Predicting a Dramatic Contraction in the 10-Year Passenger Demand

Daniel Y. Suh
Megan S. Ryerson
University of Pennsylvania

6/29/2018
8th International Conference on Research in Air Transportation
Outline

• Introduction
• How we are planning for airports currently
• What are we doing wrong?
• Alternative approaches
• Predicting demand uncertainty
• Discussion of Results
• Implications
• Application
Airport Master Plans

- Guide future airport growth and development
- Airfield facilities (runways, taxiways)
- Terminal facilities (gates, concourses, pedestrian walkways)
- Landside facilities (access roads, parking, rental car facilities)
Airport Master Plans

Source: ATL Airport Master Plan (2015)
Airport Master Plans

Source: ATL Airport Master Plan (2015)
Demand Uncertainty and Airport Expansions (St. Louis Airport)

STL Airport (Source: ACRP Report 76)
Systematic Optimism in 10-year Forecasts (top 64 airports, 1995-2005)
Alternative Airport Planning Frameworks

- Theoretical frameworks
- No empirical evidence of efficacy
- High costs of implementation
- Missing areas of inquiry in the technical evaluation and improvement in airport planning techniques

• Dynamic Strategic Planning (De Neufville, 2000)
• Flexible Strategic Planning (Burghouwt, 2007)
• Adaptive Policy-Making (Kwakkel, 2010)
• Adaptive Airport Strategic Planning (Kwakkel et al., 2010)
Systematic Optimism in 10-year Forecasts (top 64 airports, 1995-2005)
Case 1: Infrastructure investments maybe justified (eventually…)

Miami Int’l Airport (MIA)  San Francisco Int’l Airport (SFO)
Case 2:
Maybe not a good idea...

**St. Louis Lambert Int’l Airport (STL)**

**Pittsburgh Int’l Airport (PIT)**
Demand Uncertainty and Airport Expansions (St. Louis Airport)

Master plan Runway completed

$1.3 Billion

Rarely Used

Source: ACRP Report 76
Research Question: What are the operational and socioeconomic characteristics of an airport on the verge of experiencing a severe contraction in passenger volumes?
Methodology

Logistic Regression

\[ \theta \]

A severe contraction in passenger volumes in the next 10 years (1)

Stable passenger demand (0)

\[ X_1 \quad X_2 \quad X_n \]
Methodology

Operational and Socioeconomic variables

Static and Dynamic variables

A severe contraction in passenger volumes in the next 10 years (1)

Stable passenger demand (0)
Methodology

Operational and Socioeconomic variables

Static and Dynamic variables

Logistic Regression

A severe contraction in passenger volumes in the next 10 years (1)

Stable passenger demand (0)

Data-driven definition
Data-Driven Definition of a Severe Contraction

• Data: Annual enplanements data (FAA) from 1995 to 2015

• Study airports: 64 major airports in the top 50 metropolitan statistical areas (MSA)

• Outcome: 10-year % change in passenger volumes

\[ P_b = \frac{E_{b+10} - E_b}{E_b} \times 100 \]

• 11 base years (1995 – 2005) for 64 airports (N = 704)
Distribution of 10-year % change in passenger volumes

Normal distribution (almost)

Multiple peaks
Distribution of 10-year % change in passenger volumes

Gaussian Mixture Model
- Assumes the data points came from a mixture of normal distributions
- Posterior probabilities of each data point belonging to each of the distributions (4)
- Assign each point to a distribution with the highest posterior probability

N = 704
Distribution of 10-year % change in passenger volumes

Gaussian Mixture Model
- Assumes the data points came from a mixture of normal distributions
- Posterior probabilities of each data point belonging to each of the distributions
  \((4)\)
- Assign each point to a distribution with the highest posterior probability

\[
N = 704
\]
## Binary Outcome Variable

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Growth Cluster (n=9)</th>
<th>Cyclical Cluster (n=559)</th>
<th>Contraction Cluster (n=136)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>167.20</td>
<td>22.40</td>
<td>-28.61</td>
</tr>
<tr>
<td>Median</td>
<td>138.70</td>
<td>18.09</td>
<td>-22.00</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>87.51</td>
<td>23.08</td>
<td>18.26</td>
</tr>
<tr>
<td>Max</td>
<td>395.70</td>
<td>99.15</td>
<td>-11.23</td>
</tr>
<tr>
<td>Min</td>
<td>110.00</td>
<td>-10.74</td>
<td>-80.79</td>
</tr>
</tbody>
</table>
Methodology

Operational and Socioeconomic variables

Static and Dynamic variables

Logistic Regression

A severe contraction in passenger volumes in the next 10 years (1)

