Simulating Airline Behavior: An Application for the Australian Domestic Market

November 15, 2018

Khan Doyme[1] (corresponding author)
Email: k.doyme@ucl.ac.uk

Lynnette Dray[1]
Email: L.dray@ucl.ac.uk

Aidan O'Sullivan[1]
Email: aidan.osullivan@ucl.ac.uk

Andreas Schäfer[1]
Email: a.schafer@ucl.ac.uk

[1]Air Transportation Systems Laboratory
UCL Energy Institute
The Bartlett School of Environment Energy and Resources
14 Upper Woburn Place, London, WC1H 0NN

Word count: 5273 words + 6 tables × 250 words = 6773 total

Submission date: August 1st, 2018

Abstract

In this paper we demonstrate the ability of a model, which simulates competition between airlines in a domestic aviation market, to accurately reproduce real world behaviour. The Australian market was chosen as a test case as it is a geographically isolated region with significant demand and complexity, including one of the busiest routes in the world, where connecting international passengers do not significantly skew the market. The model is based on an n-player noncooperative game, where each airline represents a player within the game. The primary assumption is that each airline attempts to maximise profits by adjusting the decision variables of airfares, flight frequency and choice of aircraft on routes within its network. The approach works iteratively, allowing each airline to respond to the decisions made by other airlines during each successive optimisation. The model is said to reach convergence when there is no significant change in any airline's profit from one iteration to the next. Once this occurs, the predictions of each airline's decision variables can be compared to real data. The model gives highly detailed predictions of airline specific airfares, flight frequencies on segments, passenger flows and airline market share, which strongly correlate with observed values.

Keywords: modelling aviation market, airline competition, airline behaviour, market optimization.

Introduction

The demand for air passenger travel is forecast to grow strongly over the next twenty years, with global revenue passenger-km (RPK) predicted to increase year on year at a rate of 4.4-4.7% [1, 2]. This will result in a doubling of RPK within 15 years. Whilst there are positive outcomes associated with this, such as enhanced economic growth and social connectivity [3], the growth in aviation also generates challenges. For example, it will place significant pressure on existing infrastructure, particularly airports, to meet this demand. According to the Boeing Global Airport Congestion Study, 15 major airports in the world operated at levels greater than 95% of capacity in 2013, a number that is expected to increase to 48 airports by 2023 [4]. As a result, airport capacity expansion programs currently exceed half a trillion USD [5] and are large in scale and often controversial. The climate impacts of aviation are another significant concern which needs to be addressed, as the sector already produces around 2.5% of all energy-related CO_2 emissions [6]. To mitigate climate impacts, new disruptive technologies (such as hybrid-electric and electric aircraft [7]) and policy measures (such as carbon taxes and offsetting schemes [8, 9]) are being developed. In order to fully understand the impacts of decisions made to address the challenges described above, we require a model capable of reproducing the observed market characteristics. Such a model could inform decision makers about the impacts of different scenarios: for example, how the expansion of capacity at one airport might have knock on impacts for other airports. Airlines are key decision makers in this system and the competition between them is a significant determinant of airfares and itinerary frequency, which in turn affects demand. For this reason several studies have modelled aviation markets by attempting to capture airline competition and behaviour, using a number of approaches discussed below.

In one of the first airline competition studies, with profit as the objective function, Hansen solved a model with flight frequency as the decision variable for a network of 52 U.S. airports and 28 airlines [10]. There have been many other frequency competition studies since. Wei and Hansen have examined how airlines make decisions on aircraft size, as well as service frequency, within a competitive environment [11]. Evans used frequency competition models to analyse flight routing network structure [12], and the implications of capacity constraints on the US market [13]. Whilst investigating airport capacity and congestion, Vaze and Barnhart found reasonable agreement between observed frequencies and the equilibrium predictions from their frequency competition model [14]. However, common to all the studies mentioned, these models cannot determine the effects of competition on airfares.

Others have extended competition models to include both frequency and airfare as decision variables, including single-stage approaches as well as two-stage frequency-fare models. Dobson and Lederer studied the competitive choice of flight schedules and route prices by airlines operating in a single hub system. Utilizing a sub-game perfect Nash equilibrium for a two-stage game, with simplifications including single aircraft size and identical hub networks for competitors, they found equilibria in a five-node network [15]. Adler developed a model framework to identify the most profitable hub-spoke networks, with the aim of classifying airports most likely to remain major hubs in Western Europe [16]. Brueckner has compared the properties of single-stage and two-stage frequency-fare games [17], and solved a simplified model of a single-stage game with frequency, seats and airfares as decision variables [18]. Adler, Pels and Nash created a frequency, seats and fares model with a view to performing cost-benefit analysis of different transport investment options [19], whilst Hansen and Liu set up a frequency-fare competition model in order to understand the effect of using different market share functions [20]. However, none of these studies were validated against a real market.

The objective of this paper is to incorporate air passenger demand and choice modelling within a frequency-airfare competition model, and to validate this model's output against data from the Australian market. To do so, we will compare predicted passenger numbers, flight frequency and airfares—for each segment of the market—against observed values. The approach used will be an extension of the frequency competition framework previously developed by Evans [12, 13].

