A Game-Theoretic Modeling Approach to Air Traffic Forecasting

Vikrant Vaze, Reed Harder

Thayer School of Engineering
Dartmouth College

ATM2017
Motivation: Modeling Airline Capacity Decisions

- Airlines allocate capacity (frequency and seats-per-flight) across a network of airports, in the face of competition.
- Capacity decisions affect operating costs as well as fare revenues, but are also interdependent with decisions of other carriers, and thus have significant implications for overall system performance.
Motivation: Modeling Airline Capacity Decisions

- Misallocation of capacity has been implicated in...
  - Costs worth billions of dollars annually (Ball et al., 2010),
  - Wastage of system resources (Morisset and Odoni, 2011),
  - Passenger inconvenience (Barnhart, Fearing and Vaze, 2014; Wittman, 2014),
  - Environmental damages (Schumer and Maloney, 2008).

- Frequency competition in particular implicated in increased airport congestion (Vaze and Barnhart, 2012a).
Motivation: Modeling Airline Capacity Decisions

- Aviation planners leverage airport traffic forecasts in evaluation of technological improvements, airport capacity expansion, workforce planning.
- Want to model these decisions and competitive dynamics, evaluate different scenarios in a changing environment.

US Airways Network, Q1 2007
Motivation: Modeling Airline Capacity Decisions

- Aviation planners leverage airport traffic forecasts in evaluation of technological improvements, airport capacity expansion, workforce planning.
- Want to model these decisions and competitive dynamics, evaluate different scenarios in a changing environment.

US Airways Network, Q4 2007
Why a game-theoretic approach?

- Capacity and fare decisions of different airlines are interdependent: interaction determines market share, and ultimately profits.

Why a two-stage game?

- Capacity decisions made under approximate knowledge of fares.
- Capacity decided in Stage I and fares in Stage II.
Literature Review

- Single-stage frequency-only game (Hansen, 1990; Vaze and Barnhart, 2012a; 2012b; 2015).

- Single-stage frequency-seat game (Hong and Harker, 1992; Adler and Berechman, 2001; Wei and Hansen, 2007).

- Single-stage capacity-fare game (Adler, 2001; 2005; Adler, Pels and Nash, 2010; Brueckner, 2010).

- Treatment of two-stage capacity-fare games more limited, largely focuses on single-market duopoly (Dobson and Lederer, 1993; Schipper, Rietveld and Nijkamp, 2003; Brueckner and Flores-Fillol, 2007; Hansen and Liu, 2015).
  - Little work that bridges analytical, computational, and empirical results.
Overview

- Introduce two-stage game model.

- Analysis of simplified model: demonstrate that it has properties that ensure a tractable and credible solution.

- Use numerical approximations of 2\textsuperscript{nd} stage payoff functions to show that these properties extend to more realistic scenarios.

- Empirical calibration and validation using data from U.S. air transportation network.
Two-Stage Game Model: Frequency versus Seats-per-Flight

- Assume seats-per-flight held constant because...
  - Limited effect on passenger demand, and
  - Shows lower variability across time and segments.

- For 10-year data across the U.S. (2005-2014)

<table>
<thead>
<tr>
<th>Airline</th>
<th>Coefficient of Variation: Across Years</th>
<th>Coefficient of Variation: Across Segments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seats-per-Flight</td>
<td>Frequency</td>
</tr>
<tr>
<td>American Airlines (AA)</td>
<td>4.7%</td>
<td>12.7%</td>
</tr>
<tr>
<td>Delta Airlines (DL)</td>
<td>6.3%</td>
<td>21.7%</td>
</tr>
<tr>
<td>United Airlines (UA)</td>
<td>5.3%</td>
<td>27.3%</td>
</tr>
<tr>
<td>US Airways (US)</td>
<td>6.3%</td>
<td>16.9%</td>
</tr>
<tr>
<td>Southwest Airlines (WN)</td>
<td>3.5%</td>
<td>16.6%</td>
</tr>
</tbody>
</table>
Two-Stage Frequency-Fare Model

