.012

24. Soyk C, Ringbeck J, Spinler S (2017) Long-haul low cost airlines: Characteristics of the business model and sustainability of its cost advantages. Transportation Research Part A: Policy and Practice 106:215-234. doi:https://doi.org/10.1016/j.tra.2017.09.023
25. Gillen D, Gados A (2008) Airlines within airlines: Assessing the vulnerabilities of mixing business models. Research in Transportation Economics 24 (1):25-35. doi:http://dx.doi.org/10.1016/j.retrec.2009.01.002
26. Wensveen JG, Leick R (2009) The long-haul low-cost carrier: A unique business model. Journal of Air Transport Management 15 (3):127-133. doi:10.1016/j.jairtraman.2008.11.012
27. Casadesus-Masanell R, Ricart JE (2010) From Strategy to Business Models and onto Tactics. Long Range Planning 43 (2):195-215. doi:https://doi.org/10.1016/j.lrp.2010.01.004
28. Daft J, Albers S (2013) A conceptual framework for measuring airline business model convergence. Journal of Air Transport Management 28:47-54. doi:10.1016/j.jairtraman.2012.12.010
29. Dobruszkes F (2013) The geography of European low-cost airline networks: a contemporary analysis. Journal of Transport Geography 28:75-88. doi:https://doi.org/10.1016/j.jtrangeo.2012.10.012
30. Klophaus R, Conrady R, Fichert F (2012) Low cost carriers going hybrid: Evidence from Europe. Journal of Air Transport Management 23:54-58. doi:10.1016/j.jairtraman.2012.01.015
31. Lohmann G, Koo TTR (2013) The airline business model spectrum. Journal of Air Transport Management 31:7-9. doi:10.1016/j.jairtraman.2012.10.005
32. Jean DA, Lohmann G (2016) Revisiting the airline business model spectrum: The influence of post global financial crisis and airline mergers in the US (2011–2013). Research in Transportation Business & Management 21:76-83. doi:10.1016/j.rtbm.2016.06.002
33. Moir L, Lohmann G (2018) A quantitative means of comparing competitive advantage among airlines with heterogeneous business models: Analysis of U.S. airlines. Journal of Air Transport Management 69:72-82. doi:https://doi.org/10.1016/j.jairtraman.2018.01.003
34. Mason KJ, Morrison WG (2008) Towards a means of consistently comparing airline business models with an application to the 'low cost' airline sector. Research in Transportation Economics 24 (1):75-84. doi:10.1016/j.retrec.2009.01.006
35. O'Connell JF (2011) The rise of the Arabian Gulf carriers: An insight into the business model of Emirates Airline. Journal of Air Transport Management 17 (6):339-346. doi:https://doi.org/10.1016/j.jairtraman.2011.02.003
36. Pereira BA, Caetano M (2015) A conceptual business model framework applied to air transport. Journal of Air Transport Management 44-45:70-76. doi:https://doi.org/10.1016/j.jairtraman.2015.02.006
37. Urban M, Klemm M, Ploetner KO, Hornung M (2018) Airline categorisation by applying the business model canvas and clustering algorithms. Journal of Air Transport Management. doi:10.1016/j.jairtraman.2018.04.005
38. Dobruszkes F, Peeters D (2017) Where do planes fly past overhead? Determining departure and arrival routes from radar traces. Applied Geography 89:173-183. doi:https://doi.org/10.1016/j.apgeog.2017.10.009
39. Lange A, Sieling J, Gonzalez Parra G (2019) Convergence in airline operations: The case of ground times. Journal of Air Transport Management 77:39-45. doi:https://doi.org/10.1016/j.jairtraman.2019.03.001
40. Budd LCS (2012) The potential use of consumer flight radar apps as a source of empirical data: A study of A320 utilization rates in Europe. Journal of Air Transport Management 25:53-56. doi:https://doi.org/10.1016/j.jairtraman.2012.08.004
41. Dunn S, Wilkinson SM (2016) Increasing the resilience of air traffic networks using a network graph theory approach. Transportation Research Part E: Logistics and Transportation Review 90:39-50. doi:https://doi.org/10.1016/j.tre.2015.09.011
42. Latora V, Marchiori M (2001) Efficient Behavior of Small-World Networks. Physical Review Letters 87 (19):198701. doi:10.1103/PhysRevLett.87.198701
43. Lin J, Ban Y (2014) The evolving network structure of US airline system during 1990–2010. Physica A: Statistical Mechanics and its Applications 410:302-312. doi:10.1016/j.physa.2014.05.040
44. Sun X, Wandelt S (2014) Network similarity analysis of air navigation route systems. Transportation Research Part E: Logistics and Transportation Review 70:416-434. doi:10.1016/j.tre.2014.08.005
45. Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99 (12):7821-7826. doi:10.1073/pnas.122653799
46. Freeman LC (1978) Centrality in social networks conceptual clarification. Social Networks 1 (3):215-239. doi:10.1016/0378-8733(78)90021-7
47. Brin S, Page L (1998) The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems 30 (1):107-117. doi:10.1016/S0169-7552(98)00110-X
48. Katz L (1953) A new status index derived from sociometric analysis. Psychometrika 18 (1):39-43. doi:10.1007/bf02289026

