DEMAND FOR AIR TRAVEL BETWEEN NEW YORK CITY AND OTHER LARGE CITIES

Vincent Su and Lucy Huffman, Baruch College, City University of New York

The object of the research is to develop and use an air travel demand model between New York City and other large cities. Demand is organized into a 3-way classification (time period, city destination, and airline) that corresponds to the 3-stage decision made by the traveler. Analysis of variance is used to test the significance of the classification. Each section of the demand model is then estimated by variables determining that part of the travel decision. This model is significant because it breaks down demand to allow an economic analysis of each section, and because more accurate extrapolation forecasting of air travel demand within each classification is possible when each section of the model is estimated.

•A NUMBER of studies that attempted to deal with the air travel problem have been done in the past. In general, they can be put into 2 major categories. The first borrows the form of Newton's law of gravity to set up the hypothesis that travel or communication between 2 cities is proportional to the "mass" of the 2 cities and decreases as the square of the distance between them increases (2, 3). The researcher then chooses the definition of the mass of a city. In general, it is assumed to be represented by the population, income, and some other characterization of the city. Therefore, air travel can be written as a multiplicative function of these determinants, and each determinant has an exponential parameter. After using simple logarithmic transformations, the researcher fits the equation by the least squares technique and estimates all parameters. The second category closely follows conventional demand analysis in which the demand of each individual for air travel is assumed to be a function of his or her income, ticket price, and other relevant factors (4). Some questions arise on these explanatory factors. Should permanent income rather than current income be used? Should price expectation rather than market price be used? When these questions are answered, total demand function can be obtained by linear aggregation. However, demand studies that simply group together all explanatory variables in a regression equation ignore the distinction of variations in demand, and the economic significance obtained from such a model is ambiguous.

THEORETICAL STRUCTURE

The decision to take a particular flight is made in stages. In general, the first decision concerns when to fly, and this depends on when the traveler can afford it. The second decision concerns determining the destination. The destination decision usually is affected by the attraction or convenience of the destination city. Obviously, the larger the city is, the greater the attraction and convenience are. Therefore, this decision is affected by the population, income, or other characteristics of the destination city. The order of the decisions is interchangeable. A passenger may decide on the destination first and then select the time to go or vice versa. The third decision is made after the first 2 decisions are made and involves selecting an airline. This

Publication of this paper sponsored by Committee on Passenger and Freight Transportation Characteristics.

is affected by the convenience provided by different airlines such as number of flights, type of airplanes, and services provided because ticket prices are the same.

The theoretical structure of demand for air travel fits well an analysis of variance over 3 classifications. A 3-way classification analysis of variance technique is first used to break down the total variation in demand for air travel into components in accordance with the 3 decisions discussed. Because the selection of time and destination are independent of each other, the time effect and the city effect are crossed classifications. However, the choice of airline is made after the destination decision is made so the classification of airline is nested in the classification of destination. However, because of the small size of our sample and the lack of randomness in selection of the air carrier routes, an overall airline effect was observed. Thus, the classification of airlines will be treated as crossed in this research. The model can be written as

$$X_{ijk} = \bar{X} + A_i + B_j + C_k + AB_{ij} + AC_{ik} + BC_{jk} + u_{ijk}$$

where

X_{ijk} = demand for air travel between the j th city and New York as experienced by the k th airline in the i th year (throughout this study, the subscripts i, j, and k indicate year, destination city, and airline respectively),
$\bar{X}$ = overall mean,
A_i = time effect of the i th year,
B_j = city effect of the j th city,
C_k = airline effect of the k th airline,
AB_{ij} = interaction of the i th year and the j th city,
AC_{ik} = interaction of the i th year and the k th airline,
BC_{jk} = interaction of the j th city and the k th airline, and
u_{ijk} = overall random error term, assuming the 3-way interaction effect to be zero.

After the analysis of variance has been applied, every significant main effect and cross interaction will be explained by a regression analysis. In each regression, of course, only those explanatory variables that are precisely related to the dependent variables will be used.