- Multinomial logit model of market share: Consistent with several prior studies. Two most common formulations:

\[
MS_{a,m} = \frac{e^{\alpha \ln(f_a) - \beta p_{a,m}}}{N_m + \sum_{i \in A_m} e^{\alpha \ln(f_{i,m}) - \beta p_{i,m}}}
\]

“S-Curve Formulation”

\[
MS_{a,m} = \frac{e^{-\varphi f_{a,m} - r - \beta p_{a,m}}}{N_m + \sum_{i \in A_m} e^{-\varphi f_{i,m} - r - \beta p_{i,m}}}
\]

“Schedule Delay Formulation”

- Payoff function \( (\pi_a) \):

\[
\pi_a = \sum_{m \in K_a} \min(M_m(\ MS_{a,m}), f_{a,m} s_{a,m}) \ p_{a,m} - c_{a,m} f_{a,m}
\]

Revenue

Operating Cost
Sub-game Perfect Pure Nash Equilibrium

- For a given frequency decision, fare decision modeled by second-stage pure strategy Nash equilibrium.
- With knowledge of fare equilibria, players choose frequency.
- Predict frequency decisions as 1st-stage outcomes of subgame-perfect pure strategy equilibrium.
- Many games are quite difficult to solve.
- Is this a reasonable solution concept in this case?

Starting assumptions (strong):
- Two airlines, single segment
- No connecting passengers
- Unlimited seating capacity
- No no-fly alternative
Analytical Results

Under these strong assumptions, for either formulation:

- **Proposition 1:** Second stage game has a unique equilibrium.
  - Supports the credibility and practical tractability of second stage equilibrium.
Analytical Results

Under these strong assumptions, for either formulation:

- **Proposition 2:** Each airline’s payoff in the first stage game is a submodular function in the overall strategy space, i.e.
  \[
  \frac{\partial^2 \pi_i}{\partial f_1 \partial f_2} < 0 \quad \text{for } i \in \{1, 2\}.
  \]

- **Corollary:** The game is supermodular: can transform strategy space such that
  \[
  \frac{\partial^2 \pi_i}{\partial f_1 \partial f_2} > 0 \quad \text{for } i \in \{1, 2\}
  \]
  Ensures the convergence of several iterative learning dynamics, and the existence of first stage equilibrium.
Analytical Results

Under these strong assumptions, for either formulation:

- **Proposition 3:** Each airline’s payoff in the first stage game is concave in its own strategy, i.e. \( \frac{\partial^2 \pi_i}{\partial f_i^2} < 0 \) for \( i \in \{1, 2\} \) in plausible parameter ranges: \( \alpha < 2.44 \) in s-curve model. Empirical range: \( \alpha = 1.3 - 1.7 \) (Belobaba, 2016)
  - Ensures that individual payoff maximization problems are easy to solve (even for large scale networks), existence of 1st stage equilibrium.
  - Not guaranteed for single stage frequency game (Hansen, 1990).
- Simple model: nice analytical properties. But these results rest on strong assumptions. What if we relax them?
Numerical Extensions

- Approach: numerically approximate second stage payoffs with polynomials, evaluate approximation function properties.
- Relax earlier assumptions and evaluate a wide variety of scenarios:
  - for 1-player, 2-player and 3-player situations
  - in the presence of a no-fly alternative
  - with finite seating
  - with connecting passengers
  - with various values of $\alpha$, $\beta$, $N$, $\varphi$, $r$ utility parameters
  - with various seating capacities
Numerical Extensions

1) Heuristic to compute 2nd stage equilibria

Player 1 fare equilibria,
$\alpha=1.29$, $\beta=0.0045$, $N=0.5$
Numerical Extensions

1) Heuristic to compute 2nd stage equilibria

\[ \pi_{a,m} = \min(M_m(M_{S,a,m}, f_{a,m}s_{a,m}) p_{a,m} - c_{a,m} f_{a,m}) \]

2) Compute payoffs

Player 1 fare equilibria, \( \alpha=1.29, \beta=0.0045, N=0.5 \)
Numerical Extensions

1) Heuristic to compute 2nd stage equilibria

2) Compute payoffs

\[
\pi_{a,m} = \min(M_m(MS_{a,m}), f_{a,m}s_{a,m}) \ p_{a,m} - c_{a,m} f_{a,m}
\]

3) Fit payoffs to polynomial function of frequency profile.

