Reliability and Compatibility of Major Journey-to-Work (JTW) Origin-Destination (OD) flows

investigating CTPP, LODES, Cellphone Signal, and Mobile Apps location data for selected zones in NYC

Presented at The Census for Transportation Planning Subcommittee Meeting on January 11, 2016

Kyeongsu Kim
Senior Planner
Louis Berger
48 Wall Street, 16th floor, New York NY 10005
Email: kkim@louisberger.com
Study Background

- Trip Origin-Destination (OD) data: a fundamental resource for transportation planning and analysis
  - e.g., regional travel demand forecast (TDF), traffic analysis and environmental impact study.

- A robust trip OD table helps better:
  - Answer when, where, how, why, how long, and how far people travels
  - Supports a policy decision-making process for future transportation investments

- Data quality Issue of existing data resources
  - e.g., CTPP, regional household travel survey (RHTS), regional TDF model OD matrix
  - Limited sample size.

- LODES, government administrative-based data. But,
  - No info of travel mode, travel time, and departure time,
  - Missing cash-paid jobs,
  - Issue of identifying actual work site for multiple worksite employers.
Study Background

- New discussion in transportation data community:
  - How to utilize massive amount of
    - Cell phone signal data &
    - Mobile device’s Global Positioning System (GPS) location data

- Strengths:
  - Based on real trips
  → supposedly better count estimates, tracking daily trips, and identifying non-work trips

- Questions:
  - Unknown data extraction methods,
  - Concern on location accuracy,
  - Reliability of trip count estimates
Study Approach

- Examined four different data
  1. CTPP,
  2. LODES,
  3. AirSage Cellphone signal data: HBW Trips,
  4. Place IQ mobile apps location data: Home dwellers’ day time location (b/w 10 am to 12 pm & 1 pm to 3 pm)

Remember that when you download almost any smartphone apps, you are asked to share your location information. The mobile apps location data is almost new to transportation data community. It is based on network ads location data that compiles smartphone-users location information, reported when (almost any) mobile apps are opened.

- Study Interest
  - Compare different data’s OD: from selected smaller zones to broader NYC community district boundaries
Selected Zone Boundary

- Data comparability issue due to boundary alignment
  - The neighborhood boundary for this study: not aligned with census boundary (CTPP and LODES).
    - The initial neighborhood boundary for AirSage data: delineated by NYCDOT and AirSage
    - PlaceIQ data: extracted based on the given gis boundary

- Extracted public data based on
  - LODES: Census Block boundary that falls in the selected neighborhood zone.
  - CTPP: TAZ boundary (almost equivalent to a single census tract in NYC)
    - Zone 1: almost equivalent to a single TAZ boundary.
    - Zone 2 and Zone 3: multiple TAZs where only a part of each TAZ boundary is included. This accounts for the higher CTPP estimates than LODES.
Selected Zone Boundary

Figure 2.1: CTPP TAZ vs. Zone

Figure 2.2: Census Block vs. Zone

Solutions for a better world
Data Sources

1. CTPP 2006-2010 Part 3: TAZ level flow (A302100 - Total Workers (1) (Workers 16 years and over)
2. LODES 2012 data for all jobs (NY_OD_MAIN_JT00_2012)
3. AirSage: based on the average of one month (February, 2015) data
4. PlaceIQ: based on the average of five months (April to August, 2015) data
COMPARISON OF EACH DATA SOURCE:
FROM A NEIGHBORHOOD ZONE TO THE REST (BROADER) OF NYC ZONES

From a public housing neighborhood in Staten Island, NYC
COMPARISON OF EACH DATA SOURCE:
FROM A NEIGHBORHOOD ZONE TO THE REST (BROADER) OF NYC ZONES

From Rego Park in Queens, NYC - mixed-use neighborhood (mostly residential)
COMPARISON OF EACH DATA SOURCE:
FROM A NEIGHBORHOOD ZONE TO THE REST (BROADER) OF NYC ZONES

From Astoria in Queens, NYC - mixed-use neighborhood (mostly residential)
FINDINGS

- Similarity of the proportional distribution of workplace location
- Difference of the total number of commute between government and private data.
  - The CTPP’s total outbound zone trip estimate is similar to LEHD.
  - Airsage provides the smaller OD estimates; in particular from Zone 1 (the smallest boundary)
    - But, we identified large number of trip OD between broader community district zones.
  - PlaceIQ’s mobile apps data provides the highest OD counts
    - Include additional typical daytime trips e.g., school and cash-paid jobs
- In summary, the result of this comparison study discloses that
  - Private data source can be great resources to build more comprehensive OD table.
  - However, “how to” has not been dealt yet, which suggests the necessity of more progressive efforts.