Journal of Air Law and Commerce

Volume 25 | Issue 3 Article 7

1958

A Study of Airline Interstation Traffic

Daniel M. Belmont

CURRENT RESEARCH

A STUDY OF AIRLINE INTERSTATION TRAFFIC*

BY DANIEL M. BELMONT

Associate Research Engineer, Institute of Transportation and Traffic Engineering, University of California

THIS report presents a simple formula describing normal airline traffic and yielding estimates that are in reasonable agreement with observed traffic.

DEFINITIONS

The following definitions apply throughout this discussion:

Total traffic (T_i) : the number of domestic airline passengers originating or terminating at a given center, i, in a given period.

Interstation traffic (T_{ij}) : the number of domestic airline passengers making the trip between centers i and j; the common portion of the total traffic of two centers.

Long distance: more than 800 miles.

Medium distance: 400 to 800 miles.

FACTORS INFLUENCING TRAFFIC

Interstation traffic is influenced by a variety of factors, of which the principal ones may be:**

1. Size—the number of potential air travelers at each station. Population, income level, characteristic economic activity, etc., may be included among the components of this complex factor.
2. Cost—of the trip between the stations, in time and in money.
3. Economic bonds between the stations.
4. Service—at the two stations, including location of facilities and relative quality of airline and competitive services.
5. General—such as economic conditions, travel desire, and willingness to fly.

A complete theory of interstation traffic would have to include all such factors and in quantitative form, a formidable task indeed. It happens, however, that on many of the most important routes the interstation traffic depends primarily on a single quantitative variable. This makes possible an approximate theory which has fairly wide direct applicability and which may serve as a useful base for future elaboration.

It will be shown that, on many routes, interstation traffic depends primarily upon the product of the total traffic of the terminal stations. Such routes are here called "normal" routes. The dominant variable is the total traffic of a center. It reflects the effective size of the center; it incorporates also, at their average values for the center, many of the other factors; it is a definite quantity.

* Research Report No. 26, Institute of Transportation and Traffic Engineering.

** For a recent discussion of factors, and for a general method and conclusion different from those of this paper, see Richmond.[1]

[1] Richmond, Samuel B., "Interspacial Relationships Affecting Air Travel," *Land Economics*, February 1957, Vol. 33, No. 1.

TABLE 1—INTERSTATION TRAFFIC ON MAJOR NONSTOP ROUTES
September 17-30, 1955

i	j	Interstation Distance	Total Traffic i	Total Traffic j	Interstation Traffic Observed	Interstation Traffic Estimated*	% Error of Estimate**
Long Distance							
Chicago	Boston	859	175,732	72,175	4,310	4,559	— 5
	Los Angeles	1,751		124,155	7,056	6,409	+10
	New Orleans	828		24,562	945	1,319	—28
	San Francisco	1,856		98,600	4,717	4,785	— 1
	Seattle	1,896		34,761	1,254	1,408	—11
Los Angeles	Dallas	1,245	124,155	37,562	1,169	1,222	— 4
	Kansas City	1,368		29,498	1,044	887	+18
	New York	2,475		327,670	14,931	11,490	+30
	St. Louis	1,596		37,791	1,259	1,111	+13
	Washington	2,331		92,721	2,668	2,706	— 1
New York	Dallas	1,381	327,670	37,562	2,920	3,627	—19
	Denver	1,632		26,822	2,062	2,287	—10
	Houston	1,430		31,573	3,043	2,921	+ 4
	Kansas City	1,104		29,498	2,215	2,998	—26
	Minneapolis	1,020		31,833	3,468	3,385	+ 2
	New Orleans	1,186		24,562	2,926	2,353	+24
	St. Louis	888		37,791	4,539	4,373	+ 4
	San Francisco	2,580		98,600	10,146	8,638	+17
Washington	Dallas	1,183	92,721	37,562	709	888	—20
	Denver	1,509		26,822	773	543	+42
	Houston	1,220		31,573	690	716	— 4
	New Orleans	971		24,562	673	587	+15

