CHAPTER 4

FORECASTING THE VALUES OF EXOGENOUS VARIABLES: TRANSPORTATION SYSTEM ATTRIBUTES

Introduction

The development and use of disaggregate travel demand models for transportation policy analyses requires auxiliary forecasts of the variables exogenous to the model system. The previous chapter described a method of developing the exogenous socioeconomic attributes for a homogenous market segment or for a representative sample of households. In this chapter a method is described that provides estimates for the various components of travel time on the various modes of travel as a function of the transportation system serving a door-to-door trip from a given origin zone to a known destination zone. The same method can be used to forecast those items of trip user costs that are defined functions of trip distance.

The two measures, travel time and travel costs and their components, are often used to completely characterize the performance of the transportation system in travel demand models. This is not to say that such measures as on-time performance, security, safety, and other system attributes are not important, but that current knowledge of them is limited. It is acknowledged, then, that more research is needed to quantify those variables but the matter is not pursued here. Another point to note is that headways of buses are taken to be exogenous variables; the feedback of demand through management is assumed not to exist.

The use of networks to compute the travel times and costs is so ingrained among transport planners that the use of anything else to represent transport system performance is considered to be a questionable shortcut or viewed as a "back-of-the-envelope" computation. Nonetheless, the use of networks has several disadvantages and shortcomings (Talvitie and Dehghani, 1976; see also Part IV, Chapter 8) that encourage the development of alternative methods to computing network independent "supply" models. These are the accuracy of network calculations in computing trip times and the aggregation of individual demands to obtain total demand. As noted by Talvitie (1973), McFadden and Reid (1975), and Westin (1975), unbiased aggregation of individual travel demands requires that the within-group variances are accounted for. However, the networks provide only one estimate of level-of-service per mode for an O-D pair; that is, the within-zone variance of level-of-service attributes is assumed to be zero. Is this a reasonable assumption, or more generally do the minimum path algorithms operating within coded networks provide accurate values for level-of-service to be used either in developing the travel demand models or in forecasting with them? An examination of this issue is undertaken in the next two sections of this chapter.

In the third and fourth sections a method and models are developed and described which permit network-independent description of travel time on various modes for access and linehaul. These models are used in a transport corridor policy study (FRX).

A Comparison of Experienced and Network Based Travel Time Measurements

To start, we need to define the way in which the two types of values were obtained. The experienced transit travel times were obtained by asking the transit agencies' information service to route travelers as if an inquiry call for a transit route were made by the traveler.[1] The experienced auto travel times were based on travel time runs (moving vehicle method) made at various times of day and by routing travelers at the minimum time path at their time of travel.

The network values were obtained through standard network models and associate either peak or off-peak values with the traveler depending on when the trip took place. These data were prepared independently of the present research and are reported in McFadden (1973b). (For a detailed account of the data preparation see articles by M. Johnson and F. Reid in that publication.) Round trip travel time and cost values are used in both sets of data. The comparison of the types of measurements will be done in two different ways by comparing the measurements directly using various indices and by comparing the models estimated with the two types of measurements.

Means, standard deviations, and correlation coefficients

The comparison of the experienced (E) and network (N) travel times may be started by listing the means and variances of the travel time components of interest. These appear in Table 49. Examination of the values in Table 49 reveals no spectacular differences; the variances in the "experienced" data cells appear to be consistently higher and the means differ somewhat. The greatest concern, on the basis of the values in Table 49 appears to be with the headway of the first bus and with the transfer times, either to work or to home. This difference can, at least in part, be attributed to headways that can vary substantially within peak hours.

---

[1]This does not guarantee that the chosen or "would be chosen" path was actually measured. A check was possible with those who reported which path they took (occasional and regular transit users) against what was measured. These paths agreed in seventy-six of the eighty-six cases where the information was available. Of the ten misses, five persons took a different bus route altogether, in the other five misses one part of the bus route was different. Thus, there is some justification to take issue with the use of the word "experienced" in this paper. Those of the readers who prefer to do so should regard this paper as a comparison of two different algorithms to choose a transit path and hence the travel time components of a door-to-door trip.

TABLE 49 Means and Variances of Travel Time Components by Mode and Type of Measurement

Type	Time Component	Auto		Bus/Walk Access		Bus/Car Access	
		Mean	SD	Mean	SD	Mean	SD
Net	On-Vehicle Time	45.2	24.6	68.5	34.7	70.5	34.9
Exp	On-Vehicle Time	50.5	28.7	77.3	35.8	80.0	36.5
Net	Walk Time	NA	NA	23.8	11.5	9.9	3.6
Exp	Walk Time	NA	NA	21.8	17.3	6.6	5.4
Net	First Headway to Work	NA	NA	12.2	10.9	As with Bus/Walk Access	
Exp	First Headway to Work	NA	NA	18.1	13.6	As with Bus/Walk Access	
Net	First Headway to Home	NA	NA	8.1	7.0	As with Bus/Walk Access	
Exp	First Headway to Home	NA	NA	14.9	12.5	As with Bus/Walk Access	
Net	Transfer Time to Work	NA	NA	2.8	3.3	As with Bus/Walk Access	
Exp	Transfer Time to Work	NA	NA	5.3	7.9	As with Bus/Walk Access	
Net	Transfer Time to Home	NA	NA	6.1	8.5	As with Bus/Walk Access	
Exp	Transfer Time to Home	NA	NA	8.0	10.5	As with Bus/Walk Access	
Net	Total Wait Time to Work (equaling Xfer time and 1/2 headway	NA	NA	8.9	6.9	As with Bus/Walk Access	
Exp	Total Wait Time to Work (equaling Xfer time and 1/2 headway	NA	NA	14.4	12.5	As with Bus/Walk Access	
Net	Total Wait Time to Home	NA	NA	10.1	8.9	As with Bus/Walk Access	
Exp	Total Wait Time to Home	NA	NA	15.4	13.8	As with Bus/Walk Access	
Net	Number of Transfers	NA	NA	1.9	1.5	As with Bus/Walk Access	
Exp	Number of Transfers	NA	NA	1.8	1.4	As with Bus/Walk Access	
Net	Cost	190.3	134.6	103.2	53.8	137.5	46.1
Exp	Cost	188.0	138.3	109.9	64.0	145.0	56.7

"Total wait time," which is the sum of cumulative transfer times and one-half of first headway, confirms that differences exist in the two types of measurement: the experienced values being higher than the network values. In summary, excepting the walk time whose mean is just two minutes higher by the network algorithm, all the mean values from the network are somewhat lower than the corresponding values experienced by the travelers. The standard deviations of the experienced travel times are substantially larger than those of the network times. It may be computed from Table 49 that between zone variance (network data) accounts for thirty to sixty percent of the total variance (experienced data) for the excess time components and about seventy to ninety percent of the on-vehicle time variances. This must have a bearing on aggregation as will be discussed later.

The values of the correlation coefficients, regression intercepts, and "slopes" in Table 50 indicate that other than the on-vehicle times, the desired values of unity and zero for correlation intercept and slope are not achieved; even for the on-vehicle times the hypothesis that the slope $\underline{b}$ equals unity must be rejected.

Root mean square errors and Theil-U coefficients

The information produced so far about the similarities and dissimilarities of objective and network measurements of travel times can be conveniently summarized using two measures: the root mean square error and Theil's U-coefficient.[1] The former is often used as an "all around" measure of "goodness-of-fit;" the latter measure is zero for perfect measurements (or forecasts) and has an upper bound of one. Theil's U-coefficient can furthermore be decomposed to three components, denoted U^M, U^S, U^C, which indicate the proportional loss in accuracy that is due to differences in means, in standard deviations, and in covariances, respectively. These useful summary measures are given in Table 51.

[1]This coefficient is computed as

$$\left(\sum_i (N_i - E_i)^2 / \left(\sum_i (N_i^2 + E_i^2)\right)\right)^{1/2} .$$

The numerator can be decomposed into three components and normalized to unity.

TABLE 50 Correlation Coefficients, Intercepts, and Slopes for Regressions Between Some of the Experienced and Network Times

Travel Time Component	Correlation Coefficient	Intercept a (std error)		Slope b (std error)	
Auto Round-Trip On-vehicle Time	.91	6.52	(1.73)	.766	(.030)
Bus Round-Trip On-vehicle Time	.89	1.83	(3.18)	.862	(.037)
Walk Time Bus/Walk Access	.45	17.28	(1.40)	.297	(.050)
Walk Time Bus/Auto Access	.32	8.45	(.45)	.213	(.050)
Headway on Bus to Work	.52	4.64	(1.32)	.416	(.058)
Transfer-Time on Bus to Work	.35	2.07	(.32)	.144	(.033)
Headway on Bus to Home	.50	3.91	(.79)	.279	(.041)
Transfer-Time on Bus to Home	.38	3.68	(.83)	.303	(.063)

TABLE 51 Root Mean Square Errors and Theil U-coefficients of Travel Time Components by Network Measurements as Compared to the Experienced Travel Times

Variable	Means (experienced)	RMSE	Theil U			
			U	U^M	U^S	U^C
On-vehicle time - auto	50.5	13.1	.17	.16	.10	.74
On-vehicle time -bus(w)[a]	77.3	18.8	.16	.22	.00	.78
On-vehicle time - bus(a)	80.0	19.3	.16	.24	.01	.75
Walk time - bus(w)	21.8	16.0	.42	.02	.13	.85
Walk time - bus(a)	6.6	6.4	.47	.27	.08	.65
Headway to work	18.1	13.6	.49	.19	.04	.77
Headway to home	14.9	12.8	.58	.28	.18	.54
Transfer-time to work	5.3	7.8	.75	.10	.35	.55
Transfer-time to home	8.0	10.9	.65	.04	.03	.93

a/ w, a denote access mode - a = auto, w = walk

The results in Table 51 are revealing. Excepting the linehaul travel times, the root mean square errors are roughly equal in magnitude to the means of the experienced travel times indicating rather large errors in measurement. The same result is conveyed by the Theil U- coefficient. Again, the U-coefficient obtains very large values for out-of-vehicle time components, even for on-vehicle time components the U-coefficient, (and the RMSE) is quite high. One wonders if travel forecasts would be as highly regarded as network travel time data if they were subject to errors of these magnitudes. Finally, the components of the U-coefficient indicate that the largest share of the error comes from the covariances between the network and objective measurements; in some cases a substantial part is also due either to differences in means or standard deviations.

Frequency plots of model variables

As a final item before actually estimating choice models using the two types of measurement, it is instructive to examine the frequency plots of some of the travel time variables. The analysis performed by McFadden and Reid (1975) tells that zonal averages will yield consistent estimates for coefficients given that the distributions of variables within a zone are not skewed. Thus, the distribution of the variables for the entire sample (envisioned as one large zone) ought not to be skewed either if good coefficients are to result from using zonal averages. The comparison of frequency distributions will also help state *a priori* expectations for the model coefficients; however, for two reasons caution must be exercised in doing so. First, because the logic model operates on differences, the distributions ought to be plotted separately by choice and the difference examined; due to the small sample size this was not possible.[1] Second, the components of the Theil U-coefficient indicated that most of the difference between the two types of measurements is due to covariances. This means that the frequency plots for the two measurements can look similar without the measurements being similar because measurements in any given interval may not pertain to the same individuals.

[1]The sample size was 142; there were 103 auto users, twenty-eight bus riders with walk access and eleven bus riders with car access.

It is natural to start with the plots of auto and transit (with walk access) on-vehicle times; these are shown in Figure 13 (auto), (bus), and (bus minus auto). A visual examination of the plots in Figure 13 suggests that there is a great deal of similarity between the two types of measurements; the only noticeable difference is the "fat tail" of the experienced auto on-vehicle times distribution. One might suspect that the lack of "fat tail" in the network times distribution is due to improper accounting of congestion effects. A χ^2 test against the null hypothesis that the distribution of network times is identical to the distribution of the experienced travel times had to be rejected, however, at the .95 level of confidence.