Model Approach

The passenger aviation industry is modelled as an n-player, noncooperative, game. Each airline represents a player within the game. The primary assumption is that each airline attempts to maximise profits by adjusting the decision variables of airfares, flight frequency and choice of aircraft on routes within its network. We attempt to find the equilibrium points of this model, which can be compared to observed data.

The annual profit, P_A, of an airline, A, is defined by a revenue term, two cost terms, and an ancillary revenue term:

$$\begin{aligned} P_A = \sum_{i \in \mathrm{ITN}_A} fare_i.pax_i - \sum_{j \in \mathrm{SEG}_A} \sum_{a \in \mathrm{CRFT}_j} opcost_{a,j}.freq_{a,j} \\ - \sum_{j \in \mathrm{SEG}_A} \sum_{a \in \mathrm{CRFT}_j} paxcost_{a,j}.pax_{a,j} + arev_A.pax_A. \end{aligned} \tag{1}$$

The revenue term consists of the product of the number of passengers per annum, pax_i, that travel on itinerary i multiplied by the average airfare charged for itinerary i, summed over the set of itineraries, ITN_A, offered by airline A. The first cost term represents flight-specific operating costs, defined by the product of the operating cost, $opcost_{a,j}$, over a flight on segment j using aircraft type a and the number of flights per annum, $freq_{a,j}$, on segment j using aircraft type a, summed over all segments, SEG_A, operated by airline A. The second cost term reflects the passenger related operating costs, given by the product of the additional cost of carrying a single passenger, $paxcost_{a,j}$, on segment j using aircraft a, multiplied by the number of passengers per annum, $pax_{a,j}$, carried on segment j using aircraft a. The last term corresponds to an additional income stream through ancillary revenue per passenger, $arev_A$, that airlines generate from related commercial activity (frequent flyer programs, advertising etc.), multiplied by the total number of passengers per annum, pax_A, carried by airline A. Itineraries can consist of one segment (a direct itinerary), or multiple segments (an indirect itinerary). Equation 1 defines the objective function that each airline attempts to maximise.

The specific decision variables for each airline are the itinerary fares ($fare_i$), the number of passengers on each itinerary (pax_i), and the flight frequency by aircraft type on each network segment ($freq_{a,j}$). The airlines are not free to choose any values for the decision variables; the model must have restrictions that reflect real limits on their operations. To capture this, we impose constraints.

The first constraint ensures that the number of passengers on any given airline segment is less than or equal to the number of available seats:

$$pax_{a,j} \leq \sum_{a \in \text{CRFT}_j} seats_a.freq_{a,j} \qquad \forall j \in \text{SEG}_A, \tag{2}$$

where $seats_a$ is the number of commercial seats on aircraft type a.

The passengers on a segment must be related to the passengers on the itineraries which use a given segment. This is captured by the following identity:

$$\sum_{a \in \text{CRFT}_j} pax_{a,j} = \sum_{i \in \text{ITN}_{j,A}} pax_i \qquad \forall j \in \text{SEG}_A, \tag{3}$$

where $\text{ITN}_{j,A}$ is the set of itineraries, offered by airline A, that use j as one of their segments.

The number of flights that an airline can operate is limited by their fleet of available aircraft. For each airline, A, and aircraft type, a, we have the constraint:

$$\sum_{j \in \text{SEG}_A} freq_{a,j}.(time_j + ground_a) \leq (365 \times 24).fleet_{a,A} \tag{4}$$

where $time_j$ is the flight time (in hours) of segment j, $ground_a$ is the average number of hours aircraft a spends on the ground per flight cycle, and $fleet_{a,A}$ is the number of aircraft of type a available to airline A.

The number of flights into and out of an airport must be restricted by the airport capacity. This is accounted for by using the inequality:

$$\sum_{j \in \text{SEG}_p} \sum_{a \in \text{CRFT}_j} freq_{a,j} \leq max_p \qquad \forall p \in \text{APT}, \tag{5}$$

where APT is the set of all airports, SEG_p is the set of all airline segments in the network that have airport p as an origin or destination, and max_p is the maximum number of aircraft movements per year at airport p. This is the first inequality that includes decision variables (segment flight frequencies) from multiple airlines. Through this inequality, the behaviour of one airline can now alter and constrain the decisions made by another.

The next constraint restricts the number of passengers on any given itinerary to be less than the itinerary's market share of overall passenger demand between the origin and destination:

$$pax_i \leq \text{MS}_i.\text{D}_i \qquad \forall i \in \text{ITN}_A, \tag{6}$$

where MS_i is the market share of itinerary i, and D_i is the annual passenger demand between the origin and destination of itinerary i.