Player 1 fare equilibria, \( \alpha = 1.29, \beta = 0.0045, \ N = 0.5 \)
Quadratic Approximations

Fit polynomial approximations of profits as functions of frequencies of all players:

- 1-player non-stop only game: \( \pi_1 \sim \beta_0 + \beta_1 f_1 + \beta_2 f_1^2 \)
- 2-player non-stop only game: \( \pi_1 \sim \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_1^2 + \beta_4 f_2^2 + \beta_5 f_1 f_2 \)
- 3-player non-stop only game: \( \pi_1 \sim \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_3 + \beta_4 f_1^2 + \beta_5 f_2^2 + \beta_6 f_3^2 + \beta_7 f_1 f_2 + \beta_8 f_1 f_3 + \beta_9 f_2 f_3 \)
- 2-player non-stop and one-stop game: \( \pi_1 \sim \beta_0 + \beta_1 f_{11} + \beta_2 f_{12} + \beta_3 f_{21} + \beta_4 f_{22} + \beta_5 f_{11}^2 + \beta_6 f_{12}^2 + \beta_7 f_{21}^2 + \beta_8 f_{22}^2 + \beta_9 f_{11} f_{21} + \beta_{10} f_{21} f_{22} \)
- Etc.
Quadratic Approximations: Results

- Model parameters varied one-at-a-time over ranges found in literature or practice: N (0 to 1), α (1 to 2), r (.1 to 1), ϕ (1 to 10), β (-0.001 to -0.01) and seats (25-250), players (1-3)
- In nearly all cases, excellent fit: $R^2$ of model > 0.9, ranging from 0.91-0.9998. Additionally, signs of all estimated parameter values ensure submodularity and concavity.

\[ \pi_1 \sim \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_1^2 + \beta_4 f_2^2 + \beta_5 f_1 f_2 \]

<table>
<thead>
<tr>
<th>Seats</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\beta_5$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>122200.0</td>
<td>18135.0</td>
<td>-17856.0</td>
<td>-494.4</td>
<td>686.1</td>
<td>-533.0</td>
<td>95.5%</td>
</tr>
<tr>
<td>225</td>
<td>122250.0</td>
<td>18130.0</td>
<td>-17861.0</td>
<td>-494.2</td>
<td>686.2</td>
<td>-532.9</td>
<td>95.5%</td>
</tr>
<tr>
<td>200</td>
<td>122400.0</td>
<td>18115.0</td>
<td>-17876.0</td>
<td>-493.9</td>
<td>686.6</td>
<td>-532.4</td>
<td>95.5%</td>
</tr>
<tr>
<td>175</td>
<td>122640.0</td>
<td>18095.0</td>
<td>-17901.0</td>
<td>-493.4</td>
<td>687.2</td>
<td>-531.6</td>
<td>95.5%</td>
</tr>
<tr>
<td>150</td>
<td>123470.0</td>
<td>18030.0</td>
<td>-17989.0</td>
<td>-492.1</td>
<td>689.6</td>
<td>-529.0</td>
<td>95.5%</td>
</tr>
<tr>
<td>100</td>
<td>129430.0</td>
<td>17885.0</td>
<td>-18838.0</td>
<td>-495.5</td>
<td>716.3</td>
<td>-514.1</td>
<td>93.9%</td>
</tr>
<tr>
<td>75</td>
<td>136710.0</td>
<td>18277.0</td>
<td>-20301.0</td>
<td>-512.8</td>
<td>773.0</td>
<td>-518.1</td>
<td>92.8%</td>
</tr>
<tr>
<td>25</td>
<td>104880.0</td>
<td>27814.0</td>
<td>-18929.0</td>
<td>-743.5</td>
<td>773.0</td>
<td>-846.6</td>
<td>96.9%</td>
</tr>
</tbody>
</table>

For 2-player s-curve, non-stop passengers case with $\alpha=1.29$, $\beta=0.0045$, N=0.5, and different seats-per-flight values.
Payoff Approximations: Further Extensions

- Results suggest the model is highly tractable and can be efficiently solved by an iterative best response heuristic.
- Can we make this more rigorous?
Payoff Approximations: Further Extensions

- **Proposition 4**: Our N-player game, with quadratic, concave, and submodular payoff functions and non-negative strategy spaces, assuming interaction terms are same within an airline, belongs to the class of quasi-aggregative games (as defined by Jensen, 2010), and the *myopic best response* algorithm converges to the set of Nash Equilibrium.