The walk time (bus-with-walk-access) frequency distribution in Figure 13 indicates that the network-coded walk time has a highly peaked distribution while the distribution of the experienced walk times both peaks earlier and is much "fatter." The appearance of the two distributions is as expected. Traffic zones are connected to the network with relatively few common values, and the experienced values show a scatter that relates to the location of individuals with respect to the bus line configuration.

The frequency plot for bus headways (round trip; directional headways summed) appears also in Figure 14. It may be noted that the network headways are shorter in duration; their distribution also has a noticeably thinner tail than that of the experienced headways. The apparent reason is that zones have been connected to trunk-line streets on which many bus lines operate and have low headway for consecutive buses, when actually the travelers' origins and destinations are dispersed within the zones, and by taking note of schedules, the travelers can gain the advantage of nearer buslines in spite of their lower service frequency.

Finally, a look at the frequency plot for transfer time indicates that the distributions for transfer time resemble each other; they are also the only distributions where the null hypotheses that network travel time distribution is the same as the distribution of experienced travel times cannot be rejected at the .99 level of confidence. It is surprising that transfer time distributions are so similar, because at least fifty-six percent of the travelers are known to take different paths; the Theil U-coefficient was also very high for the transfer times: however, most of the error was due to covariances that may explain the outcome.

FIGURE 13

Frequency Plots of On-vehicle Time

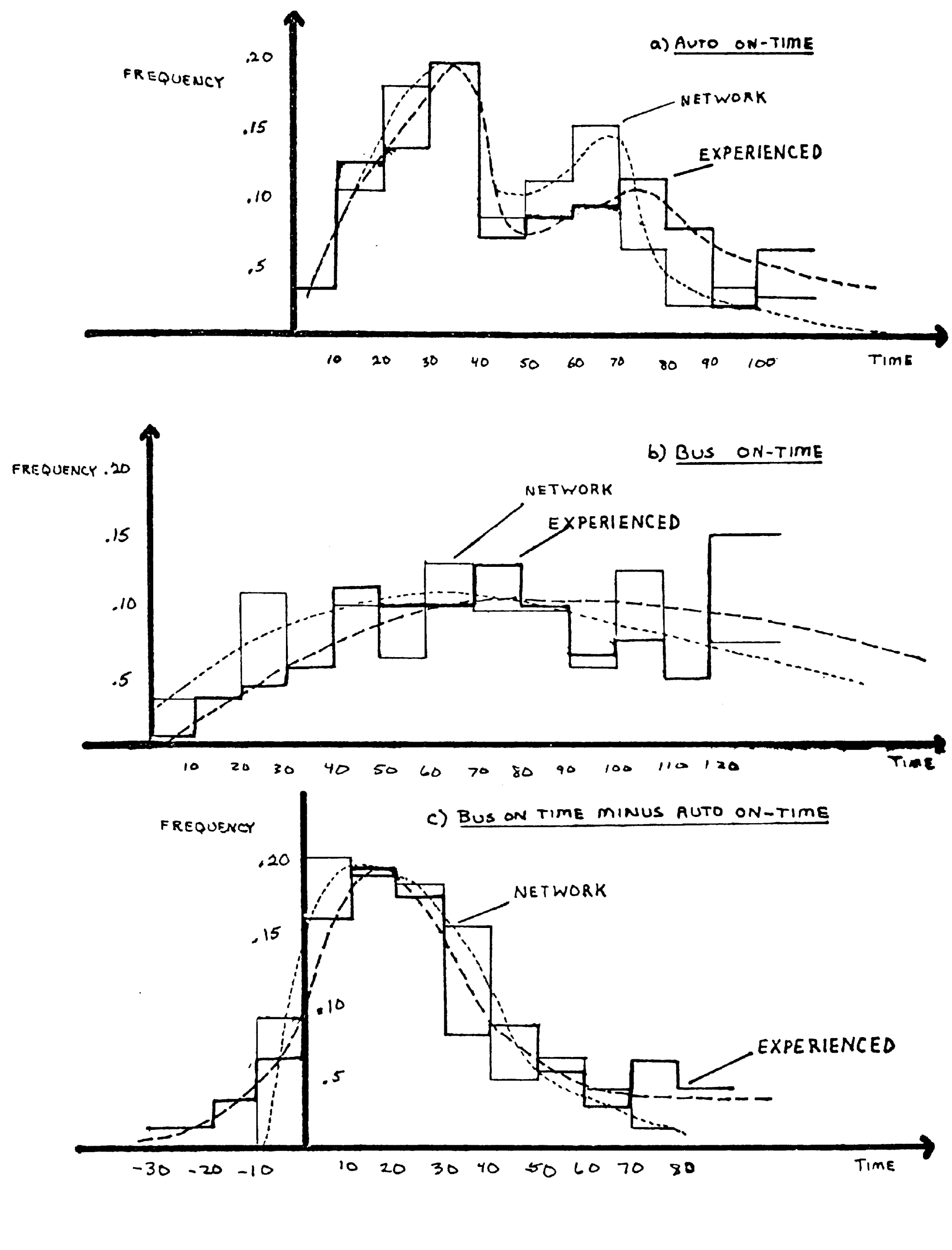
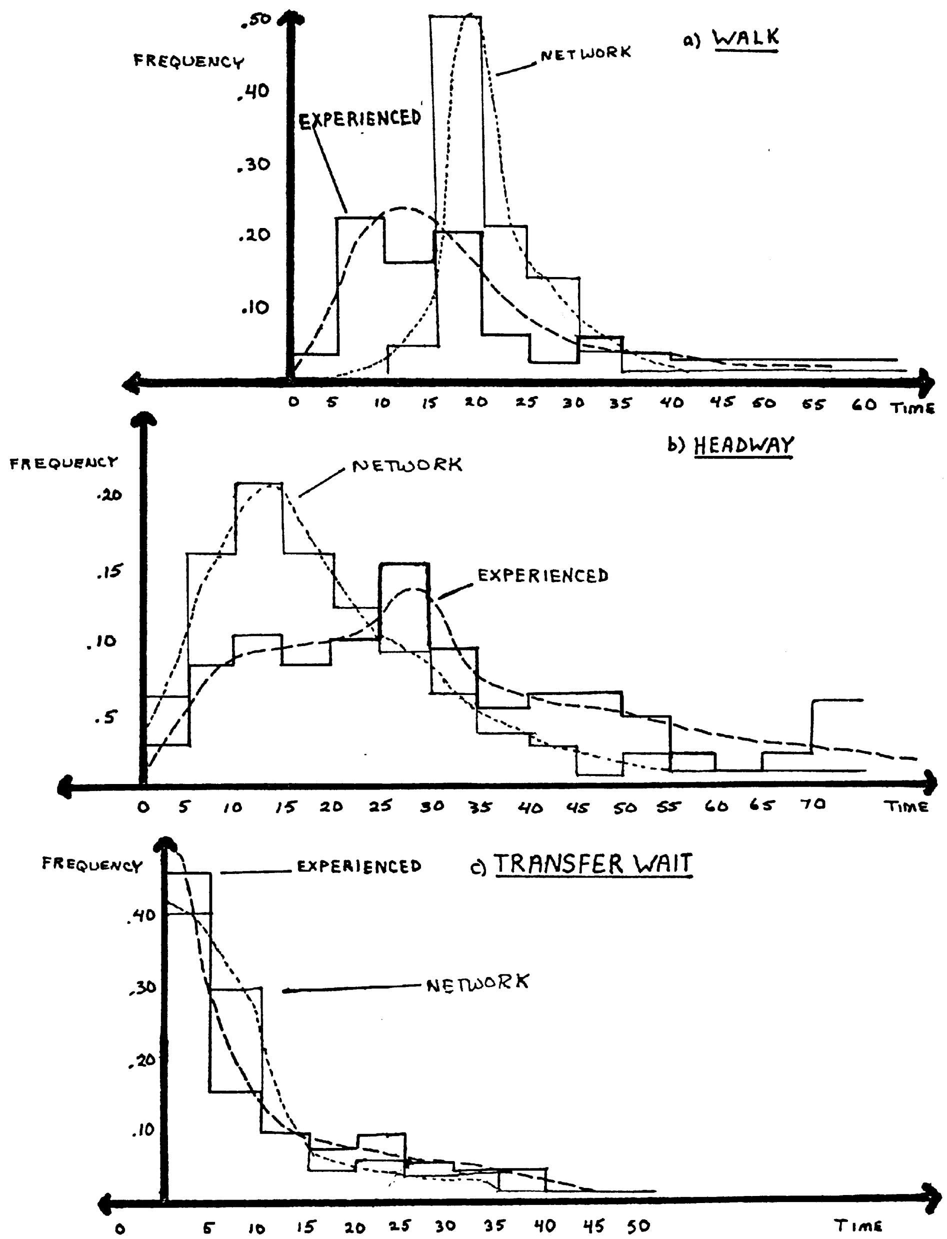


FIGURE 14

Frequency Plots of Walk Time (a), First Headway (b), and Transfer Time (c) All Round Trip



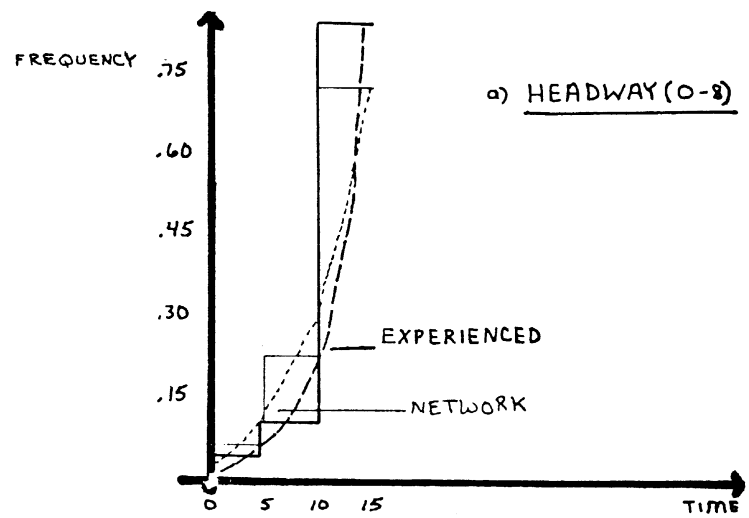
In order to capture the non-linearity of travelers' response to transit headways in modeling choice, to be reported in the next section, the headways are segmented into two components (see also Chapters 1 and 2, Part II). The first component of the headway is up to a maximum of eight minutes, and the second component is the remaining headway time. In order to do justice to both legs of a round-trip, these two components have been computed separately for both directions and then summed. The plots of the two headway segments appear in Figure 15. It may be noted that the distributions for the truncated portion of the headway are quite similar (note, however, the scale on the vertical axis) while the tail sections of the headway have quite different distributions; the latter is already evident from Figure 15.

