Forecasting High Speed Rail Ridership Using Aggregate Data: A Case Revisit of High Speed Rail in Taiwan

Yu-Ting Hsu, Wei-Ren Lin, Yung-Cheng (Rex) Lai, and Tsung-Chung (TC) Kao

Department of Civil Engineering, National Taiwan University
Railroad Transportation & Engineering Center (RailTEC)
High Speed Rail (HSR) to address continuously increasing global vehicular demand

The trend of HSR development
- Looking backward to the past 50 years
  First HSR system: Shinkansen between Tokyo and Osaka in 1964
- Recent trend (by November 2013)
  - 21,472 km of HSR lines in operation
  - 13,964 km under construction (70% in developing countries)
  - 16,347 km planned
Background: The Challenges

- **Failure of ridership forecasting (Flyvbjerg et al., 2002)**
  Investigation of 27 railway projects
  - Ridership overestimation for more than 90% of them
  - 67% of them were overestimated by over two-thirds

- **The fundamental questions to be answered:**
  - Do we need to build a HSR system? And why?
  - Can sustainable fare revenue be expected?
  - How do we plan and design a HSR system based on the forecasted ridership?
Background: Taiwan High Speed Rail (THSR)

- **Basic information:**
  - Running over 345 km (214 mi) with eight stations
  - Signalization & train control: European specifications
  - Train & track: technology of Japan’s Shinkansen
  - Maximum speed of 300 km/h (186 mph)

- **Brief of THSR History:**
  - Planning started in the late 1980’s
  - Tendered as a BOT project in 1996 (by THSRC with US$15 billion construction cost)
  - Start of operation in Jan. 2007 (targeting at Oct. 2005 originally)
  - THSRC facing bankruptcy now (restructuring or government takeover?)
Demand Forecast for THSR

Operation since 2007

- Actual Ridership
- Sofrerail (1991)
- Transmark (1993)
- MVA (1993)
- IOT (1993)
- MVA (1997)
- MVA (2005)
Demand Forecast for HSR Systems

- **O-D demand table prediction**
  - First two steps of the four-step transportation planning procedure (trip generation + trip distribution)
  - Direct demand model

- **Determination of modal shares for predicted O-D demand**

- **Trip assignment**

- **Induced demand** *(induced by increased mobility)*
  - Empirical multiplier
  - Multiplier determined by the gravity model
  - Directly characterizing the new transportation mode in total demand
Sofrerail Approach

Total travel demand prediction

Demand distribution over existing modes for each O-D pair

Ridership diversion

Air

Conventional railway

Bus

Automobile

VOT distribution

Binary logit model

Diverted ridership to THSR

Induced travel demand

Total THSR ridership
Sofrerail Approach: O-D Demand Prediction

- **Total travel demand prediction**
  
  Linear regression model based on islandwide population and GDP
  
  \[ \Rightarrow \text{Total travel demand over the western corridor} \]

- **Demand distribution**
  
  - Distributing total travel demand over each O-D pair along the western corridor (based on the proportion in the historical data)
  
  - Distributing O-D travel demand onto existing modes: air, automobile, bus, conventional train (based on the proportion in the historical data)

  \[ \Rightarrow \text{Travel demand on each transportation mode of each O-D pair} \]

  along the western corridor
Sofrerail Approach: Ridership Diversion

- **Value of time (VOT) distribution**
  - Generalized cost for mode $k$ : $C_k = F_k + h_k TT_k$
  - Balanced VOT between modes:
    $$h_{k_1/k_2} = \frac{F_{k_1} - F_{k_2}}{TT_{k_2} - TT_{k_1}}$$

VOT distribution approximated by household income
Sofrerail Approach: Induced Travel Demand

- Gravity model to capture changing mobility

\[ T_{ij} = \frac{a \cdot (P_i P_j)}{(I_{ij})^b} \]
\[ \hat{T}_{ij} = \frac{a(P_i P_j)}{\left(\hat{I}_{ij}\right)^b} = \left(\frac{I_{ij}}{\hat{I}_{ij}}\right)^b T_{ij} \]

