Modeling Variability as “Expected Lateness”

Joel P. Franklin
KTH Royal Institute of Technology
Stockholm, Sweden
Outline

1. Motivation/Research Questions
   - A Specific Problem Setting
2. Modeling Framework
   - Basing our model on scheduling theory
3. Brief Look at Some Data
   - A challenge to model consistently
4. Estimated Model for “Expected Lateness”
   - Significant, but Low Explanatory Power
5. Assessment, Limitations, Future Work
Motivation

• Transport investments & policies affect average travel times, **and** variations in travel times
  - Reliability improvements might be **worth** 10-15% that of travel time savings

• Suggests different kinds of policies, e.g. incident management, traveler information

• What is the trade-off from users’ perspective?

• How can we use this in Benefit-Cost Analysis?
How to answer these questions?

• Need forecasts of the effects of policies on travel time distributions (supply side)
  – Or at least, the important features of variation

• How much are reliability improvements worth? (demand side)
  – Or, what are the important preference paremeters?

• My focus:
  – Can we predict “mean lateness” (which seems to be important from a scheduling perspective), and use it for benefits assessments?
Problem Setting

• We Have:
  – User Preferences (vis-a-vis the Scheduling Model)
  – Observations of Today’s Travel Time Distributions
  – Model Predictions of Travel Times, etc., for Future Scenarios

• We Want:
  – Value of Reliability Benefits (or losses) in Future Scenarios
  – Relative Valuations for Different Policy Alternatives

• We’ll Need:
  – Estimates of mean lateness for Future Scenarios
Problem Setting

• Estimating Reliability Measures for Future Scenarios
  – **Standard Deviation** (e.g. Eliasson, 2006; other examples): statistical model (log-log) on (mean time / freeflow time), scaled by link length, queuing phase, and speed limit.
  – **Expected Lateness** (*this work*): similar form to Eliasson’s model
Modeling Framework: Two Approaches

1) Variability as Standard Deviation

\[ U = \gamma C + \omega \mu + \rho \sigma \]

- Cost
- Mean Time
- Std. Dev. Time

- Uses Observable (and Predictable) Data
- No Behavioral Theory behind Preferences
2) Variability as Earliness and Lateness
(Noland & Small, 1995)

\[ U(t_h) = \gamma C + \omega \mu + \lambda SDE + \delta SDL + \theta I[\text{Late}] \]

- Cost
- Mean Time
- Early Delay
- Late Delay
- Indicator of Being Late

- Strong for Identifying Preferences
- Not trivial to predict “delay” statistics
Modeling Framework

Unified Approach

• Take travel time $T = \mu + \sigma X$, $X \sim \Phi(0,1)$, in std. time

• Noland & Small, 1995: For specific travel time distributions, one can relate the lateness parameter ($\nu$) to std. dev. ($\sigma$): $\nu \cdot H \cdot \sigma$

• Fosgerau & Karlström, 2009: For any distribution, can compute $H$:
  − Take travelers’ departure time to be optimal, given travel time distribution and early/late preferences
  − $H$ is actually the (standardized) expected lateness, among late arrivals
Modeling Framework
(Fosgerau & Karlström, slightly modified)

- Optimal Departure Time:

\[ \text{Chosen Departure Time} \quad -D \]
\[ \text{Preferred Arrival Time} \quad 0 \]

\[ F(t) \quad \lambda \quad \frac{\lambda}{\nu} \quad \nu \]

Marginal Utilities

Area: \( K \)

Actual Arrival Time

\( T-D \)
Modeling Framework

- Take $T = \mu + X$, $X \sim F(0, \sigma)$, in real time units
- Utility can be expressed:
  \[ E(U) = (\lambda + \omega) \mu + \nu K \left( \frac{\lambda}{\nu}, F \right) \]
  \(\text{Value of Time}\)  \(\text{Value of Reliability}\)
- Where $K$ ("expected lateness" in minutes) is:
  \[ K \left( \frac{\lambda}{\nu}, F \right) \equiv \int_{1-\frac{\lambda}{\nu}}^{1} F^{-1}(x) \, dx = H \sigma \]
Data