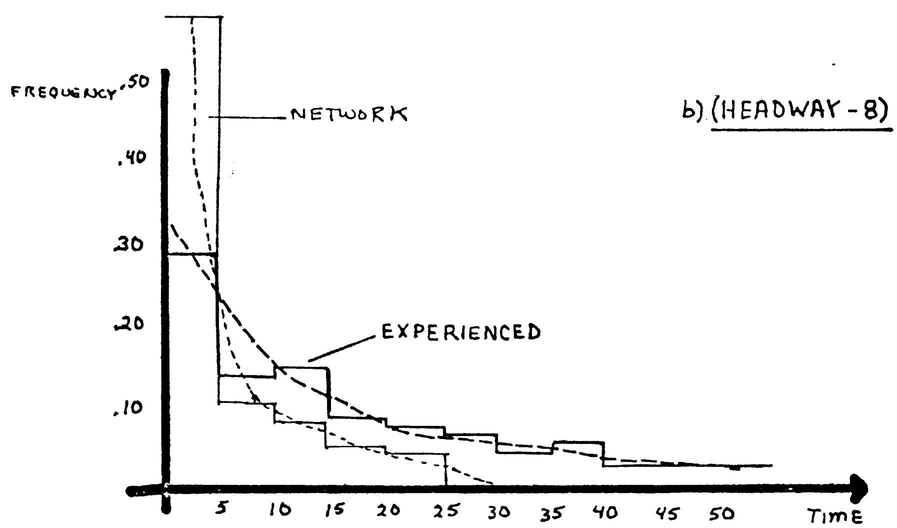
Discussion

This section may be concluded by noting that the two types of measurements--experienced and network--of travel time variables are certainly different. On the basis of the frequency plots one would expect that similar coefficients can, nevertheless, be estimated for on-vehicle time, transfer time, and first component of the headway regardless which type of measurement is used. However, one must keep in mind that for these variables most of the error was due to covariances and cannot be seen in the plots; thus, if the experienced travel times are strongly correlated with socioeconomic variables or with each other, then the anticipation of similar coefficients will fail. This latter is most certainly true with the headway variable; the first component must correlate highly with the second component because well over one-half of the travelers had round-trip headways greater than sixteen minutes. Correlations between socioeconomic and travel time variables will also exist: transportation folklore suggests that bus walk times and socioeconomic variables are highly correlated. Finally, it may be noted that the experienced travel times have distinctly skewed distributions; this, it may be recalled, violates one of the conditions for obtaining consistent coefficient estimates using zonal averages measured by the network models. The comparison of mode choice models using the two types of supply variable measurements follows next.

FIGURE 15

Frequency Plots of Headway Components 0 - 8 Minutes (a) and Headway Minus 8 Minutes (b)





A Comparison of Mode Choice Models Developed with Two Types of Supply Measurements

The "basic model" specifications used as a benchmark here is the one reported by McFadden and Train (1975).[1] Table 52 shows the coefficients for that model (labeled Model Ax; x = N for network or x = E for experienced) and another model, labeled B , whose specification includes the second, untruncated portion of the headway and no transfer time, which was deleted from the B model due to the coefficient instability and closeness to zero. The two models are estimated using both experienced and network travel time measurements.

The discussion of these models will be focused on the transportation system performance variables: travel times and travel costs. The socioeconomic variables, which are common to both (N and E) type models and pertain to individuals rather than zones, are discussed only insofar as their deletion or inclusion affects the coefficients for the travel time or cost components.

A number of points are worth noting in comparing the xN and xE type models. First, the coefficient for travel cost is stable across the models, regardless of the type of measurement. Second, the on-vehicle travel time coefficient is twice as high in the models estimated with experienced travel times as compared to the model estimated with network travel time measurements. On the other hand, the coefficient for walk time behaves exactly in the opposite way. It is as if these two variables have exchanged coefficients in the two types of models; and there is no reporting error here. The values of time computed from these models are, naturally, similarly reversed. In the E series models the value of the on-vehicle time is around seventy percent of the wage rate and the value of walk time nearly forty percent of the wage rate. In the N series models these two figures are around forty and eighty-five percent, respectively. The behavior of the traveling time coefficients, especially that of walk time, is such that one is led to suspect the underlying correlations to be secretly at work. For example, from the distribution of the walk times in Figure 14 one is willing to surmise that the coefficient of experienced walk times will be higher than the one estimated for network measured walk time. This is because the distribution of experienced walk times is located to the left of the distribution of network walk times. More will be said about walk times later.

---

[1]For specification of the variables and the coefficients of the McFadden, Train model, see Table 2, Part II, Chapter 1.

TABLE 52 Coefficients (t-values) of Work Trip Mode Choice Models Estimated Using Both Experienced and Network Travel Times

	Model AE Experienced Travel Times		Model AN Network Travel Times		Model BE Experienced Travel Times		Model BN Network Travel Times	
INC 1	.000413	(1.4)	.000239	(.80)	.000339	(1.1)	.000279	(.90)
INC 2	.000955	(1.8)	.000684	(1.4)	.0001075	(1.9)	.000592	(1.2)
INC 3	-.000743	(2.6)	-.000643	(2.5)	-.000748	(2.6)	-.000624	(2.4)
Residence	.178	. (3.0)	.193	(2.9)	.176	(3.0)	.195	(3.0)
Population Density	-.643	(2.1)	-.449	(1.5)	-.713	(2.1)	-.480	(1.6)
Parking	-.433	(1.2)	-.308	(.92)	-.384	(1.0)	-.316	(.90)
Age	-.759	(1.1)	-.711	(1.02)	-.720	(1.0)	-.735	(1.1)
Child	-1.575	(2.3)	-.969	(1.50)	-1.676	(2.4)	-1.049	(1.6)
Drivers	1.321	(2.8)	1.182	(2.7)	1.248	(2.6)	1.186	(2.7)
Cost/Wage	-.0427	(3.1)	-.0468	(3.1)	-.0448	(3.1)	-.0473	(3.1)
On Time	-.0293	(1.8)	-.0132	(.60)	-.0304	(1.9)	-.0175	(.90)
Walk Time	-.0157	(.74)	-.0418	(1.3)	-.0185	(.90)	-.0379	(1.1)
Transfer-Time	-.0395	(1.0)	-.0346	(.80)	---		---	
Number of Transfers	-.165	(.50)	-.0387	(.20)	-.321	(1.3)	-.134	(.6)
Headway 1 (≤ 8 min)	-.242	(2.1)	-.246	(2.3)	-.137	(1.1)	-.216	(2.4)
Headway 2 (> 8 min)	---		---		-.0647	(1.9)	-.0427	(1.8)
D_a	-6.399	(2.3)	-4.806	(1.9)	-5.159	(1.9)	-4.640	(1.8)
D_{ba}	-2.806	(3.2)	-2.989	(3.4)	-2.724	(3.1)	-2.932	(3.3)
Likelihood ratio index	.601		.575		.613		.578	
% Right	81.7		83.1		82.4		84.5	
Value of On-Time	69% of wage		28% of wage		68% of wage		37% of wage	
Value of Walk Time	37% of wage		89% of wage		41% of wage		80% of wage	
Ratio of Walk/On-Time	.54		3.17		.61		2.17	

The third outstanding item in Table 52 has to do with the coefficients for the number of transfers, headways, and transfer time. The number of transfers has about a three times higher coefficient in the models using experienced route information; the two headway components also have different weights in the two types of models. If only the truncated portion of the headway is included in the model--as was done in Models AE and AN--then this headway coefficient in the two types of models has just about the same value. This is an interesting outcome (it can be surmised by looking at the frequency distributions of these variables in Figure 15) and will be discussed later. Model estimations also showed that the coefficient for transfer time became very unstable and close to zero when the tail section of the first headway was added to the set of independent variables. This outcome may be partially due to the fact that when transfers are involved, the transfer headway is often equal to the first headway in the return trip. Thus, with truncation of the first headway, something is the matter with transfer time; however, when the first headway is not truncated the effect of the transfer time seems to fall off. (See also discussion in Part II, Chapter 2.)

The fourth and final item evident from Table 52 is the low statistical significance of most of the coefficients and the equal forecasting accuracy of the models, as judged from the likelihood ratio and percent right indices. Given the low statistical significance of the system variables, one can rightly ask are the models developed using the two different supply measurements statistically different? The answer appears to be no. This result was arrived at by performing a sort of Chow-test on the two models (McFadden, 1973). The coefficients were first estimated using the combined sample (284 observations) and restricting the coefficients of two types of measurements to be equal and then relaxing this restriction and performing the likelihood test on the results.

The χ^2 distributed test-statistic with six degrees of freedom had a value of 6.8, which is well below the .95 level critical value of 12.5. Note however that observations in the two subsamples are not independent, putting in question the statistical validity of the test.

<u>Further discussion of results</u>

In evaluating the outcome of the statistical test conducted above it is good to keep in mind that the sample size used here was quite small and that most of the system variables did not obtain statistical significance in the original models of Table 52 at the level used in the Chow-test test. In fact, the results of that table give grounds for pursuing the matter of network measured attributes and experienced attributes a little further. A convenient starting point for doing this is provided by estimating a simpler model than that in Table 52; the discussion promised earlier regarding the walk times and headways is appropriately conducted here also.[1]

It is seldom the case that all the socioeconomic variables used in the models applied here are available for the model development of travel forecasting in particular. A question then arises about the effect of dropping some of them from the models; that is, what is the effect of model misspecification on the results. Specifically, let us assume that variables "length of residence," "population density" (as defined here), "parking index," "age," "child-dummy," and "the number of drivers" are unavailable for model estimation; and let us also eliminate the number of transfers from the model and truncate the headway (to form a wait time); these are all quite reasonable assumptions and customarily made. The models estimated with the remaining variables using the two types of supply measurements appear in Table 53. The results show that, excepting the transfer time, all the coefficients bear statistical significance at reasonable levels of confidence (one-tailed tests) in the model estimated using experienced travel time information. In the network-based model, on-vehicle time and the one-income component in addition to transfer time have low statistical significance. Particular attention should generally be paid to values of time that have climbed up substantially due to a nearly fifty percent decrease in the cost coefficient and a 100 percent increase in the walk time coefficient. Also of interest is the low coefficient (and statistical significance) of the on-vehicle time in the network based model; this dilemma is almost daily experienced by travel demand modelers working with network-based LOS information. On the other hand, the experienced on-vehicle time coefficient has not substantially changed from the results obtained earlier.

These changes, especially in the walk time coefficient, must be attributed to their correlation with the socioeconomic variables which were eliminated from the model. These correlations between the socioeconomic and transportation system variables are related to the locational and transportation choices that

---

[1] It is pertinent to remind the reader that the walk and headway times (and transfer times) were previously the variables that "caused trouble" in validation, see Part II.

people make under their own circumstances.[1] The "experienced" LOS values preserve these correlations and are likely to yield unbiased coefficients and demand elasticities (given a good model specification) while the network calculations do not appear to preserve these correlations, and, by simple logic, must yield biased coefficients.

The effect of the variable specification also deserves examination. In particular, attention is drawn to the coefficients of the headway variables. It may be noted from Tables 52 and 53 that if only the truncated part of the headway (in this case up to eight minutes, one way) is included in the model its coefficient remains very nearly the same, regardless of the type of measurement. It may be recalled that the frequency plots of this variable were also very similar for both types of measurement. The implication is that the construction of the variable has procreated the nearly identical coefficients.

---

[1]Past data indicate that people adjust to changes in their circum stances (and vice versa)--for whatever reason--reasonably quickly. Of about 100 working travelers in 1972 only about twenty percent had the same origin and destination in 1975 (UTDFP panel travel survey).

