Update on Passenger Delay Analysis

Douglas Fearing
MIT Global Airline Industry IAB/AIC Joint Meeting
October 29, 2009

Collaborators:
Cindy Barnhart, Amedeo Odoni,
Nitish Umang, Vikrant Vaze
FAA Total Delay Impact project

- Published estimates of costs of delays to airlines and passengers vary from $14 billion to $31 billion
- Indirect costs to the U.S. economy are even harder to quantify
- Have NEXTOR apply a rigorous methodological approach to calculate costs of delays
  – For airlines, passengers, and the U.S. economy
Published passenger delay cost estimates

• Air Transportation Association estimates the costs of passenger delays at $4 billion for 2008
  – $37.18 per hour times flight delays
• U.S. Congress Joint Economic Committee estimates the costs at $12 billion for 2007
  – $37.60 per hour (including schedule padding)
• Who is right?
Passenger flow data

• Planned flight schedules
  – ASQP on-time performance data

• Flight seating capacities
  – Schedule B-43 airline inventory, ETMS ICAO aircraft codes, T-100 monthly segment demands

• Aggregate passenger demand data
  – T-100 monthly segment demands, DB1B quarterly 10% coupon samples (one-way itinerary routes)

• Proprietary ticketing / booking data
  – Two major carriers, one quarter each
Passenger delay calculation

1. Determine ASQP flight seating capacities
2. Generate potential passenger itineraries based on planned ASQP flights
   – Non-stop and one-stop (over 95% of passengers)
3. Allocate passengers to generated itineraries
   – This is where most of our work has been...
4. Determine disrupted passengers based on ASQP flight delays and cancellations
5. Re-accommodate disrupted passengers
Flight seating capacities

1. Match ASQP flights against Schedule B-43 airline inventories
2. Use average T-100 seating capacities when the variation is small
3. Determine ICAO aircraft code from ETMS and flight offering data  
   – Lookup seating capacities in Schedule B-43s
4. For remaining 1.5% of flights, default to T-100
Generated itineraries

- Match ASQP flights against ASQP flights
- Filter carrier routes based on DB1B
  - DB1B contains multi-carrier routes, so we do not explicitly consider code shares
- Allow 30 minute to 3 hour connection times
  - Longer connections are less likely to be disrupted
Passenger allocation approaches

1. Deterministic optimization allocation
   - Linear program assigns passengers to itineraries to minimize deviation from aggregate demand statistics

2. Sampled discrete choice allocation
   - Calibrate parameters of discrete choice itinerary shares model using proprietary data
   - Sample passenger allocations from calibrated model to disaggregate passenger demand
Problems with optimization based assignment

- Difficult to incorporate secondary factors
  - E.g., connection time and short vs. long haul
- Too many degrees of freedom
  - Basic feasible solutions tend to the extremes
Discrete choice sampling

• Train discrete choice itinerary shares model using proprietary airline bookings data
  – Initial features include time of day, day of week, and connection time

• Sample passenger counts for generated itineraries based on estimated proportions:

\[
P(i) = \frac{e^{\beta x_i}}{\sum_{i} e^{\beta x_i}}
\]
Discrete choice allocation examples

Example #1

<table>
<thead>
<tr>
<th>Day of Week</th>
<th>Departure</th>
<th>Connection</th>
<th>Weight</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>7:00 AM</td>
<td>Non-stop</td>
<td>1.00</td>
<td>21%</td>
</tr>
<tr>
<td>Monday</td>
<td>10:00 AM</td>
<td>Non-stop</td>
<td>1.01</td>
<td>22%</td>
</tr>
<tr>
<td>Monday</td>
<td>2:00 PM</td>
<td>Non-stop</td>
<td>0.94</td>
<td>20%</td>
</tr>
<tr>
<td>Monday</td>
<td>6:00 PM</td>
<td>Non-stop</td>
<td>0.88</td>
<td>19%</td>
</tr>
<tr>
<td>Tuesday</td>
<td>7:00 AM</td>
<td>Non-stop</td>
<td>0.83</td>
<td>18%</td>
</tr>
</tbody>
</table>

Example #2

<table>
<thead>
<tr>
<th>Day of Week</th>
<th>Departure</th>
<th>Connection</th>
<th>Weight</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>7:00 AM</td>
<td>30 min.</td>
<td>1.11</td>
<td>24%</td>
</tr>
<tr>
<td>Monday</td>
<td>7:00 AM</td>
<td>1 hour</td>
<td>1.35</td>
<td>29%</td>
</tr>
<tr>
<td>Monday</td>
<td>7:00 AM</td>
<td>2 hour</td>
<td>1.18</td>
<td>25%</td>
</tr>
<tr>
<td>Monday</td>
<td>7:00 AM</td>
<td>3 hour</td>
<td>1.04</td>
<td>22%</td>
</tr>
</tbody>
</table>
Evaluating the two approaches

