

# 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 parameters?
- 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

$$U = \gamma \underbrace{C}_{\text{Cost}} + \omega \underbrace{\mu}_{\text{Mean Time}} + \rho \underbrace{\sigma}_{\text{Std. Dev. Time}}$$

- Uses Observable (and Predictable) Data
- No Behavioral Theory behind Preferences

# Modeling Framework

## 2) Variability as Earliness and Lateness (Noland & Small, 1995)

$$U(t_h) = \underbrace{\gamma C}_{\text{Cost}} + \underbrace{\omega \mu}_{\text{Mean Time}} + \underbrace{\lambda SDE}_{\text{Early Delay}} + \underbrace{\delta SDL}_{\text{Late Delay}} + \underbrace{\theta I [\text{Late}]}_{\text{Indicator of Being Late}}$$

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

# Modeling Framework

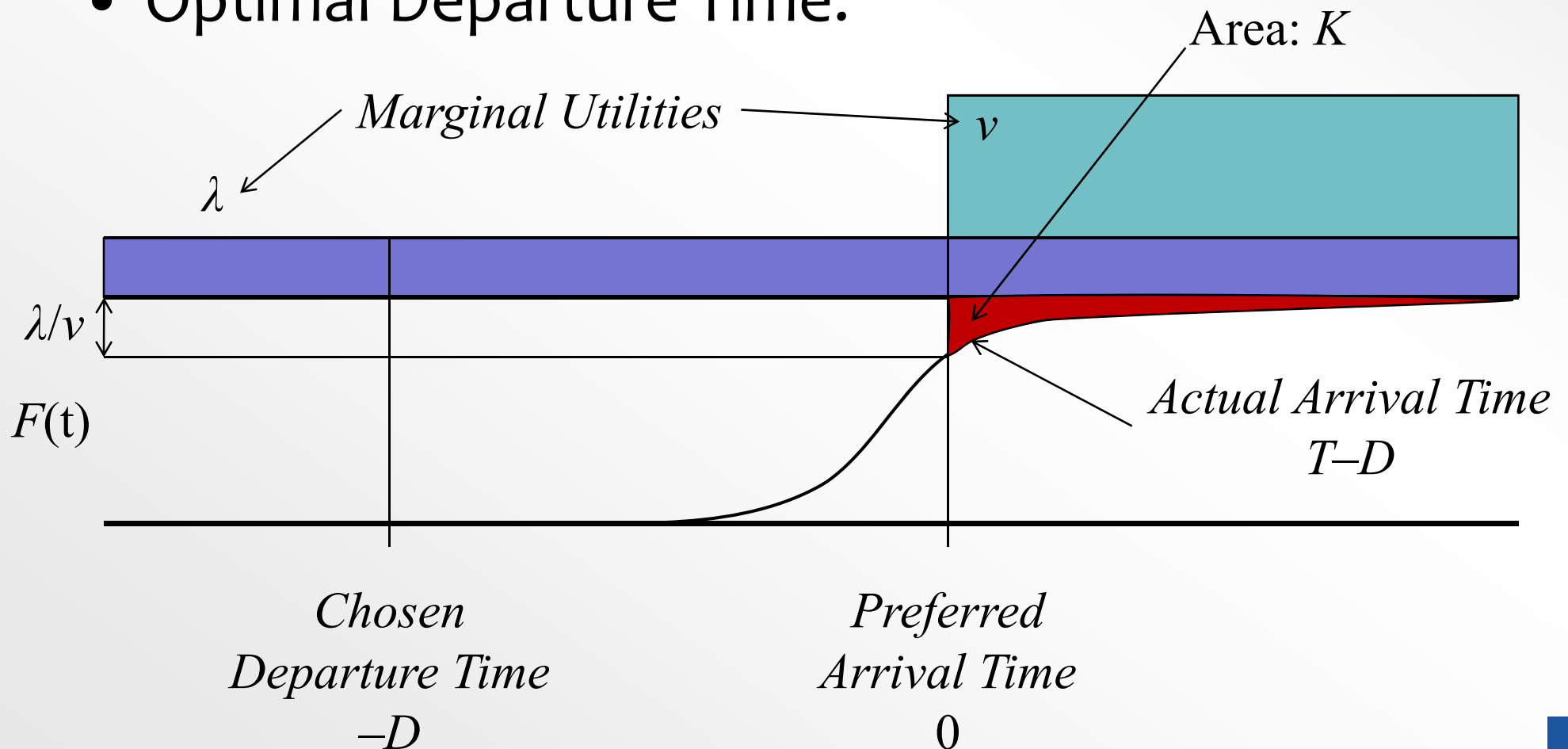
## Unified Approach

- 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 ( $v$ ) to std. dev. ( $\sigma$ ):  $v^*H^*\sigma$
- *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:



# Modeling Framework

- Take  $T = \mu + X$ ,  $X \sim F(0, \sigma)$ , in *real* time units
- Utility can be expressed:

$$E(U) = \underbrace{(\lambda + \omega)}_{\text{Value of Time}} \mu + \underbrace{v K\left(\frac{\lambda}{v}, F\right)}_{\text{Value of Reliability}}$$

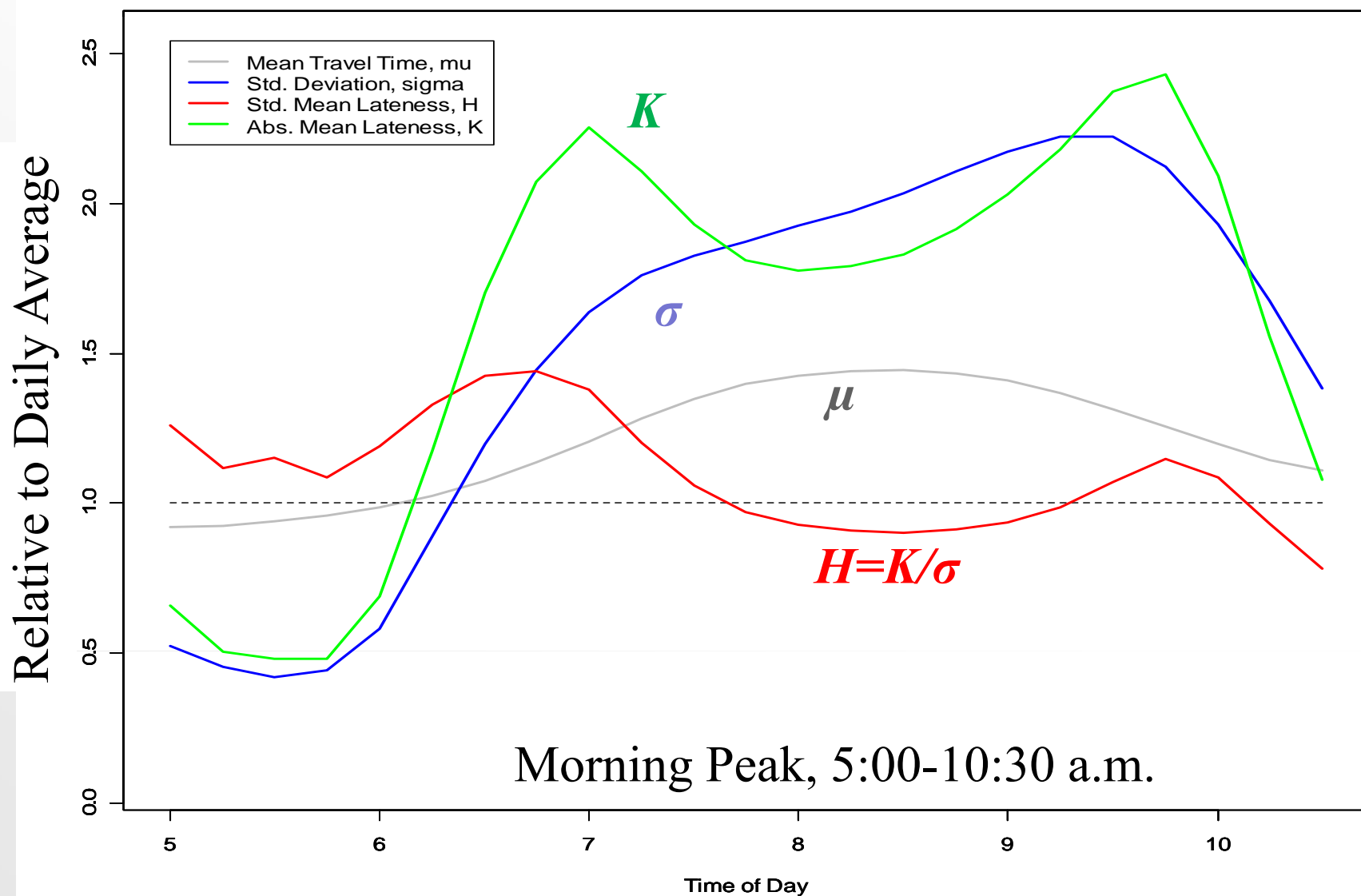
- Where  $K$  (“expected lateness” in *minutes*) is:

$$K\left(\frac{\lambda}{v}, F\right) \equiv \int_{1-\frac{\lambda}{v}}^1 F^{-1}(x) dx = H \sigma$$

# Data

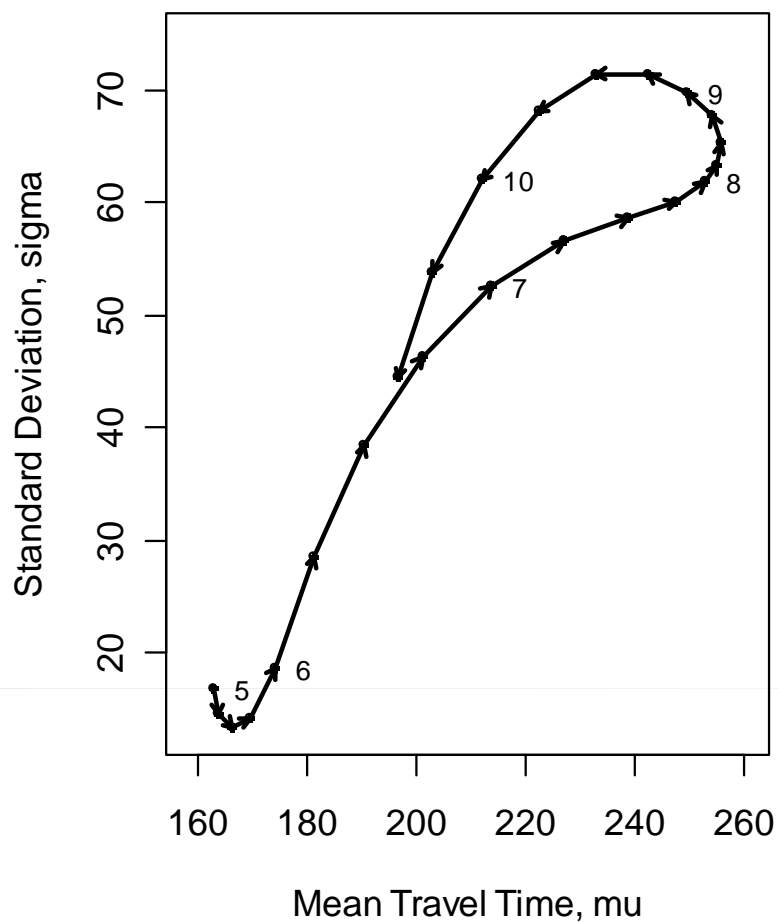
- Automatic License-Plate Matching
- 92 (One-Way) Roadway Segments in Stockholm
- Three Autumns, 2005-2007, Morning Peak
- Extensive filtering for “strange” data
  - Idiosyncratic driver behavior
  - Equipment failures
  - Major changes in roadway conditions
  - Narrowed to 58 Segments

