An Econometric Analysis of US Airline Flight Delays with Time-of-Day Effects

Chieh-Yu Hsiao
Mark Hansen

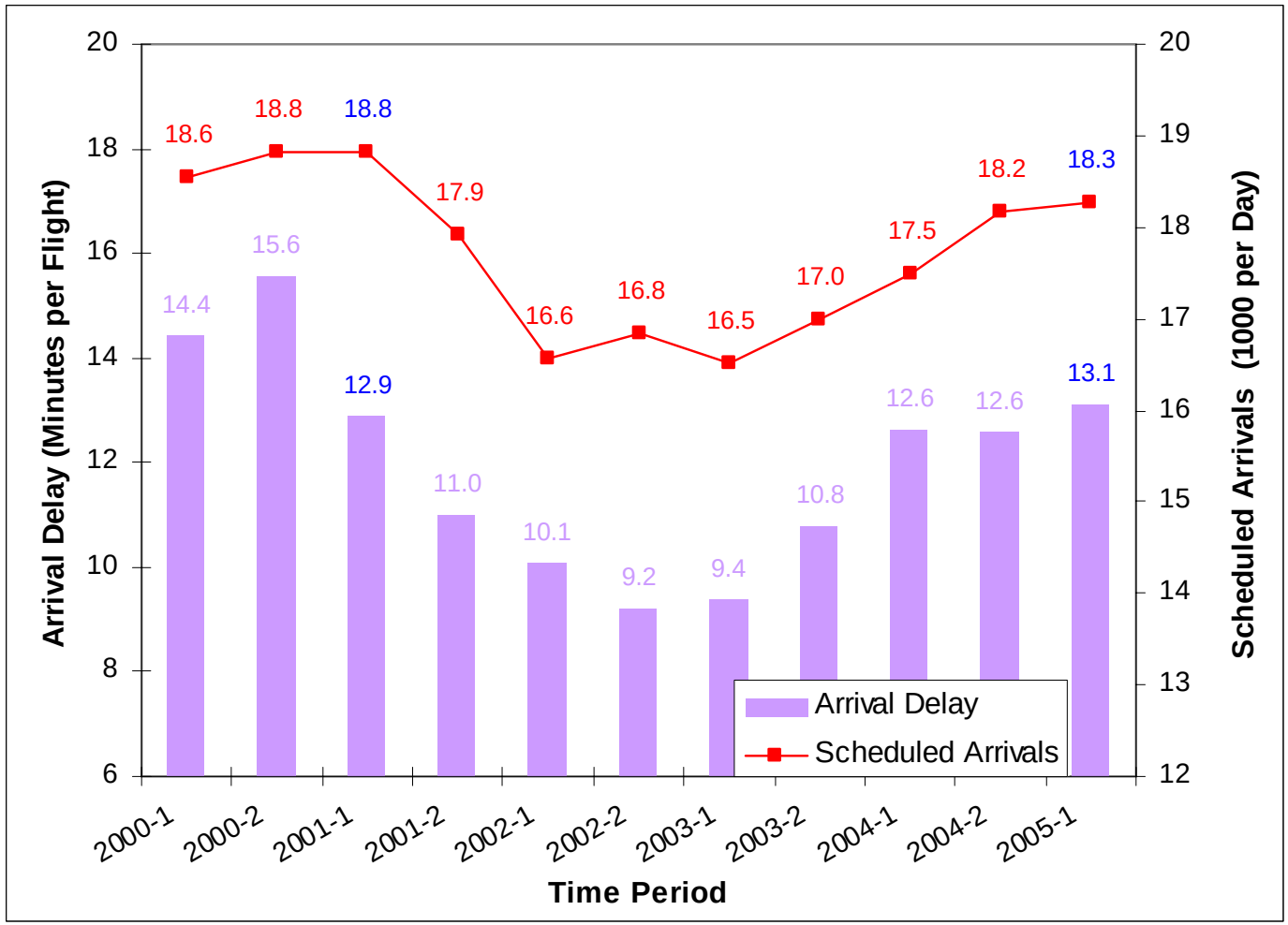
National Center of Excellence in Aviation Operations Research (NEXTOR)
Department of Civil and Environmental Engineering
University of California at Berkeley

01/20/06

Outline

- Introduction
- Model
  - Factors
  - Data
- Estimation Results
- Delay Change Decomposition
- Conclusions

