A Model for Joint Choice of Airport and Ground Access Mode

Surabhi Gupta, Peter Vovsha, Robert Donnelly; PB Americas
1 Penn Plaza, 2th Floor, New York, NY 10001, 212-465-5796, guptas@pbworld.com

1. INTRODUCTION

Current practice for regional travel models is to treat airports in a simplified manner, considering airports as employment centers (attracting trips to work), along with an additional *special generator* type estimation of non-work (air passengers) travel. Specific features of trip distribution (e.g., a large share of visitors staying in hotels) and mode choice (e.g., higher willingness to pay) are rarely analyzed or explicitly modeled. As a result, regional models usually provide little help in analyzing policies involving changes/improvements to the airports (and their services) and/or ground access transportation.

Also, it is known that the air passenger trips to/from airports have distinctive characteristic with market segments that includes non-residents (visitors), along with willingness-to-pay that is higher than for regular trips. In addition, the modal split is very different from other trip purposes, characterized by a significantly higher share of pick-ups/drop-offs, taxis, and special shuttle bus services. Inaccuracy in modeling travel generated by airports can lead to significant distortions in general regional model results, particularly, for sub-areas adjacent to airports.

For the purpose of this study[1], New York Region provided an ideal example to analyze the air passenger preferences with respect to choice of both airport and ground access mode. Unlike most US cities, air passengers flying to and from the New York region face a wide choice of airports and ground access modes, including transit options. The region has 3 major commercial airports -- John F. Kennedy (JFK), Newark (EWR), and La Guardia (LGA); and 6 smaller airports (SWF, ACY, ABE, HPN, ISP and TTN) with service to domestic destinations only. There are multiple ground access options available in New York to most airports such as auto drop-off/pick-up, park at airport, taxis/limos, shared van services, rental cars, rail, local and chartered buses.

The rest of the paper is organized as follows: Section 2 provides an overview of the findings from prior studies on airport and ground access mode choice modeling. Section 3 discusses the model structure developed in this research, and Section 4 describes the data assembly for model estimation including data sources, sample preparation and sample description. Section 5 discusses the variables considered in model estimation, and Section 6 presents the final empirical results for specified models. Finally, Section 7 identifies limitations and suggests recommendations for future work.

2. LITERATURE REVIEW

A vast literature is available on choice models application in travel demand modeling area. However, comparatively few studies have been done for analyzing the choices made for airports and travels to/from airports. The first studies on airport choice modeling were done using a MNL

[1] This paper is based on analyses done by the authors at PB as part of the *Phase I: FAA Regional Air Service Demand Study (2006)* for the PANYNJ, NYSDOT, and DVRPC. It focuses on a discussion of the technical issues and methods used in the development of a preliminary mode of access and airport choice model for the region. It is not meant to represent or document any specific findings or recommendations of that study.

model for the Baltimore-Washington area (Skinner, 1976) and San Francisco Bay area (Harvey, 1987). These studies found that the flight frequency and ground accessibility significantly impact the airport choice.

Some of the recent studies in US were done for the San Francisco Bay area. Pels et al. (1998) developed a combined airport-airline-access mode choice model for the residents of San Francisco Bay area for business and leisure trips. The airport and access mode are chosen simultaneously at the upper level and airline at the lower level. In line with other studies, they found that access time and cost were significant for the airport-access mode choice. In another study, Pels et al (2001) used a NL model for airport and airline preferences for business and leisure travelers in the SF bay area and found it statistically better than the regular MNL model. Basar and Bhat (2004) tried a different approach where the choice set for each individual may or may not include all airports. The choice set formation is determined based on threshold where an airport is included in the choice set if the consideration utility is greater than threshold utility. Hess and Polak (2005) identified significant heterogeneity for in-vehicle access time, frequency and access cost coefficients; however, the gains in model fit was marginal, meaning that mixed MNL model performs only slightly better as compared to MNL.

