Introducing the Idea of the "K Distribution" to Transportation Patterns

YACOV ZAHAVI, Kolin and Zahavi, Transportation and Traffic Engineers, Tel-Aviv, Israel

The "Distribution of Factor K" is an empirical distribution that has been found to operate in many research fields. The application of this distribution by the transportation engineer may contribute considerably to the definition of some important aspects of transportation. This empirical law should help solve, for instance, the problem of defining an index of the pattern of concentration of traffic, and of the relationship between the traffic-attraction powers of the various traffic zones in a city. Another problem is that of defining the scope of the significant data essential to planning. The transportation planner often is confronted with an abundance of data relating to numerous zones. A great deal of time and expense could be saved were it possible to confine the scope and limit to the most significant data. It is the object of this paper, then, to present a possible approach to these and other questions by applying an extended form of the "K Distribution" to some transport behaviors.

• IN VARIOUS RESEARCHES and in different spheres, the following empirical phenomenon has been noted. When ranking certain observations in a decreasing order of their value, where the rank of the observation is plotted along the x-axis and its value along the y-axis, a straight line is formed on a double logarithmic scale. This straight line can be formulated as follows:

$$r = KP^{-q}$$

where

r = the rank of the observations,
P = the value of the observations,
q = an exponent, the value of which nears 1 as the slope increases, and
K = a constant factor, the value of which is close to the highest observation.

The formation of such a straight line means that the values of the observations decrease in a regularity that is bound to a certain relationship between the larger observations and the smaller ones. Values decrease sharply at the beginning; later, the decrease becomes smaller and smaller. Such a hyperbola-like distribution provides the points on the top of the straight line with much greater weight than the ones at the bottom of the line.

The slope of the line expresses the degree of the relationship between the observations, or the level of differentiation in their weights of activity. A moderate slope, while still possessing its hyperbola-like distribution, denotes a smaller such differentiation.

Paper sponsored by Department of Urban Transportation Planning.

On the log-log scale, some of the observations are situated at the margin of the straight line; their values are low and they do not continue the trend of the straight line. Thus, a breaking point of the straight line is located that terminates the above described specific regularity of the observations.

The K distribution (KD) thus briefly described has been observed in a number of demographic, sociologic, and biologic behaviors. Certain alternative explanations have been given to it, like those offered by Zipf, Rashevsky, Simon, and others. This paper is not going to offer another one but, as stated, is going to apply its empirical operation for practical uses in some transportation issues.

Before this application is discussed, it would be useful to note the operation of the KD in some other fields and to evaluate its meaning concerning those specific variables.

The KD has been repeatedly noted in the size distribution of settlements. Table 1 and Figure 1 show these distributions for the United States and Israel (1). In both cases a straight line is formed and broken at the level of the smaller settlements. The KD is applicable to the settlements for different countries and thus has served as a basis for some significant demographic conclusions.

1. The straight line indicates that the settlements situated along it are developing in a specific regularity, different from the regularity governing the rest of the settlements. Their varying size is related to a specific order so that their power to attract new inhabitants is in proportion to their relative size. Figure 1 shows that this regularity is common to both very large and very small countries.

2. The slope of the line defines the "level of urbanization" of a country. When the slope is steep, like that for the United States and Israel, a major part of the population is concentrated in urban settlements. A moderate slope, on the other hand, signifies a dispersion of the population and expresses a less extreme differentiation between the settlements and thus a low level of urbanization. These moderate slopes are typical of agricultural or nomadic countries. Another characteristic of the slope is that the steeper it gets, the fewer the settlements along the straight line and the more below the breaking point.

