Paper No. 14-4399

Comparison of Mode Cost by Time of Day for Non-driving Airport Trips to and from New York City Penn Station

Ci Yang, M.S.
Graduate Research Assistant, Dept. of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
96 Frelinghuysen Road, Piscataway, NJ 08854
Tel: (631) 880-1028
Email: jessie_yang06@yahoo.com

Ender Faruk Morgul, M.S.
Graduate Research Assistant, Dept. of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
96 Frelinghuysen Road, Piscataway, NJ 08854
Tel: (732) 445-6243
Email: morgul@eden.rutgers.edu

Eric J. Gonzales, Ph.D. (corresponding author)
Professor, Dept. of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
96 Frelinghuysen Road, Piscataway, NJ 08854
Tel: (848) 445-2868, Fax: (732) 445-0577
Email: eric.gonzales@rutgers.edu

Kaan Ozbay, Ph.D
Professor, Dept. of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
96 Frelinghuysen Road, Piscataway, NJ 08854
Tel: (732) 445-2792 Fax: (732) 445-0577
Email: kaan@rci.rutgers.edu

Word count: 6,379 + 6 (Figures and tables)*250 = 7,879

Paper Submitted for Presentation and Publication at the
Transportation Research Board’s 93rd Annual Meeting, Washington, D.C., 2014
Submission Date: November 15, 2013

ABSTRACT

In this study we develop a novel methodology using taxi GPS data and high-resolution transit schedule information to compare travel times and travel fares of the two main non-driving travel modes for airport ground access: taxi and transit. Five origin-destination (OD) pairs between Penn Station in New York City (NYC) and three airports in Greater NYC area are used as an example to demonstrate our methods. The total cost analysis considers both travel time and fare spent. A binary logit model is used to model modal choice of travelers. The results indicate that the transit mode is the more likely choice most of the time except the midnight period when transit service has longer headways. A sensitivity analysis shows the relationship between the value of time and total cost for different numbers of passengers traveling together and different times of day. The higher the value of time and the number of passengers, the more likely that taxi will be chosen for airport trips. The attractiveness of one mode relative to the other varies spatially and temporally according to the travel time and price. This paper focuses on understanding temporal variation of total cost of each mode and the effect that this is likely to have on mode share.

INTRODUCTION

Transportation planning for airport ground access has attracted increasing attention from both planners and researchers in recent years. Approximately 65% of airport trips are made by private vehicles in the US and Europe *(1)*. The remaining 35% of the journeys depend on alternative airport access modes. Several studies have been conducted to measure the effectiveness of commercial airport ground transportation services other than private automobile such as taxi cabs, buses, or trains, and travel time reliability analysis of the alternative modes indicates the effectiveness of these systems *(2-3)*. Another challenging issue is to understand how passengers decide on the travel mode for their airport trips. Trips to and from airports should be treated as a separate trip category in regional behavioral travel models due to the their unique characteristics (e.g., different value of time for passengers) and a number of influencing factors on mode choice such as trip purpose (e.g., business or leisure) and time-of-day *(4)*. The trip cost of an alternative mode relative to driving has been shown to be one of the major factors affecting the mode choice of passengers along with factors such as frequency of service and luggage capacity *(5)*.

The generalized cost of travel for different modes of transportation varies significantly depending on several factors related to traffic conditions and public transportation schedules. The use of public transportation might save money, but in return people give up the convenience of a door-to-door ride directly to their home or work place. Willingness to pay for a more reliable transportation service for airport access is estimated to be considerably higher than for regular daily commutes since scheduling constraints are more binding *(4).* The estimates of value of time for airport trips are found to be significantly different than the estimates which were made for overall network travelers *(4)*. Moreover shorter travel times are valued more for business trips compared to non-business trips to avoid the risk of missing the flight *(6)*.

This study focuses on comparing airport trips made by public transit (e.g., train, subway, and bus) and by taxi or limousine, which are the most common transportation modes used by non-driving passengers in the New York City (NYC) region *(4)*. Three major airports serve the region: John F. Kennedy International Airport (JFK), Newark Liberty International Airport (EWR) and LaGuardia Airport (LGA). These constitute the largest airport system in the United States. A number of public transit options exist for each airport and taxi traffic is also very busy, for example 37% of the passengers leaving the JFK airport are estimated to use taxis *(7)*.

This study is motivated by the importance of realistically comparing travel costs for alternative modes for airport ground access using revealed data sources. The majority of studies in the literature calculate the transit trip costs using previously determined schedules for certain types of transit services which cannot address all available transit options for airport ground access in cities with highly complex transit networks such as NYC. Different transit options for different times-of-day should be considered in the analysis, because the speed and reliability of each mode varies over the course of a day. While some of the existing methodologies consider waiting times as a part of total travel time, rough estimations for possible transfer times at train or bus stations can lead to biased results. It has also been demonstrated that about 40% of the population in the US uses smart phones, and there is a great potential for these devices to disseminate reliable real-time transit travel information *(8)*. Taxi travel times, on the other hand, are usually calculated per mile,with the assumption that the vehicle travels on the shortest path to the airport. Traffic conditions on the network by time of day have not been addressed adequately in previous studies although traffic congestion affects both the taxi fare and travel time.

In this study we use a novel methodology utilizing big data to compare travel times and fares for airport ground access by taxi and transit. For the analysis we use historical taxi GPS data recordings in NYC and transit schedule information from a web-based application we developed using Google Maps developer API sources. Transit travel time data gives us the opportunity to compare the travel times for a transit passenger who uses his/her smart phone to plan an airport trip. Taxi travel times, on the other hand, are the average values of the observed trips which are extracted from a dataset including more than 80 million trip records in NYC. Therefore the comparison in this study aims to evaluate the travel options for a well-informed passenger, who has perfect knowledge about the expected taxi fare and travel time.

This paper is organized as follows. The literature review section gives a summary of findings from previous studies on airport travel mode choice and related ground access studies. The next section describes data collection methods followed by a section explaining the methodology used for the comparison analysis. A conclusion section summarizes the findings and describes directions for future work.

LITERATURE REVIEW

A significant body of research has been conducted to understand how individuals choose their mode for airport trips. Harvey (*9*) was one of the first studies to demonstrate the factors influencing airport travel mode choice of departing airline passengers based on a travel survey in the San Francisco Bay Area. The analysis using a multinomial logit model shows that travel time and travel cost are two strong explanatory variables. Business travelers are found to be more sensitive to airport access travel time than leisure travelers, and values of time for most of the individuals are estimated at least as high as the average wage. Extra luggage, which is defined as more than one piece per person, deters passengers from choosing transit *(9)*. Psaraki and Abacoumkin *(10)* analyzed mode split for Athens International Airport in Greece to predict future mode shares and found that international passengers are more likely to use taxis or be dropped off by private cars. Pels et al. *(11)* studied mode choice in the San Francisco Bay Area and reported that business travelers have higher value of time and higher access time elasticity compared to leisure travelers. The authors also reported that access time has a larger influence on mode and airport choice compared to access cost. However, in their analysis they calculated travel times for each alternative mode as following: public transit travel time estimates are drawn from train and bus schedules, and for taxi travel times are based on distances from the center of the zip code to the airport. Therefore their estimations do not account for walking and transit transfer times for public transit or the effect of traffic congestion for taxis or private vehicles.

