CHAPTER 3

THE TRIP TIMING DECISION FOR TRAVEL TO WORK BY AUTOMOBILE

Introduction and Summary

The travel time for a specific trip generally depends on the time of day at which the trip is made. During peak periods there is not only a substantial increase in travel time due to congestion, but also a greater variability in travel time and thus a more uncertain time of arrival. The trip may therefore be rescheduled away from the peak, at the expense of reaching the destination earlier or later than the desired time.

Data for the Urban Travel Demand Forecasting Project (UTDFP) sample have been used to investigate whether this effect is measurable in the case of trips to work by automobile. Subjects were asked to give their usual time of arrival at work, as well as their official work start time (if any). In some two-thirds of the sample, trips are timed so that the subject arrives somewhat before his official work start time. The decision as to how much leeway to allow will depend, at least in part, on observable supply variables and, perhaps, also on some observable characteristics of the subject. For automobile trips, in particular, we expect a key determinant to be the amount of traffic congestion that the subject has to contend with at different times. A measure of this congestion is the variation of on-vehicle time for different arrival times; this data had been obtained for the UTDFP sample from network computations augmented with data from floating car runs, which made it possible to parameterize the travel-time curve for each subject (see Reid, Reinke, and Small, 1975).

This data was used to estimate the parameters of a simple empirical model for the probabilities of rescheduling by various amounts. The model may be viewed as expressing the subject's internal tradeoff of time "wasted" through arriving earlier than necessary, versus the increased travel time in the peak (which includes effects such as the unpleasantness of driving in dense traffic) and the probability of arriving late. The estimated parameters can then be used to predict the extent to which travelers will reschedule their trips if congestion time changes, due, for example, to an increase in highway capacity.

We find that there is indeed a measurable, though generally small, effect due to highway congestion times: if congestion times increase, work trips will tend to be rescheduled away from the peak, subject, of course, to the constraint of getting to work on time. The smallness of the effect, and the difficulty in obtaining a reasonably precise estimate of its magnitude, may perhaps be attributed to the following: (1) arrival times for work trips are relatively inflexible, because the worker usually cannot arrive late and there are limited opportunities for other activities if he arrives early, (2) most subjects in the sample traveled over relatively uncongested roads--for which congestion delay in the peak was less than five minutes; (3) the considerable flexibility in empirically selecting a suitable parametric form of the probability function to be estimated. Qualitatively similar results were, however, obtained from several different model specifications.

Although the overall effect of rescheduling is small when averaged over the entire sample, one can identify and analyze special cases where this effect should be significant. Such cases occur at highway bottlenecks---tunnels and major bridges, for example--where queues are formed when demand exceeds capacity. We therefore give a short discussion of a deterministic queuing model, of the type due to May and Keller (1967), but taking into account changes in demand induced by congestion (in this analysis, through rescheduling, rather than through switching to other transportation modes). First, assuming the validity of the trip-timing model, the queuing model predicts a maximum rate of increase in congestion time that is consistent with observed travel time curves. Second, one can investigate the effects of, say, an increase in highway capacity at the bottleneck: we reach the somewhat surprising conclusion that the benefit to travelers (i.e., average increase in utility) is <u>greater</u> than that given by the May-Keller analysis, which assumes an unchanged level of demand. The reason is as follows: although rescheduling into the peak tends to restore congestion delay to what it was before widening the highway, this is more than offset by the gain in utility achieved by travelers arriving closer to their preferred arrival times.

We note that rescheduling can be considered a form of schedule delay (usually considered in connection with public transportation) applied to automobile trips. In part, this schedule delay involves the internal tradeoff made by the traveler, as described above; it becomes significant, however, only when it also involves physical constraints on the system--if travel demand at peak hours tends to exceed highway capacity, then equilibrium can be reached only if congestion becomes so severe that some travelers reschedule their trips out of the peak.