TABLE 53 Coefficients (t-values) of Purposefully Underspecified Mode Choice Models Estimated Using Both Experienced and Network Supply Data

Variable	Experienced Supply Data		Network Supply Data	
INC 1	.000207	(1.5)	.000154	(.8)
INC 2	.000961	(2.1)	.000693	(1.7)
INC 3	-.000046	(1.9)	-.000364	(1.7)
Cost/Wage	-.0237	(2.2)	-.0259	(2.6)
On-vehicle Time	-.0239	(1.8)	-.00394	(.2)
Walk Time	-.0377	(2.0)	-.0724	(2.3)
Transfer Time	-.0228	(1.0)	-.0408	(1.1)
Headway 1	-.257	(2.7)	-.277	(4.1)
D_a	-5.42	(2.8)	-4.84	(3.0)
D_{ba}	-1.29	(2.6)	-1.82	(3.1)
Likelihood ratio index	.484		.482	
% Right	76.8		76.8	
Value of On-time	101		15	
Value of Walk Time	159		280	
Value of Xfertime (as % of wage)	96		158	
Ratio Walk/On-time	1.58		18.4	

There is good reason to take this implication a few steps further. It may be observed from Table 52 that the relative magnitudes of on-vehicle and walk time coefficients are between two to three for the network-based models. This is very close to the relative weight (two; plus penalties for waiting and a limit on transfers) of walk time used in building the transit paths. The obvious hypothesis then is that the conventions used in building the paths and coding the networks procreate the choice models based on these types of supply data.[1] Thus, if networks in two or more cities are coded using similar conventions and without regard to correlations with socioeconomic variables, and if paths are built using similar weights, and variables created using the same type of rules (e.g., wait time is one half of the head way up to ten minutes of headway and one-fourth thereafter; this is quite akin to the eight minute maximum used here), then with a normally low percentage of transit users the resulting choice models for those cities should indeed be nearly identical. The models so obtained are not really behavioral nor transferable--in a true sense--travel demand models. They are not likely to predict correctly travelers' behavior when service changes occur.

If the hypothesis that network-based supply measures procreate the choice model coefficients is true, then it logically leads to the unfortunate conclusion that the accumulated evidence from numerous studies on travel demand elasticities with respect to travel time components cannot be taken as mutually supporting statements of truth.

Is the hypothesis made true? The statistical test performed earlier tells that the null hypothesis of the network-based model of not being "cooked" could not be rejected. However, two other pieces of information are relevant here. First, logic suggests that unless there is only one path or many paths with similar attributes the hypothesis must necessarily hold. Because, at least in some cases, there are many paths--this has already been shown to be true--the network-based model must be "cooked" at least to some extent. Second, a statistical question can be asked: is it possible that the coefficients obtained from one set of data, say from the manually coded supply data, can with equal likelihood be obtained from the network-based supply data (or vice versa)? If we constrain the six system

---

[1]The perceptive reader may ask, then why does the on-vehicle time have such a low coefficient in the network-based model in Table 53? A hypothesis can again be made that if we omit most of the socioeconomic variables, then the choice will be attributed to the system variables; however, in building the paths using the network models such a low weight is given to the linehaul time that linehaul time really does not matter in choosing a path; and a result which is a statistical artifact will follow. The hypothesis in the main text is preferred, however, because it is more general and can encompass both types of situations encountered.

variables to obtain the coefficients estimated with the "experienced" supply data (Model BE, Table 52) and use the network data to re-estimate the same model, the $\chi 2$ distributed test-statistic is equal to 5.58 and the critical value at .05 level with six degrees of freedom is 12.59. By taking the network-based model coefficients (Model BN, Table 52) as the base, the test-statistic is equal to 48.06 with the critical value remaining unchanged. These results tell us that, with the data used here, it is possible that the coefficients obtained using the "experienced" supply data can also be obtained from the network data with no loss in statistical significance of the model. The reverse is not true, the coefficients obtained using the network data cannot be obtained from the "experienced" supply data. A possible interpretation here is that when using the network data the likelihood function is quite flat and practically any set of coefficients is possible.[1]

<u>Implication of results to aggregation</u>

It certainly has not escaped the reader's attention that the forecasting accuracy of the models is nearly identical regardless of the type of data used. Furthermore, the network-based models seem to have simple aggregation properties. Koppelman's carefully done in-depth study on aggregation (1976) shows that predictions with zonal averages seem to perform remarkably well. The same type of results can be read from Atherton and Ben-Akiva (1976), and Talvitie (1975). There are two reasons that cause this to be the case. First, networks ignore the within-zone variances, the source of aggregation bias. Thus, using the networks, there is not much left to aggregate as far as the non-linehaul LOS variables are concerned. Second, assume that the network travel times and costs are "errors-in-variables" or

(1) $$X = Z + v \quad ,$$

---

[1]The problem with this test is that the samples are not independent because of the common socioeconomic attributes. If the socioeconomic attributes are deleted and the models run without time, then the independence problem should disappear (but misspecification will appear); the respective χ^2 statistics are 16.8 and 1164.6, with eight degrees of freedom (alternative-specific dummies are included); the critical value is 15.5. Thus, the hypothesis that the coefficients obtained from one data set could also be obtained from the other must be rejected.

where Z = network values;

X = true values;

v = a random error.

Let us then assume that X and v are independently and normally distributed with means m_x and zero, and variances of σ_x^2 and σ_v^2 . These are reasonable assumptions; any time when a trip is taken and the trip time is not known exactly, it is a random variable; and this random variable is independent of the traveler's location within the traffic zone. The hypothesis in disaggregate travel demand models is that the choices of travelers depend on the true values X or, in the sense of regression,

$$V = \alpha + \beta X + e \quad ,$$

where V is the demand or choice. In predicting, we do not know the true value X , but we know the network value Z , and thus we need to obtain $E(X \mid Z)$, where E is the expected value operator, equal to (Benjamin and Cornell, 1970)

(2) $$E(X \mid Z) = \frac{\sigma_v^2 \cdot m_x + \sigma_x^2 \cdot Z}{\sigma_v^2 + \sigma_x^2} \quad ,$$

and the estimated expected demand is

(3) $$E(V \mid Z) = \hat{\alpha} + \hat{\beta} E(X \mid Z) = \hat{\alpha} + \hat{\beta} \left(\frac{\sigma_v^2 \cdot m_x + \sigma_x^2 \cdot Z}{\sigma_v^2 + \sigma_x^2} \right) \quad ,$$

where $\hat{\alpha}$ and $\hat{\beta}$ are consistent errors-in-variables estimators for α and β . On the other hand, the least squares predictor is

(4) $$V = \bar{v} + b(Z - \bar{Z}) \quad ,$$

where b is just an OLS estimator of V on network values Z . It has been shown (Johnston 1972) that

$$b = \beta / (1 + \sigma_v^2 / \sigma_x^2) \quad , \quad \text{or}$$

(5) $$V = \overline{V} + b(Z - \overline{Z}) \quad ,$$

$$V = \alpha + \beta \cdot m_x + \beta(Z - \overline{Z}) / (1 + \sigma_v^2 / \sigma_x^2) \quad .$$

Because $E(v) = 0$, $\overline{Z}$ is an unbiased estimate for m_x , and

(6) $$V = \alpha + \beta \left(\frac{\sigma_v^2 \cdot m_x + \sigma_x^2 \cdot Z}{\sigma_v^2 + \sigma_x^2} \right) \quad .$$

But this is exactly what was obtained using the consistent errors-in-variables coefficients α and β , equation (3).

This discussion can also be interpreted in another way. The network coding conventions and practices can be viewed as a process where the true values X and the coded values Z have a joint normal distribution. By working out $E(X \mid Z)$ that way the same result reached above is obtained.

Thus, even though the coefficient b in Equation (4), obtained using the network values directly, are not unbiased, they yield unbiased forecasts. Therefore, for prediction purposes, the network-based models, whether aggregate or disaggregate, can be used with success, provided that conventions for network coding and path building are not changed and normal caution is exercised in out-of-range predictions. Note that incremental forecasts cannot be made because the coefficients are not unbiased. Clearly, the usefulness of such a model is limited, particularly in policy analyses.

Summary and conclusions

In the two previous sections, experienced and network-based travel times were compared and mode-choice models were estimated using both types of data. The implications of the results of this work are obvious. On the demand side, incremental forecasts should be avoided unless the demand elasticities implied by individual model coefficients can be supported by real-world experience. This is because the coefficients estimated with the two types of data were not numerically similar even though the statistical inference regarding the two sets of coefficients was inconclusive. This outcome of the statistical test was felt to be due partially to the small sample size and partially to the transportation system components having low or no statistical significance in the models.

On the supply side the burden is clear: data errors in model estimation can be costly. On the basis of the comparisons made one must express the concern as to whether the models based on contemporary network models are nothing but procreated constructs of the adopted network-coding and path-building conventions aided by suitable variable definitions.

In sum, the results of the comparison and the discussion on aggregation provide strong motivation for developing alternatives to the network system for obtaining and forecasting the values of the exogenous system attributes. One such method is described in the next two sections.

Development of Parametric Models to Forecast Level-of-Service Variables for Access

The results of the previous two sections suggest that it is desirable to replace the networks by statistically estimated parametric equations. How such a model might look, and how it could be developed and used, are best demonstrated by an example. Note that the example below does not incorporate all the variables we might desire in a more general specification.

The example of parametric supply models is a model for the zonal "drive time to a station."[1] The explanatory variables (options) are: area of the zone (A), arterial spacing (S), number of lanes on arterials (L), trip-end density (DE), distance to the station from the zone centroid (DI), and signal synchronization (SD = 0 if yes; = 1 if no). The model was created in the following way: a set of fifty random drawings was made for each of the over 150 different settings of the options, the mean and the variance of travel time was computed for each of the 150 settings of the options. The means and standard deviations were then regressed against the values of the options. The models had the following coefficients and other relevant statistics.

Model form: $y = ae^{bx}$

(7) Mean (Time) =1.08 + .04A + .76S - .35L + .06DE + .28DI + .04SD

t-values	(6.9)	(3.3)	(6.3)	(13.0)	(18.4)	(4.1)

$R^2 = .90$ Coefficient of Variation = .10 F = 193.5

(8) Std Dev = .27 + .05A + 1.045 - .51L + .08DE + .27DI + .l0SD

(8.8)	(17.4)	(8.8)	(7.7)	(17.8)	(11.2)

R = .88 Coefficient of Variation = .31 F = 185.4

[1]This model is from A. Talvitie and T. Leung, "A Parametric Access Network Model" (to appear in the TRB Record). A simplified version of the model was developed by Talvitie and Dehghani (presented in OSRA/TIMS meeting in Las Vegas, 1975) and applied to Chicago data by Talvitie and Templeton (paper to appear in TRF Proceedings, 1976).

The null hypothesis that "drive time to a rail station" is lognormally distributed with mean given by equation (7) and standard deviation by equation (8) could not be rejected at the .95 level-of-confidence.

When models such as in equation (7) are used for computing travel times no networks need to be coded. Travel time is derived directly from the definition of the transportation system and its operation--from the options. Because these equations can be developed for every component of a door-to-door trip, access modes can be explicitly considered. Travel times obtained from the equations also cover entire trips or independent components of a multimode trip. Thus, if the models are developed using observed data on trips, the data incorporates the interactions between consecutive links. Such an assumption, made by network models, is untenable in congested conditions.

Transportation policy analyses, including transportation system management policies, can be performed in a natural and reasonably rapid manner using supply models based on equations.[1] This is because the coefficients of the supply models indicate how much travel times would be improved in response to certain measures: e.g., signal synchronization; the demand effects of travel time improvements can then be traced using travel demand models.

The supply model in the example is also sensitive to the size of traffic zones. This facilitates the use of large traffic zones in planning and policy analysis. The effect that zoning practices (i.e., intensity and type of land use) may have on travel time, and hence to travel demand, can also be analyzed through the trip-end density variable. The supply model in the example also is the result of the statistical estimation allowing an explicit assessment of its accuracy.

Finally, it may be noted that a distribution of driving time is attached to the model; this permits the use of the model not only in making a traditional aggregate travel forecast but also in making a "disaggregate forecast" (Domencich and McFadden, 1976). Because the distribution and its parameters are defined, disaggregate values of driving time to station can be associated with each individual by drawing their values randomly from the distribution.