3. Demographers have defined settlements along the straight line as "urban settlements." Thus, the breaking point of the line separates the larger settlements ranked as urban settlements of the country, from the settlements which, at that time, exert a weaker power of attraction and constitute a smaller proportion of the population. The point at which a settlement becomes urban lies on different levels in different countries. In the United States or Israel, for instance, the breaking point is at a high level; a settlement requires some 10,000 inhabitants (Fig. 1) to be urban. Within the hierarchy, however, there are countries in which the settlements become urban at a much lower

TABLE 1

POPULATION IN LOCALITIES BY SIZE-CLASS

Size of Locality (number of inhabitants)	United States of America 1.4.1960		Israel 22.5.1961	
	Number of Localities	Population	Number of Localities	Population
Total		179,323,175		2,179,491
In localities	19,790	125,808,073	873	2,148,310
500,000 and over	21	28,595,050	—	—
100,000-499,999	111	22,418,307	3	736,526
50,000-99,999	201	13,835,902	2	144,841
20,000-49,999	632	19,400,682	15	471,048
10,000-19,999	934	13,118,216	14	175,311
5,000-9,999	1,394	9,779,714	20	145,026
2,000-4,999	3,048	9,577,903	50	164,797
1,000-1,999	3,575	5,049,869	51	75,983
500-999	3,267	2,341,061	121	75,932
200-499	4,153	1,389,190	411	137,862
Under 200	2,454	302,179	186	20,984
In others		53,515,102		31,181

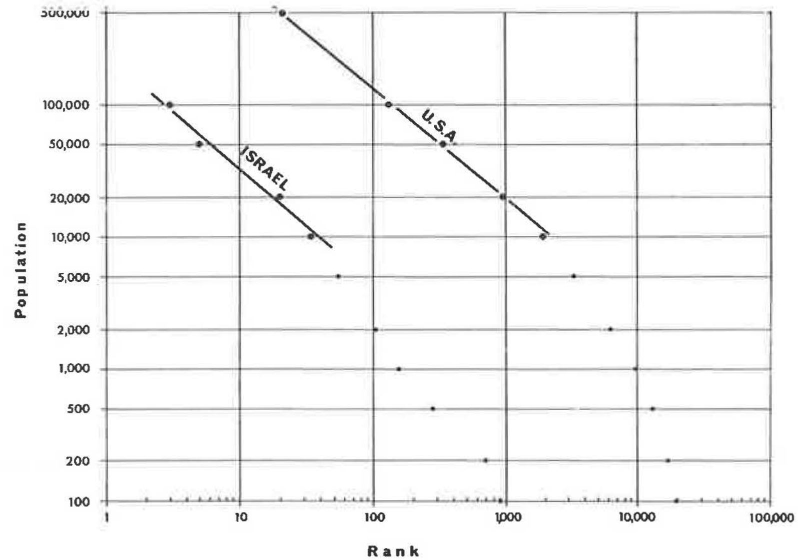

Figure 1. Distribution of settlements in the United States of American and Israel.

level. The breaking point marks the level above which settlements can be considered as dominant ones, while those beneath it can be considered as minor within the distribution.

Conclusions of this kind, which have been reached upon applying the KD to demographic studies, have been verified in other fields as well. To give another example, the KD can be found even in some quite unexpected spheres, as in the frequency distribution of the alphabetic letters in the Bible. In the 10th century, Rabbi Saadya Gaon calculated the frequency at which the different letters appear in the Old Testament, as given in Table 2. Plotting these frequencies by the method just discussed forms once again a straight line with the typical breaking point, as shown by Figure 2. Similar conclusions can be drawn from the distribution of the letters of the alphabet regarding the concentration of the use of the various letters, the regularity of their appearance, the dominant nature of certain letters, etc.

It is possible, therefore, that some important findings might emerge concerning a variable known to be K distributed. Consequently, the KD will be of special interest to transportation engineers, if it turns out that it operates for certain transportation distributions as well. Our findings show that this happens to be true.

Investigation of a series of data concerning trip-ends in different towns of various sizes showed that their distributions plotted on a log-log scale assume the typical form of the KD with surprising accuracy. From a large number of similar cases, three examples of such distributions are shown by Figure 3.