Some studies have been investigating the factors that influence airport ground access mode choice including demographics, trip cost, travel time, travel time reliability and accessibility. Gupta et al. *(4)* developed a ground access mode choice model for NYC using an air passenger survey and a nested logit model. In that analysis they prepared a transit Level of Service (LOS) data exogenously using online schedules, and waiting times are also taken into account. Similar to the previous studies demographic characteristics, trip cost, travel time and trip purpose are found to be the most significant variables in passengers' mode and airport choice. Tam et al. *(12)*

investigated how travel time reliability affects the mode choice of passengers using a combined dataset from revealed and stated preference surveys. The authors stated that increasing reliability can attract more passengers to use bus services. Luken and Garrow *(13)* studied the airport choice problem in NYC. Their analysis, based on online ticketing data, showed that the accessibility of the airport significantly affects the airport choice. They stated that analyzing peak and off-peak driving times can be an improvement for the model they presented.

This study aims to provide insight into understanding the accurate cost information, which other works have suggested will affect the mode choice behavior of passengers by utilizing realistic data sources for taxi and transit options. Neufville *(14)* argues that planning for low-cost airports include a significant focus on public transportation options so more attention should be drawn to non-driving passengers.

The analysis presented in this paper is unique in several ways. First, this study considers costs by time-of-day, which has not been considered in detail in the existing literature. Real data for peak and off-peak time periods provides useful information about how the travel costs of different alternatives change by time of day. Second, this study uses two novel datasets for ground access mode choice which are 2-year taxi GPS-data records for an accurate distribution of taxi travel times and Google Maps API based transit travel times based on transit schedules. The analyses are made from the perspective of a passenger who considers his/her options in dollar and time terms before taking an airport trip.

METHODOLOGY

The objective of this study is to develop a data-oriented method to compare the generalized cost for different non-driving modes to get to and from airports, and to understand whether transit or taxi yields a better utility at different times of the day. While the results of this study may be useful to individuals trying to choose the best way to travel, the method proposed in this study can also help policymakers understand the factors that affect peoples' mode choice and to plan for improved airport ground access.

As web services and information technology become more advanced, it is easier for people to acquire complete transit and taxi information such as transit timetables, transit fares, and taxi rates for them to make travel decisions. Assuming that passengers make these decisions on the basis or money costs and the estimated travel time, the relative attractiveness of one mode over the other may change as transit schedules, fares, and taxi travel times vary for different times of day and days of the week.

Google Maps Transit service can estimate real-time transit route and travel time for any origin-destination (OD) pair through its online services. The data has been made publicly available to developers since June 2012, so minute by minute travel information for transit can be acquired by using a JavaScript to extract the information from Google Maps. This web-based application enables us to extract scheduling and routing information for transit trips as often as every 5 minutes. This allows us to compare the characteristics of transit trips with taxi trips serving the same OD pair across all times of the day and night.

To quantify the cost, including both time and money at different times of day, the total generalized cost for an individual in units of money can be computed for each mode i at time j is denoted TC_{ij}, calculated as follows *(15,16)*:

$$TC_{ij} = \alpha \times T_{ij} + \frac{F_{ij}}{n} \tag{1}$$

where

$TC_{ij} = total\ cost\ for\ each\ travel\ mode\ i\ at\ hour\ j\ (\$)$
$T_{ij} = average\ travel\ time\ for\ each\ tavel\ mode\ i\ at\ hour\ j\ (hour)$
$F_{ij} = average\ fare\ paid\ for\ mode\ i\ at\ hour\ j(\$)$
$\alpha = value\ of\ time\ (\$/hour)$
$n =\ number\ of\ passengers\ sharing\ taxi\ service, for\ transit\ service\ n = 1$

The units of TC_{ij} in equation (1) are dollars, but we can also express the total generalized cost in units of hours by dividing the total cost by the value of time, α. The utility of each travel mode in hour j is defined based on the generalized cost as

$$U_{tran,j} = b - \beta \times TC_{tran,j} = b - \alpha\beta \times T_{tran,j} - \beta \times F_{tran,j}$$

$$U_{taxi,j} = b - \beta \times TC_{taxi,j} = b - \alpha\beta \times T_{taxi,j} - \frac{\beta \times F_{taxi,j}}{n} \tag{2}$$

where b is the benefit for each individual of completing a trip to or from the airport, and β is the equivalent utility of a dollar. In this paper, we assume that benefit is the same for both choices as long as the OD pair is fixed. The choice between two modes, such as transit and taxi, is typically modeled with a binary logit model based on the difference of utilities between the choices *(17)*. The probability that an individual will choose one mode over the other is shown in equation (3):

$$P_{tran,j} = \frac{\exp(-\beta \times TC_{tran,j})}{\exp(-\beta \times TC_{tran,j}) + \exp(-\beta \times TC_{taxi,j})}$$

$$P_{taxi,j} = \frac{exp(-\beta \times TC_{taxi,j})}{exp(-\beta \times TC_{tran,j}) + exp(-\beta \times TC_{taxi,j})} \tag{3}$$

where $P_{tran,j}$ and $P_{taxi,j}$ are the probabilities of choosing transit and taxi at each hour j, respectively. The traffic flow for each mode choice is the product of the probability from equation (3) and the total travel flow. In the following sections, comparisons of transit and taxi trips in NYC are performed based on their travel time, total cost and their corresponding choice probability.

CASE STUDY

We compare two major types of public transportation services for airport access in NYC: transit (including train, air train, subway, and bus) and taxi between NYC Pennsylvania Station (Penn) and the three international airports in NYC (Figure 1); JFK, EWR and LGA. There are two reasons that we select Penn in our case study. First, the number of airport trips ends in Penn ranks top among almost all origin and destination points in NYC. About 18% trips from Penn to outside NYC are to EWR, and about 1% of all trips from or to Penn Station end in LGA or JFK.

Second, as Penn is a transit hub located in the center of NYC, both taxi and transit services can be accessed conveniently at Penn or the three airports, which provides an individual with good availability for both transportation modes. The accessibility through transit, taxi, or car for all airports is listed as follows *(18)*:

- JFK, located in Queens, NY, is easily accessible from AirTrain, buses, and taxis. AirTrain JFK provides access to both the Long Island Rail Road (LIRR) and Metropolitan Transportation Authority (MTA)'s New York City subway and bus system, with connections at both the Jamaica and Howard Beach station.
- EWR is located in Newark, NJ, and is accessed from, Lower Manhattan and the West Side of Manhattan via the Holland and Lincoln Tunnels if traveling by car/taxi. AirTrain Newark provides access to New Jersey Transit trains into Newark, New York City and points on Amtrak's national rail network.
- LGA, located about four miles from Manhattan, can be access by taxi. LGA also has a variety of transportation options to NYC's five boroughs and beyond by public buses, NYC subways and other regional rail lines.