DATA

Our interest in this project is the domestic demand for air transportation in and out of the New York City Standard Metropolitan Statistical Area. Thus, the metropolitan areas of other large cities are paired with New York City, and the demand for air travel between these cities and New York City is defined as the number of passengers transported between them. Demand data were obtained from the Civil Aeronautics Board: Domestic Origin-Destination Survey of Airline Passenger Traffic, 1959-1970. The aggregate annual data obtained from a 10 percent sample of tickets is broken down by city of destination and airline. Transfer passengers who are changing airlines at the transfer city and do not consider this transfer city their destination are therefore not included, and the multiple destination passengers who are visiting a number of cities in 1 trip are adjusted for by determining a single destination from the ticket.

The scheduling variable is the average number of daily flights offered by a trunk line to a city within the year. The fact that not all airlines fly to all cities in every time period generates a "missing data" problem in the analysis of variance. In order to avoid this problem we reduced our sample size. We studied only 10 cities, 4 airlines, and 11 years. The annual data taken are from 1959 to 1970. The 10 cities are Boston; Baltimore; Chicago; Philadelphia; Cleveland; Columbus, Ohio; Detroit; Hartford; Washington, D.C.; and Pittsburgh. The 4 airlines are American, Eastern, TWA, and United. Air travel service provided by other airlines not included in this study is grouped in "Other."

EMPIRICAL RESULTS

Analysis of Variance

A straightforward application of a complete 3-way classification analysis of variance to our data indicated that AB_{ij} and AC_{ik} were not significant at all. This finding suggested that these 2 insignificant interactions should be dropped from our original model. The model was therefore modified to

$$X_{ijk} = \bar{X} + A_i + B_j + C_k + BC_{jk} + u_{ijk}$$

The analysis of variance results are given in Table 1. According to the F tests, all 3 major effects and the city-airline interaction were highly significant at the 99 percent level. The percentage of total variation in X_{ijk} explained by this analysis of variance was 88.3 percent.

The individual effects in each classification are given in Table 2. All of them were in the same unit as the dependent variable taken from the origin-destination survey—10 percent of the number of passengers. Each individual effect is measured as the deviation from the class mean, so that the sum of effects in each category is equal to zero. For example, the time effect of 1960 was -5,292, indicating that the demand for air travel in 1960 was lower than the overall mean, which was 12,229, by 5,292.

From the significance tests, certain conclusions can be drawn.

1. As anticipated, the difference among years, the difference among cities, and the difference among airlines all significantly influence the demand for domestic air travel in and out of New York City.
2. The demand for air travel to a certain city in a particular year and the demand for travel by a certain carrier in a particular year are not significant. In other words, passengers do not prefer one city over another in a given year, and the passengers do not prefer one airline over another in a given year.
3. However, the interaction between city and airline is highly significant. This means that as soon as a destination city is determined, passengers prefer to travel on a particular airline.

Regression Analysis

The main effects due to A_i in the 11 years studied were more or less dominated by time trend. This reflects the increasing acceptance of air travel as a mode of travel during this time. Because the desire of a passenger to travel by air in a particular year is affected by income and cost of travel during that year, constant dollar gross national product (GNP) and average price per mile are also used to explain the variation in time effect. The average annual yield per passenger-mile, which is the total industry revenue divided by total passenger-miles, is used as a proxy variable for an index of price. However, a multiple regression including all 3 explanatory variables is impaired by the collinearity between trend and GNP. The final equation used to explain time effect is as follows (T-ratios are given in parentheses):

$$A_i = 2{,}264 - \underset{(-1.40)}{1{,}584.7}\ \mathrm{YIELD}_i + \underset{(11.52)}{1{,}032.1}\ \mathrm{TREND}_i$$

$$R^2 = 0.94 \qquad s = 922.95$$

This coefficient of YIELD simply indicates that, if the price of air travel drops 1 cent per mile, average increase in demand will be 15,847 passengers per year because the data are from a 10 percent sample. In general, the demand will increase by 10,321

Table 1. Analysis of variance.

Source of Variation	Sum of Squares	D.E.	Mean Square	F
Time effect	655,576,300	100	65,557,610	7.17
City effect	649,932,200	9	722,146,500	79.01
Airline effect	171,688,400	4	429,221,000	46.96
City-airline interaction	185,560,300	36	515,445,100	56.40
Error	3,290,333,000	360	9,139,814	

Table 2. Individual effects calculated by analysis.