- Estimated coefficients are consistent with guaranteed unique first-stage equilibrium according to Rosen’s diagonal strict concavity condition (Rosen, 1965), with a few exceptions (high $\alpha$, $>1.7$).
Airline Network Case Study: Description

- 11-airport network in the Western U.S.
  - In Q1 2007: 4 carriers, 68 carrier-segments (each with one frequency value to be predicted).
- Actual BTS (Bureau of Transportation Statistics) data.
Airline Network Case Study: Description

- Solved using iterative optimization of quadratic payoffs.
- Payoffs initialized with fitted functions above, transformed in each market with cost/demand data.
- Additionally, enforce aircraft availability constraints.
Convergence Results

- Model converges within 6-7 iterations (four optimizations each, one per airline) in <1 second.
Model Calibration and Results

- Minimize the Mean Absolute Percentage Error (MAPE)
  \[ MAPE = \frac{\sum_{c\in CM} |\hat{f}_{cm} - f_{cm}|}{\sum_{c\in CM} f_{cm}} \]
  between predicted and observed frequencies (similar to Vaze and Barnhart, 2012a; Cadarso et al. 2015)

- Segment-carriers grouped as: ‘3-player’, ‘2-player hub-to-hub’, ‘other 2-player’, and ‘1-player’.
  - Calibrate 11 total coefficients, transformed by demand/cost
  - Adjusted using a stochastic gradient approximation algorithm.
  - MAPE = 18.4%
  - 49% abs. errors < 1, 78% < 2
Out-of-Sample Testing: Example

- Train coefficients on past data, test on future data, with new demand, cost, market and player attributes.
- E.g.: Train 11 coefficients on 2007 Q1 data. Test on 2007 Q4.
- Out-of-Sample Testing **MAPE = 20.6%**
  - 47% abs. errors < 1, 73% < 2
- Error Adjusted **MAPE = 11.2%**
  - 72% abs. errors < 1, 92% < 2

Note: Compare to single airport MAPEs in 14%-20% range from previous research as benchmark (Vaze and Barnhart, 2012).

- Market PDX-SFO does not exist in 2007 Q1, only in Q4.
- Using Q1-trained coefficients, predict frequencies of this new market:

<table>
<thead>
<tr>
<th>PDX-SFO</th>
<th>True Frequency</th>
<th>Estimated Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>3.02</td>
<td>3.52</td>
</tr>
<tr>
<td>UA</td>
<td>6.11</td>
<td>7.28</td>
</tr>
</tbody>
</table>
Multi-Year Validation Results, 2007-2015

- Average MAPE < 20% for up to a couple years of look-ahead.
- Increases almost monotonically as the look-ahead period increases.
- Reasonable prediction tool for short-to-medium term horizon.
Out of Sample Aggregated Results

<table>
<thead>
<tr>
<th>Aggregated Measure</th>
<th>MAPE calculation</th>
<th>Adj. MAPE (Avg. Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>$\frac{\sum_{a \in A} \sum_{m \in K_a}</td>
<td>f_{a,m} - \sum_{m \in K_a} f_{a,m}</td>
</tr>
<tr>
<td>Coefficient category</td>
<td>$\frac{\sum_{\text{coef \in COEF}} \sum_{m \in \text{CM}}</td>
<td>\hat{f}<em>{\text{coef},cm} - \sum</em>{m \in \text{CM}} f_{\text{coef},cm}</td>
</tr>
<tr>
<td>Market</td>
<td>$\frac{\sum_{m \in K} \sum_{a \in A_m}</td>
<td>f_{a,m} - \sum_{a \in A_m} f_{a,m}</td>
</tr>
<tr>
<td>Airport</td>
<td>$\frac{\sum_{a \in AP} \sum_{m \in \text{CM}}</td>
<td>\hat{f}<em>{ap,cm} - \sum</em>{m \in \text{CM}} f_{ap,cm}</td>
</tr>
</tbody>
</table>