Introduction:
Arrival Delays & Scheduled Arrivals



Goals

- Develop statistical model to forecast delay in the NAS
- Certain advantages over simulation models
  - Estimated over large number of days
  - Reduced data needs
  - Instantaneous calculation
  - Future year prediction based on application to a full year of days rather than a few sample days
- Compare results with simulation

Objectives

- Model daily average arrival delay
- Assess impacts of
  - Queuing
  - Volume
  - Weather (storm and terminal conditions)
  - Seasonal Effects
- Investigate time of day effects for queuing delay

Model

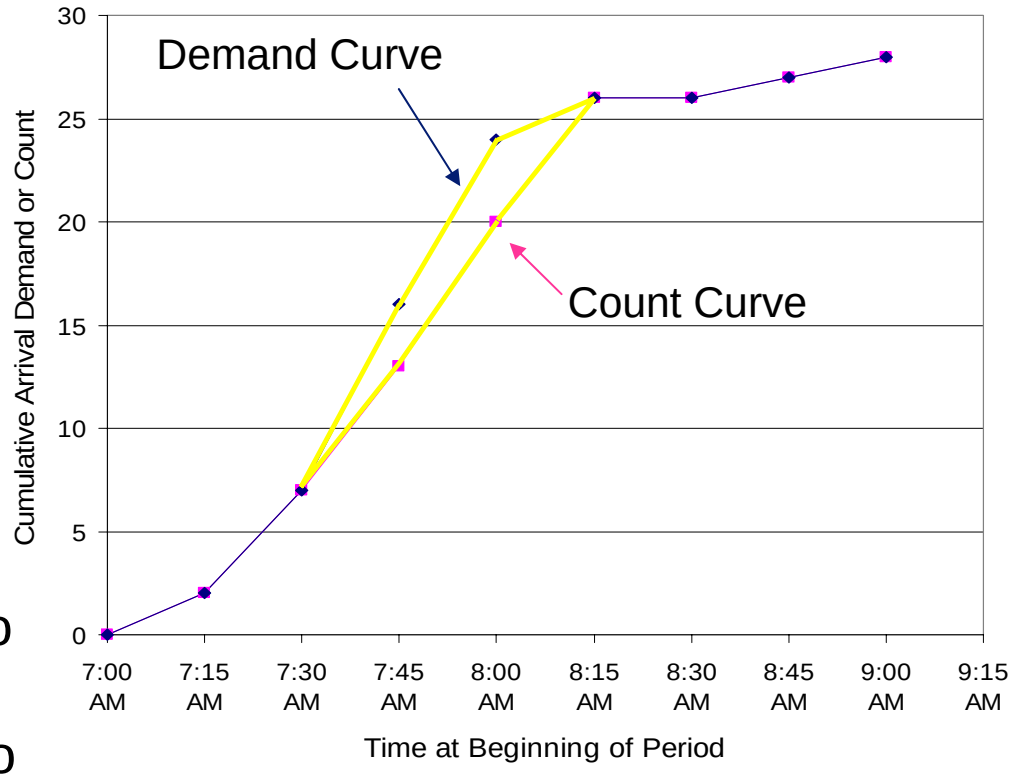
$$AvDelay(t) = f(\overrightarrow{Queuing}(t), SFlights(t), GAFlights(t), MFlights(t), IFR(t), Windspeed(t), \overrightarrow{THX}(t), \overrightarrow{Events}(t), \overrightarrow{Season}(t), \overrightarrow{Period}(t)) + v_t$$

Queuing

- Calculate deterministic queuing delay on a daily basis for each of 32 DOT airports
- Based on
  - Quarter-hr demand profile (based on schedule and adjusted for cancelled flights)
  - Quarter-hr AAR