Stable passenger demand (0)
Predictors

• Static (point-in-time) socioeconomic and operational variables in base year values
  • Population of Philadelphia MSA in base year 2000
• Corresponding dynamic (change-over-time) variables in 5-year average annual % change values up to base year
  • Average annual % change in population of Philadelphia MSA from 1995 to 2000
## Predictors

<table>
<thead>
<tr>
<th>Variables in base year numbers</th>
<th>Unit</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passengers</td>
<td>Persons (millions)</td>
<td>8.42</td>
<td>7.95</td>
<td>FAA</td>
</tr>
<tr>
<td>Airport competition</td>
<td>Unitless</td>
<td>3.74</td>
<td>5.57</td>
<td>FAA</td>
</tr>
<tr>
<td>Connecting passenger share</td>
<td>Proportion</td>
<td>0.47</td>
<td>0.11</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>Avg. number of seats per aircraft</td>
<td>Seats</td>
<td>118.40</td>
<td>26.87</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Avg. ticket price</td>
<td>Dollars</td>
<td>227.70</td>
<td>53.01</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>HHI</td>
<td>Unitless</td>
<td>0.35</td>
<td>0.20</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Population</td>
<td>Persons (millions)</td>
<td>3.56</td>
<td>3.44</td>
<td>Census</td>
</tr>
<tr>
<td>Per capita income</td>
<td>Dollars (thousands)</td>
<td>45.87</td>
<td>7.91</td>
<td>BEA</td>
</tr>
<tr>
<td>Service sector employment</td>
<td>Persons (millions)</td>
<td>0.92</td>
<td>0.91</td>
<td>Census</td>
</tr>
</tbody>
</table>

- **9 Static Predictors**
- **9 Dynamic Predictors**
## Predictors

<table>
<thead>
<tr>
<th>Variables in base year numbers</th>
<th>Unit</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passengers</td>
<td>Persons (millions)</td>
<td>8.42</td>
<td>7.95</td>
<td>FAA</td>
</tr>
<tr>
<td>Airport competition</td>
<td>Unitless</td>
<td>3.74</td>
<td>5.57</td>
<td>FAA</td>
</tr>
<tr>
<td>Connecting passenger share</td>
<td>Proportion</td>
<td>0.47</td>
<td>0.11</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>Avg. number of seats per aircraft</td>
<td>Seats</td>
<td>118.40</td>
<td>26.87</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Avg. ticket price</td>
<td>Dollars</td>
<td>227.70</td>
<td>53.01</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>HHI</td>
<td>Unitless</td>
<td>0.35</td>
<td>0.20</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Population</td>
<td>Persons (millions)</td>
<td>3.56</td>
<td>3.44</td>
<td>Census</td>
</tr>
<tr>
<td>Per capita income</td>
<td>Dollars (thousands)</td>
<td>45.87</td>
<td>7.91</td>
<td>BEA</td>
</tr>
<tr>
<td>Service sector employment</td>
<td>Persons (millions)</td>
<td>0.92</td>
<td>0.91</td>
<td>Census</td>
</tr>
</tbody>
</table>

\[ \sum \text{Enplanements for neighboring airport (< 100mi)} \]

Distance to neighboring airport
Predictors

Herfindahl-Hirschman Index (HHI)

Measure of competition among firms (airlines)
In an industry (airport)

\[ HHI_\alpha = \sum_i m_{ai}^2 \]

where \( m_{ai} \) is a proportion of seats provided by airline \( \alpha \).

Lower HHI = greater competition
Higher HHI = lower competition, dominance of market share among few firms (airlines)
Modeling Framework

Training Data (n = 556) → Binary Logistic Regression → Test Data (n = 139) → Model → Prediction
Modeling Framework

- Training Data (n = 556)
- Binary Logistic Regression
- Test Data (n = 139)
- Model
- Prediction
ROC Curve

Best cutoff = 43.9%
84% True Positive Rate
23% False Positive Rate
## Final Model Output

|                                      | Odds ratio | $P > |z|$  |
|--------------------------------------|------------|--------|
| (Intercept)                          | 0.1200     | 0.000*** |
| Airport competition % change (5AAC)  | 0.6121     | 0.000*** |
| Connecting passenger share           | 1.5547     | 0.000*** |
| Connecting passenger share % change (5AAC) | 0.9652 | 0.005**  |
| Avg. number of seats per aircraft    | 0.7087     | 0.000*** |
| Avg. ticket price                    | 0.6123     | 0.000*** |
| HHI                                  | 2.2339     | 0.004**  |
| HHI % change (5AAC)                  | 1.3456     | 0.003**  |
| Population % change (5AAC)           | 0.2010     | 0.000*** |
| Per capita income                    | 1.5385     | 0.001**  |
| Service sector employment            | 0.4056     | 0.001**  |