Figure 3 shows the manner in which the attracted trip-ends are distributed in Chicago (2), Pittsburgh (3), and Tel-Aviv (4). Typical to a KD, the main traffic zones or districts, which attract most of the trips, are situated around a straight line in all three towns. In all these cases, three

TABLE 2

THE FREQUENCY OF APPEARANCE OF ALPHABET LETTERS IN THE BIBLE

Letter	Frequency	Rank	Letter	Frequency	Rank
Alef	42,377	7	Mem	77,778	1
Beth	38,218	10	Nun	41,696	8
Gimel	29,537	13	Sameh	13,580	21
Daleth	32,530	11	Ayin	20,175	20
He'h	47,754	6	Pe'h	22,725	17
Vav	76,922	2	Zadik	21,822	19
Zayin	22,867	16	Kof	22,972	15
Heth	23,447	14	Resh	22,197	18
Teth	11,052	22	Sheen	32,148	12
Yod	66,420	3	Taf	59,343	4
Kaph	48,253	5	Total	815,330	22
Lamed	41,517	9			

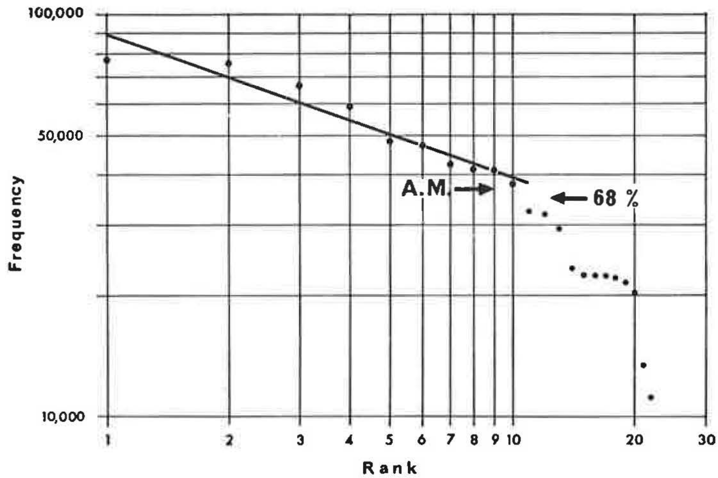

Figure 2. Frequency distribution of alphabet letters in the Bible.

different slopes can be observed. At relatively low levels of trip attractions a breaking point occurs in all three towns, though the number of trip-ends it signifies is different, of course, for each.

Examining the distribution of traffic volumes on the road network by the hours of the day showed that they too are K distributed. In Figure 4, public transport passenger trips in Tel-Aviv (4) are plotted on a log-log scale. Once again, characteristic to a KD, the hourly trips form a straight line with a specific slope, which is clearly broken at a low level of passenger trips.

These are only two examples of traffic distributions that can be considered to be K distributed. The deductions that the traffic engineer may arrive at by applying the characteristics of the KD to the attracted trip-ends distributions and to the distribution of traffic volumes by the hour cannot be overestimated. Before these deductions are dealt with, however, one characteristic of the KD, i.e., the breaking point, merits some special attention.

THE DOMINANT OBSERVATIONS

Whereas the empirical KD has been recognized and dealt with in numerous publications, no definition or location of the breaking point has been offered, and its significance has not yet been fully appreciated.