# Example: Klarastrandsleden

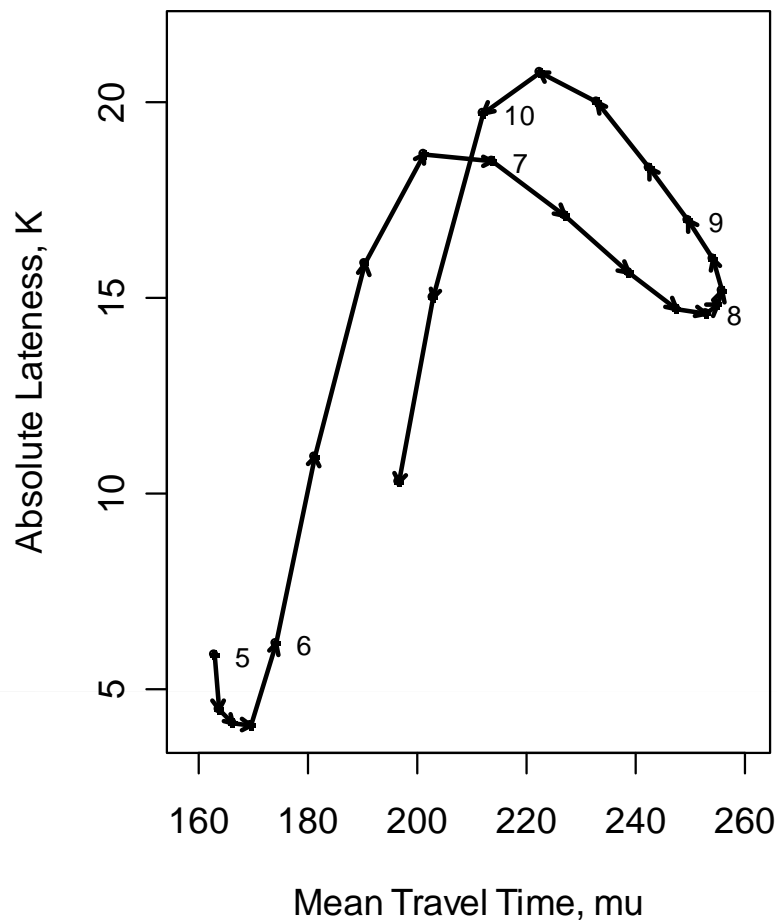


# Example: Klarastrandsleden

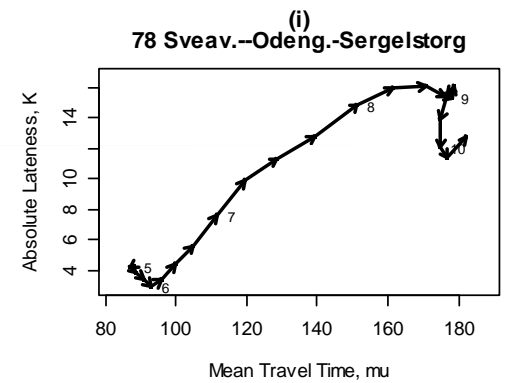
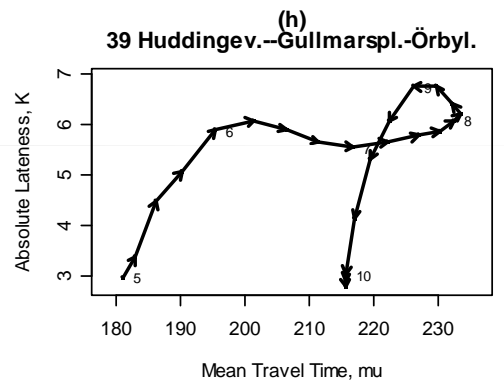
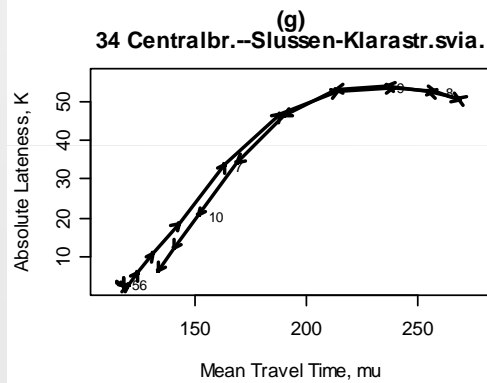
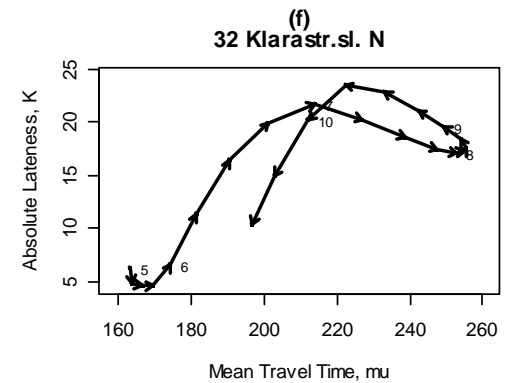