3. MODEL STRUCTURE

For the purpose of this study, a nested logit (NL) model with airport choice at the upper level, and mode choice at the lower level, was considered for statistical estimation. The utility for each choice includes airport characteristics (such as travel impedance, attractiveness of the airport, and size variable) and mode characteristics. The utility also constitutes person socio-economic characteristics and trip origin related variables.

The utility V_{ijmq} associated with choosing airport j and access mode m by an individual q in location i can be written as sum of utilities of airport and access mode as shown in (1)

$$V_{ijmq} = \underbrace{\beta_a I_{ij} + \gamma' A_j + \ln(S_{1ijq} + \sum_{k>1} \delta_k S_{kijq})}_{\text{Airport Utility}} + \underbrace{\beta_m^0 + \beta_m^t T_{ij} + \sum_k \beta_{mk}^c C_{ijk} + \sum_p \beta_{mp}^l L_{ijpq}}_{\text{Access Mode Utility}} \qquad (1)$$

where, I_{ij} is the travel impedance between origin location i and airport j, A_j is a vector consisting of attraction attributes for airport j and S_{ijq} is the size variable which makes the airport choice utility non-linear (in parameters). β_m^0 is the access mode specific constant, T_{ij} is the access mode specific travel time, C_{ijk} represent various access mode specific costs and parking cost at airport, L_{ijpq} is the air passenger and trip characteristics (e.g., gender, party size, resident status and others).

The model is estimated using a maximum likelihood procedure in the ALOGIT software package. The nesting coefficient for the adopted structure proved to be slightly large than 1.0; thus the structure was eventually restricted to be joint MNL.

4. DATA SOURCES AND ASSEMBLY

The primary data source for this study is the 2005 originating air passenger survey conducted by the Federal Aviation Administration (FAA), Port Authority of NY and NJ (PANYNJ), New York Department of Transportation (NYSDOT), and the Delaware Valley Regional Planning Commission (DVPRC) in the greater New York region. This survey was carried out at 9 airports in the 54-county region. The survey questionnaire included trip information such as purpose of travel, origin location, destination, mode of transport to airport, size of traveling party and person socio-demographic attributes.

A rich database with 19,127 observations was built based on the survey with 5,812 business travel records, and 13,315 non-business records. It was augmented by the data on the airport characteristics, as well as level-of-service variables for all 9 airports and 8 ground access modes.

Table 1 shows the sample description for the two data sets: business and non-business. The male-female ratio for full sample is close to half-half, but there are more males in business trips and more females in non-business trips. The ratio of residents to visitors is nearly half for both trip purposes. The sample represents similar percentages of income group categories for both travel types except for high income group individuals. There is a higher percentage of high income group people travel on business travel. However, these percentages do not show the actual breakdown because a large percentage (nearly 25%) of individuals refused to tell their income. Sample represents that people travel alone more frequently on business trips (88%) as compared to non-business trips (30%) which is as expected.

The secondary data source is a level-of-service (LOS) file with estimated information for travel times, distance and costs for ground access modes for the 54-county study region to each of the nine airports. Given budget and other limitations in this phase of the study, various methods were adopted, including some simplified ones, to create a set LOS data that could support the estimation of a preliminary model. New York BPM model outputs were used to calculate the travel times and costs ("skims") data for the 28-county region at TAZ (4,000) level, to each of the 9 airports. For the remaining 26 counties, county to airport skims were prepared based on MapQuest and BPM outputs.

The tertiary data source was obtained by Landrum and Brown which included airport characteristics such as number of flights, average gauge, average yield, average airport delay and probability of delay for all the 9 airports. Other information about airport parking rates and available ground access modes were obtained from the official airport websites.