FIGURE 1 Location of NYC Penn Station and three airports

Taxi Data

A 10-month dataset (February 1, 2010, to November 28, 2010) of global position systems (GPS) data are available for taxi trips from the NYC Taxi & Limousine Commission (TLC). This data set includes trip information for all taxis travelling from Penn Station to each of the three airports. NYC taxi trips have also been used in several studies such as travel time reliability for inner-city traffic *(19)* or commercial vehicle delivery time estimation *(20)*.

The data consists of 147 million taxi trips across the city for 301 days, from which we are able to extract trips between Penn Station and the three airports. The dataset has 30 data fields for each taxi trip with detailed taxi-specific information (e.g., medallion number, shift number, driver name) temporal and spatial information (taxi pick-up and drop off date, time, location), fare (e.g., toll, tip, total fare amount), and distance travelled. The boundary of locations for selecting those trips are restricted to 500 feet radius of the center of Penn Station, 1 mile radius of the center of EWR and within the census tract of JFK (Census tract ID: 36081071600), and LGA (Census tract ID: 36081033100). As NYC taxies are not allowed to pick up passengers from EWR, no taxi trips are available from EWR to Penn Station during this period, so only trips from Penn Station to EWR are considered for comparison. The remaining two airports have taxi trips in both directions, thus each airport has two OD pairs for comparison analysis. In total, 5 OD pairs are considered and summarized in Table 1.

TABLE 1 The taxi trips extracted from 10-month GPS data and 1-week transit trips

	Taxi					Transit	
OD pair	**No. of Obs**	**Passenger No**	**Total Amount ($)**	**Trip Time (min)**	**Trip Distance (mi)**	**Trip Time (min)**	**Fare ($)**
		Mean *(SD)*	**Mean *(SD)***	**Mean *(SD)***	**Mean *(SD)***	**Mean *(SD)***	**Mean *(SD)***
Penn-JFK	5624	1.81 *(1.29)*	52.34 *(5.26)*	47.79 *(17.57)*	17.24 *(1.89)*	52.51 *(10.34)*	12.47 *(1.69)*
JFK-Penn	2691	1.87 *(1.27)*	51.69 *(4.33)*	45.11 *(12.74)*	17.72 *(1.79)*	60.96 *(10.03)*	12.65 *(1.37)*
Penn-LGA	9697	1.63 *(1.21)*	34.85 *(5.64)*	31.43 *(10.30)*	10.23 *(1.38)*	60.85 *(6.32)*	4.92 *(3.57)*
LGA-Penn	3630	1.65 *(1.22)*	35.07 *(5.58)*	32.06 *(9.35)*	10.21 *(1.71)*	61.48 *(7.57)*	6.91 *(3.23)*
Penn-EWR	1445	1.80 *(1.28)*	67.30 *(10.48)*	32.09 *(9.49)*	17.08 *(2.05)*	58.87 *(19.90)*	17.25 *(9.54)*

SD: Standard Deviation

The total fare amount is the usually a flat rate between most of Manhattan and JFK or EWR airports plus any tolls, tips, and surcharges. The fare between Manhattan and LGA has some variability because trips are charged the normal metered rate. The trip time has largest variability among all four variables listed for each airport, which indicates variability of traffic conditions. The trip distance is mostly stable for all airport trips, but the slight variability indicates that various alternative routes might be taken to for the same OD pairs.

Transit Data

For transit travel time data we used Google Maps API Transit Directions Service, which offers free transit route guidance with a daily request limit. The information was gathered in XML format using a web-based JavaScript code.[1] We developed an application that extracts travel time information based on schedules for a set of OD pairs within certain time intervals throughout the day and stores that information in a database format. The routing information provided by Google is assumed to be the optimal transit option for the requested time and OD pair since web-based routing service compiles all available scheduling information for different transit modes. The transit travel times used in this study realistically reflect the information available to an individual, because the extracted schedules and routes are the same as what an individual would find by using Google Maps on a mobile device. The transit travel duration of each trip includes waiting time, transfer time, and in-vehicle travel time.

One week (including weekdays and weekends) of transit service information has been obtained every 5 minutes from Google Maps API for each of the five OD pairs between Penn Station and the three airports. Data for approximately 2,016 transit trips have been collected for each OD pair. The transit travel time distribution for all 5 OD pairs are shown in Figure 2, indicating that travel time in weekdays are very close to each other, which could be resulting from the fact that transit schedule are almost the same for weekdays (Monday through Friday), it is necessary to analyze Saturday and Sunday separately according to the figure. The average travel time of transit at weekends is higher than that of weekdays because service headways are longer while average travel time of taxi at weekends is lower than that of weekdays because traffic is less congested. The day of the week will affect the costs that travelers face on each mode.

[1] The developed JavaScript source code is available by request.

