Modeling Taxi Demand with GPS Data from Taxis and Transit

Eric J. Gonzales, Ph.D., Ci (Jesse) Yang, M.S., Ender Faruk Morgul, M.S., and Kaan Ozbay, Ph.D.

MNTRC Project 1141

July 2014

SJSU Research Center
210 N. Fourth St., 4th Fl.
San José, CA 95112

Tel // 408.924.7560
Fax // 408.924.7565

transweb.sjsu.edu

Typically transit is more competitive during the day when the frequency of service is high.

Taxis provide an alternative to conventional public transit services in many cities, and understanding the demand for taxis requires consideration of the role that taxis serve in the greater transportation system. The primary objective of this study was to identify the factors that drive taxi demand and to understand how this varies by location and time of day. A secondary objective was to demonstrate how emerging big data from taxis and transit systems can be integrated with demographic, socioeconomic, and employment information to develop useful demand models. Ultimately, developing methods to systematically analyze and extract meaningful information from these large data sources will help improve the way transportation systems are monitored and managed.

Study Methods

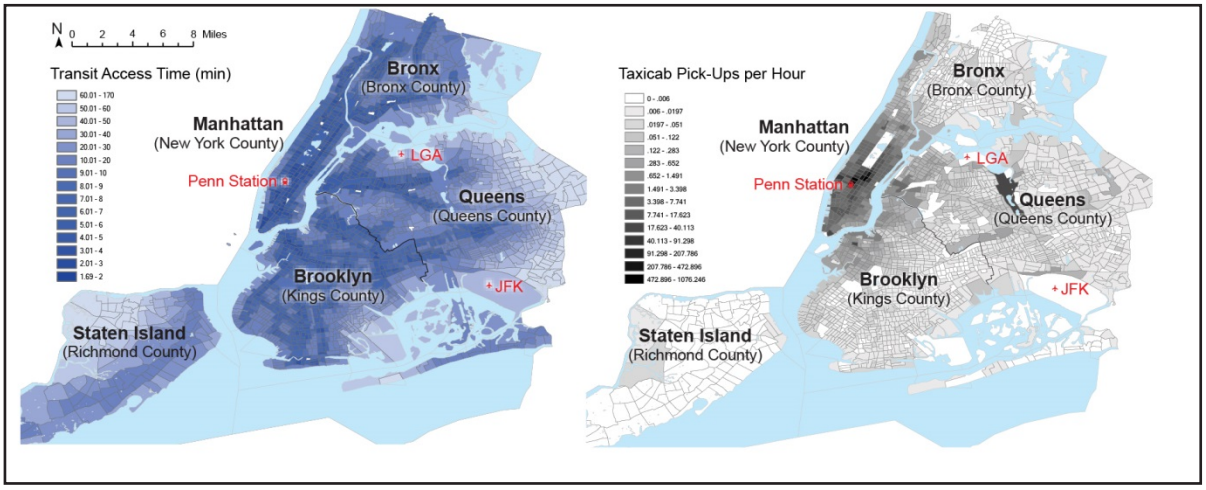
The study focuses on developing demand models for taxi trip generation and mode choice that explicitly account for the characteristics of transit service in the neighborhoods where trips are made. The dataset includes records of every taxi trip in New York City over a 10-month period. The data was tracked by Global Positioning System (GPS) receivers installed in each licensed taxicab. Additional data sources included detailed transit schedule and routing information. Demographic, socioeconomic, and employment data were obtained from the U.S. Census Bureau at the spatial resolution of census tracts. The analysis was used to develop models that provide insights into the factors that determine the number of trips made by taxis. This study demonstrates the modeling approach for NYC, but the methods are general and can be applied to cities around the world where similar data is collected and available.

The study was conducted in two parts. First, a trip generation model was developed to identify location characteristics that determine the number of taxi trip origins (pickups) and taxi trip destinations (drop-offs) during each hour of the day. In addition to demographic data, a novel measure of transit accessibility based on the time that it takes to walk to and wait for a transit departure is developed to account for the effect of transit service on taxi demand. Second, a mode choice model was developed and analyzed to determine how the competitive appeal of taxi travel versus transit changes by time of day as the travel cost for each mode varies. The mode choice model uses taxi and transit trip data between Pennsylvania Station and the major airports in the New York region.

Findings

The trip generation models that have been developed from these data reveal that there are six characteristics of a census tract that have the greatest explanatory power for estimating taxi demand: transit accessibility, population size, median age, percent of population educated beyond bachelor's degree, median income per capita, and number of employment opportunities. An additional, detailed investigation of taxi demand within Manhattan shows that there are certain types of employment opportunities that are more correlated with taxi trips than others. The number of employees working in retail, accommodation and food service, and healthcare are the strongest determinants of taxi demand. It appears that taxis and transit sometimes operate in

competition and at other times are complements, because both modes follow and influence the levels of activity in neighborhoods across the city.



Transit accessibility and the number of taxicab pick-ups per hour by census tract for the evening peak, 5:00 P.M. to 6:00 P.M.

The mode choice analysis, based on a comparison of trip costs by taxi and transit, show how the likelihood of travelers choosing one mode or the other changes over the course of the day. A sensitivity analysis is particularly useful in showing the tipping points at which the number of passengers traveling together in a group, or the value they place on their time, makes the additional expenditure for a taxi worthwhile. Typically transit is more competitive during the day when the frequency of service is high, especially during the morning and evening peaks when traffic congestion also slows taxis. Taxis are more competitive in the evening hours when traffic moves quickly and less frequent transit service imposes longer waiting times on travelers.

Policy Recommendations

The study shows how detailed data from GPS devices in taxicabs can be used to understand how the demand for taxi service varies across time and space. The authors recommend that planners and decision makers consider the role of taxi service in different parts of the city for different kinds of trips when making decisions about regulating the taxi market. This will become increasingly important as competing services seek to serve segments of the market that are not adequately served by conventional taxicabs.

About the Authors

Eric J. Gonzales, Ph.D., is an Assistant Professor in Civil and Environmental Engineering at the University of Massachusetts Amherst, Jessie Yang is a Graduate Assistant in Civil and Environmental Engineering at Rutgers University, Ender Faruk Morgul is a Graduate Research Assistant and Kaan Ozbay, Ph.D., is a Professor in Civil and Urban Engineering at the Polytechnic Institute of New York University.

To Learn More

For more details about the study, download the full report at ***transweb.sjsu.edu/project/1141.html***

MTI is a University Transportation Center sponsored by the U.S. Department of Transportation's Research and Innovative Technology Administration and by Caltrans. The Institute is located within San José State University's Lucas Graduate School of Business. WEBSITE transweb.sjsu.edu