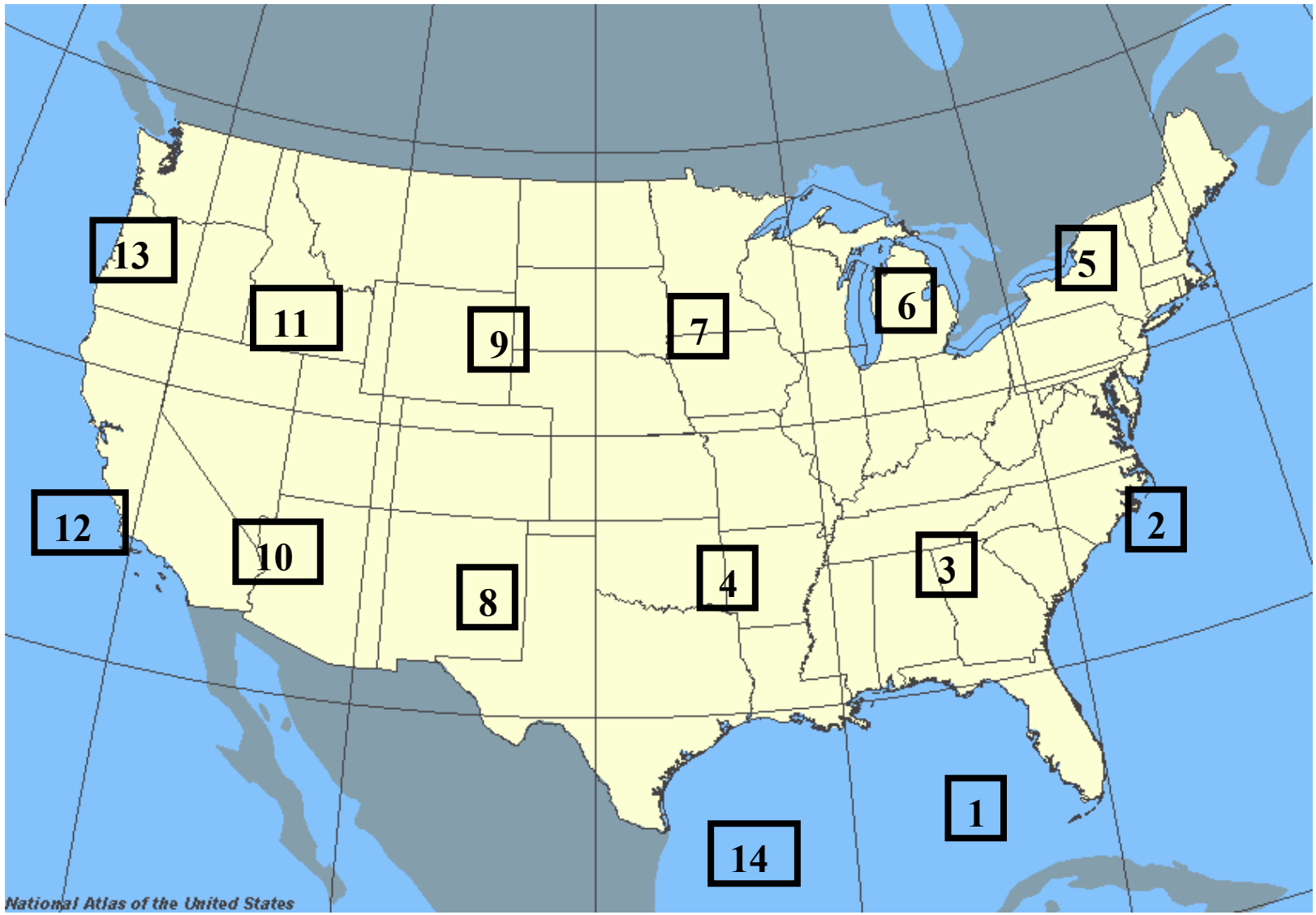
Queuing: Procedure

- Airport level
  - Construct Cumulative Arrival Demand Curve
  - Construct Cumulative Arrival Count Curve
  - Calculate Delay as Area Between Demand and Count Curves
- System level
  - Summed airport delays to get total delay
  - Divided by total arrivals to get average delay



En Route Weather (Storms)

- Based on Surface Summary of the Day from NOAA (National Oceanographic and Atmospheric Administration)
- Daily summary from ~1500 US weather stations
- Each station reports binary (yes/no) thunderstorm variable
- Used proportion reporting thunderstorms as storm metric
  - Construct area-specific thunderstorm metrics on lat-long quandrangles
  - Include each metric as separate explanatory variable


13
11
9
7
6
5
12
10
8
4
3
2
1
14
National Atlas of the United States

Terminal Weather

- Proportion of the day under IFR
  - Airport IFR proportion: proportion of time under IFR
  - System IFR proportion: weight airport IFR proportion by # of operations
  - Higher IFR proportion, expected higher delay due to lower AAR
- Wind speed
  - Airport wind speed: avg. wind speed of the day
  - System wind speed: weight airport wind speed by # of operations
  - Higher wind speed, expected higher delay