The form of the passenger demand function, D_i, for air travel between an itinerary's origin, o_i, and destination, d_i, is

$$\text{D}_i = e^{\eta}(P_{o_i}P_{d_i})^{\alpha}(I_{o_i}I_{d_i})^{\beta}\bar{fare}^{\gamma}_{o_i,d_i}\bar{time}^{\delta}_{o_i,d_i}drive^{\epsilon}_{o_i,d_i}e^{(\zeta.special_{o_i,d_i})} \tag{7}$$

where P_{o_i} is the population within the greater metropolitan area of the origin o_i of itinerary i, P_{d_i} is the population within the greater metropolitan area of the destination d_i of itinerary i, $\bar{I}_{o_i}$ and I_{d_i} are the household incomes per capita for the origin and destination areas respectively, $\bar{fare}_{o_i,d_i}$ is the average fare paid by passengers on itineraries between the origin and destination, $\bar{time}_{o_i,d_i}$ is the average flight time, $drive_{o_i,d_i}$ is the journey time by car from origin to destination, $special_{o_i,d_i}$ is a binary variable indicating whether both origin and destination cities are special (capital cities, major business or tourist destinations etc.) and η is a constant. This is an example of a gravity model, which is often used in air transport studies to model passenger demand [21, 22, 23]. The coefficients, α-η, are determined using linear regression.

The market share of an itinerary, MS_i, is predicted using a multinomial logit (MNL) model. Between any two metropolitan areas, passengers have a choice of which airport to depart from and arrive at, which airline to fly with, and which route (direct, indirect etc) to take. The market share function allocates passengers to each itinerary option based on the utility it provides. This method has been applied extensively within aviation research in order to model passenger decisions and choice [24, 25, 26, 27]. The market share for each itinerary, i, is

$$\text{MS}_i = \frac{e^{U_i}}{\sum\limits_{j \in \text{ITN}_{o_i d_i}} e^{U_j}}, \tag{8}$$

where $\text{ITN}_{o_i d_i}$ is the set of all itineraries with origin, o_i and destination, d_i. The utility, U_i, is

$$U_i = \theta.fare_i + \kappa.time_i + \lambda.ln(freq_i) + \mu.nseg_i + \mathit{ffx}_{A,i}, \tag{9}$$

where $fare_i$ is the airfare of itinerary, i; $time_i$ is the journey time from origin to destination airport of itinerary i; $freq_i$ is frequency of itinerary i; $nseg_i$ is the number of segments the make up itinerary i; ffx_{A_i} is a constant whose value is dependent only on the airline, A_i, operating the itinerary i. The coefficients, θ-μ, and fixed effect constants, ffx_A, are found by taking the logarithm of the ratios of itinerary market share and using OLS regression on the resultant linear equation [10, 27].

The frequency of an itinerary, $freq_i$, used in equation 9, is yet to be defined. It will depend on the flight frequency between each of the itinerary's segments:

$$freq_i = \min\{freq_j; j \in \text{SEG}_{i,A_i}\} \qquad \forall i \in \text{ITN}_A, \tag{10}$$

where $\text{SEG}_{i,A}$ is the set of segments that constitute itinerary i of airline A_i, and $freq_j$ is the total flight frequency on segment j ($freq_j = \sum_a freq_{a,j}$).

Given the non-linear objective function (equation 1) and constraints (in particular equation 6), determining the globally optimum equilibrium for large networks and multiple competitors is challenging. Even the existence of such an equilibrium is not guaranteed, as the market share function can mean an airline's strategy set is non-convex [28]. We circumvent these difficulties by linearising the objective function and constraints with respect to the decision variables, and then searching for an increase in the profit near some starting point. For example, given some initial state of the airline's decision variable's, $\{fare_{i,0}, pax_{i,0}, freq_{a,j,0} | i \in \mathrm{ITN}_A, j \in \mathrm{SEG}_A, a \in \mathrm{CRFT}_A\}$, the linearised form of the profit function about this point is:

$$
\begin{aligned}
P_A \approx & \sum_{i \in \mathrm{ITN}_A} fare_{i,0}.pax_i + \sum_{i \in \mathrm{ITN}_A} fare_i.pax_{i,0} - \sum_{i \in \mathrm{ITN}_A} fare_{i,0}.pax_{i,0} \\
& - \sum_{j \in \mathrm{SEG}_A} \sum_{a \in \mathrm{CRFT}_j} opcost_{a,j}.freq_{a,j} - \sum_{j \in \mathrm{SEG}_A} \sum_{a \in \mathrm{CRFT}_j} paxcost_{a,j}.pax_{a,j} \\
& + arev_A.pax_A.
\end{aligned} \tag{11}
$$

The constraints are similarly linearised. The linearised form of the equations are only applicable in the neighbourhood of the initial starting point of the decision variables. We therefore add constraints which restrict the amount that the decision variables can be changed:

$$
\begin{aligned}
fare_{i,0} - \Delta_{fare} \leq & fare_i \leq fare_{i,0} + \Delta_{fare}, \\
pax_{i,0} - \Delta_{pax} \leq & pax_i \leq pax_{i,0} + \Delta_{pax}, \\
freq_{j,0} - \Delta_{freq} \leq & freq_j \leq freq_{j,0} + \Delta_{freq}, \qquad \forall i, j.
\end{aligned} \tag{12}
$$

The constants Δ_{freq}, Δ_{pax} and Δ_{fare} are chosen to be small enough such that the linear approximation captures the correct behaviour of the non-linear model for the entire range of allowed values.