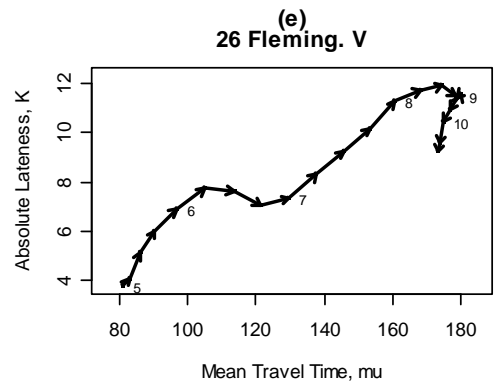
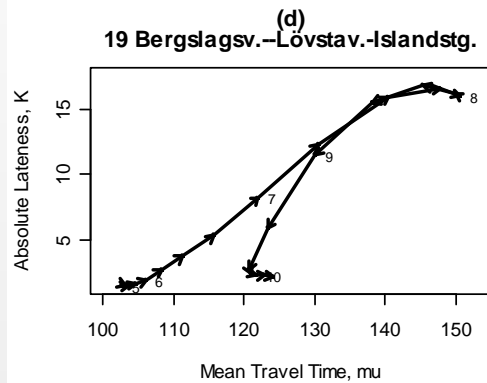
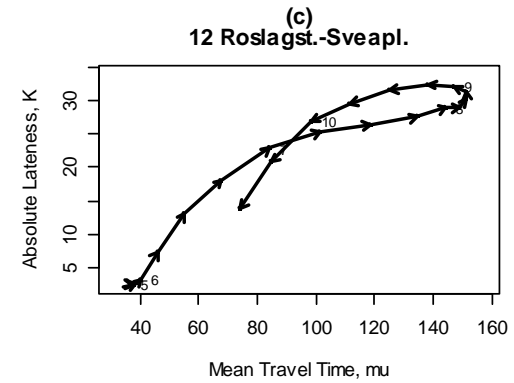
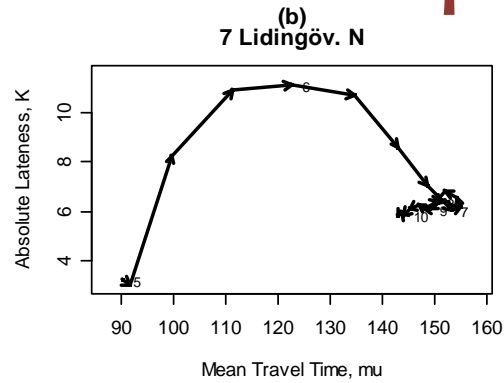
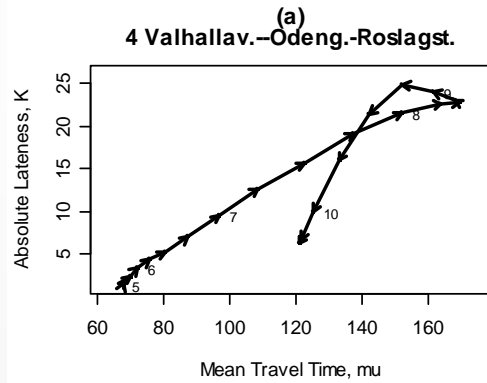
Mean vs. Std. Dev.,  $\sigma$