- \( T_{ij} \): Travel demand between cities \( i \) and \( j \)
- \( I_{ij} \): Travel impedance between cities \( i \) and \( j \)
- \( P_i \): Population of city \( I \)
- \( a, b \): Parameters
Issues

- **Total demand prediction**
  - Only GDP and population considered
  - Distributing aggregate demand over each O-D pair

- **Change of socioeconomic backgrounds**
  - Export to China: 2.5% (2000) ⇒ more than 25% (2014)
  - More than two million Taiwanese people settle in China or frequently travel across the strait

- **Ridership diversion**
  - Restrictive ridership diversion pattern
  - Air and conventional rail in the same group
  - VOT based on income level

- **Induced demand**
Proposed Methodological Framework

Total travel demand prediction for each O-D pair (including induced travel demand)

Mode split (over all available modes)

Total travel demand (for each O-D pair)

Multinomial / nested logit model

THSR  Air  Conventional railway  Bus  Automobile

Total THSR ridership
Proposed Framework: Total Travel Demand

- Direct demand prediction based on linear regression and ridge regression modeling

- Variable selection
  - Intercity activities
    - Commuting trip
    - Business trip
    - Tourism
    - Going back to hometown / visiting family members or relatives
  - Effects of changing mobility conditions

- GDP
- Employed population
- Car ownership
- Travel time (average travel time – minimum travel time)
- Travel cost (average travel cost – minimum travel cost)
- Residential population
- Aging population
- Tourism market
<table>
<thead>
<tr>
<th>Independent variable (explanatory aspect)</th>
<th>Segment 1 (20 million ~ )</th>
<th>Segment 2 (8 million ~ 20 million)</th>
<th>Segment 3 (3 million ~ 8 million)</th>
<th>Segment 4 (~ 3 million)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
</tr>
<tr>
<td>GDP</td>
<td>76.08</td>
<td>1.85</td>
<td>21.32</td>
<td>4.33</td>
</tr>
<tr>
<td>Resident population</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Car ownership</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Employed population</td>
<td>2.43</td>
<td>1.13</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Aging population</td>
<td>3.41</td>
<td>2.88</td>
<td>0.20</td>
<td>4.19</td>
</tr>
<tr>
<td>Travel cost</td>
<td>--</td>
<td>--</td>
<td>35.49</td>
<td>5.79</td>
</tr>
<tr>
<td>Travel time</td>
<td>--</td>
<td>--</td>
<td>-793.28</td>
<td>-8.71</td>
</tr>
<tr>
<td>Constant</td>
<td>-125.06</td>
<td>-0.27</td>
<td>28.93</td>
<td>1.02</td>
</tr>
<tr>
<td>Observations</td>
<td>72</td>
<td>90</td>
<td>81</td>
<td>243</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.96</td>
<td>0.48</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Average VIF</td>
<td>1.02</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Proposed Framework: Mode Split

- Distinct choice sets of available modes across O-D pairs (which are modeled separately)
- Nested logit model to capture similarity/correlation between alternate modes
- Variable selection
  - Travel time
  - Travel cost
  - Car ownership
  - Flight frequency
- Predication accuracy: average MAPE within 5%

Including time and cost associated with station access, waiting, and transfer
Result Analysis and Discussion

- Early period of demand adaption
- Induced travel demand
- Diminishing market of air transportation

Intricacy of socioeconomic evolution needs to be better accounted for

Data availability:
- In-depth understanding of travel demand pattern based on detailed travel survey
- More delicate stated preference survey to better capture inter-city mode choice behavior
- More comprehensive data collection/acquisition sources
Thank You for Listening

Questions or Comments?

Yu-Ting Hsu, Wei-Ren Lin, Yung-Cheng (Rex) Lai, and Tsung-Chung (TC) Kao

Department of Civil Engineering, National Taiwan University
Railroad Transportation & Engineering Center (RailTEC)