TABLE 1—INTERSTATION TRAFFIC ON MAJOR NONSTOP ROUTES
September 17-30, 1955

i	j	Interstation Distance	Total Traffic i	Total Traffic j	Interstation Traffic Observed	Interstation Traffic Estimated*	% Error of Estimate**
San Francisco	Dallas	1,531	98,600	37,562	592	858	—31
	Kansas City	1,507		29,498	641	652	— 2
Dallas	Cincinnati	854	37,562	24,381	235	214	+10
Medium Distance							
Chicago	Atlanta	592	175,732	37,397	1,861	2,469	—25
	Dallas	795		37,562	2,049	2,141	— 4
	Kansas City	405		29,498	3,708	2,188	+69
	New York	724		327,670	27,963	28,500	— 2
	Pittsburgh	420		47,360	3,985	3,743	+ 6
	Washington	600		92,721	5,895	7,071	—17
New York	Atlanta	762	327,670	37,397	3,977	4,602	—14
	Cincinnati	585		24,381	3,385	3,115	+ 9
	Cleveland	417		52,418	8,942	8,728	+ 2
	Columbus	479		17,968	2,381	2,369	+ 1
	Dayton	549		17,691	2,512	2,200	+14
	Detroit	511		76,067	12,730	12,390	+ 3
	Indianapolis	659		19,501	1,981	2,287	—13
	Milwaukee	739		20,511	2,836	2,315	+23
Washington	Boston	399	92,721	72,175	3,691	2,965	+24
	Detroit	415		76,067	2,299	3,102	—26

* by formula (5)

** $\frac{\text{Observed-Estimated}}{\text{Estimated}}$

Restriction of this study to normal routes avoids the difficulty of quantifying the most troublesome factors. This is at the cost, of course, of ignoring some very interesting routes—those on which traffic is strongly influenced by special intercommunity ties or by exceptional quality of air service. There remain for analysis, however, a large number of routes covering very wide ranges of traffic and distance. There is, moreover, reason to suppose that the abnormal routes may best be studied in terms of deviations from the normal pattern.

The cost factor turns out to be rather unimportant for long-distance routes (i.e., differences in cost account for little of the differences in traffic), and of only secondary importance for medium-distance routes. The cost factor is apparently well enough and very conveniently represented by the interstation distance.

REGRESSION ANALYSIS OF 1955 DATA

The travel data used in this analysis are from the domestic section of the Civil Aeronautics Board survey of 1955.[2] The data are for tickets sold but are here taken to represent passengers carried. Chosen for analysis were the long-distance routes between the 25 largest centers, and the medium-distance routes included in the 100 routes ranking highest in interstation traffic. From these, certain routes were excluded for reasons discussed below. Remaining as the basis for mathematical analysis were 41 routes involving 23 centers. These are listed in Table 1.

Shown for each station-pair are the interstation distance, the total traffic of each station, and the interstation traffic. These data were used to find a formula relating interstation traffic to total traffic and interstation distance. The resulting estimates of interstation traffic are shown in Table 1, together with their relative deviations from the actual interstation traffic. These results are shown also in Fig. 1.

In a previous paper,[3] based on 1954 data, it was found that, on long-distance nonstop routes, interstation traffic tended to be independent of distance and to be roughly of the form

$$T_{ij} = k\, T_i T_j \qquad (1)$$

where T_{ij} is the interstation traffic, T_i the total traffic at i, and T_j the total traffic at j. This formula showed enough agreement with the data to suggest that for such routes, an expression could be found of the form

$$T_{ij} = k\, (T_i T_j)^p \qquad (2)$$

which would represent the data with useful adequacy.

The analysis of the 1955 data covered medium-distance routes as well as long-distance routes, and it was found that distance was here a significant factor. To include the effect of distance, D, formula (2) was generalized to

$$T_{ij} = \frac{k\, (T_i T_j)^p}{D^q} \qquad (3)$$

In fitting the formula to the data, it was convenient to use the logarithmic equivalent

$$\log T_{ij} = c + p \log (T_i T_j) - q \log D \qquad (4)$$

This form has the important advantage also of permitting the minimizing of relative rather than absolute deviations.

---

[2] Civil Aeronautics Board, *Origin-Destination Airline Revenue Passenger Survey, Sept. 17-30, 1955.*

[3] Belmont, Daniel M., "A Pattern of Interstation Air Travel," *Transactions,* American Society of Civil Engineers, 1957.