Variable	Effect	Variable	Effect
Time		City-Airline (contd.)	
1960	-5,292.0	Baltimore-Eastern	-7,308.0
1961	-4,819.6	Boston-Eastern	70,737.8
1962	-3,715.6	Chicago-Eastern	-33,576.9
1963	-1,979.7	Philadelphia-Eastern	-9,448.4
1964	-878.4	Cleveland-Eastern	-18,054.6
1965	367.8	Columbus-Eastern	-12,071.6
1966	1,126.5	Detroit-Eastern	-20,290.6
1967	3,130.4	Hartford-Eastern	-8,824.4
1968	4,225.7	Washington-Eastern	57,130.5
1969	4,820.7	Pittsburgh-Eastern	-18,293.9
1970	3,014.4	Baltimore-TWA	3,691.8
City		Boston-TWA	-24,972.2
Baltimore	-9,021.3	Chicago-TWA	2,455.3
Boston	20,634.0	Philadelphia-TWA	5,256.2
Chicago	12,319.6	Cleveland-TWA	-786.7
Philadelphia	-10,996.3	Columbus-TWA	13,187.3
Cleveland	-3,323.9	Detroit-TWA	-5,043.1
Columbus	-9,313.6	Hartford-TWA	5,691.5
Detroit	-1,098.0	Washington-TWA	-19,700.2
Hartford	-10,012.5	Pittsburgh-TWA	20,220.0
Washington	13,896.1	Baltimore-United	56.5
Pittsburgh	-3,084.0	Boston-United	-30,033.2
Airline		Chicago-United	17,853.7
American	3,532.6	Philadelphia-United	1,839.0
Eastern	9,176.3	Cleveland-United	22,649.0
TWA	-5,832.5	Columbus-United	-55.8
United	-2,506.8	Detroit-United	-169.9
Other	-4,369.6	Hartford-United	490.2
City-Airline		Washington-United	-21,479.3
Baltimore-American	-5,186.4	Pittsburgh-United	8,850.0
Boston-American	-13,479.7	Baltimore-Other	8,746.0
Chicago-American	27,573.1	Boston-Other	-2,252.8
Philadelphia-American	-4,106.3	Chicago-Other	-14,305.2
Cleveland-American	-1,144.4	Philadelphia-Other	6,459.5
Columbus-American	-3,189.9	Cleveland-Other	-2,663.2
Detroit-American	19,246.9	Columbus-Other	2,130.0
Hartford-American	-2,567.3	Detroit-Other	6,256.8
Washington-American	-5,435.1	Hartford-Other	5,210.0
Pittsburgh-American	-11,710.9	Washington-Other	-10,516.1
		Pittsburgh-Other	934.9

persons every year because of the increasing acceptance of air travel and increasing income.

A comparison of the magnitude of city effect in Table 2 and the magnitude of city population suggests that city effect correlates with size of population. In fact, the larger a city is, the more its people travel to a nearby metropolitan city for business and pleasure. Because the analysis of city effect is a cross-section study, the city population (POP_j) used as an explanatory variable is the average of the 11 sample years.

Because business trips are made out of corporate income, and personal trips are made out of disposable income, both of these incomes should be included in the determination of air travel between cities. Unfortunately, these breakdowns in income are not available at the city level. Thus, the per capita city income (YPC_j) is used as the explanatory variable.

Distance between cities was initially included in accordance with the "gravity" theory, which states that the squared distance between 2 cities has a negative effect on transportation between them. Our results do not support the theory, at least over the range of distances in the sample, probably because distance is a proxy for competitive modes of travel over short distances.

Of the 10 cities chosen, Baltimore, Boston, Hartford, Philadelphia, and Washington are roughly within a 200-mile (322-km) range of New York. The demand for air travel over short distances is heavily affected by other competitive transportation modes such as the railroad and private automobile. Indeed, 4 out of these 5 cities lie on the main line of the Amtrak-Penn Central Railroad system. Of course, as travel distance is extended, consideration of other modes of transportation fades away. Thus, a dummy variable of short-distance city (DUM1) is used to signify the difference between the short-distance city and other cities.