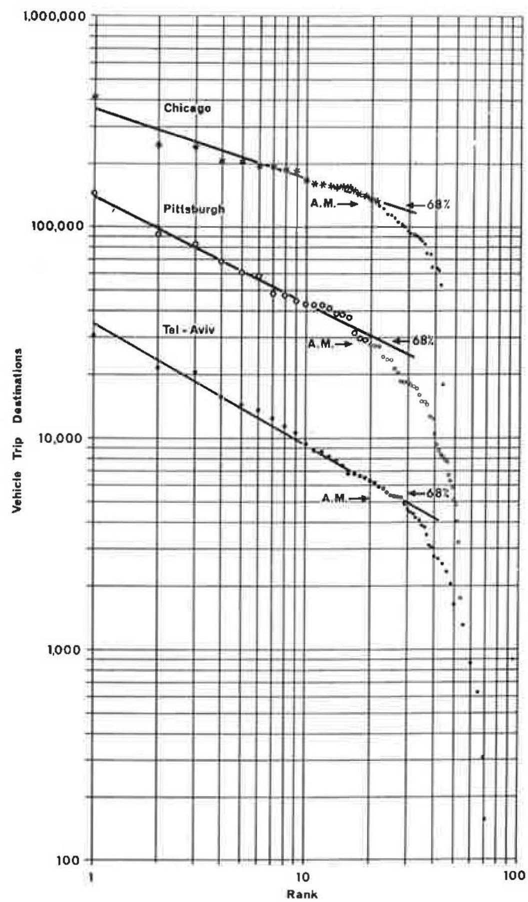

Figure 3. Vehicle trip distribution in Chicago, Pittsburgh, and Tel-Aviv.

In the framework of the Transportation Master Plan for Tel-Aviv (4), in which the operation of the KD in certain traffic distributions was found, the essential nature of the breaking point was thoroughly investigated. Although different sorts of trip-end distributions of various towns were examined, an interesting possibility emerged for a clear definition of the characteristics of the breaking point. This has been found to be true for all the K-distributed variables examined. Comparisons made to this end marked two different parameters by which the breaking point may be defined: (1) by the arithmetic mean of the observations; and (2) by the sum of observations, the value of which is close to 68 percent of the total.

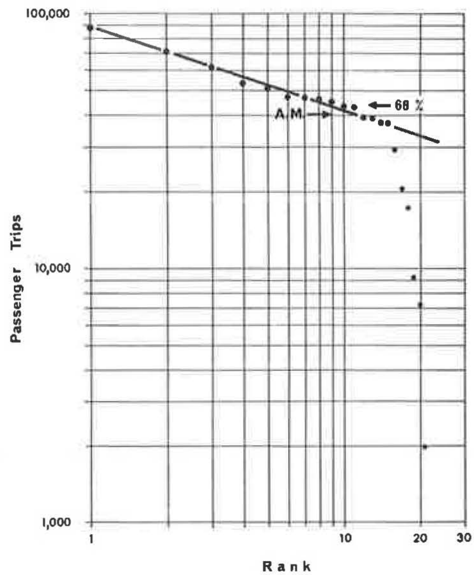

Figure 4. Distribution of bus passenger trips in Tel-Aviv, 1965.

1. It has been found that summing up the values of all observations and dividing their total by the number of observations, an average value is obtained which, surprisingly enough, marks the breaking point of the straight line. The average size of the settlements, for instance, has been found to indicate quite closely the point at which the settlements cease to follow the initial trend of the straight line. The traffic zone that attracts the average number of trips, as another example, marks in most cases the point at which the straight-line trend of the traffic zones that attract most of the trips is broken. This average stratifies the zones, then, into the dominant ones and others. Figures 2 through 4 show this relationship between the breaking point and the arithmetic mean (AM) of the observations. This established an empirical correlation that appears to be incidental but may well be reasonable and logical. One of the explanations describes the observations as being distributed around an average condition in a steady-state environment. The observations above this average are, therefore, more active than those below it. Moreover, the empirical distribution shows that the interrelationship characterizing the dominant zones, i.e., those above the average, is entirely different from that characterizing the observations below the average.

2. In most of the KD it has been noticed repeatedly that the value of the observations above the breaking point amounts to about 68 percent of the total value of the observations. This percentage reminds us of the observations that deviate from the mean, in a normal distribution, by one standard deviation. This might indicate that those observations that are dominant in character range by only an average deviation from the highest observation. This recurrent relationship seems, too, to be casual; essentially, however, the relationship demonstrates that the dominant and the active observations (the active towns or traffic zones) are those close in character to the highest and most active observation (the main town or traffic zone). Figuratively, by borrowing the statistical normal curve concepts, we can say that in a KD only those units having a total value that does not exceed a normal deviation are along the straight line and above the critical point.