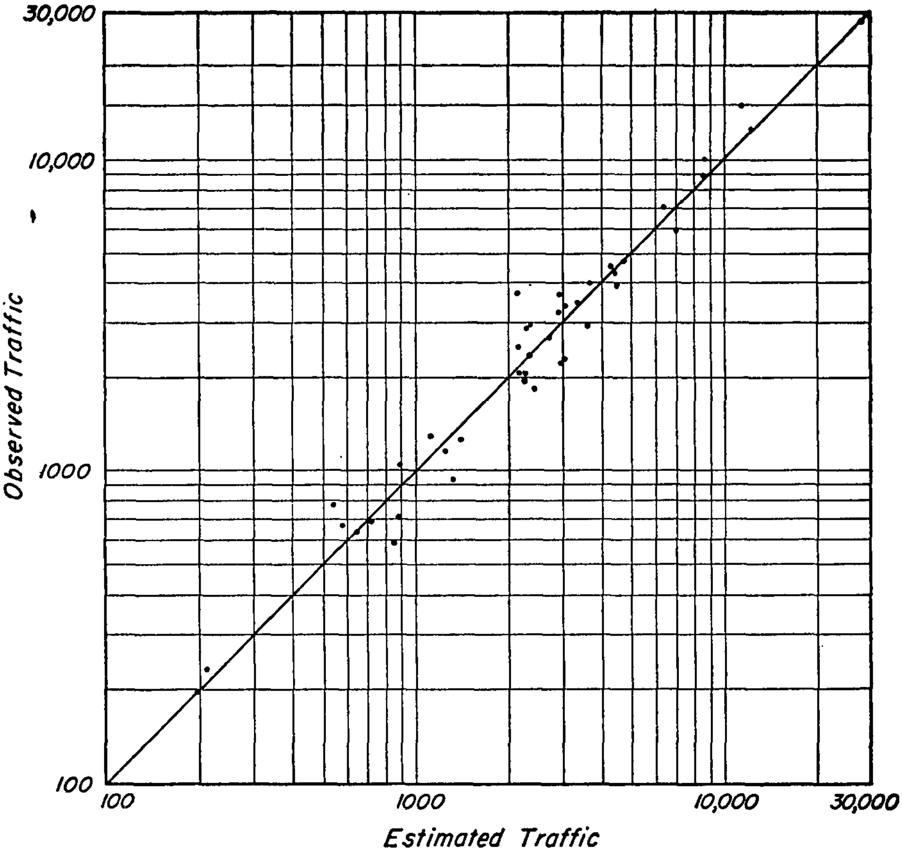

Fig. 1-OBSERVED vs ESTIMATED INTERSTATION TRAFFIC
41 NONSTOP ROUTES, SEPT. 17-30, 1955.

Formula (4) was fitted to the 41 routes of Table 1 by the usual method of least squares to obtain

$$\log T_{ij} = -6.90755 + 1.164780 \log (T_i T_j) - .40950 \log D \qquad (5)$$

with the correlation index $R^2 = .962$.

The estimates of interstation traffic generated by this formula are those shown in Table 1 and in Fig. 1. Their average error is 14.7%. In view of the great range of interstation traffic (200 to 28,000) and of distance (400 to 2,600 miles), the substantial differences in quality of air service, the existence of special economic ties on some routes, and the inevitable random traffic fluctuations, it is remarkable that the formula yields such accurate estimates.

For the routes analyzed, the size factor (the total traffic of a center) is dominant in determining interstation traffic; distance is of significance only for the medium distance routes. (A statistical examination of the significance of the distance factor is given in the Appendix.) All other factors combined account for a comparatively insignificant part of the variation in interstation traffic.

It may be worthwhile to emphasize the extraordinarily simple primary structure of long-distance interstation traffic. The long-distance routes listed in Table 1 may be described by formula (2):

$$T_{ij} = k\,(T_i T_j)^p$$

This formula, in its best-fitting logarithmic form, has the correlation index $R^2 = .964$, indicating that it accounts for practically all the observed variation in interstation traffic. If we write

$$C_i = k^{1/2}\,(T_i)^p$$

we obtain $$T_{ij} = C_i C_j$$

exhibiting the structure of long-distance interstation traffic in its simplest form. The formula asserts that each station has a characteristic traffic index, C_i, such that normal long-distance interstation traffic is just the product of the traffic indices of the terminal stations.

There is a serious shortcoming in the formula: it is dimensionally unpleasant. The value of k varies with the period in question and with changes in the relative values of the T_i's in not too obvious a manner. A closer examination of the index C_i and a discussion of its use in forecasting is proposed for a future paper.

Routes Excluded from Analysis

The routes of Table 1 were selected from the major domestic routes to meet the following conditions:

1. *Nonstop service.* The traffic figures used included also traffic with one or more stops, but the existence of at least some nonstop flights establishes a reasonably satisfactory minimum service level, preventing too wide a spread in the quality of air service offered on the several routes. Substantial differences do occur in frequency and convenience of flights, and these no doubt account for some of the differences between observed and estimated traffic.
2. *Interstation distance of at least 400 miles.* The 400-mile limit may not be critical. At shorter distances, however, it is known that special community of interest between centers becomes more probable and that the effects of competing modes of transport complicate the relationship between distance and air travel.
3. *Normal interstation traffic.* The concept of normalcy may best be explained negatively. Certain routes, described below, were expected to carry traffic well above the normal pattern. The remaining routes were considered to be normal. It was not known in advance how well these normal routes could be described by formula (3). (The two "normal" routes most poorly fitted by the formula, Chicago-Kansas City and Washington-Denver, might reasonably have been considered abnormal. Their exclusion would not seriously affect the parameters of the best-fitting formula.)