Out of sample predictions of daily flights deployed at all airports in network, Q4 2007. Calibrated on Q1 2007 data.
Out of Sample Aggregated Results

<table>
<thead>
<tr>
<th>Airport</th>
<th>Observed Flights Q1</th>
<th>Observed Flights Q4</th>
<th>Predicted Flights Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAX</td>
<td>137</td>
<td>128</td>
<td>132</td>
</tr>
<tr>
<td>SJC</td>
<td>60</td>
<td>64</td>
<td>61</td>
</tr>
<tr>
<td>LAS</td>
<td>154</td>
<td>167</td>
<td>168</td>
</tr>
<tr>
<td>SAN</td>
<td>101</td>
<td>108</td>
<td>110</td>
</tr>
<tr>
<td>SMF</td>
<td>70</td>
<td>71</td>
<td>70</td>
</tr>
<tr>
<td>SEA</td>
<td>98</td>
<td>105</td>
<td>104</td>
</tr>
<tr>
<td>PDX</td>
<td>33</td>
<td>41</td>
<td>43</td>
</tr>
<tr>
<td>SFO</td>
<td>67</td>
<td>92</td>
<td>99</td>
</tr>
<tr>
<td>ONT</td>
<td>61</td>
<td>61</td>
<td>62</td>
</tr>
<tr>
<td>PHX</td>
<td>159</td>
<td>159</td>
<td>153</td>
</tr>
<tr>
<td>OAK</td>
<td>88</td>
<td>88</td>
<td>85</td>
</tr>
</tbody>
</table>

Out of sample predictions of daily flights deployed at all airports in network, Q4 2007. Calibrated on Q1 2007 data.
Out of Sample Aggregated Results

Airport level aggregated MAPE values (adjusted for calibration error) for varying look-ahead durations

<table>
<thead>
<tr>
<th>Airport</th>
<th>Observed Flights Q1</th>
<th>Observed Flights Q4</th>
<th>Predicted Flights Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAX</td>
<td>137</td>
<td>128</td>
<td>132</td>
</tr>
<tr>
<td>SJC</td>
<td>60</td>
<td>64</td>
<td>61</td>
</tr>
<tr>
<td>LAS</td>
<td>154</td>
<td>167</td>
<td>168</td>
</tr>
<tr>
<td>SAN</td>
<td>101</td>
<td>108</td>
<td>110</td>
</tr>
<tr>
<td>SMF</td>
<td>70</td>
<td>71</td>
<td>70</td>
</tr>
<tr>
<td>SEA</td>
<td>98</td>
<td>105</td>
<td>104</td>
</tr>
<tr>
<td>PDX</td>
<td>33</td>
<td>41</td>
<td>43</td>
</tr>
<tr>
<td>SFO</td>
<td>67</td>
<td>92</td>
<td>99</td>
</tr>
<tr>
<td>ONT</td>
<td>61</td>
<td>61</td>
<td>62</td>
</tr>
<tr>
<td>PHX</td>
<td>159</td>
<td>159</td>
<td>153</td>
</tr>
<tr>
<td>OAK</td>
<td>88</td>
<td>88</td>
<td>85</td>
</tr>
</tbody>
</table>

Out of sample predictions of daily flights deployed at all airports in network, Q4 2007. Calibrated on Q1 2007 data.
Takeaways

- Two-stage frequency-fare games are behaviorally consistent with the actual airline decision process.
- For simple cases, the analytical properties indicate very well-behaved games: existence, uniqueness, convergence, stability.
- These nice properties extend to more complex cases too.
- Model predictions approximate actual airline decisions reasonably well within a short-to-medium term horizon.
Takeaways

- Theory-motivated framework for informing 2-stage game-theoretic predictive model.
- Refinements of model could serve as a scenario analysis tool to aid in planning and policy-making decision-support.
- Future work: extension to larger networks, more flexible payoff parameterizations to capture market and airline characteristics, integration with statistical forecasting models, deeper understanding of schedule competition.
References (1)

References (2)

References (3)


Thayer School of Engineering at Dartmouth
14 Engineering Drive
Hanover, NH 03755