Table 1: Survey Sample Description by Trip Purpose

	Business		Non Business		Total	
	Count	%	Count	%	Count	%
Total	5,812	*100*	13,315	*100*	19,127	*100*
Gender						
Male	3,747	*64%*	5,119	*38%*	8,866	*46%*
Female	2,046	*35%*	8,144	*61%*	10,190	*53%*
Missing	19	*0%*	52	*0%*	71	*0%*

	Business		Non Business		Total	
	Count	%	Count	%	Count	%
Resident	2,853	*49%*	7,143	*54%*	9,996	*52%*
Visitor	2,959	*51%*	6,172	*46%*	9,131	*48%*
Income Groups						
Low (<60K)	701	*12%*	3,590	*27%*	4,291	*22%*
Medium (>60K & <140K)	2,188	*38%*	4,148	*31%*	6,336	*33%*
High (>140K)	1,723	*30%*	1,996	*15%*	3,719	*19%*
Refused	1,200	*21%*	3,581	*27%*	4,781	*25%*
Age Group						
Less than 35 yrs	1,463	*25%*	4,288	*32%*	5,751	*30%*
35 yrs to 55 yrs	3,261	*56%*	4,757	*36%*	8,018	*42%*
Greater than 55 yrs	1,013	*17%*	3,990	*30%*	5,003	*26%*
Unknown	75	*1%*	280	*2%*	355	*2%*
Party Size						
Single Person	5,118	*88%*	3,951	*30%*	9,069	*47%*
Group of 4 or less	2,839	*49%*	8,513	*64%*	11,352	*59%*
Group of 5 or more	134	*2%*	851	*6%*	985	*5%*
Time of Travel						
Peak	2,548	*44%*	5,582	*42%*	8,130	*43%*
Off-Peak	3,264	*56%*	7,733	*58%*	10,997	*57%*

5. ALTERNATIVES, VARIABLES AND SEGMENTATION

Alternatives Specification

Airports: The 9 airports in the 54 county study area are:

1. John F Kennedy (JFK)
2. La Guardia (LGA)
3. Newark Liberty International (EWR)
4. Stewart International (SWF)
5. Long Island Islip Macarthur (ISP)
6. Westchester County (HPN)
7. Atlantic City International (ACY)
8. Lehigh Valley International (ABE)
9. Trenton Mercer (TTN)

Ground Access Modes: For the purpose of this study, some ground access modes were grouped together. There are 8 mode alternatives specified for airport access as shown below:

1. Auto Drop off
2. Auto Park
3. Taxi and Limos
4. Shared Vans, Shared Limos and Hotel Courtesy Vehicles

5. Rental Cars
6. Rail (Subway, PATH, Commuter Rail, Amtrak and others)
7. Local Buses
8. Chartered Buses

There are total of 72 alternatives (9 airports × 8 access modes), out of which 68 were available and 65 were observed in the survey. For example, local bus and shared ride vehicles are not available for Trenton Mercer airport.

Explanatory Variables: Components of Choice Utilities

The explanatory variables considered in the choice model for analysis can be classified into variables for airport choice and mode choice. The following explanatory variables were statistically tested:

Airport Choice

1. Highway distance
2. Average domestic yield ($)
3. Probability of Delay (%) and Average Delay (minutes)
4. Number of domestic[2] and international airports served
5. Dummy for River crossing – the river crossings in the New York region are categorized into three:
    a. Hudson River
    b. East River and Harlem River
    c. Delaware River and others
6. Size Variable
    a. Number of domestic flights per day
    b. Number of international flights per day
    c. Domestic Gauge
    d. International Gauge

Ground Access Mode Choice

1. Travel time – it includes wait time (weighted by 2) for rail and local bus
2. Travel cost including transit fare ($)
3. Daily airport parking rate
4. Number of domestic and international flights per day
5. AirTrain connectivity to airport
6. Manhattan origin dummy
7. Resident or Visitor
8. Gender – male or female
9. Age group - Less than 35 years, 35 to 55 years and above 55 years
10. Income groups
11. Travel party size

[2] All flights to Canada and Mexico are considered domestic.

The first five variables may assume specific numerical values; while the rest are dummy variables, i.e. are categorical classifications (nominal) with value 1 if true and 0 otherwise (e.g., if a person is female then Female =1 else 0).

This study uses a measure of average delay at airports, which addresses the unreliability of flight departure time, in airport choice modeling[3].