It should be remembered that the KD itself is empirical and does not operate in all cases with mathematical precision. Therefore, the above definition of the breaking point should not be expected to pinpoint its exact location, but serves to define it with sufficient accuracy for actual planning purposes. In most of the K distributions examined, this relationship has been found to exist with surprising accuracy. Figures 2 through 4 show this relationship.

The two empirical factors somewhat extend the accepted notion of the KD and provide a better position from which to elaborate the significance of applying the KD to some transportation patterns.

THE K DISTRIBUTION—A DEVICE FOR TRANSPORTATION PLANNING

This section details some of the significances arising from applying the KD to some transportation distributions. The extended concept of the KD developed in the previous chapter has some new bearings on other fields, too, especially the demographic field. This section, however, will deal only with its significance concerning transportation issues.

One of the most important transportation distributions is that of the spatial dispersion of the produced and attracted trip-ends between the various traffic zones in a city or a metropolitan area. The significances of this application are discussed as follows, especially in connection with these distributions, although the significances are general in character.

Levels of Traffic Concentration

A major parameter characterizing the travel pattern of a city is its level of traffic concentration. If the distribution of trip-ends is said to be K distributed, then the slope of its straight line may serve as an indication of the traffic concentration in that city. A steep slope indicates that a great number of trips are attracted to some selected active zones whereas the rest of the zones attract considerably fewer trips. The steep slope is an expression of a concentrated pattern of trips or of a high level of "traffic urbanization" in that city. A moderate slope, or a low level of traffic urbanization, indicates that each of the many zones attracts relatively few trips and that there are no outstanding strong zones of destination.

Obviously, planning for such different conditions of travel will take quite different directions. The slope of the line, therefore, may assist in determining the specific planning approach to a specific travel pattern.

Development Trends

The slope of the straight line may be compared between periods of time. Such comparisons should point out the direction of the development trends of traffic urbanization of a metropolitan area. Before conducting such comparisons, however, we should review the essential nature and characteristics of the unique variations in the slopes of K distributions.

In early stages of a country's development, as is the case in agricultural or nomadic countries, the slope is quite moderate. In the next stage of development, with an outburst of growth of some settlements to the level of big towns, the slope becomes increasingly steeper and the level of urbanization increases accordingly. Surprisingly enough, in the next phase of development, the slope of the line may get moderate again. The reason is immediately apparent: At that stage the main growth continues in the rest of the satellite metropolitan towns because the central city has already approached a certain saturation level with high population density. The level of concentration, therefore, is decreasing; thus the slope returns to a more moderate form. An example of such a development is the growth that is taking place in the New York-Baltimore belt.

Parallel variations in the slopes are observed in the distribution of trips to zones of destination in a metropolitan area. At the beginning, the slope is moderate because the distribution of the trips is rather dispersed. The level of trip concentration increases as the main town of a metropolitan area develops and creates some major zones of destination. At that stage the metropolitan area is said to have a high level of traffic urbanization. At the next stage of variation in the travel pattern, a process of trip dispersion sets in again: The other satellite metropolitan towns start to attract an increasingly larger number of trips; but the main city, after reaching a certain level of saturation, does not develop at such a pace.

Bearing in mind the nature of the variations of the slope in a KD, we can gain considerable knowledge concerning the trends of traffic urbanization in a certain area by comparing its slopes for different periods of time.