It is appropriate to conclude this section by noting that the method used to create models in equations (7) - (8) may be deceptively simple. The model

---

[1]The word "supply" will be used from now on to characterize the technological relationship between travel time (or cost) and the system definition. It is not the industrial supply curve of economists, but its use in transportation literature is a convention.

certainly needs improvement, if not in accuracy, then at least in complexity. Some of these complications will be discussed and introduced in the sections that follow. There is one important problem--perhaps a whole area of problems–that hinders the development of aggregate supply models such as the one described above. This is the lack of statistical analysis of how much individual trip travel times are improved as a result of various transportation improvements. Too much of the traffic engineering measurement and analysis is based on spot speed, spot volume, and single intersectlon studies. There exists a real need for traffic engineering studies where travel times over reasonably long distances are examined in various road, volume, and traffic control conditions. The models, not only for access but in particular for linehaul, suffer from the lack of such studies.

Method and Assumptions in Detail

Line choice model needed first

An essential prerequisite for the development of supply models is a line/path choice model. Such a model might look as follows:

(a) Prob (line) = g_1 (Headway, Walk time, On-vehicle time, Price, No. of Xfers, S) + Unobserved Reliability, Seat Availability, Security, Safety, etc.

where S denotes socioeconomic attributes.

This model can be estimated using logit/probit techniques. Other than data problems, the main difficulties with the model are associated with the definition of choice set and with the assumption of independence from irrelevant alternatives. For example, does the choice set for line choice include transit lines from both rail and bus modes, or can the choice between lines be cast in a recursive relation to the mode choice? Is "express bus" a "mode" in its own right or is it just a "line" within a mode? If "express bus" is a line, is it a line within "bus" or "rail" modes? That is, is express bus more similar to conventional bus than rail or vice versa? For the purposes of supply model development, a position is taken here that bus and rail modes exist and are commonly understood to preserve the traditional mode choice problem. If it turns out that scheduled transit is a collection of competing lines, this will not cause any great harm to the supply models. The important thing is the concept of line choice from among competing lines.

Within a bus mode, the alternative lines are likely to be quite similar. Thus, the observed attributes should contain the main causes underlying the choice and the unobserved attributes should be restricted to a few randomly occurring causes which to a great extent are specific to lines and thus independent between alternatives. Equation (9) is believed to be such a model. Armed with such a model, the development of the models for the distributions of the access attributes may be done in the following way.[1]

[1]Even if the line-choice model does not exist, a model developed in the context of traditional auto-bus mode choice problem has most of the coefficients needed for the line choice model; such a model is used here.

Method in detail

Consider a traveler who has an origin (residence) at 0 and destination (work, shopping) at D . The traveler at 0 going to D is served with four bus lines Hl, H2, H3, and H4 that are identical, with the exception of their headways Hi (see Figure 16). There are four feasible paths from 0 to D : 0-1-4-5-D; 0-2-6-D; and 0-2-6-5-D (path 0-6-etc. is dominated by other paths). The choice among these paths depends on the headways H1, the walk distances to and from bus lines, riding times on bus on the various lines, fare, the number of transfers and socioeconomic attributes of the traveler. A model embodying a decision rule for this type of choice was shown in equation (9).

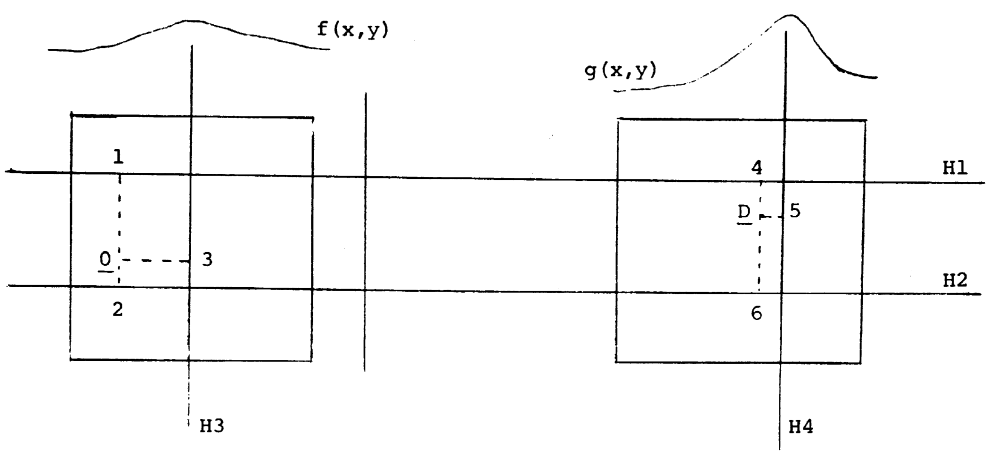


FIGURE 16 Hypothetical Example of Bus Line Choice

Once we know the path the individual traveler will take, we can observe his travel time components, headways, and so forth. By repeating this experiment for several randomly drawn individuals from the population distribution f(x,y) to employment distribution g(x,y) and given an identical transport system between them, we can construct the distributions of the travel time components for travel (by bus) between these two zones. By changing the transportation system or its operation between these zones and repeating the experiment for another random set of individuals (in effect, this is integration using the Monte Carlo technique), these distributions can be related to the underlying transportation system and the development densities within the origin and destination zones. These distributions may then be used in forecasting travel because they are functions of the underlying transportation system; by changing the transportation system the descriptors of these distributions (e.g., mean variance) will change, and so will travel demand.[1] There are many difficulties in implementing the above approach; simplifications and assumptions made in this work are described next.

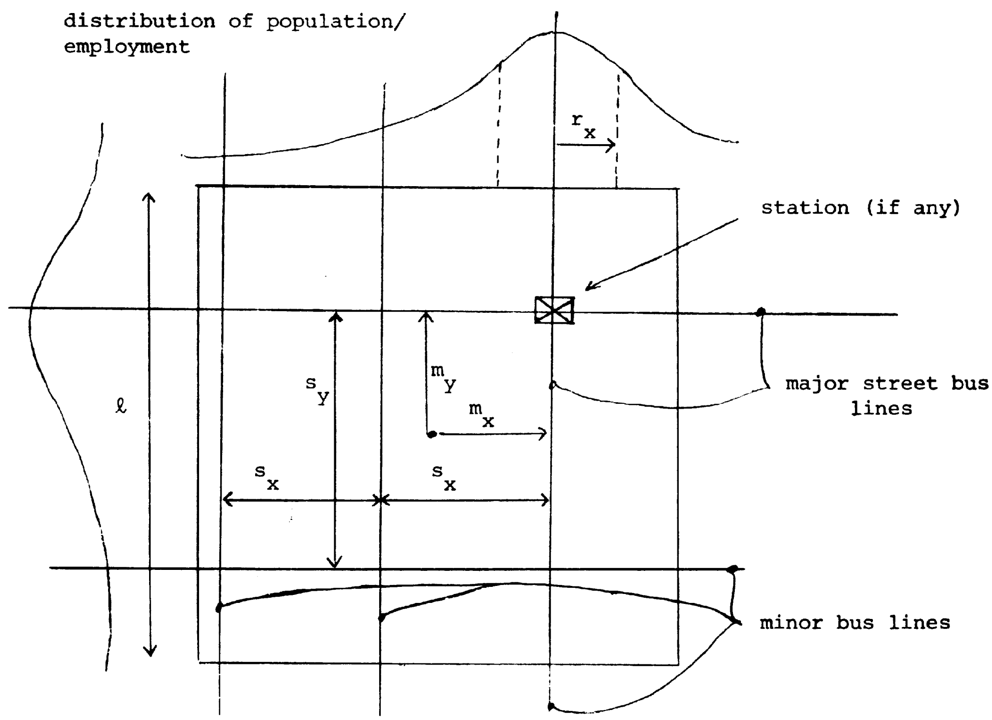
Simplifications and assumptions

The first simplification is implied by the title of this section, and it is the separation of access and linehaul components. For the purpose of access supply model development an abstract "basic traffic zone" is used. The assumptions regarding the basic zone are the following. The "basic zone" is one to sixteen square miles in size and within it (or in its immediate vicinity) there exists at least one major street. There can be one major street in both x and y directions but this is not necessary. The coordinate distances from the zone centroid to these major streets are m_x and m_y. If there is a guideway (BART) station within or near this zone, it is located in the intersection of these major streets. The spacing (s_x , s_y) of bus routes within the "basic zone" is uniform; it can be different to x and y directions. The spacing of bus stops must be equal for a given direction. The frequency of buses can be different in x and y directions, but the headways on all minor street bus lines must be equal in a given direction; for the model all headways are expressed as a difference to bus headways on the major street in the

[1] The technique described here has enormous potential. Consider that the random points of origin (residence) and destination (jobs) for work trips were given by a land use model (S), and the destination points for non-work trips by the travel demand models, one could then develop parametric supply models for entire cities. These models would facilitate a rough comparison of different city forms in a model environment where all the submodels are interdependent. The technique can also be used to create supply models for such technologies as taxi and dial-a-ride, and many other purposes.

y direction. The population/employment is distributed about the intersection of the major streets. This distribution is assumed to be independent normal in both x and y directions; a uniform distribution can also be used, as will be explained later. There also can be a "hole" in the population/employment distribution of radii r_x , r_y about the center of the distribution (m_x , m_y). Finally, all the traffic originating from the zone is assumed to go into a point, such as the (BART) station. This means that the choice of transit line at one end of a trip is independent of what happens at the other end. Figure l7 describes the situation. Before discussing these assumptions, initial model specifications of the transit access models will be presented.

FIGURE 17 Schematic Diagram of Basic Traffic Zone



Variable Definitions and Initial Model Specifications

The following variables will be used to define the "basic zone" and its transportation system; some combinations of them will be used as independent variables in the access supply model. The range of the variables in the simulation is also given below.

TABLE 54 Definitions of Policy Variables and Their Ranges

ℓ -	side length of the square zone; range 1-4 miles
m_x, m_y -	distances from the zone centroid to the major streets x and y directions respectively; range 0.0 - (ℓ/2+1) miles
v_x, v_y -	standard deviations of population/employment distributions within the zone, with mean m_x, m_y ; range 0.5 - 6.0 miles
r_x, r_y -	radii of a possible empty "cylinder" about the center of the distribution; range 0.0 - 0.6 miles
s_x, s_y -	spacing of bus routes in x and y directions; if no buses to either direction use min(2ℓ,3) as the spacing in that direction, do not use zero; range 0.5 - 3.0 miles
H_x -	headway differences on bus lines operating on major street to x direction (minutes; H_y assumed to equal zero); range 0 - 15 minutes
h_x, h_y -	headway differences in minor street bus lines with respect to H_y ; range 0 -30 minutes
b_x, b_y -	distance between bus stops in bus routes; range 0.125 - 0.5 miles
dummy -	1 if station outside of zone

Four models will be developed for transit access using the method and assumptions described earlier. They are: Walk Time to Bus Line; Walk Time to Guideway (BART) Station; Drive Time to Guideway Station; and Ride Time in Bus to Guideway Station. (In general, "guideway station" can be regarded as the center of the population/employment distribution, not necessarily as a station.) The models can be applied at several different levels of complexity, depending upon the completeness of available information, by using "default" values or averages for the missing information. Three different levels of complexity can be envisioned:

- LEVEL 1 (complex) Information exists regarding:
  - location of guideway stations, bus route spacings, and bus stop intervals
  - bus headways
  - population/employment density distributions within zones

- LEVEL 2 (intermediate) Information exists regarding:
  - location of guideway stations, bus route spacings, and bus stop intervals
  - bus headways

- LEVEL 3 (simple) Information exists regarding:
  - location of guideway stations, bus route spacings, and bus stop intervals

The following general mathematical form will be used for estimating the coefficients of the access supply models.

$$DV = a_0 + a_1CO + a_2ST + a_3SP + +a_4L + a_5HD + a_6HX + a_7BMI + a_3SD + a_9HO + a_{10}CE + a_{11}LHT\ a_{12}DU \quad . \qquad (10)$$

The variables of equation (10) are defined in Table 55.