Each airline is now given the opportunity to maximise its profit function, equation 11, by changing their decision variables within the constraints of equations 2, 3, 4, 5, 6, 10 and 12. Due to the linearised objective function and constraints, this is straightforward to solve using a linear programming algorithm, such as IBM CPLEX. After this is performed, each airline will have a new set of values for their decision variables, $\{fare_{i,1}, pax_{i,1}, freq_{a,j,1} | i \in \mathrm{ITN}_A, j \in \mathrm{SEG}_A, a \in \mathrm{CRFT}_A\}$. There is no guarantee that these new values truly maximise each airline's profit, because the extent to which the decision variables could be changed was capped in equation 12. However, the values for the decision variables can be used as a new starting point, and the process iterated. A 'quasi-equilibrium'[10] is reached when, from iteration to iteration, the profit for each airline ceases to change significantly. When this occurs, the resulting revenue per passenger, flight frequencies and passenger numbers for each segment in an airline's network are obtained as the model's predictions.

Data

Information on itinerary passenger numbers, airfare, frequency of service and journey time for the 2014 Australia market was obtained from the Sabre Market Intelligence dataset [29]. An example of the data, for one itinerary, is given in table 1. We consider only direct and one-stop itineraries, as this simplifies the model. Passengers traveling on itineraries with more than one stop represent less than 1% of the Australia market.

Airport landing charges, passenger fees and en-route charges were obtained from the RDC database [30]. These feed into the passenger cost, $paxcost_{a,j}$, and flight operating cost, $opcost_{a,j}$. Airport passenger numbers were calculated from the Sabre data. Movement capacity was inferred

Table 1: *Example of itinerary data obtained from Sabre and used within the model, for a direct Qantas itinerary from Melbourne International to Sydney airport.*

Airline	Origin Airport	Connecting Airport	Destination Airport	Avg. airfare (USD)	Passengers p.a.	Frequency of Service p.a.	Journey time (minutes)
QF	MEL	-	SYD	$183.56	1632758	10734	90

from the number, type and layout of runways. Each airport is assigned to a metropolitan area, indicated by an area code, and it is possible for multiple airports to be within the same metropolitan area. An example of the data for each airport is given in table 2.

Table 2: *Example of airport data used within the model, specifically of Melbourne International airport.*

IATA Code	Met. Area Code	Latitude	Longitude	Departing Passengers	Max. Movements per Hour	Pax. charge (USD)	Landing Charge (USD) for aircraft of size class 0;1;2;3;4;5;6;7;8
MEL	67	-37.66	144.83	16870988	53	16.05	133.66;210.26;384.47;476.23;618.48; 1780.78;2013.06;2709.88;3697.05

Household income and population statistics for each metropolitan area associated with an airport were obtained from the Australian Bureau of Statistics [31]. Together with the metropolitan area's special status, this forms the data required for each airport's location in Australia. Only Sydney and Melbourne have special status. The metropolitan area of Melbourne is given as an example in table 3. Data on the drive time between airports was collected from the Google Maps API [32].

Table 3: *Example of data on the metropolitan areas to which airports are allocated.*

Metropolitan Area Name	Area Code	Population	Household Income per Capita (USD)	Special Status
Melbourne	67	4665904	31100	True

Using the information on population and income statistics, itinerary data and drive time it was possible to estimate the coefficients of the demand model (equation 7) for the Australia market. The same data, together with itinerary frequency and number of segments, was used to estimate the coefficients of the market share model (equations 8 and 9). Results are given in table 4. However, the coefficient for airfares, θ, within the market share function is determined to be insignificant. This implies that airfare is correlated with the other market share parameters, and its effects difficult to determine via linear regression. The coefficient was instead obtained by selecting the value which reduced the error between the overall model's predicted airfares and the observed data. Using this approach, a value of $\theta = -0.01$ was found to most accurately reproduce airfares.

There are four main airline brands that operate in the Australian domestic market: Qantas (QF), Virgin Australia (VA), Jetstar (JQ) and Tigerair Australia (TT). Determining their behaviour is the focus and scope of this work. The fleet of aircraft available to each airline was determined using the

Table 4: *The coefficients of the passenger demand (equation 7) and market share models (equation 8), determined using linear regression. Demand:* $R^2 = 0.82$*, no. of observations = 436. Market share:* $R^2 = 0.85$*, no. of observations= 508.*

Coefficient	**Estimate**	**p-value**
α	0.54	0
β	0.96	1.1×10^{-5}
γ	-0.49	3.6×10^{-3}
δ	-2.92	0
ϵ	1.82	0
ζ	0.87	2.7×10^{-4}
η	-18.01	6.4×10^{-6}
θ	-0.00015	0.87
κ	-0.00885	4.6×10^{-13}
λ	0.841	0
μ	-1.672	7.9×10^{-11}
ffx_{Qantas}	-0.2912	2.0×10^{-6}
ffx_{Virgin}	-0.2975	1.29×10^{-11}
$ffx_{Jetstar}$	-0.5365	2.4×10^{-12}
$ffx_{Tigerair}$	-0.3676	9.2×10^{-3}

FlightGlobal fleet database [33] and the Australian Civil Aviation Safety Authority aircraft register [34]. Based primarily on seat numbers, each aircraft was assigned to one of nine size classes. A stereotypical aircraft within each class, together with the seat numbers of that class size, is given in table 5. The number of aircraft within each airline's fleet is also given in table 5.

