MBTA Fare Model Reconstruction

Mali Akmanalp
Nathaniel Burbank
Xinyuan Wang
Introduction
What’s MBTA’s current fare model?

MBTA's current fare model uses a single average fare for all trips across all modes of transit.
However, both costs and rider behaviors vary significantly by route, time, and mode of transit.

Our goal - provide a more precise model of passenger behavior

a more accurate view of the efficiency of MBTA's current allocation of resources.
Datasets
Datasets

Swipe Dataset
"Boardings" for subway and bus in 2016

GTFS Dataset
MBTA’s schedule and trip planning information

MBTA Cost Dataset
Trip count, total distance, and total duration for all routes
Datasets - Swipe Dataset

275 million+ “boardings” subway + bus data in 2016

Over 12 million unique Charlie Card hashes

Data fields:

- **passenger id**: Charlie Card hash
- **cents**: fare paid
- **t**: boarding time (e.g. 2016-01-20 13:30:00)
- **time period**: time category (e.g. Early morning)
- **fare type**: 7 Day Link, Monthly Link, Student, ...

| medium | id | t  | fareboxid | routeorstation | cents | faretype | timeperiod | bookcanc |
Datasets - GTFS dataset

General Transit Feed Specification

Industry-standard format that describes:

- Stops
- Transit Routes
- Trips for each route
- Holidays / Schedule exceptions

Trip schedules change:

MBTA makes historical GTFS data available

We use this to calculate operations for a given hour / day in 2016
Datasets - MBTA cost data

12 bus trip information data sets for the whole year

- **Season:** Spring, Summer, Fall, Winter
- **Day type in a week:** Weekday, Saturday, Sunday

In each trip data set, we have total trip count, total distance, total duration for all routes on a day of that day type (e.g. Spring-Weekday).

**Labor cost:** Hourly wage rate

**Bus cost:** Maintenance cost, Fuel cost for different vehicle types
Data Preprocessing
01
We removed cancelled transactions.

02
For pass users, we imputed the cost of each transaction by distributing the revenue of daily, weekly and monthly passes across each user’s transactions.

03
We grouped transfers together and re-split revenue across all portions of a given trip, reassigned these values back.

04
We corrected the route names in the dataset with the MBTA official route list.
Data Preprocessing

MBTA Monthly or Weekly Pass vs. Single Fare Boardings, 2016

Pass Users: 64%
Non-pass Users: 36%
**Data Preprocessing**

**Assumptions** about transfer grouping and segment weighting can affect revenue by more than half a million dollars per month.

<table>
<thead>
<tr>
<th>Four ways of dividing passenger revenue between transit types</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rebalanced by trip, weighted by transit type</strong></td>
</tr>
<tr>
<td>$8.70 m (24.95%)</td>
</tr>
<tr>
<td>$26.16 m (75.04%)</td>
</tr>
</tbody>
</table>

| **Per segment, weighted by transit type**                   |
| $9.24 m (26.49%)                                           |
| $25.62 m (73.50%)                                          |

| **Rebalanced by trip, unweighted by transit type**          |
| $9.12 m (26.15%)                                           |
| $25.75 m (73.86%)                                          |

| **Per segment, unweighted by transit type**                 |
| $9.34 m (26.80%)                                           |
| $25.52 m (73.20%)                                          |

---

**Estimated mean passenger revenue per month**

- Local and express bus revenue
- Subway revenue
Assumptions about transfer grouping and segment weighting can affect revenue by more than half a million dollars per month.
Model Construction
Model Construction

GTFS Data

Trips
per route
per hour

Cost
Data

Cost
per trip

Swipe Data

Cost
per route
per hour

Cost
per route
per hour

Revenue
per route
per hour

Revenue
per route
per hour

Profit
per route
per hour
Analysis
Bus Revenues vs. Costs in each week

Commuter trends dominate the hourly revenue patterns, but the total required subsidies remain relatively consistent over a week. (Interactive version)
Estimated profits are affected by passenger loads, the length of the route and the type of vehicle used.
Analysis

The average bus revenue per trip increases from $1.15 to $1.28 due to the system-wide fare increases in July of 2016.
The average bus revenue per trip increases from $1.15 to $1.28 due to the system-wide fare increases in July of 2016.
Analysis

Mean Rev. Per Half Trip vs. Mean Passengers Per Half Trip

Mean Rev. Per Half Trip vs. Mean Rev. per Passenger
Monthly pass users tend to bring more revenues overall, but discounted pass users are associated with lower revenues per half trip.
This map of MBTA Bus system, colored by normalized-revenue-per-half-trip, highlights the relative popularity of express buses that serve the urban core as well as the increased prevalence of some lower-revenue routes in South Boston. (Interactive Version)
Positive factors for revenue:
- Inner_express,
- Outer_express,
- Latenight & weekend
- Northern routes

Negative factors:
- Latenight, Weekend,
- Percentage of discounted Pass users

Analysis

|                                | coef | std err | t     | P>|t|  | [0.025] | [0.975] |
|--------------------------------|------|---------|-------|------|---------|---------|
| Intercept                      | 0.1824 | 0.002 | 75.872 | 0.000 | 0.178 | 0.187 |
| is latenight[T.True]           | -0.6322 | 0.003 | -236.008 | 0.000 | -0.637 | -0.627 |
| is wkend[T.True]               | -0.2830 | 0.003 | -102.835 | 0.000 | -0.286 | -0.278 |
| is north[T.True]               | 0.2712 | 0.002 | 129.044 | 0.000 | 0.267 | 0.275 |
| C(bus_type, Treatment(reference='local'))[T.inner_express] | 0.5221 | 0.004 | 126.834 | 0.000 | 0.514 | 0.530 |
| C(bus_type, Treatment(reference='local'))[T.outer_express] | 1.6532 | 0.010 | 162.454 | 0.000 | 1.633 | 1.673 |
| is latenight[T.True]:is wkend[T.True] | 0.2227 | 0.006 | 39.758 | 0.000 | 0.212 | 0.234 |
| prct_d_passsrsa                | -2.0181 | 0.016 | -127.685 | 0.000 | -2.049 | -1.987 |
| prct_cash                      | -0.9168 | 0.013 | -66.171 | 0.000 | -0.943 | -0.890 |
Interactive Interface
We built an interactive interface that can be used by both the MBTA and the public. By simply toggling, hovering, zooming, and brushing, users can explore both aggregate trends and route-by-route comparisons.
Future Work
Future Work

01
If provided with more granular transaction data, we could investigate similar questions at a per bus-stop level, which could help inform decisions around new route locations.

02
We could analyze passenger profiles in more detail, creating customer archetypes for commuters, tourists, and infrequent transit users.

03
Another natural avenue is doing more scenario analysis, for example to figure out how changing the fare structure could change revenues.
Thanks!