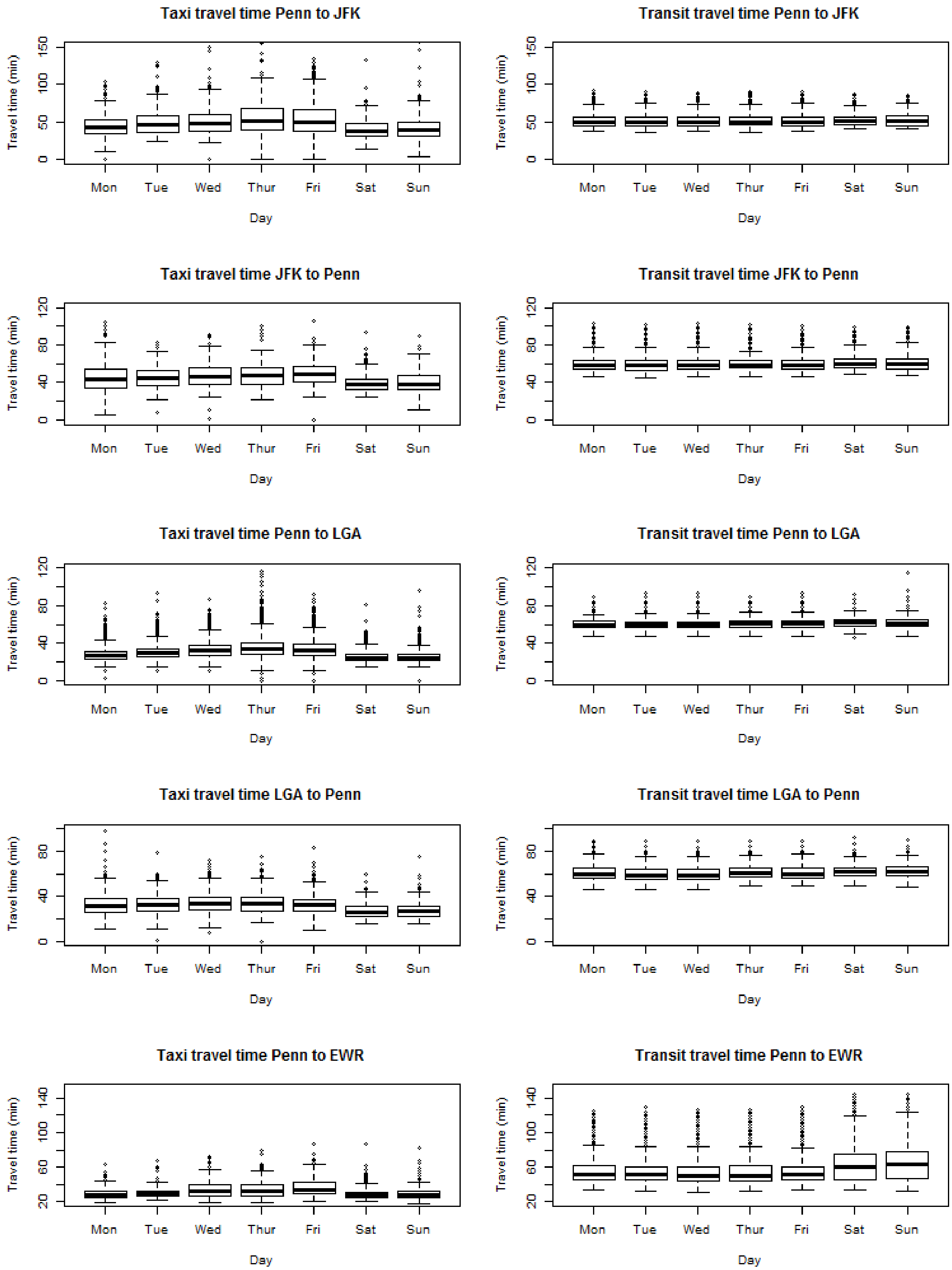

FIGURE 2 Boxplot of transit travel time by day of the week for 5 OD-pairs.

Determination of value of time, α

The value of time for airport trips varies considerably from person to person and it can be considered as a continuous variable that is distributed probabilistically across the user population *(21)*. Value of time for business trips can be higher than value of time for leisure trips *(4, 9)*. The distribution of values of time for airport trips is also likely to differ from that of other trip purposes, which makes estimation of this value from other sources difficult.

The UK department of transportation suggests £47.95 (equivalent \$76) as the value of time for Taxi/Minicab passenger and £39.65 (\$63) for rail passenger in 2010 *(22)*. Since we are considering trips to/from Penn Station, the income in Manhattan is used as reference *(9)*. The 5-year (2007-2011) American Community Survey (ACS) estimate per capita income in Manhattan is \$61,290 *(23)*, which is \$29.5 per hour if working as full time job with 40 hours per week. Gupta et al. *(4)* has suggested considering airport trips as one leg of long-distance travel, thus the travel time is valued higher than a regular trip. Travelers may also be willing to pay more to avoid missing of a flight. They suggest \$42/hour for leisure trips and \$63/h for business trips. Since we do not know the value of time for passengers that made airport trips in NYC we select one reasonable fix value of time \$40/h to represent for everyone and anytime based on above references as an illustration. In the sensitivity analysis that follows, we provide analysis for a wide range of values of time.

Calibration of coefficient β

Equation (2) describes the relationship between cost and utility. The coefficient β influences the outcome of the binary logit model. In order to estimate mode choice, it is necessary to determine β, which can be done by comparing the total number of transit and taxi trips.

We used JFK ridership information and taxi GPS data to estimate a single β. There are two reasons why we use one value of β for all three airport trips. First, β is the marginal utility of total cost, and it is likely that the average value is similar for passengers using each of the airports. Second, we lack data to estimate a specific β for LGA and EWR. LGA does not have AirTran service available and most public transportation to and from LGA are bus and subway, so we are not able to acquire LGA ridership counts. Taxi data for EWR only contains trips from Penn to EWR but no trips from EWR to Penn, so we are not able to estimate total taxi trips. Paid ridership of JFK AirTrain was 5.3 million in 2010 *(24)*, which accounts for nearly all of the transit trips to and from JFK airport. In the same time period, there were 3.386 million taxi trips to and from JFK extrapolated from the complete 10-month records in taxi GPS data assuming the 2 month without records in 2010 has the average taxi ridership of the 10-month data. Based on the trips counts above, the overall mode share between taxi and transit is 39% taxi and 61% transit for trips to and from JFK.

Without more detailed transit ridership data, we assume that the overall mode share for all trips to and from JFK during 2010 is the same as mode share for trips between Penn Station and JFK. Therefore, we calibrate the logit model by selecting the β value that makes the model estimates over the course of the day match this observed mode share. Figure 3 shows the relationship between β and the aggregated probability of taking a taxi using the average travel time and fare for all the taxi and transit data at each hour, j, between Penn Station and JKF (both OD pair

Penn-JFK and JFK-Penn), and *n*=1, *α*=$40/hour. At each hour, the relationship between number of taxi trips ($n_{taxi,j}$) and the estimated number of transit users ($\hat{n}_{transit,j}$) is shown in equation (4). When *β*=0.012, the expected probability of people choosing taxi to the airport is 0.39, which matches the data from 2010. This value of *β* is applied to the cost data for all the airport trips in order to estimate the mode share by taxi and transit.

$$\frac{n_{taxi,j}}{P_{taxi,j}} - n_{taxi,j} = \hat{n}_{transit,j}$$

$$P_{taxi,j} = \frac{\sum_j n_{taxi,j}}{\sum_j \hat{n}_{transit,j} + \sum_j n_{taxi,j}} \tag{4}$$

Assumptions and Data Processing

In this study, the main challenges are the data collection and data processing. Some assumptions made in order to perform the comparison of transit versus taxi are as follows:

1) Taxi fare is calculated per person, so the value that applies in equation (1) divides the fare by the number of passengers, but the travel time is experienced by each passenger regardless of the group size.
2) There are always a sufficient number of taxi cabs available at each airport for passengers to hail, so there is no time spent for walking and waiting for a taxi at any airport.
3) The origin and destination of the transit trips are very close to the transit stops, so the walking distance is negligible in our analysis.
4) The average travel duration in one hour is assumed to represent for the travel time at that hour of the day for both train and taxi, provided that we have a large number of trips per hour.
5) All passengers are able to buy the tickets before boarding transit to avoid additional fees. MTA subway and bus riders pay a flat rate of $2.50 per trip. Discount fares such as senior citizen ticket, weekly tickets, or monthly tickets are not considered.
6) Travelers to all airports are assumed to have the same average utility preferences, so *β* is the same for all trips.

