

Hurricane Sandy: Notes on LODES/OnTheMap Data Flows

November 5, 2012

The following are useful background notes for users of the County Flows maps produced using LODES data from the OnTheMap application. Further documentation on the data product and analysis tool can be found at <http://lehd.ces.census.gov/onthemap/> and <http://lehd.ces.census.gov/datatools/onthemap.html>. Additionally, the production of LODES data is reliant upon outputs from the data processing for the LED Partnership's Quarterly Workforce Indicators (QWI). Information on the preparation of the QWI can be found at <http://lehd.ces.census.gov/datatools/qwiapp.html>.

Creation of the Flows Map

This is the first time that this type of map has been produced by the LED Partnership. Consistent with our goal of providing useful data to the public, the design has the potential to evolve based on needs of our user community.

The county flow map of the coastal New York and New Jersey areas affected by Hurricane Sandy was created by querying county-to-county home/employment relationships within the OnTheMap application. This consisted of a series of Destination Analyses using Private Primary Jobs for All Workers in 2010 applied to all counties within DR-4085 and DR-4086.¹ The data from these analyses were tabulated and overlaid on special maps produced using ArcMap. A minimum cutoff of 10% was established for a flow to be displayed, meaning that a flow was displayed if it represented at least 10% of the origin county's residential workforce or at least 10% of the destination county's employment. For example, Staten Island (Richmond County, NY) is shown as sending almost 43,000 workers to Manhattan (New York County, NY). This represents only 2% of Manhattan's employment but is 29% of the employed workers residing in Staten Island.

Job Definition

For the purposes of LODES and OnTheMap, the LED Partnership defines a **job** as a link between a worker and a firm at which the worker has been employed during the reference quarter *and* during the quarter *prior* to the reference quarter. The reference quarter for LODES is Quarter 2 (April-June) of the year of interest. This definition of "job" is sometimes called a "Beginning of Quarter" job because it is assumed

¹ A guide to running a Destination Analysis in OnTheMap can be found at <http://lehd.ces.census.gov/led/datatools/doc/OnTheMapSampleDestAnalysis.pdf>.

that the worker was employed at that firm on the first day of the reference quarter. The latest year of data available in LODES/OnTheMap is 2010.

If a worker is employed at more than one job during the referenced period and those jobs are covered by the core datasets, then all of those jobs will be captured in the dataset. A **primary job** is defined as the one job for each worker that provides the most earnings. By tabulating primary jobs, users are seeing "one job per worker," whereas tabulating "All Jobs" users are seeing all the jobs falling within a spatial selection. Primary jobs are calculated across the full universe of data. So if a New Jersey resident has two jobs – one in New Jersey and one in New York – and the job in New York earns her more income, then the New York job will be defined as the primary job.

Job definitions in LODES and OnTheMap also include **ownership** information. Two ownership types are reported explicitly: all ownership and private ownership. Public ownership is available implicitly by subtracting private from all. **Private primary jobs** are those primary jobs (defined within the full job universe of LODES) that are held at private-sector firms. These are the jobs shown in the Flows Map for New York and New Jersey. Due to issues with the reporting of multi-establishment firms within the public sector, only private sector information was used for this analysis.

Coverage

LODES data are available for jobs whose employment locations are in one of the 50 states/territories currently producing data with the LED Partnership. The eligible states and territories *not yet* in LODES and OnTheMap are Massachusetts, Puerto Rico, and the U.S. Virgin Islands. **Data are available in all years (2002-2010) except for the following combinations:**

- Arizona (2002 and 2003)
- Arkansas (2002)
- The District of Columbia (2002-2009)
- Mississippi (2002 and 2003)
- New Hampshire (2002)
- Massachusetts (All years)
- Puerto Rico (All years)
- U.S. Virgin Islands (All years)

In general these gaps in the coverage should have very little effect on the New York-New Jersey area and the data shown in the Flows Map.

The LED Partnership builds its data infrastructure based upon several core datasets provided by our state partners. These include Unemployment Insurance wage data and the Quarterly Census of Employment and Wages. With the release of Office of Personnel Management (OPM)-sourced data, LODES and OnTheMap now includes most Federal Civilian Employees (in 2010 only, see Data Notice #3

at <http://lehd.ces.census.gov/datatools/onthemap.php?name=DataNotices> for more information). Coverage under these datasets currently excludes several groups of workers. These include:

- Uniformed Military
- Self-Employed Workers
- Informally Employed Workers

As a result, jobs in these areas will not be included in any calculations from the LODES data and will not appear in the Flows Map. A project is currently underway to add Self-Employed Workers to the LED data infrastructure. For further updates on this project, please visit the LED main page at <http://lehd.ces.census.gov/>.

Understanding Locational Detail and “Flows”

The key sources of data for LODES and OnTheMap are administrative records and this results in two data system features of which users should be aware. First, for firms with multiple establishments in the same state, an imputation model is used to place workers at specific establishments since this exact information is not recorded by states’ UI wage record systems (except in Minnesota). Secondly, employment location is not always the same as a work site. For example, a construction worker or a truck driver may be employed and recorded at the main office for reporting purposes, but they actually perform their work elsewhere.

The result of the first effect is that there may be some uncertainty associated with the specific employment locations of workers in firms with multiple establishments in the same state (New York or New Jersey for the purposes of the Flows Map). The results of the second effect is that the flows shown in the map are upper bounds for actual commuting of this specific population. Because some people telework or work at non-establishment, offsite locations, not every home-employment location should be considered a commute. Research on these spatial differences between the LED partnership’s administrative record data system and the American Community Survey (which asks where a person worked) is ongoing.

LODES makes use of synthetic data methods to protect confidential information about the residential locations of workers. A synthetic dataset is one that has similar statistical properties to an original dataset but has been created in order allow for the release of data in the public domain without compromising confidentiality in cases where the original dataset cannot be released. In general it would be preferable to work with the original dataset if possible; however, synthetic data are created for OnTheMap because the original data *cannot* be released. More information on the confidentiality protection of workers’ residential location can be found at <http://lehd.ces.census.gov/datatools/onthemap.php?name=ConfidentialityProtection>.