No flat airport-specific constants were estimated or calibrated in the airport utility component However, other studies such as Pels et al. (1998) and Hess and Polak (2005) use airport specific constants in their model utilities. Hess and Polak (2005) has also specified cross coefficients to capture the effect of past experience and found it significant. . Since the model is intended for understanding and forecasting the air passenger response with respect to mode of access and airport choice to changes in ground access times and costs, along with changes in airport service, the authors were concerned that inclusion of such variables can lead to potential problems due to endogeneity.

In addition, the previous studies have confirmed the high explanatory power of flight frequency; however, due to unavailability of flight frequency data, this measure was not used in this study.

In all model segments, auto passenger drop-off serves as the reference mode utility component with all constants equal to zero. Therefore, other ground access mode coefficients/constants express attractiveness of the mode relative to auto passenger drop-off.

Segmentation

The different segments considered for the model were

1. *Travel Purpose (Business vs. Non- Business)* - the model is fully segmented by travel purpose since significant behavioral differences were found.
2. *Destination (International vs. Domestic)* - the destination effect is used to partially segment the model. It is difficult to clearly distinguish behavioral differences for destination since some international and domestic destinations are comparable in price or distance.
3. *Traveler (Resident vs. Visitor)* – the model is also partially segmented by residents or visitors. The differences in access mode choices between residents and visitors seems to be dictated by restricted choice sets (e.g., visitor can not auto park) than by fundamental behavior. However, some mode choices for visitors might be affected due to non-familiarity with the place.

6. EMPIRICAL ANALYSIS

The estimated coefficients for the adopted model structure are shown in **Table 2**. The travel time, distance and cost coefficients are generic across all modes. The negative sign on these coefficients shows that an increase in impedance between origin and destination increases the disutility associated with the airport and mode choice. However, the coefficient on distance powered by 1.5 has a positive sign, which shows that marginal disutility associated with distance reduces for longer distances. This does not mean that for very long distances the utility is positive because the net effect of the composite distance term will always be negative within the modeled range of distances.

---

[3] Authors are not aware of any previous study that uses average delay at airports

The estimated value of time (VOT) for business trips (63$/hr) is higher than for non-business trips (42$/hr). These are considerably higher than the VOT values for New York regional model of $15.8/hr (commuting) and $10-$12/hr (non-commuting) in 1997 dollars. Typically, VOT for long distance trips are higher than everyday short trips. The trips to/from airport could be considered as one leg of the long distance travel. Also, the total cost of air travel is much higher than cost of everyday travel; therefore, the travelers might be willing to pay more to reduce the risk of missing a flight. Higher willingness to pay is also reported in other studies, for example in case of business trips, Harvey (1987) reports $41.6/hr, Furuichi & Koppelman (1994) report a value of $72.6/hr and Hess and Polak (2005) report a value in the range of $93- $155/hr.

The river crossing dummy shows a strong negative influence as expected. Crossing the Hudson River that separates New York and New Jersey has a higher disutility as compared to crossing East River/Harlem River. This could be a consequence of more crossings on East River as compared to Hudson River, or due to the fact that the New York City spreads on both sides of East River.

Table 2: Airport and Ground Access Mode Choice Model

Variables	**Business**		**Non Business**	
	Coeff	**t-stat**	**Coeff**	**t-stat**
VOT ($/hr)	63		42	
Impedance				
Distance	-0.064	-8.0	-0.068	-15.3
Distance ^1.5	0.002	3.2	0.003	7.5
Attraction Measures				
Average Yield (Domestic, in $)	-5.368	-3.6	-16.128	-25.0
Average Delay (min)	-0.007	-2.8	-0.005	-4.0
Number of Domestic Airports Served				
Number of International Airports Served	0.002	1.3	0.000	
River Crossing				
Hudson	-1.207	-24.4	-0.918	-29.7
East River/ Harlem River	-0.109	-1.6	-0.008	-0.3
Delaware River	-0.905	-4.1	-0.660	-4.3
Airport Size Variable				
Domestic Flights	1.000		1.000	
International Flights	1.000		1.000	
Domestic Gauge			0.167	1.1
International Gauge				
Mode Specific LOS Variables				
Time	-0.008	-6.4	-0.007	-10.6
Cost/Occupancy	-0.008	-9.7	-0.011	-16.9
	-0.015	-2.5	-0.057	-9.8