The deficiencies in GPS data, mostly because of satellite errors, receiver noise errors, coordinate transformation errors, and errors made by the driver *(25)* need to be filtered. The taxi GPS data is processed to minimize the influence of outliers. Some false records are eliminated, for example, records which have total fare amount equal to zero or travel distance is less than the straight-line distance between the origin and destination. Sometimes we use more than two criteria (fare amount, distance between Penn and airport, and travel time) to determine whether a data point should be removed. Ultimately, less than 2% of the original taxi records were eliminated through this filtering process. This data selection procedure requires familiarity with information such as fare amount, distance between Penn and each airport, travel time, and the rate codes.

Transit data collected from Google every 5 minutes for all three airports are kept without filtering because Google estimates are already based on clean schedule data. The transit fare is calculated according to the routes that Google Maps provides at different times of the day.

Results

On weekdays, the average travel time for taxis is less than that of transit at all times. When considering both the time and money spent on the trip, the total cost indicates that even if the passenger is travelling alone, taxi has a cost advantage only in middle of the night (12 a.m. to 6 a.m.). The taxi travel times vary significantly during weekdays, and the longest travel times are usually observed during morning peak (6 a.m. to 10 a.m.) and afternoon peak (2 p.m. to 6 p.m.) as shown in Figure 3.

In order to consider the variability of travel costs, we also calculate the standard error (SE) for the total cost of trips (Figure 3). Some taxi data has few observations at midnight, which results in relatively higher SE and a wider 95% confidence interval (CI) for the mean values at each hour (approximately equals to mean+/-1.96*SE). On the other hand, the transit data show relatively small variance within each of the 24 hours. The main difference in transit travel times is arising from the waiting times and transfer times for the next available train or bus. The transit travel time and cost are less variable than taxi travel time and cost, which depend on the traffic condition at different times of day and from day to day *(12)*.

In our analysis, which is limited between Penn Station and the three airports, the probability of mode choice at each time of day is calculated using the binary logit model as explained in the methodology section. The results are plotted in Figure 4. Based on the difference of total cost for taxi and transit, the mode share can be expected to change for different times of the day. For example, transit tends to be more competitive during rush hours when traffic congestion makes taxi trips slower. On the other hand, taxis are more competitive in late night hours when transit headways are long.

Since these OD pairs are just a partial set of all trips that go to and from the airports, it can only reflect the costs at those locations. Different locations may have a totally different trend based on the travel time and cost. Time and money are not the only things that people consider when making travel choices, but the literature suggests that these are the most important factors. Furthermore, it is possible that some people just use transit for everything without even comparing it with taxi.

Some factors that likely influence people's choice on taxi or transit other than travel time and fare are that taxis provide a more personalized door to door service with additional benefits such as assistance with luggage. Some of this value is captured in the tips that are included in the taxi data and the calculated total fare. In reality people may experience addition penalty for transit because they need to walk a certain distance to get transit service. We are not considering these additional benefits or penalties in order to focus the analysis on the effects of money cost and travel time on the competitiveness of each mode.

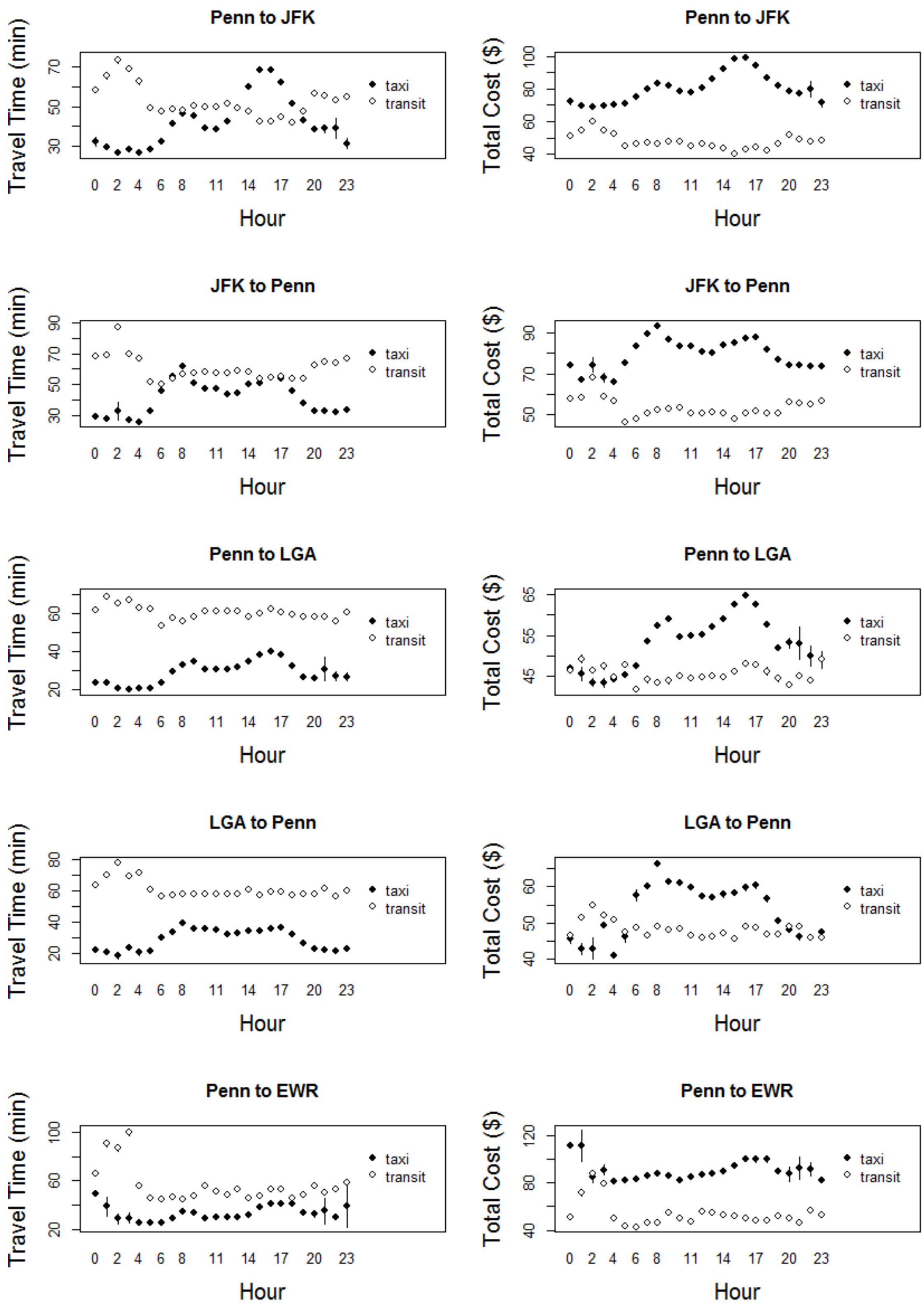

FIGURE 3 Weekday taxi and transit travel time and cost comparison mean and mean +/- Standard Error.