TABLE 55 Definitions of Explanatory Variables

Variable	Mnemonic	Definition	
CO	coordinates	$m_x + m_y$:	sum of the stations or major bus line coordinates (measured from the center of the zone), miles
ST	stops	$b_x + b_y$:	sum of bus stop intervals in each direction, miles
SP	spacing	$s_x + s_y$:	sum of bus route spacings in each direction, miles
L	ride	ℓ :	side length of zone miles
HD	headways	$h_x + h_y$:	sum of minor bus line headways , minutes
HX	headway	H_x :	major bus line headway to x-direction ($H_y = 0$), minutes
BMI	bus miles	$60 * \ell^2 (\frac{1}{s_x h_x} + \frac{1}{s_y h_y})$:	bus miles of service within zone (over and above offered by major line H_y)
SD	standard deviations (of development density)	$v_x + v_y$:	sum of the development density standard deviations
HO	hole	$\frac{r_x}{v_x} + \frac{r_y}{v_y}$:	note, if station outside the zone $r_x = \frac{m_x - \ell/2}{v_x}$ and $r_y = \frac{m_x - \ell/2}{v_y}$
CE	centrality	$\frac{\ell/2 - m_x}{v_x} + \frac{\ell/2 - m_y}{v_y}$:	this variable is zero if the station is right on zone boundary; it is positive inside the zone and negative outside of the zone
LHT	linehaul time	difference in on-vehicle time between the (best) major bus line and (best) minor bus line[1]	
DU	dummy	0 if station inside the zone; 1 if station outside the zone	

[1]The linehaul times were computed using the methods of linehaul time computation to be explained in the next section. Briefly, it was assumed that 5 passengers boarded at every stop and that the average loading time was 3 seconds; a bus speed of 10 mph was assumed.

Note that the summation of the variables simply means that the coefficients of the summed variables are constrained to be equal. In order to approximate a uniform population/employment distribution, set the standard deviation equal to $v = \ell + 1$ where v is the standard deviation of the population/employment distribution ($v = \ell + 2m/\ell + .5$, where m is the station coordinate, is an even better formula but is a little more complicated). Either one of these *ad hoc* formulae ensures a nearly uniform distribution, as the reader may verify by plotting the distributions.[1] Thus, in general, the standard deviation of population/ employment distribution should take the value of $v = \min(v; \ell + 1)$ or $v = \min(v; \ell + 2m/\ell + .5)$; do not use v greater than six.

An explanation may be in order for some of the independent variables in equation (10). Basically equation (10) remains simply a linear combination of the "planning options." There are, however, three variables that are constructions. They are: bus miles of service (BMI), "hole" or undeveloped area around the BART stations (HO), and centrality of the population/employment distribution (CE). Of these, bus miles of service is intuitively clear and requires no explanation; "hole" (HO) and centrality (CE) can be justified and understood when it is recalled that the individuals are drawn randomly from the normal distributions for employment/population and that nobody could be drawn either out of the traffic zone or from the empty area--if one was specified--about the BART station. The effect of those excluded can be accounted for by entering the normalized excluded (or included) area of the density function into the model.

---

[1]There is a good reason for adopting an *ad hoc* formula for standard deviation instead of blindly assigning a large value of, say, six for standard deviation to simulate uniform distribution in any size zone. The reason is that for one square mile zone $v = 3$ or $v = 6$ are practically equally uniform; thus zone size-independent choice would have meant loss in accuracy and sensitivity with respect to population/employment distribution.

Discussion of the Model Assumptions

Two types of assumptions were made in developing the transit access supply models: those relating to the "basic traffic zone" and those to the specification and mathematical form of the model.

The major assumptions regarding the "basic traffic zone" were the following: assumed grid system of bus routes and uniform spacing within the zone; equal headways on "minor bus routes"; omission of transfers, fare, and egress system from bus line choice considerations, and the shape and center of population/employment distributions.

The grid system is assumed not only because it simplifies analysis but also because it appears to correspond to the "on the ground" conditions more than other types of abstractions (e.g., radial). A visual review of the Bay Area bus routes, where unusual topographic conditions should encourage "funny" routing, reveals that, despite some diagonal routes and loops at the end of the line, the grid pattern is surprisingly realistic. The same holds for the uniformity of spacing when traffic zone sizes are of reasonable size, e.g., four to sixteen square miles.

The assumption of equal headways on "minor" street bus lines creates somewhat of a problem. But, given that zones are not unreasonably large, the assumption seems realistic. Transfers and fare and egress were omitted from the line-choice model in the interest of reducing model complexity. Fares on competing bus lines are often equal and cancel out. Mode-choice models developed so far have been unable to generate a plausible coefficient for the number of transfers by increasing the values for appropriate headway; however, it is possible to account for transfers, if they are known to exist. Finally egress was omitted not only in the interest of keeping the model reasonably simple but also because destination within zones was assumed to be independent of origins in any case.

The (normal) population/employment distributions are centered on the intersection of the major arterial streets (or guideway station) serving the zone. This is somewhat unrealistic, especially in the suburban areas. These distributions can, however, have a "hole" in the middle (through r_x, r_y), which alleviates the problem. The reason for installing such variables as (r_x, r_y) is that in a typical BART station there seems to be "an empty hole" (due to parking or commercial activity around the station); thus access travel times computed with the "hole" left in would be less than actual travel times. The assumptions also imply that there may be a bus line running parallel and next to the guideway. It may sound odd, but this appears to be true very often, even though the bus line may be a block or two removed from the guideway. This bus line may be eliminated by assigning it a large headway, such as sixty minutes.

The most important contribution of the inclusion of population/employment densities into the model is that it enables the examination of transportation consequences of zoning practices. Such capabilities are wholly lacking currently. It should be noted that UMTA's proposed decision criteria with regard to financing major mass transportation projects[1] mentions "densifying selected corridors" and other land use development actions as important considerations. The problem with the inclusion of development densities into the model has, therefore, little to do with whether they ought to be in the model; rather it is that data do not exist to suggest distributions for within-zone population/employment densities or to test the accuracy of the results obtained. The Normal distribution was adopted here because it is well known and understood, and because it is easy to manipulate.

The assumptions regarding the access models' specification concern mainly the linear-in-parameter mathematical form and the exclusion of certain simultaneous relationships from the equation.

The linear-in-parameters mathematical form of the supply equations falls short of producing elegant approximations of the exact travel times which could be obtained by integrating over the zone. As a simple example, consider the walk time to station:

(11) $$WT_x = \frac{k_c}{v_w} \int_x \int_y \sqrt{x^2 + y^2}\; f(x,y)\,dx\,dy$$

k_c = circuitry of sidewalk network

The strength of the linear equations lies in the ease with which they can be estimated and in their simplicity--(much simpler than what equation (11) would yield for the average travel time (Kocur and Ruiter, 1975).

---

[1] Proposed Decision Criteria Governing UMTA Commitments to Construction Financing of Major Mass Transportation Projects, DOT/UMTA, March 1976.

Correlations that may (and probably do) exist between the location of individuals with respect to transit services and their socioeconomic attributes and tastes are not incorporated in the models. In addition, there probably are simultaneous relationships also between development densities and the transit level of service. While these elements are present in the models, and as such represent an advance in the state of the art, they are independent of each other. It is believed that this independence is the most serious drawback of the described models.

In conclusion, we venture to suggest that the proposed access supply models do capture the salient features involved in determining the distribution of access times within traffic zones. These models accomplish that in a manner that is sensitive both to the transportation system serving the zone and to the distribution of activities within the zone. The shortcomings of the models can be eliminated to a large extent when the various simultaneous relationships are better understood; for now, they can be alleviated by a careful use of the models.

The results of the access supply model estimations follow.

The Estimated Models

Walk and drive time to guideway station

Zonal walk and drive travel times to a given station are similar in the sense that the distance to be covered is, by and large, the same. For this reason distance, rather than travel time, is regressed against the planning options. To convert distance to travel time, an appropriate traveling speed must be used. For walking, three mph is a good default value; car values may be chosen from the CUTS manual (1975), if local studies have not been conducted.

The following model specification was first tried for walk or drive time to guideway station.

$$\text{Walk/Drive Time} \cdot \text{Speed} = a_0 + a_1 CO + a_2 L + a_3 DU + a_4 SD + a_5 CE + a_6 HO \quad , \tag{12}$$

where the variables are as explained earlier.

In estimating model (12) it was quickly realized that centrality of the station (CE) and the dummy variable (whether station is inside or outside of the zone boundaries) are, to a great extent, measuring the same effect. For this reason, if one is entered, the other is not; also, and for the same reason, two different models are developed. The centrality variable (CE) is theoretically more pleasing, but the dummy variable (DU) is easier to use and appears to be more powerful.

The estimated models for the mean and standard deviation of walk/drive distance are given in Table 56 at each level of complexity. It may be observed that both models have statistically highly significant coefficients and possess about equal forecasting ability ($R^2 > .85$ and coefficient of variation $< .20$). It does appear interesting that the development density variables (SD, HO) have a strong bearing on walk/drive access times in addition to the contribution made by such traditional variables as location of the station and size of the zone. For example, in a four square mile zone with station in the center, the average zonal walk time to the station is eighteen minutes given uniform density of origins/destinations in the zone. In this case twenty-five percent of the origins and destinations are within a one mile zone centered about the station. If we increase the development densities in such a way that, say, fifty percent of the origins/destinations are within such a one square mile zone (standard deviation of the density is dropped to 1.0) then the average walk time drops to approximately seven minutes. Thus the density of development can have a dramatic effect on the access walk (or drive or bike) travel times to a station.

TABLE 56 Walk/Drive Time to Station•Speed = f(variables)

	BASIC				MODIFIED			
Variables	MEAN coeff	MEAN t-value	STDEV coeff	STDEV t-value	MEAN coeff	MEAN t-value	STDEV coeff	STDEV t-value
CONST	-.241	(3.3)	-.0470	(1.7)	-.271	(4.6)	-.0271	(1.1)
CO coordinates	.381	(34.6)	.166	(48.6)	.272	(32.1)	.175	(48.4)
SD std dev	.126	(8.6)	.0211	(3.9)	.138	(11.2)	.0159	(3.0)
HO hole	.488	(7.1)	-.198	(7.8)	.298	(4.7)	-.176	(6.6)
CE centrality	-.154	(6.5)	.0369	(4.2)				
DU dummy					.563	(12.6)	-.0949	(5.0)
L side	.318	(34.6)	.166	(48.6)	.272	(32.1)	.175	(48.4)
F-value	426.5		703.4		593.8		720.3	
R^2	.85		.90		.89		.91	
CV	.19		.16		.16		.15	
n	300		200		300		300	
Std error of estimate	.28		.11		.24		.10	
Mean of the dep var	1.53		.67		1.53		.67	

Finally, it may be noted that the model can also be used to obtain biking times to a station by using an appropriate biking speed and, in some situations, it can be usefully applied to calculate "walk to bus stop" travel times; a discussion on this last application is deferred the section dealing with such a model.

Walk time to bus stop

In developing the walk to bus stop model it was assumed that the travelers walk at a constant speed of three mph and choose their bus line according to the mode choice model developed in Part II, Chapter 2 (the maximum utility line being chosen by the traveler).

Experimentation with the general specification equation (1) showed that standard deviation of the development density did not materially affect the walk times to bus stop. It was also found that both the dummy variable (DU) whether the ultimate trip destination, such as a BART station, is in or out of the zone, and the on-vehicle time difference between competing bus lines tend to measure the same effect. This finding makes (limited) intuitive sense; its consequence is that the development density distributions do not affect Walk Time to Bus Stop models. The following specification was finally adopted for the model:

$$\text{(13)} \quad \text{Walk Time to Bus Stop} = a_0 + a_1\{\text{LHT or DU}\} + a_2\text{BMI} + a_3\text{HD} + a_4\text{HX} + a_5\text{ST} + a_6\text{SP} \quad ,$$

where the variables are defined in Table 55.