Other Effects

- Volume
  - Scheduled arrivals
  - GA operations
- Monthly fixed effects
  - Based on month
  - Effects of winds and other weather factors
- Yearly fixed effects
  - Year-to-year trends not otherwise accounted for
  - May reflect FAA performance among other factors
- UA strike effect
- 9/11 Effect

Estimation

- Data
  - Daily data: Jan. 2000- June 2005
  - Excludes 12/20-25, 9/11-30/01, and standard/daylight savings transition days
- Estimations
  - OLS model: Heteroskedasticity and autocorrelation
  - GARCH model
    - Model form

$$y_t = \mathbf{X}_t'\boldsymbol{\beta} + v_t$$

$$v_t = \varepsilon_t - AR1 \cdot v_{t-1}$$

$$\varepsilon_t = \sqrt{h_t} \cdot e_t$$

$$h_t = ARCH0 + ARCH1 \cdot \varepsilon_{t-1}^2 + GARCH1 \cdot h_{t-1} + HET1 \cdot Q(t)$$

$$e_t \sim IN(0,1)$$

Estimation Results

Category	Variable (parameter for category GARCH)	Baseline		Time-of-Day		Interaction		Full	
		Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
	Intercept	-14.156	2.351	-15.405	2.268	-9.663	2.357	-9.532	2.265
Queuing	Queuing	1.457	0.202			1.448	0.197		
	Queuing (00:00--08:00)			0.895	0.377			**0.861**	0.375
	Queuing (08:00--12:00)			3.304	0.607			**3.099**	0.599
	Queuing (12:00--16:00)			2.236	0.446			**2.286**	0.443
	Queuing (16:00--24:00)			1.029	0.207			**0.994**	0.205
	$Queuing^2$	-0.079	0.025	-0.074	0.024	-0.082	0.025	-0.072	0.024
Volume	Scheduled arrivals (000)	0.964	0.157	1.009	0.153	0.771	0.163	0.675	0.159
	GA operations (000)	0.098	0.030	0.131	0.030	0.073	0.030	0.118	0.029
Terminal Weather	IFR ratio	14.475	0.824	13.524	0.840				
	Wind speed	0.490	0.071	0.481	0.070				
	Sch*IFR ratio					0.828	0.047	0.800	0.048
	Sch*Wind speed					0.029	0.004	0.029	0.004
En Route Weather	Storms in region 1	1.644	0.658	1.684	0.657				
	Storms in region 2	5.627	0.814	5.636	0.816				
	Storms in region 3	2.269	0.661	2.271	0.653				
	Storms in region 4	3.946	0.475	3.939	0.468				
	Storms in region 5	9.736	0.853	9.753	0.850				
	Storms in region 6	10.153	0.788	10.061	0.789				
	Military Operations (000)	-0.246	0.035	-0.314	0.035				
	Sch*Storms in region 1					0.120	0.037	0.106	0.037
	Sch*Storms in region 2					0.358	0.046	0.345	0.046
	Sch*Storms in region 3					0.137	0.037	0.125	0.037
	Sch*Storms in region 4					0.240	0.027	0.229	0.026
	Sch*Storms in region 5					0.529	0.047	0.520	0.045
	Sch*Storms in region 6					0.605	0.044	0.583	0.044
	Sch*Military Operations (000)					-0.015	0.002	-0.017	0.002
Event	UA strike dummy	1.820	0.631	1.720	0.628	1.684	0.626	1.460	0.618
	2nd half 2001 after 9/11dummy	0.042	0.883	0.281	0.875	-0.114	0.879	0.394	0.864
R-Squared		0.720		0.724		0.727		0.729	