• Evaluate by assigning aggregate passengers and comparing to proprietary data
  – Sum absolute deviation between passenger counts for matching itineraries
  – Report as % of allocated demand

<table>
<thead>
<tr>
<th></th>
<th>Optimization</th>
<th>Discrete Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error %</td>
<td>61.2%</td>
<td>25.5%</td>
</tr>
</tbody>
</table>
Comparing flight load factors

**Optimization**

**Discrete Choice**

October 31, 2009
Measuring passenger delays

- Recover disrupted passengers for each airline
  - Using Bratu & Barnhart Passenger Delay Calculator
  - Greedy re-accommodation of passengers based on scheduled arrival time
- Example results for Continental and JetBlue for the week of October 21\textsuperscript{st} – 27\textsuperscript{th}
## Continental passenger delay estimates

<table>
<thead>
<tr>
<th></th>
<th>10/21</th>
<th>10/22</th>
<th>10/23</th>
<th>10/24</th>
<th>10/25</th>
<th>10/26</th>
<th>10/27</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Num. passengers</strong></td>
<td>79,204</td>
<td>79,324</td>
<td>68,232</td>
<td>75,007</td>
<td>81,529</td>
<td>82,903</td>
<td>58,461</td>
</tr>
<tr>
<td>Delay &gt; 15 min.</td>
<td>13%</td>
<td>34%</td>
<td>26%</td>
<td>31%</td>
<td>24%</td>
<td>24%</td>
<td>16%</td>
</tr>
<tr>
<td><strong>Num. disrupted</strong></td>
<td>175</td>
<td>797</td>
<td>792</td>
<td>1217</td>
<td>528</td>
<td>776</td>
<td>237</td>
</tr>
<tr>
<td>Cancellations</td>
<td>0%</td>
<td>16%</td>
<td>57%</td>
<td>70%</td>
<td>42%</td>
<td>53%</td>
<td>0%</td>
</tr>
<tr>
<td>Misconnections</td>
<td>100%</td>
<td>84%</td>
<td>43%</td>
<td>30%</td>
<td>58%</td>
<td>47%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Avg. delay min.</strong></td>
<td>7</td>
<td>23</td>
<td>18</td>
<td>24</td>
<td>19</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>Cancellations</td>
<td>0%</td>
<td>2%</td>
<td>14%</td>
<td>12%</td>
<td>1%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Misconnections</td>
<td>9%</td>
<td>11%</td>
<td>5%</td>
<td>6%</td>
<td>3%</td>
<td>5%</td>
<td>8%</td>
</tr>
</tbody>
</table>
## JetBlue passenger delay estimates

<table>
<thead>
<tr>
<th></th>
<th>10/21</th>
<th>10/22</th>
<th>10/23</th>
<th>10/24</th>
<th>10/25</th>
<th>10/26</th>
<th>10/27</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Num. passengers</strong></td>
<td>47,694</td>
<td>43,954</td>
<td>38,429</td>
<td>40,460</td>
<td>45,817</td>
<td>46,535</td>
<td>41,077</td>
</tr>
<tr>
<td><strong>Delay &gt; 15 min.</strong></td>
<td>11%</td>
<td>5%</td>
<td>20%</td>
<td>50%</td>
<td>26%</td>
<td>43%</td>
<td>43%</td>
</tr>
<tr>
<td><strong>Num. disrupted</strong></td>
<td>84</td>
<td>125</td>
<td>508</td>
<td>267</td>
<td>157</td>
<td>529</td>
<td>222</td>
</tr>
<tr>
<td><strong>Cancellations</strong></td>
<td>0%</td>
<td>69%</td>
<td>87%</td>
<td>0%</td>
<td>0%</td>
<td>39%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Misconnections</strong></td>
<td>100%</td>
<td>31%</td>
<td>13%</td>
<td>100%</td>
<td>100%</td>
<td>61%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Avg. delay min.</strong></td>
<td>7</td>
<td>4</td>
<td>16</td>
<td>42</td>
<td>18</td>
<td>44</td>
<td>27</td>
</tr>
<tr>
<td><strong>Cancellations</strong></td>
<td>0%</td>
<td>19%</td>
<td>33%</td>
<td>0%</td>
<td>0%</td>
<td>6%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Misconnections</strong></td>
<td>6%</td>
<td>7%</td>
<td>3%</td>
<td>6%</td>
<td>8%</td>
<td>8%</td>
<td>6%</td>
</tr>
</tbody>
</table>
Next steps

• Consider other factors, such as short vs. long haul
• Complete estimates for all ASQP carriers for 2007
• Perform multiple iterations to test sensitivity to sampling of passenger allocations
• Analyze results to look for patterns in passenger delays (e.g. scheduling, network structure, etc.)
• Develop airline disruption response simulator to evaluate passenger impacts of Traffic Flow Management
Conclusion

• Described two approaches for simulating historical passenger itinerary flows
• Demonstrated that discrete choice sampling outperforms the optimization approach
• Provided sample delay results for two airlines
• Discussed next steps and ongoing research plans