Table 5: *Examples of a typical aircraft for each size class within the model, together with seat range for the size class. The number aircraft within each airline's fleet is also given.*

Size Class	**Typical Example Aircraft**	**Number of Seats**	**Qantas no. in fleet**	**Jetstar no. in fleet**	**Virgin no. in fleet**	**Tigerair no. in fleet**
Class0	Bombardier DASH Q300	30–75	28	0	14	0
Class1	Embraer E190	76–112	17	0	14	0
Class2	Airbus A319	113–144	0	0	2	0
Class3	Airbus A320	145–164	0	0	0	0
Class4	Boeing 737-800	165–210	76	53	75	18
Class5	Boeing 787	211–255	1	19	0	0
Class6	Airbus A330	256–294	28	0	6	0
Class7	Boeing 777	295–403	10	0	5	0
Class8	Airbus A380	404–550	12	0	0	0

Operating costs on each segment, and for each size class, were calculated using the properties of a typical aircraft within each class. For the non-fuel components, such as staff salaries and maintenance, a model previously developed with U.S. form 41 data was used [35]. Fuel burn for each size class, on each route, was derived using Piano-X software [36] and used to calculate fuel costs. Together with airport fees and en-route costs, this provides all inputs needed to determine

$opcost_{a,j}$ and $paxcost_{a,j}$ for each segment and aircraft type.

Ancilliary revenue per passenger, $arev_A$, was estimated by obtaining the total ancilliary revenue from company reports [37, 38] and dividing this by the total number of passengers per annum. It is found to be approximately \$40 (USD) per passenger in the Australia market, for both Qantas and Virgin groups, and is derived primarily from frequent flyer membership schemes.

Validation and Results

The model was validated against the 2014 Australian domestic market. Australia was chosen as it is relatively isolated geographically, and interference from the international aviation market is therefore minimal. It has a large domestic market, with around 40 million passengers per year served by the four main airline brands. Figure 1 displays the Australian aviation network. Each line represents a segment, with the thickness of the line proportional to the number of passengers carried on that segment per year. The model contains all metropolitan areas served by an airport with more than 5000 departures per year (14 per day). All airports and segments considered within the model are shown in figure 1. This captures 95% of the Australian domestic market by RPK.

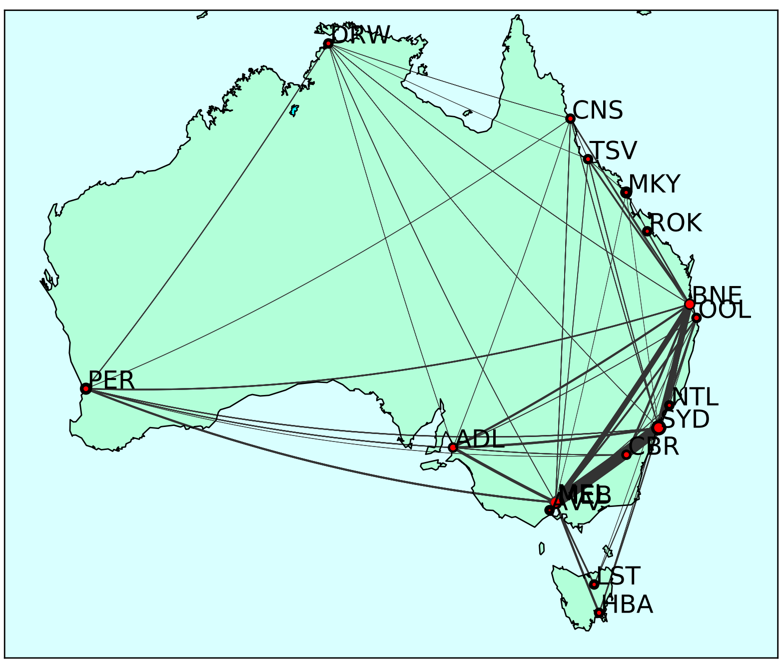

Figure 1: *A representation of the Australian domestic aviation network. Each line is a segment operated by one (or more) of the major airlines. The thickness of each line is proportional to annual passenger traffic.*

We compared predicted and actual segment flight frequency, airfares, and passenger numbers—on each network segment and for each airline—for the year 2014. Results for these comparisons, together with the coefficient of determination (R^2), can be found in Figures 2, 3 and 4 respectively.