Variables	Business		Non Business	
	Coeff	**t-stat**	**Coeff**	**t-stat**
Mode Specific Constants				
Auto Drop-off				
Auto Park	-1.830	-9.7	-1.695	-13.7
Rail	-1.947	-8.9	-1.842	-13.7
Taxi	-1.937	-4.9	-3.069	-11.0
Rental Car	1.145	13.1	-0.879	-6.4
Chartered Bus	-2.688	-10.3	-2.755	-17.8
Shared Van	-2.819	-3.4	-4.850	-7.7
Local Bus	-15.151	-3.0	-19.961	-7.7
AirTrain Present - Rail	2.166	10.6	1.517	14.3
Manhattan Origin - Rental Cars	-2.492	-15.7	-1.889	-12.3
Logarithm -Number of Flights				
Taxi	0.546	8.7	0.605	13.7
Shared Van	0.413	3.1	0.670	6.6
Logarithm -Number of Domestic Flights				
Local Bus	1.970	2.4	3.084	7.3
Resident				
Auto Park	2.177	12.5	1.625	15.8
Rail	-0.988	-8.4	-0.600	-9.0
Taxi	-0.792	-11.1	-0.496	-11.4
Rental Car	-3.433	-20.7	-2.754	-19.7
Chartered Bus	-1.469	-4.8	-1.265	-8.4
Shared Van	-1.884	-12.8	-1.107	-14.3
Local Bus	1.106	2.7	0.340	2.0
International Flight				
Auto Park	-0.609	-3.2	-0.712	-6.1
Rail			0.176	2.2
Taxi	0.176	1.8	0.417	7.5
Rental Car	-0.788	-3.7	-0.396	-2.3
Chartered Bus	0.908	2.8	1.168	8.2
Shared Van	0.259	1.5	0.551	6.4
Local Bus			-0.258	-0.9
Log Likelihood Constants only	-16375.5		-37144.8	
Log Likelihood for Final Model	-12664.7		-30502.3	

Note: the table does not show the impact of socio-demographic variables and group size. However, results are discussed below.

The Manhattan dummy variable proved to be a strong factor in disfavoring car rentals irrespective of trip purpose. This could be attributed to congested highway times, driving inconvenience, and limited and expensive parking available in Manhattan. Also, Manhattan provides a pedestrian friendly environment and dense transit services. These factors are not fully quantified by the LOS variables.

The estimated coefficient for the AirTrain dummy highly favors use of the rail service to airport. In a congested urban area like New York, providing a convenient and reliable rail service to the airport proved to be a successful measure in the past. Significant investments were made to build AirTrains connecting the Newark airport to the NJ Transit rail line (introduced in 2001) and connecting the JFK airport to the subway/LIRR station (opened in 2003). Both AirTrain projects have a visible success in attracting passengers.

Taxi and shared vans are favored to airports with higher number of flights and local bus is favored to airports with higher number of domestic flights. Here, number of flights is a proxy for airport size, thus capturing the affect of frequency and reliability of these services to bigger airports.

Gender bias is also observed for ground access mode choices. Females are less likely to take auto-park and rental cars for both business and non-business trips. This shows that females are less inclined to drive to the airport. For non-business trips, females are also reluctant to take transit options. One reason could be that females are often accompanied by kids on non-business trips which make transit options inconvenient.

International passengers are less likely to auto-park or use a rental car as compared to domestic passengers. Typically, international trips are longer in length which makes the parking costs 4 times higher. Also, international travelers might be uncomfortable with driving in the foreign country which reduces the attractiveness of rental cars. As expected, international travelers prefer taxis, shared limos and chartered buses which could be attributed to convenience and carrying more luggage.