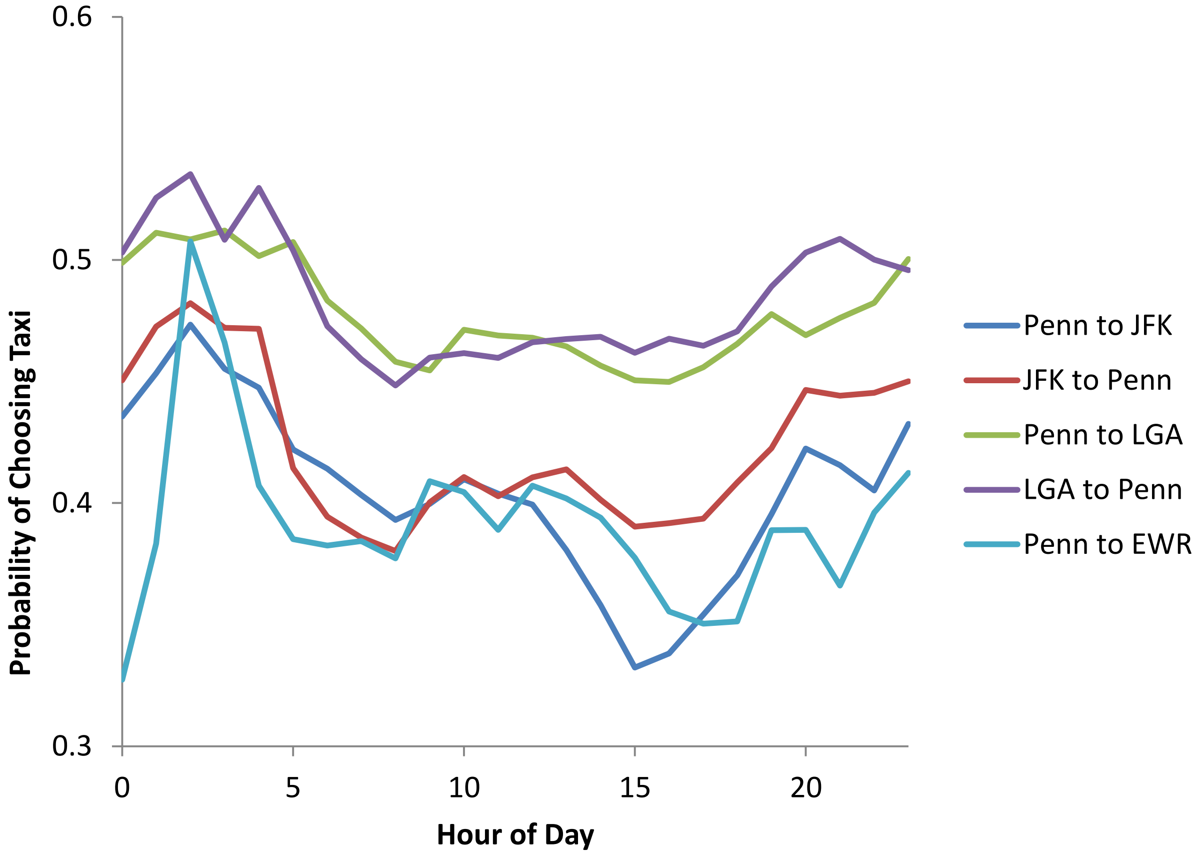

FIGURE 4 Weekday probability of choosing taxi for all airports.

Sensitivity Analysis

Figures 3 and 4 have shown the relationship between the total cost and the probability of mode choice at each hour of the day. The higher the cost of one mode, the lower the probability that a person selects this mode. The total cost is also influenced by the value of time and the number of passengers. From the taxi data, we can see that on average there is more than one passenger taking taxi together to go to or from the airports. This fact is evidenced for all three airport taxi data by an average of passenger count from 1.6 to 1.8 (Table 1). Therefore the probability of an individual's travel mode choice depends on both the value of time and the number of passengers in the group. A sensitivity analysis is performed to investigate the effects of these two factors in detail.

We consider the average travel time and fare of all records for taxi and transit as travel time and fare for each OD pair. For the sensitivity analysis we change both value of time, α (in range \$10/hour to \$70/hour), and passenger count, n (in the range 1 to 3), to see how much influence these factors have on total cost as shown in Figure 5a-5f. If the value of time is fixed, changing the number of passengers will only affect the taxi fare per person because the transit fare per person is always fixed. The slope of each mode is the travel time, and the intercept is the fare per person according to Equation (1). Intersections of taxi cost and transit cost are found for all OD pairs except trips from Penn to JFK (Figure 5a). The intersection indicates a value of time when the cost of taxi and transit are the same for different numbers of passengers. We can tell whether a taxi or transit has a higher cost beyond the intersected value of time by comparing slopes of cost for each mode. For example, JFK-Penn transit cost intersects with JFK-Penn taxi cost at $57/h for n=2, indicating that the total cost of taxi is higher when value of time is less than $57/h because the slope for transit exceeds the slope for taxi (Figure 5b). This means that if two passengers are travelling as a group, it is better to choose transit if the value of time is lower than $57/h, otherwise it is more cost-effective to choose taxi.

On average, there is no way that a trip from Penn Station to JFK will be less costly by taxi in the assumed range of values of time and number of passengers (Figure 5a). For the reverse direction, from JFK to Penn Station, taxis do become competitive for sufficiently high values of time and passenger occupancies (Figure 5b).

For trips to and from LGA (Figure 5c-5d), if traveling alone, the taxi costs are higher than transit costs when the value of time is less than $66/h (Penn-JFK) or $61/h (JFK-Penn), however, if traveling with more than two people, taxi is more competitive than transit if the value of time is larger than $27/h. This is reasonable because there is not direct transit service between Penn Station and LGA. Since the distance is shorter than any of the other airports, taxi is more competitive when people are travelling in groups.

The trip cost from Penn to EWR seems similar to JFK to Penn, except that the intersection points are slightly different. Transit costs more if the value of time is higher than $47/h (n=2) or $17/h (n=3). As $63/h represents the value of time for business trips in NYC and $42/h as the value of time for leisure trips *(4)*, it is likely that a business trip will use a taxi for trips from JFK-Penn or Penn- EWR if traveling with more than 2 people, but a leisure trip will use taxi only if travelling with at least 3 people.