The results in Table 57 indicate that the three estimated models are statistically highly significant. The R^2 's are approximately .45 and the coefficient of variation about .25 . In general, the models for the mean of the walk time are better than for the standard deviation of the walk time. Although the R^2 's are lower for these models than for the walk/drive time to station models in Table 52, the predictive accuracy of the models is not worse in terms of root-mean-square error; the models can therefore be used with reasonable confidence.

TABLE 57 Walk Time to Bus Stop = f (variables)

VARIABLES	BASIC		MODIFIED BASIC A		MODIFIED BASIC B	
	MEAN	STDEV	MEAN	STDEV	MEAN	STDEV
LHT	.0697 (.88)	.0952 (1.7)	----	----	----	----
DU	----	----	.562 (1.6)	.846 (3.4)	.549 (1.5)	.815 (3.0)
BMI	.00129 (1.7)	.00311 (5.9)	.00137 (1.9)	.00323 (6.3)	----	----
HD	.0933 (9.6)	.0477 (7.0)	.0931 (9.8)	.0477 (7.2)	.1017 (12.2)	.0678 (10.9)
HX	.0409 (2.4)	.0126 (1.0)	.0455 (2.7)	.0190 (1.6)	.0473 (2.8)	.0233 (1.8)
ST	5.541 (5.9)	1.005 (1.5)	5.472 (6.0)	.922 (1.5)	5.447 (5.9)	.864 (1.3)
SP	1.086 (6.6)	.557 (4.8)	1.158 (7.7)	.656 (6.3)	1.114 (7.5)	.552 (5.0)
CONST	.891 (1.1)	.711 (1.2)	1.004 (1.4)	.830 (1.6)	1.179 (1.6)	1.241 (2.3)
F-value	41.4	31.6	41.9	34.0	49.1	28.7
R^2	.49	.42	.49	.44	.48	.35
CV	.22	.30	.22	.30	.22	.32
n	267	267	267	267	267	257
Std error of estimate	2.20	1.56	2.20	1.53	2.21	1.64
Mean value of dep variable	10.1	5.1	10.1	5.1	10.1	5.1

Finally, two reasons for caution. First, while the dependency of bus walk time and bus headways is an important link to establish, the modeler needs to be careful with applications because bus headways are affected by bus volumes (Morlok, 1974) and because traditional forecasting models operate on a single zonal headway figure. Another reason for careful use of the models has to do with the previously noted, and somewhat troublesome result that development density distributions do not affect bus walk times. According to the models in Table 57, a zone in which all activity is located along a single street where a bus line also operates should have walk time equal to that within a zone, served by a single bus line, where activities are evenly dispersed. This type of counter-intuitive result points to the need to apply the models carefully; for example, in the extreme case of one bus line along a densely developed street, one might consider using the equations in Table 56 instead.

Bus ride time to guideway station

The choice of bus line(s) along which the distance to the guideway station is measured was based on the same mode choice model as in the walk time to bus stop model.

The following model specification was found appropriate in this model.

$$\text{(14)} \qquad \text{Ride time} = a_0 + a_1 SD + a_2 \begin{Bmatrix} CE \\ \text{or} \\ DU \end{Bmatrix} + a_3 HD + a_4 HX + a_4 L + a_6 CO \quad .$$

The estimated coefficients and the goodness-of-fit statistics for the ride time model appear in Table 58. All the models are statistically highly significant; the same applies to most of the independent variables. For example, the larger the standard deviation of development density, the longer the bus ride distance; and, the larger the headways on the minor bus lines with respect to the major (base) bus line in the zone, the shorter the bus ride distance. These both are plausible results. Large standard deviations mean more dispersed activities and longer rides. Negative coefficients on headways indicate that bus travelers are willing to walk a longer distance to gain shorter headways--and probably gain in the linehaul distance at the same time.

TABLE 58 Bus Ride Time to Station*Speed = f(variables)

VARIABLE	BASIC				MODIFIED			
	MEAN		STDEV		MEAN		STDEV	
SD	.117	(5.9)	.0335	(4.0)	.150	(8.9)	.0401	(5.2)
CE	-.0985	(4.1)	-.0211	(2.1)	-----		-----	
HX	-.00198	(.71)	.00172	(1.5)	-.00217	(.84)	.00164	(1.4)
HD	-.00691	(4.8)	-.000581	(1.0)	-.00620	(4.7)	-.000554	(.91)
DU	-----		-----		.942	(8.2)	unstable	
L	.252	(7.2)	.209	(14.2)	.281	(10.0)	.188	(17.6)
CO	.471	(18.2)	.132	(12.1)	.275	(7.4)	.147	(17.6)
CONST	-.0578	(.54)	-.0820	(1.8)	-.282	(2.9)	-.0991	(2.2)
F-value	241.3		218.0		293.0		257.4	
R^2	.85		.83		.87		.83	
CV	.25		.20		.23		.20	
n	267		267		267		267	
Std error	.37		.15		.34		.15	
Mean of dep variable	1.49		.79		1.49		.79	

The statistical indicators give equally high R^2's (about .80) and equally low standard errors of estimate (coefficient variation about .25) to all the models. In the statistical sense, the models are equivalent. In applications, however, caution should be exercised when using the models with the in/out zone dummy variable (DU); the dummy indicates that, no matter how far out of the zone the station is, bus ride distance immediately increases by one mile when the station moves out of the zone. It is the authors' feeling that this discontinuous jump is too sharp. That, however, is the result that was obtained.

Development of Models to Forecast the Travel Times as Linehaul

Models for auto travel times

In principle, the auto travel time models could be estimated statistically in the same manner as the access travel time models. However, time constraints did not permit such an undertaking. Instead, the driving time models are based on existing sources of information.

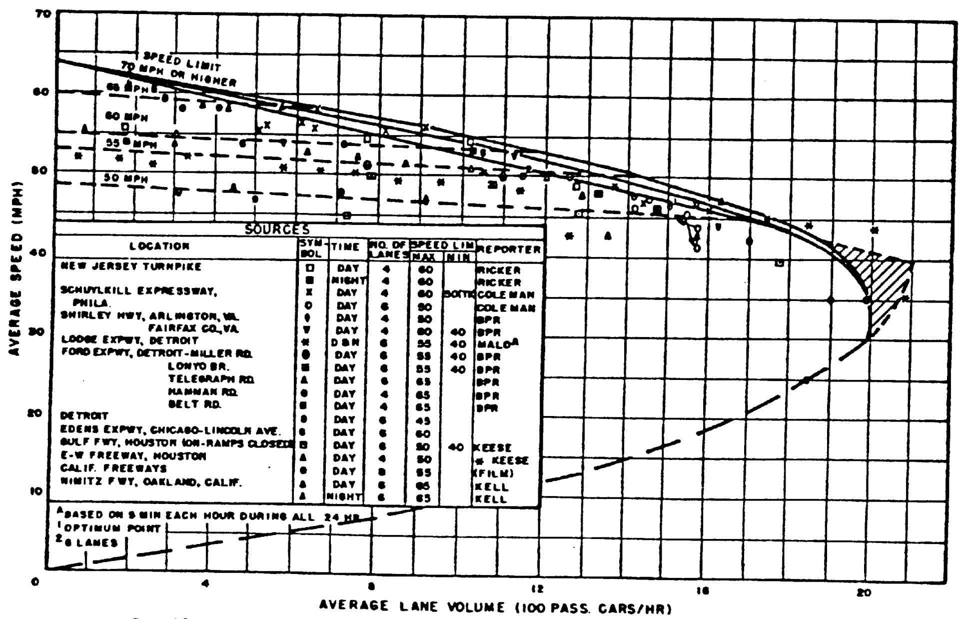
For a given capacity, speed-volume relationship is one of the basic concepts in traffic engineering; an example of such a curve is shown in Figure 18.

The speed-volume curve in the figure, developed before the fifty-five mph speed limit law, implies that travel times in non-congested conditions vary only marginally. More important, the traffic engineer's speed-volume curve is valid only for a uniform stretch of highway and for conditions when the volume is below the capacity of the road.[1] Given that below capacity the speed of travel does not differ substantially from the current fifty-five mph speed limit, and that the speed-volume relationship is invalid for situations where the volume exceeds capacity and where large delays can occur because of queueing, it is appropriate to seek other means of representing driving time. This is particularly true when policies for priority treatment of high occupancy vehicles are being considered, which in many cases impose considerable congestion delays upon the non-priority vehicles.

---

[1]A good discussion of the shortcomings of the speed-volume curve is given by Small (1976). He also presents results of an attempt to estimate speed-volume curve for a non-uniform stretch of road. Even though the obtained showed somewhat greater sensitivity of travel time to volume/capacity ratio, it still showed considerable scatter near the capacity point and, of course, did not apply to situations when $v|C > 1$.

FIGURE 18

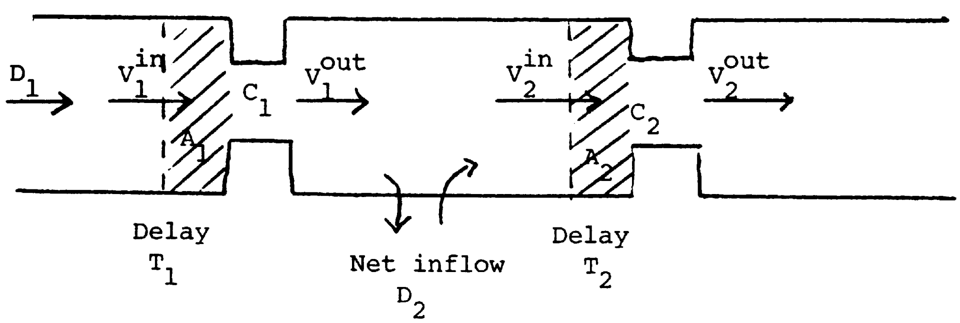


Specific reported speed-volume relationships per lane in one direction of travel under interrupted flow conditions on freeways and expressways.

Point bottleneck model

The point bottleneck model, formulated by May and Keller (1967), takes the approach that the effects represented by the upper part of the speed-volume curve can be ignored altogether and concentrates on the queueing delays that occur when capacity is exceeded. They assume also, quite appropriately, that the "arrivals" to the queue are non-random, as is the rate of service (or capacity). This type of model can be characterized as a deterministic queueing behind a bottleneck; it is graphically illustrated in Figure 19 for the case of no priority operations. This example is borrowed from Small (1976); another version appears in May and Keller (1967).[1]

FIGURE 19 Point Bottleneck Model



[1] For detailed assumptions included in this model see Small (1976) and May (1968).

The road section is assumed to consist of a small number of sections, each having a uniform speed v_0 and infinite capacity with each section ending at a bottleneck of capacity C .

B_i = net demand inflow between sections i - 1 and i (auto-equiv./hour);

V_i = actual vehicle flows, as shown (auto equiv./hour);

A_i = queue length (auto-equiv.);

T_i = queueing delay (minutes);

T = total travel time (minutes);

Dist = length of section (miles);

v_0 = free speed (miles/hour);

t = time of day (minutes past midnight).