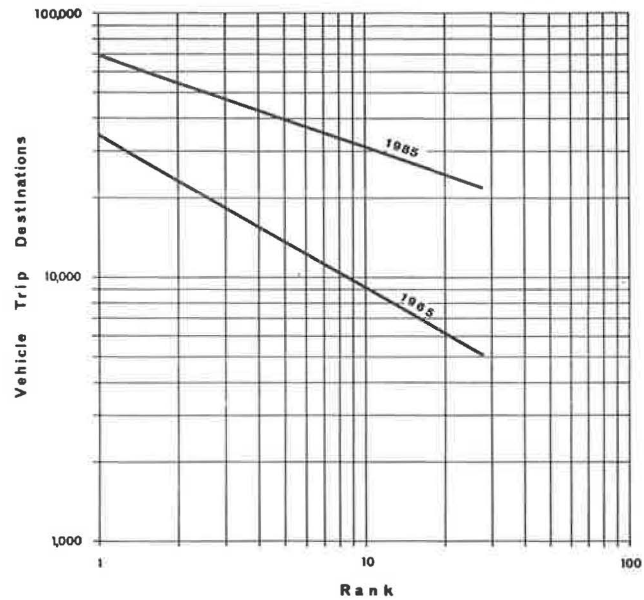

Figure 5. Vehicle trip distribution in Tel-Aviv, 1965 and 1985.

Figure 5 shows the trend of concentration in the distribution of trip-ends in the Tel-Aviv metropolitan area. The figure shows the distribution for 1965 and the one forecast for 1985. The change in the slopes is clearly noticeable. An examination of the meaning of the slope's movements leads to a clear conclusion concerning that change. The level of traffic urbanization, or traffic concentration, in the Tel-Aviv area is apt to decrease. The zones at the outskirts of the main city are going to develop considerably in a process of differential growth so that the dispersion of population, land uses, and trips will increase. A marked change in the weight of trip attraction will occur between the central and the satellite towns. This trend should have some significant bearings on future plans for the road network in the area.

It should be stressed at this point that for the purpose of comparisons between different towns or time periods, an equivalent definition of traffic zones should be maintained in each town or period. Different criteria for defining the zones may produce different slopes for the same area. Comparisons such as those shown in Figure 5 produce the best results because they refer to exactly the same zones in both periods.

Comparisons of Distributions

The slope of the straight line may be compared between different distributions as well. The comparison, for instance, between the attracted trip-ends for private vehicles and for public transport might produce some interesting results. The distributions in Tel-Aviv for 1965 and 1985 respectively were compared. The results were as follows:

1. The slope of the line for the public transport distribution was considerably steeper; the exponent of the slope had a value of 0.67, whereas the exponent for the private vehicles for the same year was 0.57.
2. With the general decrease in concentration by 1985, both distributions became less concentrated, but to a lesser extent for the public transport. The exponents for 1985 were 0.46 for public transport and 0.35 for private vehicles.

One of the most important conclusions to be drawn from these results is that a quite significant difference exists between the two distributions. The concentration for transit trips is much more notable than for private vehicles; this trend increases relatively over the years, despite the general trend toward a greater dispersion of trips.

Upon applying models of trip distribution, by the way, it would be most desirable and useful to employ two separate models, one for each mode. This is warranted by the different pattern of centralization of the two modes. The method of employing different models for the different modes seems vastly preferable to the method that applies a single model for the total person-trips, at the risk of having to split it into two spatial ones, i.e., one for the central business district and one for the rest of the metropolitan area. The need for different models is more acute for towns with a high percentage of public trips. Separate examination of the trip distributions for the two different modes, in addition to an investigation of their trends by means of the KD, therefore, should provide a better understanding of the interrelationship between the two modes and should introduce a more reliable method of forecasting the modal-split ratios.

Variations in Traffic Zones

Examination of the traffic zones along the straight line may indicate, in addition to the general trend of trip concentration in a city, the variations in individual traffic zones as well. Traffic zones situated along the straight line may drop below the breaking point, and vice versa. This will be true with some traffic zones in Tel-Aviv by 1985. Such developments should have been expected because, as has been pointed out already, changes occur with time in the relative weight of zones. A change in the location of certain zones along the straight line indicates such a change in their power of trip attraction. A knowledge of the regularity with which progression or regression of the zones located along the straight line occur together with the knowledge of the direction and strength of the changes in the location of some of them may contribute considerably to the understanding of the process of relative change affecting the zones with the lapse of time.

Locations of Dominant Traffic Zones

One of the most useful consequences of applying the KD is that it facilitates the location of the breaking point in trip-ends and other transportation distributions. As has been seen, the observations plotted along the straight line were those of above the average value, those constituting a greater proportion (about 68 percent) of the total trip-ends. This makes possible an immediate and simple definition of the "dominant traffic zone." This definition can serve as a sufficient basis for a great many transportation plans; other zones may be disregarded because their data are not essential for the planning process.