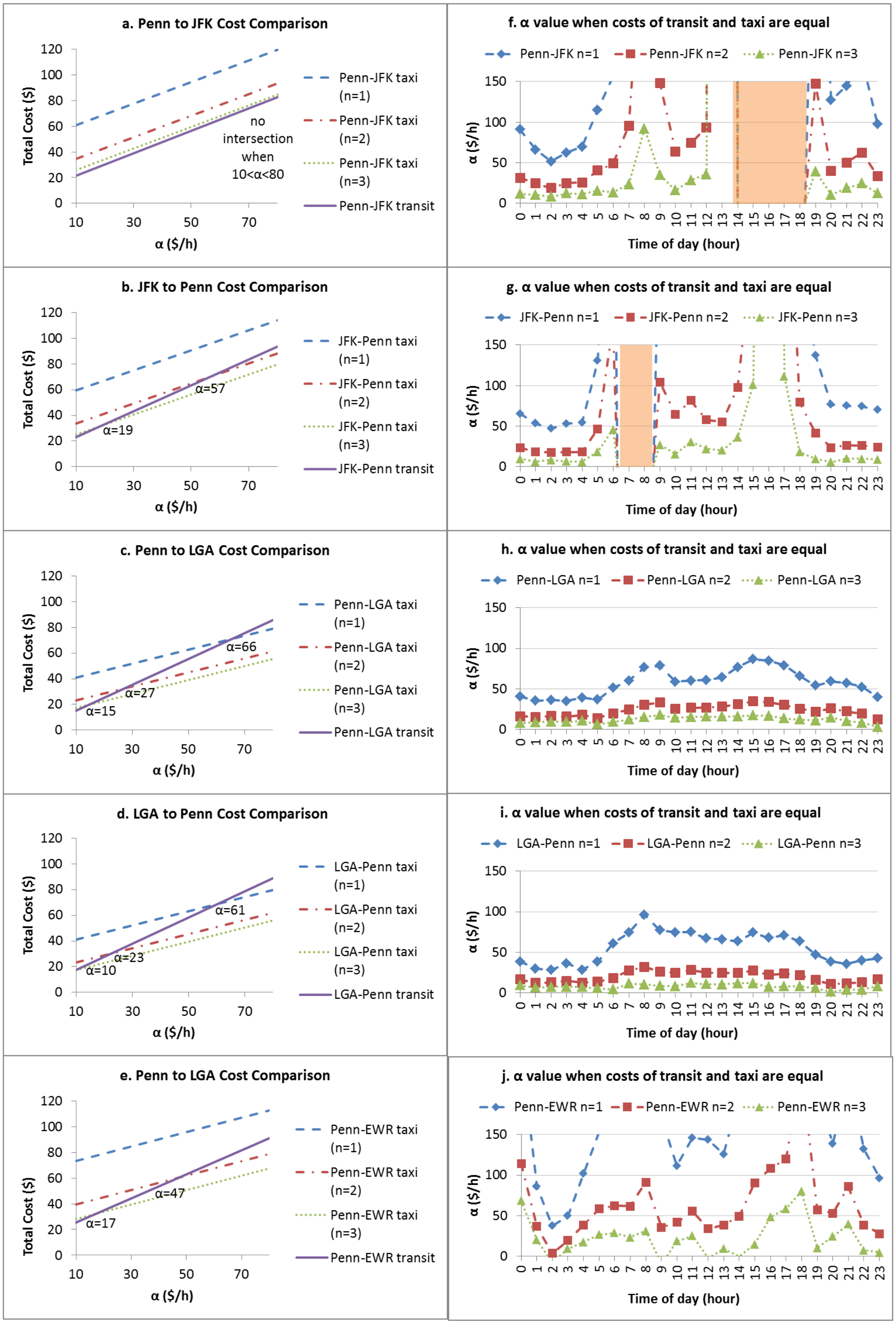

FIGURE 5 Sensitivity analysis on value of time (α) and passenger count (n).

In order to account for the effect on the variation of travel time throughout a day, within each hour we combine the two other factors by displaying the value of time when taxi cost is equal to transit cost to show how this tipping point value of time changes over the course of a day (Figure 5f-5j). For most cases, travel time is longer by transit than by taxi (slope for transit exceeds the slope for taxi), so values of time greater than the tipping point are associated with more cost-effective taxi service, and values of time less than the tipping point are associated with more cost-effective transit service. For a couple of time periods, transit is actually faster than taxi (slope for taxi exceeds the slope for transit), and the interpretation switches, so in the shaded areas of Figure 5f and 5g, all trips are more cost effectively served by transit, regardless of the value of time.

Using OD pair Penn to JFK at 5 a.m. as an example (Figure 5f), if the value of time is set to \$40/h, and we consider taxi when n=3, the value of time exceeds the tipping point value (\$16/h) taxi is the more cost-effective mode. When n=2, \$40/h is right at the tipping point. When n=1, the value of time is well below the tipping point (\$115/h), so transit is more cost-effective for traveling alone.

The relatively low tipping point values for LGA compared to EWR and JFK show that taxi is more competitive than transit for that airport, appealing to a wider range of values of time. There is also a pattern at all airports that taxi is more competitive in the early hours of the morning (around 2 a.m.) when transit service is also less frequent. These results have policy implications, because they show how airports differ in the competitiveness of ground access modes, and how this changes by time of day.

Results for EWR (Figure 5j) suggest that transit is more competitive to and from Penn Station to EWR, but for midnight trips taxi have a lower total cost than transit since the frequency of service is lower which results in longer waiting times. However, because of the relatively long distance between EWR and Penn, it is possible that taxi is more likely to be chosen based on the factors like convenience and comfort, which we did not consider in this study.

CONCLUSION

This paper presents a methodology to compare the total cost for two modes of transportation (transit and taxi) using taxi GPS data and high-resolution transit schedule information. Trips between NYC Penn Station and three New York area airports (JFK, LGA and EWR) at different times of day are used to illustrate the methods. As shown in the total cost analysis and mode choice models, transit is found to be more cost-effective than taxi for most times of the day if passengers are traveling alone and the value of time is \$40/hour, except some midnight periods when transit service has long headways that contributes a significant amount of time to wait or transfer.

The sensitivity analysis suggests that people are more likely to choose taxi to travel from Penn Station for airport trips if: 1) they have high value of time, 2) they are traveling with a large group of people, or 3) they are traveling in late night hours. It is also found that if people are traveling for business trips taxis become a less costly choice for airport access. For both JFK and EWR, because of the long distance, the taxi fare is very high, making transit a more competitive mode most of the time (especially when n=1), even though taxis have advantage in travel time.

However, as LGA airport is closer to Penn Station, the relatively low taxi fare and low travel time make taxi a better choice.

The results show that the taxi total cost or travel time always has two peaks for those airports, the morning peak (between 6 a.m. to 10 a.m.) and afternoon peak (between 1 p.m. to 6 p.m.). Taxi data provides an indication of traffic conditions in NYC *(19)*, so the use of this data to calculate the travel cost could be incorporating both the temporal and spatial effects of traffic congestion in the city. However, this study is limited to the temporal analysis of the 5 most popular OD pairs for the airports, which all include trips to and from Penn Station. This can create bias if used to estimate total costs for the entire city. Future applications could be expanded to consider the spatial dimension as well. We will need to include multiple OD pairs distributed all over the city.