Then, with the convention $V_0^{out} = 0$, the equations describing the system are:

$$V_i^{in}(t) = V_{i-1}^{out}(t) + B_i(t) \quad ;$$

$$dA_i/dt = \begin{cases} 0 \quad , & A_i(t) = 0 \quad ; \\ (V_i^{in}(t) - C_i)/60 \quad , & A_i(t) > 0 \quad ; \end{cases}$$

$$V_i^{out}(t) = \begin{cases} V_i^{in}(t) \quad , & A_i(t) = 0 \quad ; \\ C_i \quad , & A_i(t) > 0 \quad ; \end{cases}$$

$$T_i(t) = 60 \; A_i(t)/C_i \quad ;$$

$$T(t) = \sum_i T_i(t) + 60L/v_0 \quad .$$

For the case of only one bottleneck or independent bottlenecks, these equations become easy to handle. May and Keller (1967) solve them for the case of trapezoidal-shaped peak-hour demand pattern. In this study a rectangular shape with zero off peak demand is assumed. Queueing delay x is shown to be, at any time t ,

$$x(t) = \begin{cases} 0 & t \le t_1 \\ (D/C - 1)(t - t_1) & t_1 \le t \le t_2 \\ (D/C - 1)P - (t - t_2) & t_2 \le t \le t_3 \\ 0 & t_3 \le t \end{cases} .$$

where t_1 and t_2 are times at the beginning of the peak period, $P = t_2 - t_1$ is the length of the peak period, and t_3 is the time at which the queue is dissipated. The average travel time during the peak is then

$$T = \frac{Dist}{N_o} + \frac{1}{P}\int_{t_1}^{t_2} x(t)dt = \frac{Dist}{N_o} + \left(\frac{D}{C} - 1\right)\frac{P}{2} \quad . \tag{15}$$

Both Small (1976) and May and Keller (1967) have used the model with apparent success despite its simplicity.

The power of this simple model is that it can be used easily in analyzing priority treatment policies for high occupancy vehicles by dividing the available capacity into two "roads" (Small 1976). The model can also be made to represent a collection of parallel and sequential roads and is thus adaptable for service in an equilibration framework which is cast as a simultaneous equation problem. These extensions of the model are discussed in Chapter 8 of this part.

Models for transit on-vehicle time

The models for transit on-vehicle time will now be described. All are based on a very simple relation: travel time on a transit vehicle is equal to the distance divided by the speed of travel, plus the time incurred in picking up or discharging passengers. The difficulty in implementing this equation in a sensible manner arises from the fact that the speed of a transit vehicle, especially that of buses, is dependent upon many factors ranging from drivers' skills to maneuver the bus under varying volume conditions to the number of cars parked on the street sections constituting the bus line.

The equations developed below are not to be used without careful transportation engineering judgment; in this they are not unlike any other model employed to abstract a complicated reality. Specifically, it is assumed that the transportation engineer specifies the type of system and the operating rules between the two points (or zones) between which the travel time is desired. If this is done well the equations should yield travel times which are comparable or even better than the times obtained from a coded network. Furthermore, the equation for bus travel time is such that it can be used for obtaining an equilibrium (volume dependent) travel time.

The travel time equations for BART are developed first, followed by travel time equations for bus.

BART travel time. Travel time over a distance is, of course, equal to the distance divided by the average speed of the vehicle. It is assumed here that the average speed on BART, $\bar{v}_r$, is a function of the following variables: maximum speed, stop frequency, and acceleration/deceleration rates. There are two cases, each obtaining a different mathematical expression for the aversge speed $\bar{V}_R$. The first is the case where the station spacing (distance between stations) is sufficient for the vehicle (train) to reach its maximum speed; in the second case the station spacing is not large enough for the train to reach its maximum speed. Denote the average speeds ln these two cases $\bar{v}_r^1$ and $\bar{v}_r^2$. The expressions for the speeds are as follows.

(16a)
$$\bar{v}_R^1 = \frac{3600S}{\frac{v_m}{2a} + \frac{v_m}{2a'} + \frac{3600S}{v_m}}$$

and

(16b)

$$\bar{v}_R^2 = \frac{3600}{\frac{7200(a + a^{/})^{1/2}}{Saa^{/}}} ,$$

$$S < \frac{v_m^2}{7200a^{/}} + \frac{v_m^2}{7200a^{/}} ,$$

where

$\bar{v}_R^1$, $\bar{v}_r^2$ = average speeds as defined earlier (mph);

a = acceleration rate (constant mphps);

$a^{/}$ = deceleration rate (constant mphps);

N_m = maximum cruising speed (mph);

S = station spacing (miles) .

Note that the above formulae are for average moving speeds. It would have been possible to obtain an expression for the average overall speed by adding dwell time to the denominator. However, it appears to be more convenient to simply add the dwell time to the running time later.

Total travel time between any pair (AB) of stations can then be computed from the following equation:

(17)

$$T_{AB} = \sum_{i=1}^{N-1} \frac{S_{i/\, i+1}}{\bar{v}_{i/\, i+1}} \cdot 60 + \sum_{i=1}^{N} t_{wi}/60 ,$$

I = 1 = A station;

t_{wi} = dwell time at station i;

I = N = B station.

Applying equation (17) to the Concord-Daly City line, for example, between Concord and Daly City stations the following results are obtained:

for $v_m = 70$ mph , $T_{AB} = 37.7 + 9.0 = 46.7$ min ;

for $v_m = 55$ mph , $T_{AB} = 44.9 + 9.0 = 53.9$ min .

The tabulated travel time between Concord and Daly City is fifty-six minutes.

It is attractive to simplify the equation (17) in two ways. First, assume that the average speed between stations A and B is a function of the average station spacing between A and B. Second, assume a flat dwell time of, say, thirty seconds (a figure often observed casually) and add sixty seconds for a transfer. With these assumptions in mind, equation (17) has been applied to various trips on BART and the results are shown in Table 59. Examination of the values in Table 59 indicates that the tabulated travel times and the travel times derived using equation (17) differ by less than two minutes, the average deviation being around one minute. This certainly is less than any BART rider is able to detect. Equation (17) is thus accepted as dependable in yielding BART travel times.

TABLE 59 Travel Times in BART Using Developed Equations, and Time Tables. A Comparison.

Maximum reachable speed (mph)	Concord Montgomery St.		Walnut Creek Fremont		Berkeley Powell		12th St Oakland Daly City		Daly City Montgomery		Pleasant Hill Rockridge		Fremont 12th St Oakland		Richmond Mac Arthur Blvd	
	Average Speed (mph) $v(\bar{s})$	Travel Time[2] T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}	$v(\bar{s})$	T_{AB}
70[1]	60.4	34.9	59.6	48.2	56.3	18.9	55.3	21.6	49.3	13.2	61.6	15.5	59.8	28.9	55.5	14.7
55[1]	50.1	40.8	29.6	55.6	47.8	21.5	47.2	24.5	43.7	14.3	50.7	18.3	49.7	33.7	47.3	16.6
Tabulated Travel Time, T_{AB}		40.0		58.0		23.0		26.0		15.0		16.0		32.0		17.0
Total Length (miles)		28.6		38.9		13.05		15.32		7.53		13.33		23.85		10.36
Avg. Spacing (miles)		2.86		2.59		1.86		1.70		1.08		3.33		2.65		1.73
Dwell Time (minutes)		5.0		7.5		3.5		4.5		3.5		2.0		4.5		3.0
Transfer Time (minutes)		1.0		1.0		1.0		0.0		0.0		0.0		0.0		0.0

[1]Acceleration/deceleration rate = 3.0
Average speed is calculated by using equation (1) or (2), using average station spacing.
[2]Transfer and dwell times are included in Travel Time. Travel times in minutes.

Bus travel times. In this section equations are developed for bus on-vehicle time in mixed traffic: that is, the highway or street capacity is shared by autos and buses, which in heavy traffic tend to get in each other's way and slow travel speeds. Besides these volume/ capacity conditions en route, bus travel speeds are dependent on traffic signalization practices, on-street parking policies, one-waying, and other factors such as the presence of pedestrians and other distractions that tend to vary with the location (e.g., residential, downtown) of the route or link within the metropolitan area.

The on-vehicle time on bus may be expressed as

(18)
$$T_B = 60 \cdot \text{Dist} / \bar{v}(s,V/C,F) + t_a \sum_{k=1}^{N} V_k/60 \quad ,$$

where

T_B = bus on-vehicle time (minutes);

Dist = linehaul distance (miles);

$\bar{v}(s, V/C, F)$ = average speed evaluated at bus stop spacing s, volume capacity ratio V/C, and in location F with given parking policy, signalization, one-way conditions (mph);

t_a = boarding/alighting time of one passenger at a bus stop (seconds);

V_k = volume of passengers alighting and/or boarding at stop k (passengers).

The parameters or functions to be determined for applying equation (18) are $\bar{v}(\cdot)$, and t_a. Examining the boarding/alighting time first, for simplicity, it appears that, on the average, passenger service times range from two seconds (single-coin) to more than eight seconds (multiple zone fares collected by the driver) per passenger (12).

The boarding times also appear to be somewhat greater than alighting times, although at least some of the total dwell time must reflect total passenger exchange and "clearance" to get the bus back to the traffic stream. For current conditions it is assumed that regardless of the type of payment the boarding/alighting time is 3.0 seconds; this appears to be in agreement with observations made in the Bay Area and also with values in Table 60.

TABLE 60 Approximate Passenger Service Time On and Off Buses

Operation	Conditions	Time (sec)
Unloading	Very little hand baggage and parcels; few transfers	1.5 - 2.5
	Moderate amount of hand baggage or many transfers	2.5 - 4
	Considerable baggage from racks (intercity runs)	4 - 6
Loading[a]	Single coin or token fare box	2 - 3
	Odd-penny cash fares; multiple-zone fares	3 - 4
	Pre-purchased tickets and registration on bus	4 - 6
	Multiple-zone fares; cash; including registration on bus	6 - 8
	Prepayment before entering bus or pay when leaving bus	1.5 - 2.5

[a]Add 1 second where fare receipts are involved.

Source: "Bus Use of Highway--Planning and Design Guidelines," NCHRP Report 155, Washington, DC (1975)

The main factor affecting bus on-vehicle time over a given distance is, of course, the average bus speed in the traffic stream. As indicated earlier, bus speed is a function of several variables--volume/ capacity, bus stop frequency, signalization, parking policies, number of lanes, etc. A search of the literature on the topic turned up one big surprise: there appears to be a total lack of statistical analysis of bus speeds under varying operating policies and volume/capacity circumstances. For this reason somewhat *ad hoc* methods are employed here to compute the average bus speed and, hence, the bus on-vehicle time.

The average bus speed is obtained from the average car speed (v_C) using equation (15), by assuming that the average car speed is the maximum speed a bus can attain in any given conditions, and that the bus stop spacing is large enough for buses to obtain this maximum speed. Both of these assumptions are reasonable. Average car speed in turn can be obtained from the Highway Capacity Manual using standard assumptions about traffic conditions (e.g., percent trucks, percent right and left turns, etc.) or from the car travel time equations developed earlier for the single "bottleneck" situation. Using the notation defined earlier, and v_c is the car speed in prevailing conditions, the bus travel time can then be expressed as

$$(19) \qquad T_B = \sum_{s \in Dist} 60 \cdot \left(\frac{1}{v_c} + \frac{v_c}{7200s} \right) + .05 \cdot \sum_{k=1}^{N} V_k \quad ,$$

where the first summation is over all intervals between bus stops covering the trip distance. This equation can be modified to express the approximate bus travel time over a distance,

$$(200 \qquad T_B = T_A + \frac{(Dist)^2}{2 \cdot \bar{s} \cdot T_A} + \frac{.05 \bar{V}_k \cdot Dist}{\bar{s}} \quad ,$$

where the new notation is

- T_A is the auto travel time over the distance (Dist) under prevailing conditions (e.g., equation (15) or Table 13);
- $\bar{s}$ average spacing of bus stops over the distance;
- $\bar{V}_k$ average number of passengers boarding/alighting at stops.

In equation (20) the second term expresses delay due to acceleration/deceleration and the third term is the delay due to boarding/alighting passengers.

The use of equation (20) was demonstrated by a test application. Equation (20) was applied to fifteen hypothetical trips between randomly selected origins and destinations. Comparisons were made between tabulated travel times (from timetables) network coded travel times, and the calculated travel times from the supply model. The results are reported in Talvitie and Dehghani (1976).