Mean vs. Lateness,  $K$



# Nine Examples



# Model Specification

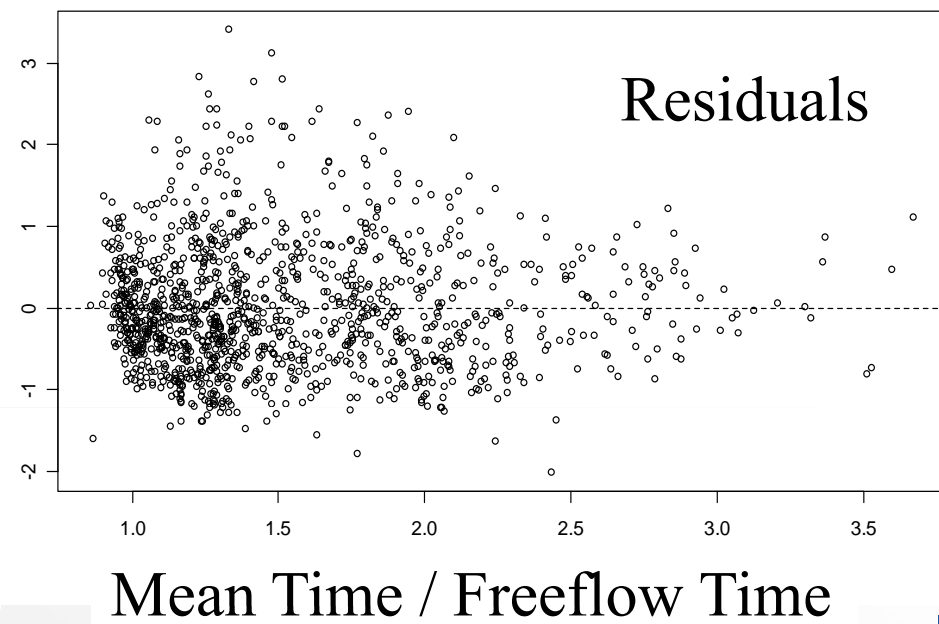
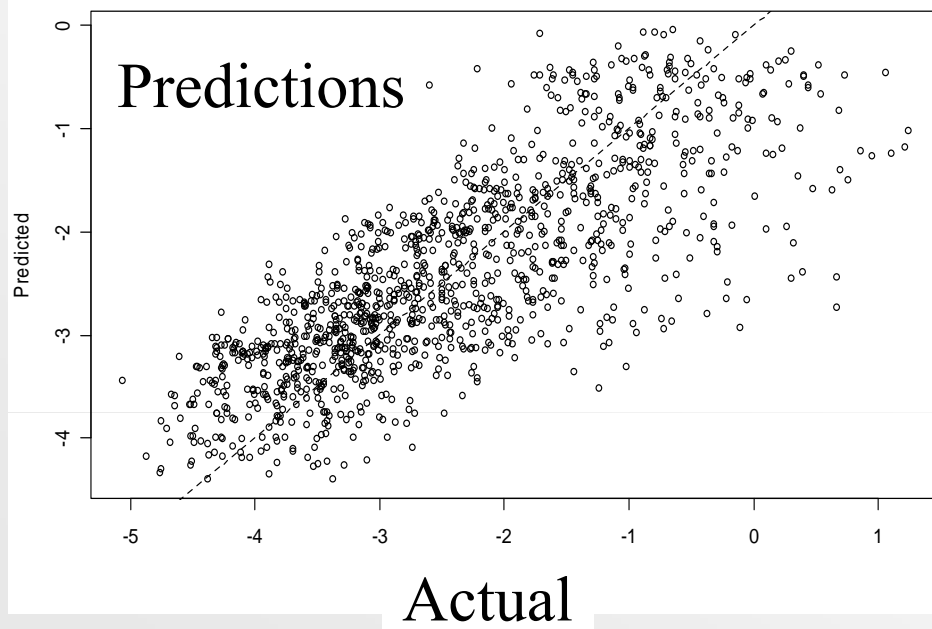
- Estimate a regression model that predicts “expected lateness”
- Model form *log-log*:

$$\log\left(\frac{\text{Lateness}}{\text{Freeflow Time}}\right) = \log\left(\frac{\text{Mean Time}}{\text{Freeflow Time}}\right) + \log^2\left(\frac{\text{Mean Time}}{\text{Freeflow Time}}\right) + (\text{Link Attributes}) + (\text{Interactions}) + \varepsilon$$