There are other factors that affect the choice of mode for trips to and from the airport, such as convenience and comfort, which are not considered in our study because they cannot be easily measured and quantified. With additional data about the number of passengers using each mode by time of day, it may be possible to gain some insights into the effect of these less tangible factors by comparing the expected mode shares from the utility functions used in this paper with the observed mode shares.

This paper could be used as an example of a practical method to estimate the travel cost including both time and money. As information and resources like travel time and fare are increasingly accessible, it is possible to design a smartphone app or a small computer program at the transit ticket vending machine to estimate total cost using our methodology. This information can be used (especially with the choice model) to understand the factors that affect the aggregate mode choice decisions of the public. This will be useful to transportation planners and policy makers to improve the quality of travel options available to people traveling to and from airports. Passengers can easily get the travel time and cost estimation from the system after inputting a value of time factor, α, the number of passengers traveling together, n, and maybe the acceptable walking distance or number of bags. This will help people make their decision to take transit or take a taxi by not only considering the fare, but also the realistic waiting and travel times.

ACKNOWLEDGMENTS AND DISCLAIMER

This material is based upon work supported by the U.S. Department of Transportation's University Transportation Centers Program under Grant Number DTRT12-G-UTC21.

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

REFERENCES

1. Humphreys, I., Ison, S., Changing Airport Employee Travel Behavior: The Role of Airport Surface Access Strategies. In Transport Policy, Vol.12, 2005, pp. 1-9.

2. Shriner, H.,W., Hoel, L.A. Evaluating Improvements in Landside Access for Airports. In *Transportation Research Record: Journal of Transportation Research Board*, No.1662, Transportation Research Board of the National Academies, Washington, D.C., 1999, pp.32-40.
3. Gosling, D.,G. Predictive Reliability of Airport Ground Access Mode Choice Models. In *Transportation Research Record: Journal of Transportation Research Board*, No.1951, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp.69-75.
4. Gupta, S., Vovsha, P., Donnelly, R. Air Passenger Preferences for Choice of Airport and Ground Access Mode in the New York City Metropolitan Region. In *Transportation Research Record: Journal of Transportation Research Board*, No.2042, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp.3-11.
5. Akar, G. Ground Access to Airports, Case Study: Port Colombus International Airport. In *Journal of Airport Management*, Vol.30, 2013, pp. 25-31.
6. Koster, P., Kroes, E., Verhoef, E. Travel Time Variability and Airport Accessibility. In *Transportation Research Part B*, Vol.45, 2011, pp. 1545-1559.
7. Conway, A., Kamga, C., Yazici, A., Singhal, A. Challenges in Managing Centralized Taxi Dispatching at High-Volume Airports: Case Study of John F. Kennedy International Airport, New York City. In *Transportation Research Record: Journal of Transportation Research Board*, No.2300, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp.89-90.
8. Gould, J. Cell Phone Enabled Travel Surveys: *The Medium Moves the Message. In Transport Survey Methods: Best Practice for Decision Making*, 2013, pp. 51-70, Emerald Group Publishing.
9. Harvey, G. Study of Airport Access Mode Choice. In *Journal of Transportation Engineering*, Vol.112, No.5, 1986, pp.525-545.
10. Prasaki, V., Abacoumkin, C. Access Mode Choice for Relocated Airports: The New Athens International Airport. In *Journal of Air Transportation Management*, Vol.8, 2002, pp.89-98.
11. Pels, E., Nijkamp, P., Rietveld, P. Access to and Competition Between Airports: A Case Study for The San Francisco Bay Area. In *Transportation Research Part A*, Vol.37, 2003, 71-83.
12. Tam, M., Lam, W.H.,K., Lo, H. The Impact of Travel Time Reliability and Ground Service Quality on Airport Ground Access Mode Choice. In *Journal of Choice Modelling*, Vol.4, 2011, pp.49-69.
13. Luken, B., L., Garrow, L. Multiairport Choice Models for the New York Metropolitan Area: Application Based on Ticketing Data. In *Transportation Research Record: Journal of Transportation Research Board*, No.2206, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp.24-31.
14. Neufville, R. Planning Airport Access in an Era of Low-Cost Airlines. In *Journal of the American Planning Association*, Vol.72, 2006, pp.347-356.
15. Zhang, G., Wang, Y., Wei, H., Yi, P. A Feedback-Based Dynamic Tolling Algorithm for High-Occupancy Toll Lane Operations. In *Transportation Research Board*, 2065, 2008, pp. 54-63.

16. Morgul, E., F., Ozbay, K. Simulation-Based Evaluation of a Feedback Based Dynamic Congestion Pricing Strategy on Alternate Facilities. In *Transportation Research Board 90th Annual Meeting*, 2011, Washington, D.C., USA.
17. Train, K. Discrete Choice Methods with Simulation. *Cambridge University Press*, 2nd edition, 2009
18. The Port Authority of New York & New Jersey website, available online at http://www.panynj.gov/ Accessed on 7/10/2013
19. Yazici, M. A., Kamga, C., Mouskos, K. Analysis of Travel Time Reliability in New York City Based on Day-of-Week Time-of-Day Periods. In *Transportation Research Record*, Vol. 2308, 2012, pp.83-95
20. Morgul, E., F., Ozbay, K., Iyer, S., Holguin-Veras, J. Commercial Vehicle Travel Time Estimation in Urban Networks Using GPS Data from Multiple Sources. In *Transportation Research Board 92nd Annual Meeting*, 2013, Washington, D.C., USA.
21. Jiang, L., Mahmassani, H.S., Toll Pricing: Computational Tests on How to Capture Heterogeneity of User Preferences. In *Transportation Research Record*, Vol. 2343, 2013, pp.105–115
22. Values of Time and Vehicle Operating Costs TAG Unit 3.5.6, UK Department of Transport, Transport Analysis Guidance (TAG), February 2013. Available online at http://www.dft.gov.uk/webtag/documents/expert/unit3.5.6.php Accessed on 10/15/2013
23. 2007-2011 American Community Survey 5-Year Estimates DP03: Selected Economic Characteristics, American Community Survey. Available online at http://factfinder2.census.gov Accessed on 11/5/2013
24. 2010 Airport Traffic Report, Port Authority of New York & New Jersey, available online at http://www.panynj.gov/airports/pdf-traffic/ATR2010.pdf Accessed on 7/22/2013
25. Zito, R., D'Este, G., Taylor, M. A. P. Global Positioning Systems in the Time Domain: How Useful a Tool for Intelligent Vehicle-highway System. In *Transportation Research Part C*, Vol.3, 1995, 193-209.