The excluded routes and their traffic as observed and as estimated by formula (5) are:

Route	*Observed Traffic*	*Estimated Traffic*
Florida-North, such as:		
Miami-Chicago	5,409	2,895
Minneapolis-Seattle	707	218
Philadelphia-Chicago	4,474	2,784
Philadelphia-Detroit	3,691	1,189
Western routes under 1200 miles such as:		
San Francisco-Portland	3,529	619
Los Angeles-Denver	2,175	965

The Florida routes were obvious candidates for exclusion; traffic is abnormally heavy to the north and abnormally light to the west.

The western routes under 1200 miles were omitted as a result of the previous study.[2] Western distances affect air traffic differently from eastern distances, largely because in the west neighboring cities with strong community of interest may be very many miles apart. Minneapolis and Seattle may be considered an extreme example of such neighbors. Only after the distance factor becomes negligible (usually beyond 1200 miles) do the western cities fall into the eastern pattern.

Philadelphia is a special case. Its traffic with Chicago and Detroit is perhaps not exceptionally heavy, but its total traffic is exceptionally small. Consequently it should be expected to be fit well into formula (5). The abnormality derives from Philadelphia's closeness to New York, as a result of which Philadelphia's total traffic receives almost no contribution from the New York route. If these cities were farther apart the increase in Philadelphia's total traffic might reasonably be great enough to bring the formula's estimate of traffic with Chicago and with Detroit into at least fair agreement with the observed figures. (The Detroit traffic would remain, however, much the heavier compared with the estimated traffic.)

DISTANCE AND AIR TRAVEL

The following considerations may help to explain why the distance factor has an insignificant effect on normal long-distance airline traffic (and only a relatively small effect on medium-distance traffic).

For short trips, there is a well-known and marked tendency for trip frequency to decline with trip length. The chief reasons for this appear to be that, for distances up to a few hundred miles:

1. Economic ties tend to be stronger between nearby centers than between more distant ones.
2. The concentration of personal acquaintances tends to diminish with distance.
3. Short trips are frequently undertaken for purposes for which longer trips would be rejected.
4. The closest destination is usually chosen for the many trips whose purpose could be equally well served at several destinations.

None of these factors apply with much force to sufficiently long-distance domestic airline travel on well-served routes. For distances greater than a thousand miles or so:

1. Business trips are motivated primarily by the size of the market rather than by its location.
2. The number of one's friends is more likely to depend upon the size of a community than upon its distance.

(There may of course be special community of interest between centers very widely separated. But such special ties are exceptional at long distances; they are the rule only at short distance.)

3. Long trips are comparatively seldom undertaken by air when still longer trips would be rejected. The difference in time is usually negligible. The higher cost of the longer trips is largely counter-balanced in its effect upon trip frequency by the greater advantage of travel by air over travel by competing modes.
4. Long-distance trips are comparatively seldom determined by the relative distance of otherwise equivalent destinations. Long trips are

seldom required for purposes which could be served equally well at alternative destinations.

CONCLUSIONS

A very simple model has been presented which usefully represents normal interstation traffic. It should serve as a guide in forecasting and perhaps as a foundation for a theory broad enough to cover also the abnormal cases. There is a tendency, of course, for "normal" to mean those cases which fit the theory. But even if there were here no better definition of "normal," the number and variety of routes to which the theory does apply would justify its presentation.

It appears that special community of interest between centers has important influence on interstation traffic only in exceptional cases. On most of the routes here considered, the terminal stations have, so to speak, an agreeable anonymity. Only their traffic indices, and perhaps their interstation distance, play a significant role in determining the interstation traffic.

APPENDIX

Formulae

$$\log T_{ij} = \log k\,(T_iT_j)^p = c + p \log (T_iT_j) \qquad \text{(a)}$$

and

$$\log T_{ij} = \log \frac{k\,(T_iT_j)^p}{D^q} = c + p \log (T_iT_j) - q \log D \qquad \text{(b)}$$

were fitted to the data of Table 1 by the usual method of least squares to obtain the following results:

	Long-distance routes	*Medium-distance routes*	*Combined routes*
Formula (a):			
n	25	16	41
c	—7.61881	—7.06679	—7.94109
p	1.106220	1.068689	1.145786
R_a^2	.9640	.8878	.9114
Formula (b):			
n	25	16	41
c	—7.30565	—5.78927	—6.90755
p	1.152823	1.122260	1.164780
q	.24592	.66133	.40950
R_b^2	.9687	.9326	.9619
F-test:			
$R_b^2 - R_a^2$	not significant at 5%	significant at 5%	significant at 0.5%

where: n is the number of routes analyzed,

c, p, and q are the constants of best fit for the formulae,

R^2 is the associated correlation index, and

F-test shows the significance of the improvement of R_b^2 over R_a^2.

It is seen that the distance factor is highly significant for the combined routes, but that its effect is largely limited to the medium-distance routes.

For the combined routes, the use of the single formula (b) was compared with the use of separate formulae of type (b) for the long- and medium-distance routes respectively. The gain in R^2 for the latter was far below significance at the 5% level.