Finally, we note certain differences between the present analysis and the investigation of trip timing for other types of trip, e.g., travel to work by transit, and non-work trips. For bus trips, a measure of reliability was available for the UTDFP sample (on an integer scale, based on subjective evaluations by persons familiar with the operation of the transit system), but not specifically of variability of arrival time or of other factors, such as overcrowding. Small rescheduling adjustments are generally not possible: one either takes the latest bus that gets one to work on time, or one takes the previous bus. In the survey sample, this corresponds (approximately) to the difference between arrival time and work start time being either greater than or less than the first headway; a preliminary investigation, however, showed no association between this choice of trip time and the reliability variable. The timing of non-work trips is expected to be much more elastic in response to supply conditions and thus lead to substantial rescheduling effects; in the absence of firm constraints on travel time, however, the nature of the tradeoff involved in rescheduling becomes much more difficult to formulate and quantify (except, perhaps, for factors such as congestion versus frequency for transit).

Data Used in the Analysis

The data sample

The initial data set consisted of the standard 991-case UTDFP sample. A number of cuts were then applied to the data, removing cases characterized by: (1) missing data on the variables required for the analysis; (2) normal mode of travel to work other than by automobile; (3) unspecified type of auto trip (e.g., drive-alone, carpool); (4) no fixed work start time; and (5) usual arrival time at work outside a one-hour window defined (below) with respect to official work start time. This reduced the sample to 360 cases, of whom 210 normally drove alone and 150 drove (or were driven) with others.

Predictions of shifts in the aggregate trip timing distribution were made using only those subjects whose work start times lay between six and ten a.m. This final subsample, unfortunately, contained only 158 who drove alone and 135 who shared rides. Considerable differences in predicted behavior were found between these two groups, as might be expected.

Definitions and sources of variables

We define the following variables:

T is the subject's arrival time at work, ranging over the choices open to him;

T_0 is the time at which the subject normally arrives at work (i.e., the subject's actual choice);

T_1 is the subject's official work start time;

$y = y(T)$ is the on-vehicle time corresponding to arrival time T ;

$$E = \begin{cases} T_1 - T & \text{if } T_1 \geq T \\ 0 & \text{if } T_1 < T \end{cases}$$; i.e., "early" arrival time at work, in minutes;

$L = L(T)$ is the probability that the subject will arrive late for work, if he plans to arrive at time T .

The times T_0 and T_1 were obtained in the original surveys. The highway travel times were available from a network-based program (Reid, Reinke, and Small, 1975) that included corrections for congestion (based on actual travel times of "floating" cars on most of the freeway links in the network).

Estimation of the model (to be described) involved approximation of the choice set by a discrete set of time intervals based on T_1. Two choice sets were used. The first consisted of six ten-minute intervals, with midpoints ranging from forty minutes before official work start time to ten minutes after official work start time. The other set consisted of twelve five-minute intervals, with midpoints from forty minutes early to fifteen minutes late. Ideally, one would like to be able to extrapolate to a continuous range of variables; it should be noted, however, that respondents almost always gave times to the nearest five minutes. The travel time $y(T)$ is therefore required at six or twelve points only, rather than as a parameterized curve.

The required values of $y(T)$ were obtained by interpolation from values previously computed by the congestion-corrected network program. (This interpolation was carried out as a low-cost alternative to re-running the network program for the additional values of T used in the present analysis.) The existing data consisted of the times $y(T)$ and slopes $y'(T)$ for four points, $T = T_1$, $T = T_1 + 30$, $T = T_1 - 30$, and $T = 480$ (i.e., eight a.m., for morning trips) or $T = 1020$ (five p.m., for afternoon trips), as well as the midday travel time. Cubic interpolation was used between these points, with modifications where necessary to ensure that the full interpolated travel time curve had no minima or secondary maxima. In all cases, a five-minute interval was assumed for the subject to park and walk to his place of work.

To calculate $L(T)$, the probability of being late, on-vehicle time was replaced by a normal random variable with mean $y(T)$ and standard deviation $\sigma = a[y(T) - y(0)]$, where $y(0)$ is the off-peak travel time and a is a constant. No compelling theoretical justification is claimed for this particular form; however, simple models (in which congestion is due to queue formation at bottlenecks) suggest that the standard deviation should be approximately proportional to excess travel time due to congestion. This applies both for random fluctuations in incoming traffic flow and for fluctuations in capacity (e.g., due to stalled vehicles). The observed travel time data is consistent with a linear relationship, but other functional behavior is not excluded (see Reid, Reinke, and Small, 1975). Several values of a were tried; we shall report results for the case $a = 0.2$.