$n = 556$

AIC = 422.66

* $p < 0.1$  ** $p < 0.01$  *** $p < 0.001$
Predictors of a severe contraction in demand in the next 10 years

More likely

- Connecting passenger share (1.6)
- HHI (2.2)
- HHI 5AAC (1.3)
- Per capita income (1.5)

Airports with high transfer activities with higher market concentration of airlines

(Hub airports dominated by few airlines)

Less likely

- Population 5AAC (0.2)
- Service sector employment (0.4)
- Airport competition 5AAC (0.6)
- Connecting passenger share 5AAC (0.9)
- Avg. number of seats per aircraft (0.7)
- Avg. ticket price (0.6)
Predictors of a severe contraction in demand in the next 10 years

More likely
- Connecting passenger share (1.6)
- HHI (2.2)
- HHI 5AAC (1.3)
- Per capita income (1.5)

Airports in MSAs with growing population and growing regional airport demand as well as strong service sector employment (Growing market)

Less likely
- Population 5AAC (0.2)
- Service sector employment (0.4)
- Airport competition 5AAC (0.6)
- Connecting passenger share 5AAC (0.9)
- Avg. number of seats per aircraft (0.7)
- Avg. ticket price (0.6)
Predictors of a severe contraction in demand in the next 10 years

**More likely**
- Connecting passenger share (1.6)
- HHI (2.2)
- HHI 5AAC (1.3)
- Per capita income (1.5)

**Less likely**
- Population 5AAC (0.2)
- Service sector employment (0.4)
- Airport competition 5AAC (0.6)
- Connecting passenger share 5AAC (0.9)
- Avg. number of seats per aircraft (0.7)
- Avg. ticket price (0.6)

**Airports in MSAs with growing population and growing regional airport demand as well as strong service sector employment**

(Growing market)

**Airports with growing share of connecting passengers, larger aircraft, and higher ticket prices**

(Diverse mix of traffic)
Demand Uncertainty and Airport Expansions
(St. Louis Airport)

$1.3 Billion

Rarely Used

Source: ACRP Report 76
STL in 1997

**MSA**
- Below average population growth in the past 5 years
- Below average service sector employment in 1997

**Airport**
- Smaller aircraft than average
- Passengers making more O-D trips and less connecting trips over the years
- A hub airline becoming more dominant at STL (high rate of growth in HHI)
STL in 1997

- Below average population growth in the past 5 years
- Below average service sector employment in 1997

MSA

- Smaller aircraft than average
- Passengers making more O-D trips and less connecting trips over the years
- A hub airline becoming more dominant at STL (high rate of growth in HHI)

Airport

Predicted probability 85%

Threshold established using a holdout sample: 44%
Demand Uncertainty and Stability

DIVERSIFIED DEMAND & SUPPLY
(connecting, international, more airline competition)

REGIONAL GROWTH
(population, service sector employment)
Implications & Applications

• Diversified demand and supply of air service
• Regional health of cities and metropolitan areas
• Supports existing literature linking air travel demand and socioeconomic characteristics
• Additional insight during planning and decision-making process
• Framework for improving forecast accuracy
  • Propensity score matching (reference class forecasting)
Reference Class Forecasting

Past Errors

Forecast
Reference Class Forecasting

Improved Accuracy
Airport’s Own Past

N = 64

![Forecast Error Chart](chart.png)
Airport’s Own Past

N = 64
Airport’s Own Past

- No statistically significant reduction in forecast errors
- Forecast errors increased by 56%

<table>
<thead>
<tr>
<th>Wilcoxon test</th>
<th>p-value = 0.5584</th>
<th>Accept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in MAPE</td>
<td>+56%</td>
<td></td>
</tr>
</tbody>
</table>

Wilcoxon test: p-value = 0.5584

Accept Change in MAPE +56%

- No statistically significant reduction in forecast errors
- Forecast errors increased by 56%
Peer Airports

N = 64
Peer Airports

N = 64
Peer Airports

<table>
<thead>
<tr>
<th>Wilcoxon test</th>
<th>p-value = 0.0000</th>
<th>Reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in MAPE</td>
<td>-25%</td>
<td></td>
</tr>
</tbody>
</table>

- Statistically significant reduction in forecast errors
- Forecast errors decreased by 25%

N = 64

Wilcoxon test p-value = 0.0000

<table>
<thead>
<tr>
<th>Underestimation</th>
<th>Overestimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.18</td>
<td>0.82</td>
</tr>
<tr>
<td>0.28</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Future Research

• Predictive accuracy improvement
  • New feature generation
  • Interaction effect
  • Sampling

• Analysis of false positives and false negatives
  • What airports do I keep missing?
  • Any patterns?

• Non-stationary trends?
Questions?

Daniel Y. Suh
dysuh03@gmail.com