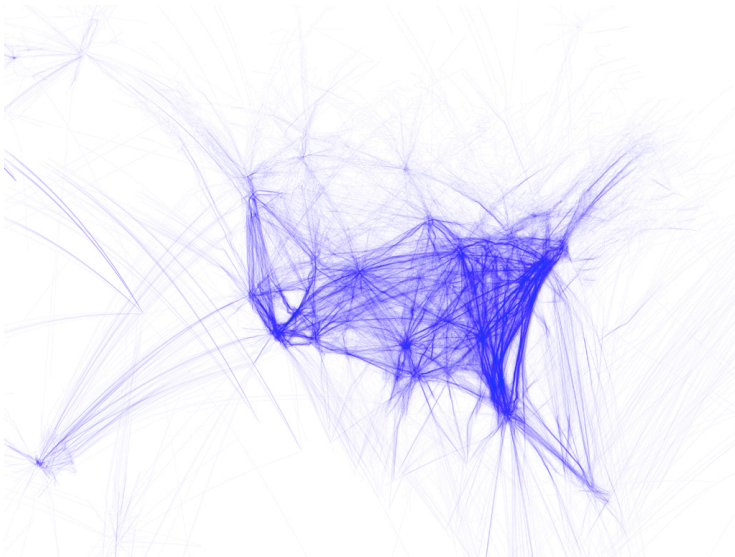# Project 2: The Plight of the Late Flight

Andrew Bernath, Heather Kitada, Ethan Edwards

Oregon State University

May 15, 2014

Contents

Overview

- Data: Bureau of Transportation Statistics (BTS)
- Last 25 years
- 30 Unique carriers
- 376 Unique origins

Overview and Question of Interest   Population-Based Findings   Sample-based Findings   Discussion, Obstacles and Solutions   Future
● ○ ○   ○ ○ ○   ○   ○

Question of Interest

Question of Interest

Are there any airlines that have shown consistent improvement in delays, across the entire country, over the 25 years of flight data?

Overview and Question of Interest   Population-Based Findings   Sample-based Findings   Discussion, Obstacles and Solutions   Future

○   ○   ○
●   ○
○   ○○
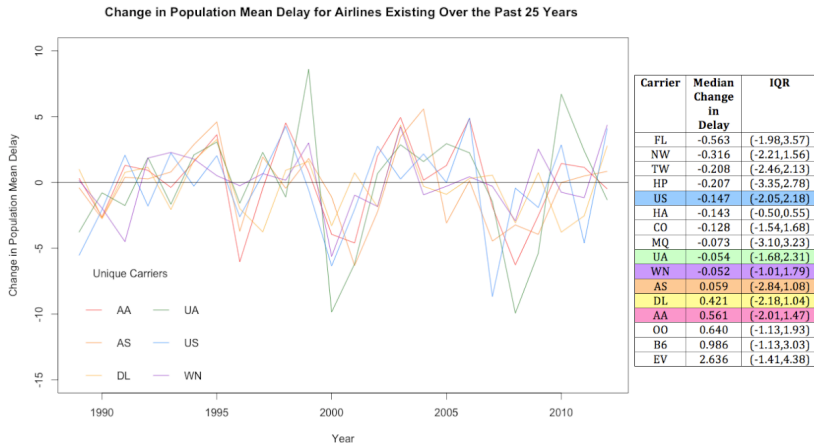
Choosing A Metric

How do we define "improvement" in delay?

- Improvement is defined as negative change in delay time, where delay can be measured using the following metrics:
    - Arrival delay
        - What customers care about
    OR
    - Carrier + Late Aircraft delays
        - What carries are able to control

- Overall improvement metric = Median change in mean delay

Overview and Question of Interest Population-Based Findings Sample-based Findings Discussion, Obstacles and Solutions Future
○   ○   ○
○   ○○   ○
●

Narrowing Scope

Narrowing Scope

- Ran all airlines and years
- Only kept airlines with 10+ years of service
    - 10+ is enougth to discern a pattern
- Aggregating to create yearly averages
    - Average over seasonal effects to compare year to year

Population-Based Findings

Change in Population Mean Delay for Airlines Existing Over the Past 25 Years

Carrier	Median Change in Delay	IQR
FL	-0.563	(-1.98,3.57)
NW	-0.316	(-2.21,1.56)
TW	-0.208	(-2.46,2.13)
HP	-0.207	(-3.35,2.78)
US	-0.147	(-2.05,2.18)
HA	-0.143	(-0.50,0.55)
CO	-0.128	(-1.54,1.68)
MQ	-0.073	(-3.10,3.23)
UA	-0.054	(-1.68,2.31)
WN	-0.052	(-1.01,1.79)
AS	0.059	(-2.84,1.08)
DL	0.421	(-2.18,1.04)
AA	0.561	(-2.01,1.47)
OO	0.640	(-1.13,1.93)
B6	0.986	(-1.13,3.03)
EV	2.636	(-1.41,4.38)

Sample-based Findings

1. Stratify by unique carrier
2. Stratify by year (1989-2013)
3. Stratify by origin airport size
   (as determined by flight traffic volume)
   - Proportional sample from strata based on number of flights