49. Conti E, Cao S, Thomas AJ (2013) Disruptions in the U.S. Airport Network. arXiv
50. Härdle W, Simar L (2007) Applied Multivariate Statistical Analysis. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/978-3-540-72244-1_9
51. Backhaus K, Erichson B, Plink W, Weiber R (2016) Multivariate Analysemethoden, vol 14. Springer Gabler, Berlin Heidelberg
52. Wen C-H, Lai S-C (2010) Latent class models of international air carrier choice. Transportation Research Part E: Logistics and Transportation Review 46 (2):211-221. doi:https://doi.org/10.1016/j.tre.2009.08.004
53. Graham M (2009) Different models in different spaces or liberalized optimizations? Competitive strategies among low-cost carriers. Journal of Transport Geography 17 (4):306-316. doi:https://doi.org/10.1016/j.jtrangeo.2009.02.004

APPENDIX:

Airline	Code	Country	Number of aircraft in dataset
Aegean Airlines	A3	Greece	47
Aer Lingus	EI	Ireland	63
Aeroflot	SU	Russia	224
Air Baltic	BT	Latvia	30
Air Europa	UX	Spain	52
Air France	AF	France	230
Air Serbia	AF	France	230
Alitalia	AZ	Italy	120
Atlasjet	KK	Turkey	24
Austrian Airlines	OS	Austria	84
Azur Air	ZF	Russia	22
Belavia	B2	Belarus	29
Blue Air	0B	Romania	30
British Airways	BA	United Kingdom	318
Brussels Airlines	SN	Belgium	51
Condor	DE	Germany	42
Eastern Airways	T3	United Kingdom	30
EasyJet	U2	United Kingdom	280
Eurowings & Germanwings*	EW/U2	Germany	78
Finnair	AY	Finland	64
Flybe	BE	United Kingdom	84
Germania	ST	Germany	29
HOP!	A5	France	86
Iberia	IB	Spain	114
Iberia Express	I2	Spain	21
Icelandair	FI	Iceland	32

Jet2	LS	United Kingdom	75
KLM	KL	The Netherlands	160
Loganair	LM	United Kingdom	28
LOT	LO	Poland	47
Lufthansa	LH	Germany	280
Lufthansa Cityline	LH	Germany	50
Norwegian	DY	Norway	142
Pegasus Airlines	PC	Turkey	74
Rossiya	FV	Russia	62
Rusline	7R	Russia	23
Ryanair	FR	Ireland	417
S7 Airlines	S7	Russia	78
SAS	SK	Sweden	162
SunExpress	XQ	Turkey	46
Swiftair	WT	Spain	38
Swiss	LX	Switzerland	77
TAP Express	NI	Portugal	22
TAP Portugal	TP	Portugal	66
Tarom	RO	Romania	25
Thomas Cook Airlines	MT	United Kingdom	60
Transavia	HV	The Netherlands	70
TUI fly	HV	The Netherlands	70
Turkish Airlines	TK	Turkey	324
Ukraine Int. Airlines	PS	Ukraine	41
Ural Airlines	U6	Russia	43
UTair	UT	Russia	68
Virgin Atlantic Airways	VS	United Kingdom	38
Volotea	V7	Spain	25
Vueling	VY	Spain	107
Wideroe	WF	Norway	41
Wizz Air	W6	Hungary	87
Yamal Airlines	YC	Russia	40

*Table A.1: List of airlines, *operated jointly at the time of data collection*