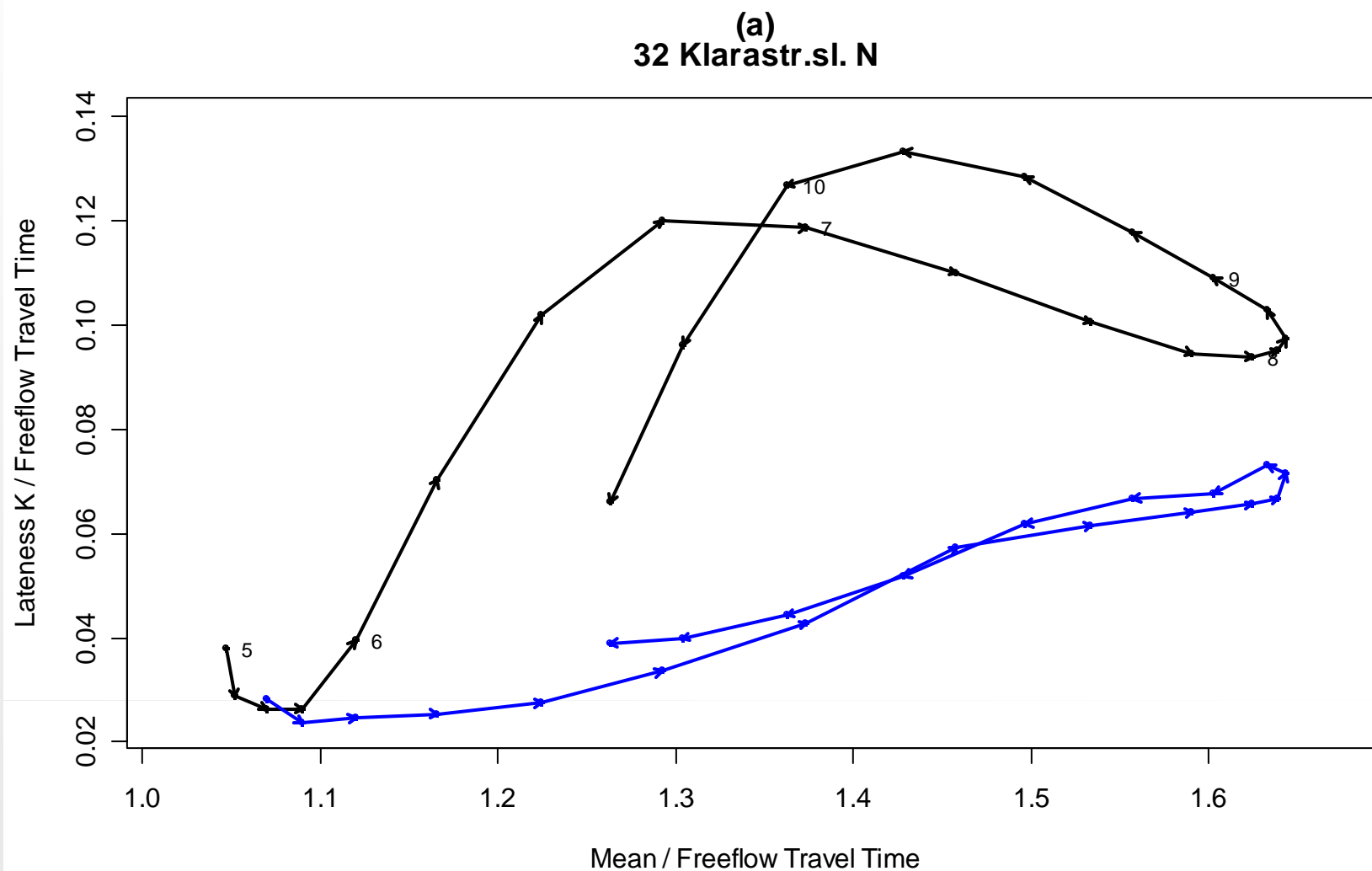
- Link Attributes:
  - length, freeflow speed, location indicator
- Basic strategy: fit polynomials, scaled to link characteristics