The air travel to Boston and Washington from New York is served by the shuttle service of Eastern in addition to other regular flights. The shuttle service, which does not require reservation and check-in and check-out of baggage, is a real convenience factor to be considered in deciding between air travel and alternative modes and makes the air traffic between New York and these 2 cities substantially heavier than would be expected from their proximity to New York. Therefore a dummy variable that indicates that the city pair is served by shuttle service (DUM2) is also included to explain city effect. The final equation used in this study to explain city effect is estimated as

$$B_j = -25{,}141 + \underset{(1.07)}{5.22\ YPC_j} + \underset{(-2.40)}{2.13\ POP_j} - \underset{(-2.40)}{7{,}503\ DUM1} + \underset{(6.62)}{25{,}649\ DUM2}$$

$$R^2 = 0.87 \qquad s = 4{,}192.3$$

All the estimated parameters have the expected sign. The per capita income of the city is related positively to air travel between it and New York. A 1,000-dollar increase in annual per capita income of a city will cause an increase of 52,200 passengers in air travel between it and New York. The low statistical significance of this estimate may be due to the increasing tendency over the last decade of lower-income people to fly. Population growth will also have a positive effect on air travel. An increase in population of 1,000 people will cause 21 additional persons to travel between New York and this city every year.

The airline effects listed in Table 2 are not related to the size of the airline. This is because the cities included in this study are all located in the East-Midwest area. The largest trunk effect is that of Eastern whose business is concentrated in the East although it is only the third largest domestic airline. The effect of United, the largest domestic trunk line, ranks third because these 10 cities are only a fraction of its total market.

The motivation for passengers to choose a particular airline must be the convenience because, on the average, the types of airplanes used are comparable and the prices charged are the same among different airlines. Because just 5 observations exist on airline effect, it need not be explained by a regression.

The interaction between city and airline in Table 2 indicates the market share of each airline on different routes. The reason that the passenger chooses to fly a particular airline to a particular city is the extra convenience provided by the airline. The average number of daily flights (FS_{jk}) scheduled on a particular route (j) by a certain airline (k) is used to indicate this convenience.

The interaction between city and airline, by definition, is measured as the deviation from the means of city effect and airline effect. Therefore, the same process was used to adjust the explanatory variable. We first calculated the average number of scheduled flights of a particular airline (FS_k) and the average number of flights scheduled between New York and a particular city (FS_j). The explanatory variable is the average number of daily flights scheduled between New York and the j th city by the k th airline minus those 2 means ($FS_{jk} - FS_k - FS_j$). The estimated equation is

$$BC_{jk} = 18{,}598 + \underset{(16.14)}{22{,}142} (FS_{jk} - FS_k - FS_j)$$

$$R^2 = 0.71 \quad s = 10{,}044$$

Again, the coefficient is positive, as expected. The equation indicates that an airline's increase of 1 flight per day to a city above the average for that city and airline will cause an increase of 22,142 passengers to that city on that airline. The difference in the significance of the fits suggests that, although the overall airline effect may have contributions from other factors, the attraction of an airline to a specific city is determined almost entirely by its schedule offering.

According to the results of the regression analysis, a summary can be made.

1. The difference in air travel due to time effects can be well explained by the price of air travel and a time trend. Other factors, such as GNP and population growth, also show high correlation with the dependent variable but cannot be included because of statistical difficulties. Perhaps these difficulties will vanish if we can increase the sample size by switching to quarterly data.

2. The difference in air travel due to city effects can be explained well by city income and city population. The distance between cities does not show a significant influence on air travel demand. But the demand for air travel to a long-range city and that to a short-range city are fundamentally different because of competitive transportation modes. The shuttle service also shows a significant impact on the city effect. This results mainly because the shuttle service is more competitive than other transportation modes.

3. The average number of daily flights explained the interaction very well. The shuttle service should also have a significant effect on this interaction, but because we have only 2 out of 40 observations that belong to this group, we did not include a dummy variable to differentiate them from the rest.

ACKNOWLEDGMENT

The research was supported by a grant from the City University of New York Faculty Research Program.

REFERENCES

1. C. A. Bennett and N. L. Franklin. Statistical Analysis in Chemistry and the Chemical Industry. John Wiley and Sons, Inc., New York, 1954.
2. R. E. Quandt. Some Perspectives on Gravity Models. In Studies in Travel Demand, Vol. 1, Mathematica, Princeton, N.J.

3. P. K. Verleger, Jr. Models of the Demand for Air Transportation. Bell Journal of Economics and Management Science, Vol. 3, No. 2, Autumn 1972, pp. 437-457.
4. K. H. Young. A Synthesis of Time Series and Cross-Section Analysis: Demand for Air Transportation Service. Jour. American Statistical Assn., Vol. 67, No. 339, Sept. 1972, pp. 556-566.