The predicted and observed flight frequencies at the individual segment level show strong agreement, with $R^2 = 0.75$. The total number of flights predicted by the model was 381,000,

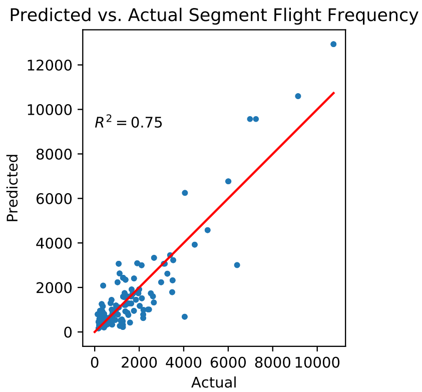

Figure 2: *The actual vs. predicted annual flight frequency, for every airline's segments within the Australia market in 2014.*

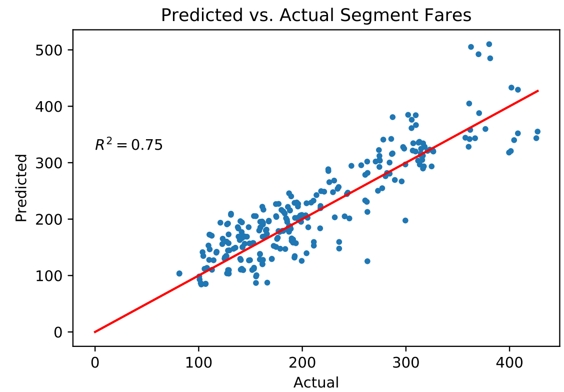

Figure 3: *The actual vs. predicted average airfares, for every airline's segments within the Australia market in 2014.*

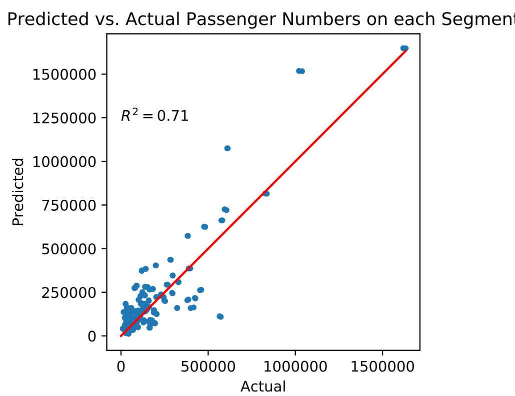

Figure 4: *The actual vs. predicted annual passenger numbers, for every airline's segments within the Australia market in 2014.*

which is remarkably close to the actual number of 379,000. With the number of flights driving the total carbon emissions, the ability to accurately reproduce this characteristic is a significant feature of the model. At the city level, looking at annual departures (see table 6), 10 of the 15 locations have predicted flight departure numbers within 25% of the recorded value. For the four largest areas by population (Melbourne, Sydney, Brisbane and Perth) the model has an error of less than 10%. This demonstrates the model's ability to predict the capacity requirements of large airports, and determine where constraints are likely to be encountered. Additionally, these outputs allow for the local air quality impacts of future growth in the market to be assessed. Whilst the percentage error in predicted departures appears to increase for airports with fewer flights, this is a consequence of absolute error in the model remaining relatively constant with respect to overall airport traffic. It suggests the model could be useful especially for large airport, market-wide and macro policy simulation, but not as reliable for specific micro-level operational decisions.

The average airfare given by the model, over the whole network, is \$191 (USD). This compares well with the average airfare of \$185 (USD), calculated directly from the Sabre dataset. The mean

Table 6: *Model predictions for the annual number of departing flights from each location within the Australia market.*

Location	Dep. flights model estimate	Dep. flights actual number	Percentage Error
Sydney	86336	83576	3%
Melbourne	85096	81723	4%
Brisbane	67745	62111	9%
Perth	16944	17020	0.5%
Adelaide	29053	24436	19%
Cairns	12518	16836	26%
Bilinga	14322	17934	20%
Canberra	20738	20152	3%
Darwin	8215	6749	22%
Townsville	8837	12215	28%
Hobart	10831	8192	32%
Mackay	5011	9507	47%
Launceston	5895	5733	3%
Newcastle	4432	5525	20%
Rockhampton	4704	6969	32%

absolute percentage error for the airfare predictions is 30%. These results suggest the model is accurately capturing operating cost and airfare competition effects; alongside demand, they are the primary determinants of airfares on a route.

The model estimates total passenger numbers of 50 million per year across the whole network, compared to the actual figure of 41 million. Considering that the model has fares that are slightly high on average, which should result in an under-prediction, this suggests that the demand model is over-predicting passenger flows. This can be corrected by adjusting the coefficient η within the demand model. Nevertheless, with a coefficient of determination of 0.71 for the model, passenger predictions across the network are highly correlated with the actual observed passenger flows.

