NBER-DOT Data Portal Update

Caitlin Gorback, NBER

Steering Committee:
Nathaniel Baum-Snow, University of Toronto
Leah Brooks, George Washington University
Edward L. Glaeser, Harvard University and NBER
Stephen J. Redding, Princeton University, CEPR and NBER
Matthew A. Turner, Brown University and NBER

New Directions in Transportation Economics
May 2022

1Thanks to the U.S. Department of Transportation, which funded this project through a grant to NBER Project on the Economics of Transportation in the 21st Century
Provide a **centralized hub** to be researchers’ go-to for transportation related data:

- **Complementary** to existing Department of Transportation (DoT) and Bureau of Transportation Statistics (BTS) datasets

- **Digitized or Edited Data:**
  - Data from digitized maps: PR-511, historical streetcar shapefiles
  - Old waves of current products: HPMS, commuting flows
  - Hand collected data: Ridesharing entry, global subway stations
  - Aggregates of Proprietary Data: Cell phone tracking, airport links

- Centrally contained and **easy to access**
  - Housed at NBER’s “Public Data Use Archive”
This project supports research on three central issues in transportation economics: the impact of the transportation sector on the economy, new transportation technologies and initiatives such as automation and ride-sharing, and the role of transportation in contributing to household welfare. It also seeks to develop new data resources to support empirical research in transportation economics. The project supports a diverse set of research studies. The findings of these studies are presented and discussed at an annual research conference, and distributed in the NBER Working Paper series.
This project supports research on three central issues in transportation economics: the impact of the transportation sector on the economy, new transportation technologies and initiatives such as automation and ride-sharing, and the role of transportation in contributing to household welfare. It also seeks to develop new data resources to support empirical research in transportation economics. The project supports a diverse set of research studies. The findings of these studies are presented and discussed at an annual research conference, and distributed in the NBER Working Paper series.
Data Resources

Since our last update: from 6 datasets to \(\sim 35\)

1. DOT sources
2. Flows of People
3. Flows of Goods
4. Travel Costs and Financing
5. Shapefiles
1. DOT Sources

- **broad summary** of DOT data products
- **comprehensive list of tables** available through DOT

---

### National Transportation Statistics

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Table Name</th>
<th>Source</th>
<th>Years of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-A-01</td>
<td>System Mileage Within the United States</td>
<td>DOT, FHWA, Highway Statistics</td>
<td>1960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AAR, Railroad Facts</td>
<td>1965</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DOT, FTA, National Transit</td>
<td>1970</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Database, Annual Data Tables</td>
<td>1975</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U.S. Army, Corps of Engineers</td>
<td>1980</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DOT, PHMSA, Office of Pipeline</td>
<td>1985</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Safety, Pipeline Statistics</td>
<td>1990-2017</td>
</tr>
</tbody>
</table>

---

- Great first stop for those with interest in transportation topics!
## Data Sources

### 2. Flows of People

<table>
<thead>
<tr>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Data</td>
</tr>
<tr>
<td>Commuting Data</td>
</tr>
<tr>
<td>For-Hire Vehicle Data</td>
</tr>
<tr>
<td>Highway Data</td>
</tr>
<tr>
<td>Mobility Data</td>
</tr>
<tr>
<td>Public Transit Data</td>
</tr>
<tr>
<td>Rail Data</td>
</tr>
</tbody>
</table>
Data Sources

3. Flows of Goods

| Commodity Flow Survey Data |
| Commodity Transportation Survey Data |

4. Travel Costs and Financing

| Congestion and Travel Time Data |
| Personal Travel Survey Data |
| Planning and Government Data |

5. Shapefiles

| Infrastructure Shapefiles |
Building our Data Archive: HPMS Example

2. Highway Performance Monitoring System, Dept. of Transportation

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- Data (1980-2008)
- Data (2011-2017)

Building our Data Archive: HPMS Example

2. **Highway Performance Monitoring System, Dept. of Transportation**

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- **Data** (1980-2008)
- **Data** (2011-2017)

2. **Highway Performance Monitoring System, Dept. of Transportation**

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- **Data** (1980-2008)
- **Data** (2011-2017)

2. **Highway Performance Monitoring System, Dept. of Transportation**

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

**Provided by Matthew Turner.** Dataset Contributor

- **Data** (1980-2008)
- **Data** (2011-2017)

2. **Highway Performance Monitoring System, Dept. of Transportation**

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- **Data** (1980-2008) Direct download
- **Data** (2011-2017)

Building our Data Archive: HPMS Example

2. Highway Performance Monitoring System, Dept. of Transportation

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- **Data (1980-2008)**
- **Data (2011-2017)** Link to DoT-FHWA for publicly available data

Building our Data Archive: HPMS Example

2. Highway Performance Monitoring System, Dept. of Transportation

Annual US road segment data providing length, number of lanes, number of vehicles per lane per day for the universe of the interstate highway system. Available from 1980.

Provided by Matthew Turner.

- Data (1980-2008)
- Data (2011-2017)


Example paper which used this data
New Data Sources: For-Hire Vehicles

1. **New York City Taxi & Limousine Commission (TLC)**

   The NYC TLC provides trip records for yellow taxis, green taxis, for-hire vehicles (such as Uber and Lyft) from 2009-present. Green taxi data available from roll-out date in August, 2013, while FHV data begins in 2015, a few years after introduction to the city.

   - **NYC TLC**

   As used in: Buchholz, Nicholas. “Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry.”

   Similar data exist for many other U.S. cities, over different time periods, for example:

   - **Chicago Taxi Trips** (2013-present)
   - **Chicago Rideshare** (2018-present)
   - **Washington, D.C. Taxi Trips** (2015-present)

2. **Uber Entry Dates**

   Uber entry dates for U.S. and Canada through 2018, collected from Uber’s newsroom as well as local news sources.

   Data provided by Jonathan D. Hall

   - **Data**
   - **Notes**

New Data Sources: Mobility

1. PlaceIQ Exposure Indices

Anonymized and aggregated mobility data based on smartphone movements provided by PlaceIQ. Available from 2020-present, updated weekly. The data provides a variety of exposure measures across and within states, counties, and CBSA’s.

Data provided by Victor Couture, Jonathan I. Dingel, Allison Green, Jessie Handbury and Kevin R. Williams.

- Device Exposure
- Location Exposure


2. Google COVID-19 Community Mobility Reports

Mobility indices tracking movement patterns for a myriad of countries, and regions within countries. For example, in the United States, mobility is broken down by country, state and county. The indices make distinctions between trips for retail and recreation, grocery and pharmacy, parks, transit stations, workplaces and residences. Available from February, 2020 – present.

- Data (as of Feb 16 2020)


3. SafeGraph Weekly Patterns

Weekly data on cell phone trips between December 31st, 2018 – present, tracking the number of visitors to a point of interest (POI) in a given week. Each POI observation also includes information on visitors’ home census blockgroup.

- Apply for Access

Moving Forward: Let’s keep this growing!

- Collecting ideas on data to include
  - Procurement
  - Shipping
  - Supply Chain
  - Vehicle Fleet Electrification
  - Water Transport

- Let us know if you would like to contribute datasets

We invite ideas for continued development!

Caitlin Gorback
gorback@nber.org

Steve Redding
reddings@princeton.edu