Many of the subjects in the survey did not have to report for work at precisely their official work start times. These subjects were asked how many minutes they could arrive late "without it mattering very much." By adding this time interval to T_1, we obtain the auxiliary variable

> T_2 : the arrival time after which the subject would definitely be considered late.

Rather than exclude these subjects from the analysis, $L(T)$ (the probability of being late) was computed with "late" defined by T_2 instead of T_1. The definition of E however, was not changed. Subjects with $T_2 - T_1$ greater than fifteen minutes were considered to have no definite work start time and were excluded from the analysis.

Formulation of the Trip Timing Model

The trip timing probability will be expressed in terms of the traveler's utility, representing his tradeoff between congestion time and schedule delay. The simplest form would be

$$(1) \qquad U(T) = -\alpha y(T) - \beta(T_1 - T) \quad ,$$

for T earlier than T_1 (i.e., no late arrivals). We may consider subjects to have three different "values of time," for (1) normal (uncongested) on-vehicle time; (2) increased travel time due to congestion; and (3) schedule delay time--that is, the difference between actual and latest possible times of arrival. A mode choice model involving times and costs results in an estimated value for some combination of (1) and (2); in the present case, we obtain an estimate of the ratio of the values of (2) and (3), represented by the parameter α/β. A typical travel time curve $y(T)$ for a heavily congested route is shown in Figure 7(a), and a corresponding utility function of the form equation (1) is shown in Figure 7(b).

A slightly more complex utility function was used, with an explicit term for the probability of being late:

$$(2) \qquad U(T) = -\alpha y(T) - \beta E - \gamma L(T) \quad ,$$

with E and L as defined above. Other forms were also used, to investigate the stability of estimates of α under changes in the specification of the model.

The actual trip timing decision is supposed to be made by maximizing $U(T)$. The probability of choosing time T (or, rather, a time in some small interval centered on T) will then depend on the form of the unobserved stochastic component of the utility function, $\varepsilon(T)$. The form of the probability will also depend on whether the parameters α and β are taken as fixed or as random variables over the population.

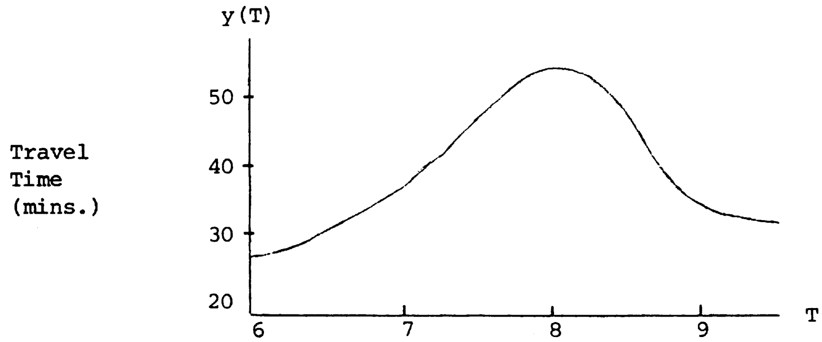

Figure 7(a): Typical Travel Time Curve (Walnut Creek to San Francisco) (Small, 1975)

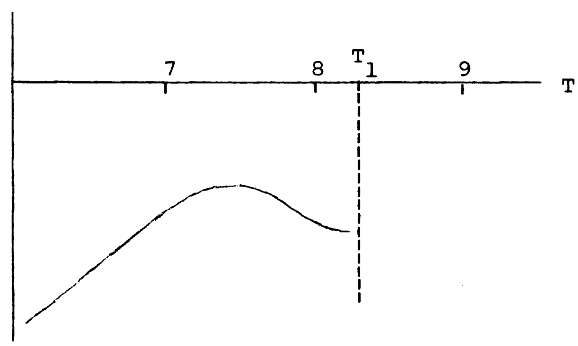

Figure 7(b): Typical Utility Function, corresponding to the travel time curve above, in the simple model [Equation (1)]