As the model has the resolution to give predictions by airline, it is possible to work out market share. For example, on the route between Melbourne and Sydney, by far the busiest in the market, the model predicts market share by flight frequency as follows: Qantas-39%,Virgin-32%,Jetstar-19%,Tigerair-10%. This can be compared to market share obtained from real data: Qantas-40%, Virgin-34%, Jetstar-16%, Tigerair-10%. This agreement between model and reality extends to the whole market in general. Over the entire aviation network, the model predicts the percentage of all flights by airline as: Qantas-41%, Virgin-35%, Jetstar-17%, Tigerair-7%. The Sabre data gives the actual market share as: Qantas-40%, Virgin-33%, Jetstar-21%, Tigerair-6%.

Conclusions

A model which simulates competition between airlines has been described in detail and validated using data from the domestic Australian market in 2014. The performance of the model has been assessed by its ability to accurately reproduce passenger flows, airfares and flight frequency on all segments of each airline's network. In our comparison and validation we found strong correlation between modelled and observed data, and obtain R^2 values above 0.7 for all key variables. This is impressive, given the complexity of the model, as other models have a lower R^2 when predicting passenger flows alone [39]. At the city level, there is good agreement between observed and estimated aircraft departures. This is particularly true for the larger city airports, where predicted departure numbers have an error of less than 10%. Predictions of the airline's average airfare and market share are also consistent with the data.

The validation of the model is an important step in demonstrating it's potential use as a simulation tool for addressing the problems faced by the aviation industry discussed in the introduction. For example, the ability to accurately determine the total flight departures at a city level allows airport capacity requirements to be forecast, and potential constraints identified as the market grows. This gives the ability to make emissions predictions based on the model outputs, not just at the overall market level but at a city and airport level, allowing for local air quality and noise impact assessments of individual airports and how these could change under proposed expansion projects. Alongside airport capacity, the landing charges and passenger fees that are input into the model can be adjusted, and the resulting changes to the system will be predicted. The model therefore has the potential to be an informative and useful planning tool for airport authorities.

Airlines are key stakeholders in the aviation system, and the competition between them has a significant impact in determining the market. However they are also subject to regulation and oversight. Mechanisms exist to encourage behaviour that governments wish to incentivise. For example, the EU's Carbon Offsetting and Reduction Scheme for International Aviation (CORSIA) attempts to influence airlines to reduce their emissions. The design of such schemes is challenging as airlines operate with tight margins, and there are numerous examples throughout history of airlines going bust. Clearly a balance must be struck between incentivising desired outcomes and excessive penalisation of specific airlines. The model's ability to accurately predict airfares and route market share at the airline level demonstrates its capability as a powerful tool for assessing the impact of such measures on individual airlines and passengers. There is also the opportunity to simulate the impact of mergers or new entrants on a market and how this could affect airfares and market share.

Finally, the model also determines which aircraft type is utilised on which route. This gives it the ability to simulate the impacts of new technologies (such as next generation aircraft, and even hybrid-electric aircraft) as they become available. In turn, we can predict which routes will adopt the technology first, together with the impact on airlines, passenger demand, airfares and emissions.

Whilst this paper has focused on the Australian domestic market, and demonstrated the validity of the model in reproducing real world behaviour, future work will involve expanding the coverage to incorporate larger and more complex regions such as North America and Europe.

Author contribution statement

The authors confirm contribution to the paper as follows: KD, built the model and is principal author; LD, compiled and organised data for the model. AOS and AS, gave advice and help specifically with demand, market share and operating cost components. All authors contributed to the writing of the paper. All authors reviewed the results and contributed to the final version of the manuscript.

Acknowledgements

This work was funded through a grant from the Engineering and Physical Sciences Research Council (ref:EP/M027031/1). Their support is gratefully acknowledged.

References

[1] Boeing. *Commercial Market Outlook* (2017). URL www.boeing.com/commercial/market/commercial-market-outlook. Accessed July 2018.

[2] Airbus. *Global Market Forecast* (2017). URL www.airbus.com/aircraft/market/global-market-forecast.html. Accessed July 2018.

[3] Button, K. & Taylor, S. International air transportation and economic development. *Journal of air transport management* **6**, 209–222 (2000).

[4] Boeing. *Global Airport Congestion Study* (2014). URL www.aci-na.org/sites/default/files/07_-_global_airport_congestion.pdf. Accessed July 2018.

[5] Construction Intelligence Center. *Huge Investment Planned for Airport Construction Projects* (2018). URL www.construction-ic.com/Insight/huge-investment-planned-for-airport-construction-projects. Accessed July 2018.

[6] International Energy Agency (IEA). *CO2 Emissions from Fuel Combustion: Detailed Estimates* (2014).

[7] Moore, M. D. Misconceptions of electric aircraft and their emerging aviation markets. In *52nd Aerospace Sciences Meeting*, 0535 (2014).

[8] Dray, L. *et al.* Airline fleet replacement funded by a carbon tax: an integrated assessment. *Transport Policy* **34**, 75–84 (2014).

[9] International Civil Aviation Organization. *Resolution A39-3: Consolidated statement of continuing ICAO policies and practices related to environmental protection – Global Market-based Measure (MBM) scheme* (2016). URL www.icao.int/environmental-protection/CORSIA/Documents/Resolution_A39_3.pdf. Accessed October 2018.