# Estimated Model of Lateness Performance Statistics

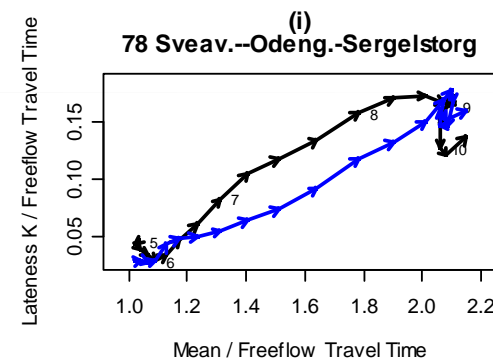
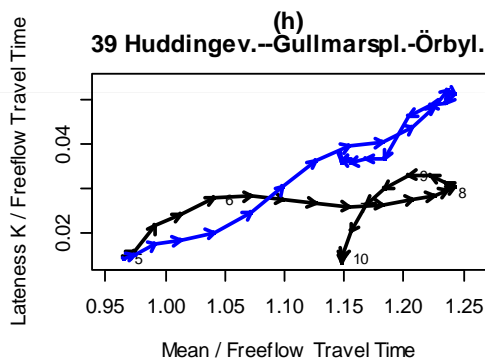
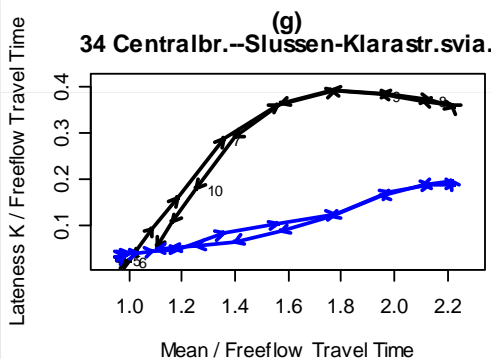
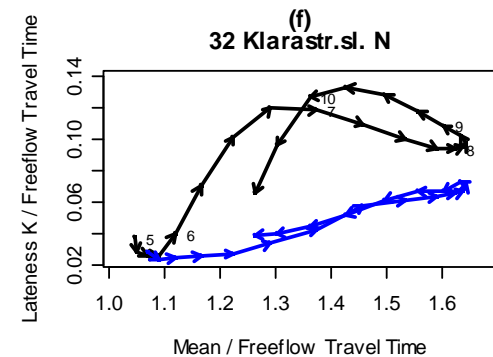
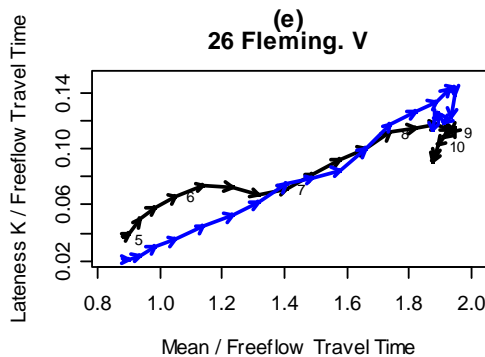
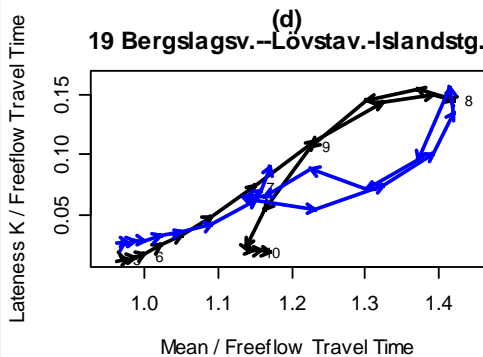
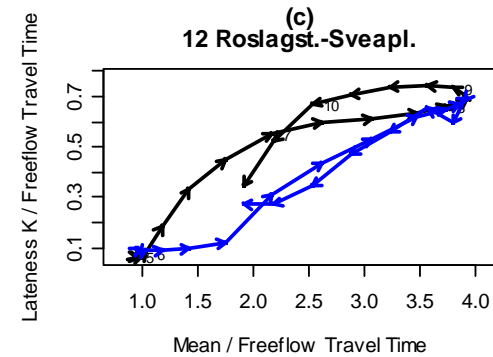
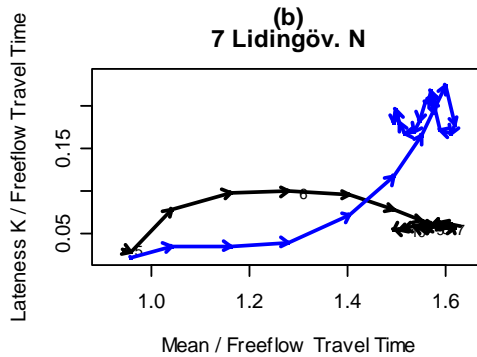
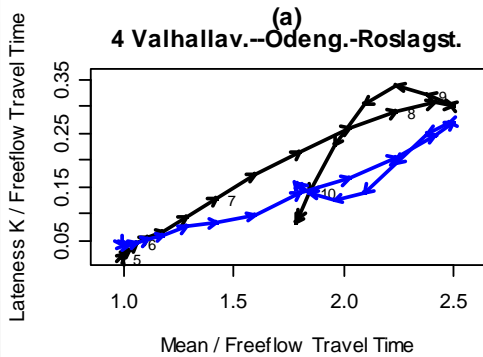
Model	Para- meters	d.f.	R2	Adj. R2	F Value
Direct Statistics	10	1206	0.610	0.606	189.8*
<i>(measured for K)</i>	10	1206	0.175	0.168	25.07*



# Predictions: Klarastrandsleden



# Predictions for Nine Examples



# 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