- Automatic License-Plate Matching
- 92 (One-Way) Roadway Segments in Stockholm
- Three Autumns, 2005-2007, Morning Peak
- Extensive filtering for “strange” data
  - Idiosyncratic driver behavior
  - Equipment failures
  - Major changes in roadway conditions
  - Narrowed to 58 Segments
Example: Klarastrandsleden

Mean Travel Time, $\mu$
Std. Deviation, $\sigma$
Std. Mean Lateness, $H$
Abs. Mean Lateness, $K$

$H = K/\sigma$

Morning Peak, 5:00-10:30 a.m.
Example: Klarastrandsleden

Mean vs. Std. Dev., $\sigma$

Mean vs. Lateness, $K$
Nine Examples

(a) 4 Valhallav.--Odeng.-Roslagst.
(b) 7 Lidingöv. N
(c) 12 Roslagst.-Sveapl.
(d) 19 Bergslagsv.--Lövstav.-Islandsg.
(e) 26 Fleming. V
(f) 32 Klarastr.slv. N
(g) 34 Centralbr.--Slussen-Klarastr.svia.
(h) 39 Huddingevo.--Gullmarspl.-Örbyl.
(i) 78 Sveav.--Odeng.-Sergelstorg
Model Specification

- Estimate a regression model that predicts “expected lateness”
- Model form log-log:
  \[
  \log\left(\frac{\text{Lateness}}{\text{Freeflow Time}}\right) = \log\left(\frac{\text{Mean Time}}{\text{Freeflow Time}}\right) + \log^2\left(\frac{\text{Mean Time}}{\text{Freeflow Time}}\right) \\
  + (\text{Link Attributes}) + (\text{Interactions}) + \epsilon
  \]
- Link Attributes:
  - length, freeflow speed, location indicator
- Basic strategy: fit polynomials, scaled to link characteristics
Estimated Model of Lateness

Performance Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>d.f.</th>
<th>R2</th>
<th>Adj. R2</th>
<th>F Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Statistics</td>
<td>10</td>
<td>1206</td>
<td>0.610</td>
<td>0.606</td>
<td>189.8*</td>
</tr>
<tr>
<td><em>(measured for K)</em></td>
<td>10</td>
<td>1206</td>
<td>0.175</td>
<td>0.168</td>
<td>25.07*</td>
</tr>
</tbody>
</table>

Predictions

Residuals

Actual

Mean Time / Freeflow Time
Predictions: Klarastrandsleden

(a) 32 Klarastr.sl. N

Lateness K / Freeflow Travel Time

Mean / Freeflow Travel Time
Predictions for Nine Examples

(a) 4 Valhallav.--Odeng.-Roslagst.

(b) 7 Lidingöv. N

(c) 12 Roslagst.-Sveapl.

(d) 19 Bergslagsv.--Lövstav.-Islandstg.

(e) 26 Fleming. V

(f) 32 Klarastr.sl. N

(g) 34 Centralbr.--Slussen-Klarastr.svia.

(h) 39 Huddingeav.--Gullmarspl.-Örbyl.

(i) 78 Sveav.--Odeng.-Sergelstorg
Conclusions/Main Points

• The missing component of a schedule-based approach to reliability, “expected lateness”, *can* be modeled with some significance and partial qualitative matches...

• ...but *not yet* with sufficient explanatory power for benefits assessments.

• Further Refinements:
  – Phase of Queue Buildup/Dissipation
  – Periodic *Change* in Mean Travel Time
  – Controls for Serial Correlation
Two Looming Issues

• Sequential Links:
  – Readily-available data is for short roadway segments…
  – … while scheduling theory is based on full Origin-Destination paths

• Assumption of Static Distribution:
  – Individuals optimize departure time as if the travel time distribution were constant…
  – … while the analysis setting here concerns a changing distribution over time
Future Directions

• Model the *causes* of travel time variation, rather than modeling variation *per se*
  – Travel times as emerging from a queuing model
  – Distinction between *supply* versus *demand* variations

• Origin-Destination travel time distributions
  – Challenge for data collection & modeling
Supported by:
Centre for Transport Studies, Stockholm

Questions to...

Joel Franklin
joelfr@kth.se