Overview and Question of Interest Population-Based Findings **Sample-based Findings** Discussion, Obstacles and Solutions Future
○   ●   ○
○   ○   ○
○   ○○

Sample Frame

Sample Frame

Assumption: Due to coordination of air traffic control efforts, flights originating from airports of similar traffic volume would have similarities in delay patterns

- Found traffic volume for each origin over 25 years
- Found average traffic volume
- Ordered and stratified based on size
  - Create subsets of carriers
  - Used %*in*% when filtering

Overview and Question of Interest   Population-Based Findings   **Sample-based Findings**   Discussion, Obstacles and Solutions   Future

Findings

Findings

Change in Sample Mean Delay for Airlines Existing Over the Past 25 Years
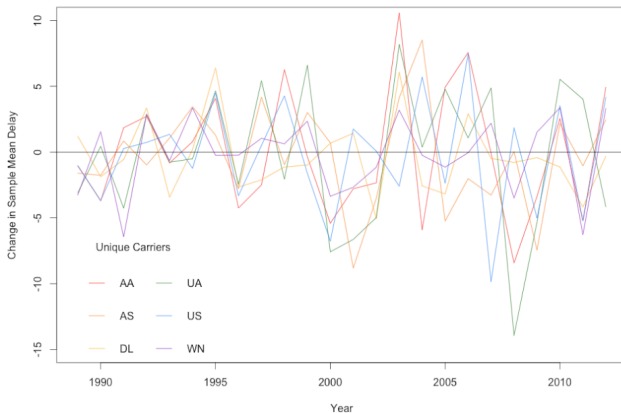
Carrier	Median Change in Delay	IQR
MQ	-1.688	(-3.85, 8.35)
TW	-1.212	(-2.86,1.23)
CO	-1.155	(-2.31,2.28)
FL	-1.073	(-3.79,5.45)
DL	-0.679	(-2.23,0.81)
AA	-0.595	(-3.46,3.06)
AS	-0.463	(-2.21,2.25)
NW	-0.218	(-2.33,2.94)
B6	-0.178	(-4.39,8.11)
WN	-0.145	(-1.54,2.23)
UA	-0.066	(-4.19,4.68)
US	0.171	(-2.78,2.27)
HP	0.702	(-3.90,3.66)
HA	0.864	(-0.47,1.11)
OO	1.478	(-2.76,2.66)
EV	2.878	(-4.35,6.33)

Overview and Question of Interest   Population-Based Findings   **Sample-based Findings**   Discussion, Obstacles and Solutions   Future

Sampling Performance

Sampling Performance: 76% Coverage

Comparing Population Mean and Stratified Sampling Mean Estimates

Overview and Question of Interest Population-Based Findings Sample-based Findings Discussion, Obstacles and Solutions Future
○   ○   ○   ○
○   ○
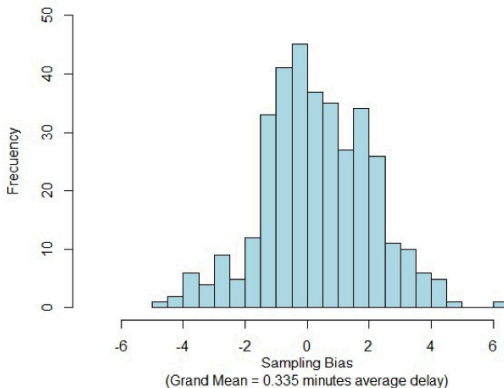○   ○●

Sampling Performance
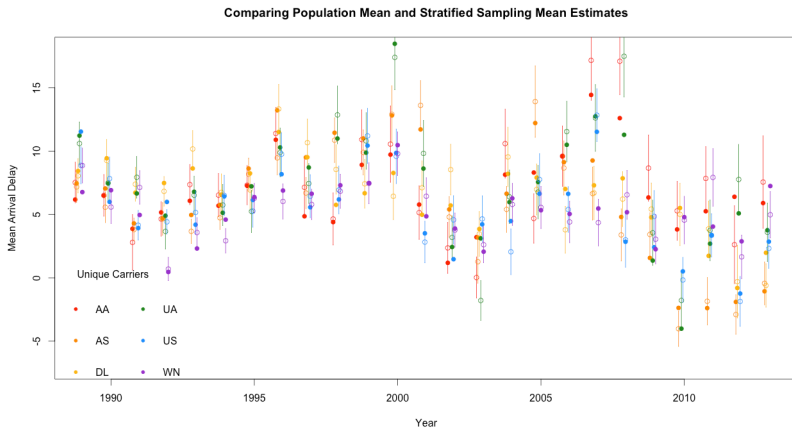
Bias

Frequency Distribution of Sampling Bias for Mean Arrival Delay per Year per Airline

The "Simple" Answer

■

Overview and Question of Interest Population-Based Findings Sample-based Findings **Discussion, Obstacles and Solutions** Future

Cognitive Time vs Computational Time

Cognitive Time vs Computational Time

■

Overview and Question of Interest Population-Based Findings Sample-based Findings **Discussion, Obstacles and Solutions** Future

Data Visualization

Data Visualization

Future Work

-

Questions

Google? Which airline is sexy?