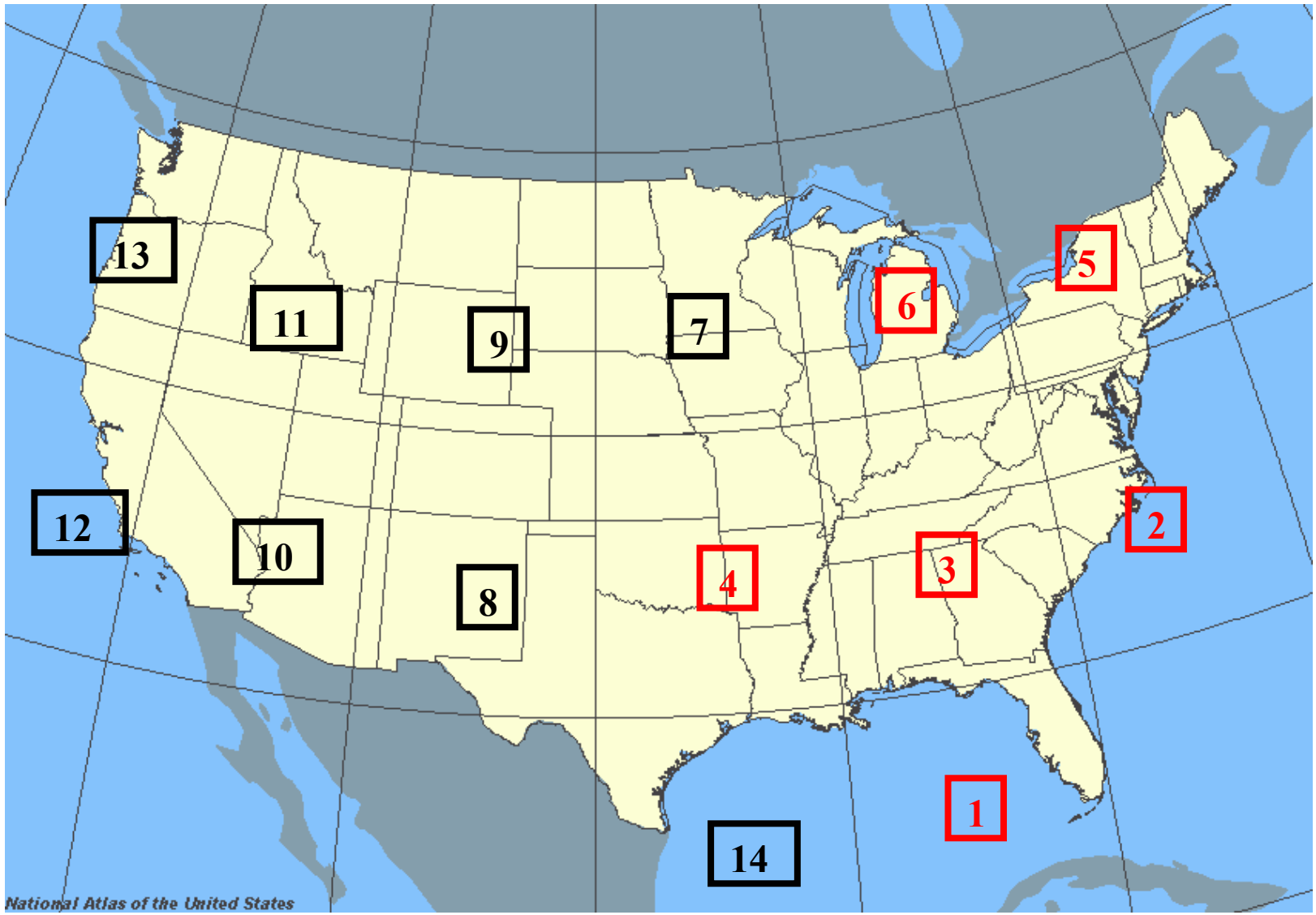
Note: "Scheduled arrivals" is abbreviated as "Sch."; Time dummies and GARCH variables are used but not shown here.

Key Findings

- Time of day queuing effects
  - Vary strongly by time-of-day
  - Greatest impact in morning (delay propagation)
  - Concave relationship
- Volume
  - Scheduled arrivals: 1000 additional scheduled flights increases average delay by 0.68-0.77 minutes on extremely good weather days
  - GA operations: 1000 additional GA operations in increases average delay by 0.1 minutes
  - Military operations: strong negative relationship with delay (!)

Key Findings (cont)

- Terminal weather effects
  - Depend on volume
  - Have impact even when controlling for queuing
- En route weather effects
  - Depend on volume
  - Wide geographic variation
- Other effects
  - 2000 UA job action had pronounced effect
  - Post 9/11 effect not significant (captured by other variables)


13
11
9
7
6
5
12
10
8
4
3
2
1
14
National Atlas of the United States

Conclusions

- Statistical models help us explain and predict delay changes
- Use of deterministic queuing increases fidelity of these models and allows time-of-day effects to be captured
- En route convective weather effects can be captured in fairly simple ways
- Substantial unexplained variation remains