For simplicity, the probabilities were taken to be of the logit form,

$$(3) \qquad P(T_i) = \exp\{U(T_i)\} / \sum_j \exp\{U(T_j)\} \quad ,$$

where T_i refers to the midpoint of time interval i . This is appropriate if the range of T can be divided into small intervals in each of which the stochastic term $\varepsilon(T_i)$ has an independent Weibull distribution; because of the special properties of this distribution, the form equation (3) remains true when these intervals are aggregated into larger intervals, at least to the extent that $U(T_i)$ is a good approximation to the maximum value of $U(T)$ in interval i . This, however, is unlikely to be a reasonable assumption: the stochastic term is supposed to represent all the unobserved components of the subject's utility that depend on T , and hence, as the time intervals become small, the terms $\varepsilon(T_i)$ and $\varepsilon(T_j)$ are expected to become highly correlated rather than remain independent.

A more realistic stochastic term could be obtained by taking the $\varepsilon(T_i)$ to be given by a stationary Markov process, and taking the limit as the size of the interval tends to zero. This appears to result, in general, in a stationary Gaussian process, and there is no simple expression for the probability that $U(T) + \varepsilon(T)$ will be maximized in an interval about T .

Another procedure is to treat the coefficients α and β as random variables over the population, with some specified distributions (a gamma distribution was tried), instead of adding a stochastic term $\varepsilon(T)$. It was not possible, however, to obtain a tractable form for the probabilities $P(T_i)$ in this model, even for the simpler utility function of equation (1) with only one coefficient α . An attempt was made to approximate these probabilities to obtain a logit form, as in equation (3), but with $U(T_i)$ replaced by some other function, say $V(T_i)$. This suggested that an appropriate form for the probabilities $P(T_i)$ might be a logit expression in which the "pseudo-utility" $V(T_i)$ contained the slope of the travel time curve, $y'(T_i)$. [Note that if $U(T)$ in equation (1) is maximized at a stationary point, then the maximum occurs at $\alpha = y'(T)$.] However, logit estimations that included the slopes $y'(T)$ yielded coefficients not significantly different from zero.

In view of these difficulties, the logit model of equation (3) was used, with $U(T)$ given by equation (2). Because of the unrealistic nature of the independence assumption, it should perhaps be regarded as an empirical fit rather than as a utility-maximizing choice model. For this reason, caution should be exercised in interpreting the estimated function $\hat{U}(T)$ as representing the traveler's utility.

In the present analysis, only supply variables and work times were included in $U(T)$, as opposed to characteristics of the subject. A potentially useful characteristic would be occupation code (or employment code), because this might help in categorizing subjects according to the seriousness of late arrival at work, and thus their response to the risk of delay.

Estimation of Parameters

The logit model described by equations (2) and (3) was estimated by standard methods, using the QUAIL program package. Table 32 presents the estimated coefficients for subjects with work start times between six a.m. and ten a.m. Subjects were divided into those who drive alone and those who share rides. The twelve-interval choice set was used. T-statistics are given in parentheses.

We see that all the parameters are significantly different from zero for those who drive alone. The relative disutility of congestion time to schedule delay, $\rho \equiv \alpha/\beta$, is 1.6 for these travelers. We note also that the response of shared-ride travelers to congestion time is much smaller and is consistent with zero. This is not unexpected and may reflect the greater difficulty in scheduling shared rides. There is also the obvious possibility of schedule delay for some of the riders that is not a result of avoidance of congestion.

Estimation for the larger subsamples, without restriction of work start times, gave estimates consistent with those in Table 32 but with larger errors.

A number of variants of this model were also estimated, to test the sensitivity of the estimated parameters to some of the assumptions made in specifying the model. A summary of these variants is as follows:

(1) Different values of a were used, where a is the constant of proportionality relating the standard deviation of travel time to the congestion time, $\sigma(T) = a[y(T) - y(0)]$. This changes the computed values of $L(T)$, the probability of late arrival. For $a = 1$ and $a = 0.5$, estimated coefficients differ from those in Table 32 by ten to twenty percent, i.e., within the estimated errors; the likelihood ratios are somewhat smaller.

(2) The model was estimated with six ten-minute intervals instead of twelve five-minute intervals. For the drive-alone group, differences in the estimated coefficients were within ten percent, and for both groups the differences were within the estimated errors. This suggests that the size of the time interval selected is not critical.