The estimated coefficients show that residents are more likely to choose auto-park and less likely to choose rental cars over other ground access modes as compared to visitors. Typically, residents have availability of a car while visitors are more likely to rent cars.

Ground access mode preferences also vary across income group categories as expected. The estimates show the lower income groups prefer low cost mode options as compared to higher income group and vice versa. Low income (< 60 K) individuals have higher disutility associated with taxis, rental cars or auto-park as compared to higher income individuals. This could be due to non-availability of car and high travel costs for taxi/rental cars. They seem to prefer public transit modes (rail and bus). The trends are similar for both travel purposes. Higher income individuals (> 140 K) are less likely to use shared ride, chartered and local buses, but they are likely to use rail transit particularly for non-business trips.

Younger individuals (less than 35 years old) are found not to auto-park or rent cars as compared to older individuals for business trips. Typically, younger people are at entry/mid level positions in their respective firms. These individuals may have budget restrictions on car rentals/auto-parking options. However, older (55 years or more) do not prefer driving modes (auto park and rental cars) and transit for non-business trips. Younger people prefer transit, taxis and shared rides as compared to older people.

Individuals make different mode choice decisions when traveling in groups. One reason is that the mode costs are shared in non-transit options. On business travel, groups tend to take chartered buses which could also be company provided buses. Groups are less likely to take transit and taxis, particularly for non-business travel and more likely to take chartered buses, rental cars and shared rides. One would expect that most of these non-business travel groups are families (including children), which makes rental car an attractive option.

7. SUMMARY AND AREAS OF FUTURE WORK

This paper presents an analysis of the preliminary development of combined airport and ground access choice model for both business and non-business travelers in the New York metropolitan region. A nested logit (NL) model was conceptualized for this work with airport choice at the upper level and access mode choice at the second level, but a multinomial logit (MNL) model was found be statistically preferred over the NL specification. Significant effects were found across the two segments – business and non-business. The choices also varied across resident and international traveler groups, however with only partial segmentation. Average yield, access time and access cost were found to be significant. Also, air passenger's socio-demographics and party size play an important role in selecting the access modes.

Currently, the model has been applied as a sample enumeration model, meaning it adjusts (or "pivots-off") observed or baseline forecast shares based on changes to either the ground access or the airport measures for a given planning scenario. This initial model has demonstrated the utility of joint choice of airport and ground access mode as a tool for analysis planning. Further improvements could be made to the model with development of additional and more refined airport related measures, such as capacity, service characteristics, and costs as well as refinement in network ground access travel times and costs. Eventually, the air passenger demand model could be incorporated as a special generator in the NY regional model.

REFERENCES

Basar, G. and C.R. Bhat (2004), A Parameterized Consideration Set Model for Airport Choice: An Application to the San Francisco Bay Area, *Transportation Research Part B*, 38(10), 889-904.

Furuichi and F.S. Koppelman (1994), An Analysis of Air Traveler's Departure Airport and Destination Choice Behavior, *Transportation Research* - A, 28 (3), 187- 195.

Harvey, G. (1987), Airport Choice in a Multiple Airport Region, *Transportation Research*, 21A (6), 439-449.

Hess, S. & Polak, J.W. (2005), Accounting for random taste heterogeneity in airport-choice modelling, paper presented at the 84th Annual Meeting of the Transportation Research Board, Washington D.C., January 2005.

Pels, E., Nijkamp, P., Rietveld, P. (1998), "Access to Airports: A case study for the San Francisco Bay Area", 38th Congress of the European Regional Science Association, Vienna, Austria, August 28 – September 1, 1998, Source: http://www.ersa.org/ersaconfs/ersa98/papers/199.pdf.

Pels, E., P. Nijkamp and P. Rietveld (2001), "Airport and Airline Choice in a Multiple Airport Region: an Empirical Analysis for the San Francisco Bay Area", *Regional Studies*, 35(1), 1-9.

Skinner, R. E. Jr. (1976), "Airport choice: An empirical study", *Transportation Engineering Journal*, vol. 102, no. 4, pp. 871-883.