The location of the breaking point in the trip-ends distribution defines those zones, the data of which would suffice for adequate planning of a public transport network, for example. It has always been clear that for a project of this sort it is sufficient to rely on the data relating to zones of a high weight of importance: The road network based on them will cover the needs of the other, less active zones; and much planning time and expense will be saved.

In the framework of the Transportation Master Plan for Tel-Aviv, the public transport network indeed was based primarily on dominant zones, defined as such on the basis of the trip-ends distribution. When travel desires were assigned to the road network, the bus lines thus planned met about 90 percent of the direct travel desired in the city, with no transfers necessary. Thus, the ability to define the dominant zones in transportation distributions is, perhaps, one of the most practical results of the application of the KD.

Changes in Rank Size of Zones

Comparisons of the straight lines between periods of time might show that some interchangeability in rank size is taking place between certain zones: Some of the dominant zones are liable to relinquish their status with the passage of time, and vice versa. By considering these changes, planners might better allocate the limited economic means to the most important ends. In zones expected to lose their dominance, investments in transportation devices should be restricted to short term ones. On the other hand, in zones apt to become dominant, long term investments should be considered, even though present conditions do not warrant such investments.

Comparisons of Characteristics

Another noteworthy aspect is the possibility of comparing the characteristics of different K distributions. The significance of comparing the slopes of the lines in different distributions already has been pointed out. Another useful comparison is that between the dominant zones of different distributions. Taking again the trip-ends distribution, for example, the dominant zones in producing trips can be compared to the dominant zones attracting them. If comparisons reveal that the two groups consist of entirely different zones, then it should be clear to the transportation engineer that he is confronted with a pattern of widely dispersed trips. The various degrees of identity between the dominant zones of origin and the zones of destination, defined by such comparison, may provide the planner with an immediate, clear picture of the travel pattern he is confronted with.

Evaluations of Surveys

Finally, the reliability of the results of various surveys or projections may be established by examining the existence of the KD characteristics. If it is known, for instance, that the distribution of land uses in a certain town is highly concentrated whereas their attracted/generated trip-ends show a dispersed pattern, then a warning is due regarding the accuracy of the findings. The same warning applies to a specific zone that is dominant in the land-use distribution but falls below the breaking point in the trip-end distribution. It is to the planner's advantage to roughly evaluate the reliability of the findings easily and quickly at a very early stage.

CONCLUSION

Studying the various applications of the KD to transportation issues leads to the following conclusion. The technique of fitting the KD to transportation functions is a single device of manifold uses that enables immediate clarification of the following important basic concepts: the level of traffic urbanization and its trends of variation over different periods of time, over different modes of travel and over different cities; the regularity of development and the interrelationship between the traffic zones that constitute the planning units in transportation; the range of dominant zones that might serve as the basis for some major transportation projects; and, finally, the trends of variations in the dominance of traffic zones and their significances.

It should be noted, nevertheless, that because the KD is primarily empirical, its application is not intended to produce clear-cut solutions to problems that cannot be solved otherwise. Its application, however, can ease immensely the work of the transportation planner by providing an additional tool for analysis.

The initial experience of applying the K distributions to certain practical aspects in the preparation of the Transportation Master Plan for the Tel-Aviv area enabled us to confirm that these outlined advantages of its use have been fully attained.

REFERENCES

1. United Nations Demographic Yearbook, 1962, pp. 383-386.
2. Chicago Area Transportation Study, Vol. 1, p. 115.
3. Pittsburgh Area Transportation Study, Vol. 1, p. 104.
4. Tel-Aviv-Yafo Transportation Master Plan, Vols. 1 and 2. Prepared for the Municipality of Tel-Aviv-Yafo, Ministry of Transport, Dan and Egged Bus Cooperatives, by Kolin and Zahavi, Transportation and Traffic Engineers.