[10] Hansen, M. Airline competition in a hub-dominated environment: An application of noncooperative game theory. *Transportation Research Part B: Methodological* **24**, 27 – 43 (1990).

[11] Wei, W. & Hansen, M. Airlines' competition in aircraft size and service frequency in duopoly markets. *Transportation Research Part E: Logistics and Transportation Review* **43**, 409–424 (2007).

[12] Evans, A. D. Comparing the impact of future airline network change on emissions in india and the united states. *Transportation Research Part D: Transport and Environment* **32**, 373–386 (2014).

[13] Evans, A. & Schäfer, A. The impact of airport capacity constraints on future growth in the us air transportation system. *Journal of Air Transport Management* **17**, 288–295 (2011).

[14] Vaze, V. & Barnhart, C. Modeling airline frequency competition for airport congestion mitigation. *Transportation Science* **46**, 512–535 (2012).

[15] Dobson, G. & Lederer, P. J. Airline scheduling and routing in a hub-and-spoke system. *Transportation Science* **27**, 281–297 (1993).

[16] Adler, N. Hub-spoke network choice under competition with an application to western europe. *Transportation science* **39**, 58–72 (2005).

[17] Brueckner, J. K. & Flores-Fillol, R. Airline schedule competition. *Review of Industrial Organization* **30**, 161–177 (2007).

[18] Brueckner, J. K. Schedule competition revisited. *Journal of Transport Economics and Policy (JTEP)* **44**, 261–285 (2010).

[19] Adler, N., Pels, E. & Nash, C. High-speed rail and air transport competition: Game engineering as tool for cost-benefit analysis. *Transportation Research Part B: Methodological* **44**, 812–833 (2010).

[20] Hansen, M. & Liu, Y. Airline competition and market frequency: A comparison of the s-curve and schedule delay models. *Transportation Research Part B: Methodological* **78**, 301–317 (2015).

[21] Harvey, D. Airline passenger traffic pattern within the united states. *J. Air L. & Com.* **18**, 157 (1951).

[22] Verleger, P. K. Models of the demand for air transportation. *The Bell Journal of Economics and Management Science* **3**, 437–457 (1972).

[23] Grosche, T., Rothlauf, F. & Heinzl, A. Gravity models for airline passenger volume estimation. *Journal of Air Transport Management* **13**, 175–183 (2007).

[24] Ben-Akiva, M. E. *Discrete Choice Analysis* (MIT press, 1985).

[25] Jamin, S., Schäfer, A., Ben-Akiva, M. E. & Waitz, I. A. Aviation emissions and abatement policies in the united states: a city-pair analysis. *Transportation Research Part D: Transport and Environment* **9**, 295 – 317 (2004).

[26] Koppelman, F. S., Coldren, G. M. & Parker, R. A. Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological* **42**, 263 – 273 (2008). A Tribute to the Career of Frank Koppelman.

[27] Hsiao, C.-Y. & Hansen, M. A passenger demand model for air transportation in a hub-and-spoke network. *Transportation Research Part E: Logistics and Transportation Review* **47**, 1112–1125 (2011).

[28] Nikaidô, H. & Isoda, K. Note on non-cooperative convex game. *Pacific Journal of Mathematics* **5**, 807–815 (1955).

[29] Sabre Corporation. *Aviation Data Intelligence database* (2016). URL www.sabreairlinesolutions.com/home/software_solutions/airports/. Accessed July 2018.

[30] RDC Aviation Ltd. *Aviation airport and enroute charges database* (2017). URL www.airportcharges.com. Accessed July 2018.

[31] Australian Bureau of Statistics. Regional income and population statistics. URL stat.data.abs.gov.au. Accessed November 2017.

[32] Google. *Distance Matrix API* (2018). URL developers.google.com/maps/documentation/distance-matrix/start. Accessed April 2018.

[33] FlightGlobal. Global fleet database. URL www.flightglobal.com. Accessed November 2017.

[34] Civil Aviation Safety Authority. Aircraft register. URL www.casa.gov.au. Accessed January 2018.

[35] Dray, L. M., Schäfer, A. W. & Al Zayat, K. The global potential for co2 emissions reduction from jet engine passenger aircraft. *Transportation Research Record: Journal of the Transportation Research Board* (2018).

[36] Lissys Ltd. *Piano-X*. URL www.piano.aero. Accessed November 2017.

[37] Virgin Australia Group. *Annual financial report* (2015). URL www.virginaustralia.com/cs/groups/internetcontent/@wc/documents/webcontent/~edisp/annual-financial-report-2015.pdf. Accessed February 2018.

[38] Qantas. *Annual report* (2015). URL www.qantas.com.au/infodetail/about/investors/2015AnnualReport.pdf. Accessed February 2018.

[39] Reynolds, T. *et al.* Modelling environmental and economic impacts of aviation: introducing the aviation integrated modelling project. In *7th AIAA ATIO Conf, 2nd CEIAT Int'l Conf on Innov and Integr in Aero Sciences, 17th LTA Systems Tech Conf; followed by 2nd TEOS Forum*, 7751 (2007).