(3) The slope of the travel time curve, $y'(T)$, was included as another variable. (See the discussion in the section, "Formulation of the Trip Timing Model," of an alternative formulation of the choice model.) The estimated coefficient of the slope term was small and consistent with zero; this remained true when different coefficients were allowed for positive and negative slopes. If the travel time $y(T)$ remained in the equation, there was very little improvement

in the fit (in terms of the likelihood ratio index), while if $y(T)$ was excluded, a poor fit was obtained.

(4) A dummy variable was included in alternative nine, i.e., for the interval corresponding to official work start time. The coefficient of the dummy variable was significantly different from zero, and the likelihood ratio index improved. However, because the coefficient of the dummy variable presumably depends, in an unspecified way, on the extent of congestion, this model may be less useful in predicting rescheduling in response to changes in travel time.

(5) The probability of late arrival, $L(T)$, was omitted. For subjects with $T_1 \neq T_2$ (i.e., a "soft" deadline T_1 and a "hard" deadline T_2), an additional linear variable,

$$E_2 = \begin{cases} T - T_1 & \text{if} \quad T_1 \leq T \leq T_2 \\ 0 & \text{otherwise} \end{cases},$$

was included. This model gave less satisfactory fits, with likelihood ratio indices of the order of 0.05 or less.

An unsatisfactory feature of the model was revealed when a special subsample was estimated, consisting of subjects who cross the San Francisco-Oakland Bay Bridge. (This subsample was selected on the basis of residence zones in Alameda County and Contra Costa County, with workplace zones in San Francisco; it may, therefore, include some who cross the Hayward-San Mateo bridge instead.) These commuters are expected to face particularly great variations in travel time (see, for example, Figure 7(a)). Although there were only eleven cases in this subsample in the drive-alone category, an apparently significant fit was obtained. The estimated coefficients are presented in Table 33. While the ratio of the coefficients of E and L has not changed appreciably, the relative disutility of congestion time has increased: the ratio $\rho \equiv \alpha/\beta$ has risen from 1.6 in Table 32 to 4.4 in Table 33. Although the small number of cases may make these latter estimates unreliable, the discrepancy does suggest a substantial degree of nonlinearity in the disutility of congestion time. This could be accommodated by including in equation (2) either (1) a term quadratic in $[y(T) - y(0)]$, or (2) a separate coefficient for the excess of congestion time above some threshold value, e.g., ten minutes.

TABLE 32

Variable	Drive-Alone		Shared-Ride	
	Estimated Coefficients			
y	-0.127	(1.70)	0.00466	(0.077)
E	-0.0782	(8.37)	-0.0548	(6.25)
L	-2.72	(8.12)	-2.61	(6.69)
LRI	0.150		0.107	
No. of Cases	158		135	

Estimated coefficients in the utility function equation (2). Numbers in parentheses are t-statistics. LRI is the likelihood ratio index. y is the congestion time, E the schedule delay, and L the probability of late arrival. Work start time restricted to 6 a.m. to 10 a.m.

TABLE 33

Variable	Estimated Coefficients (Drive Alone)	
y	-0.582	(2.15)
E	-0.132	(2.22)
L	-4.16	(2.19)
LRI	0.222	
No. of Cases	11	

Estimated coefficients for the subsample of Bay Bridge commuters.

Effects of Changes in Congestion

The estimated parameters give a choice probability function $P_{ij}(y)$, which expresses the probability that individual i will plan to arrive in time interval j. It is a known function of y, i.e., the on-vehicle times for every time interval for this individual. Thus one can calculate the estimated change in, say, the aggregate cumulative distribution of arrival times, for a specified change in congestion times. This is shown in Table 34 for the case where congestion times $(y - y_0)$ have been increased uniformly by twenty-five percent. (The differences, though small, are significant, even though the estimated distributions are of course not significant to the quoted four figures.) The smallness of the shift may be attributed to the small proportion of the sample who experience any substantial degree of congestion delay. Note that our sample is not, in fact, a random sample of commuters, in view of the original sampling procedure and the cuts imposed on the initial sample.

A single summary statistic may be more helpful than a distribution in some circumstances, and we have defined a "rate of shift" of start time and an effective "elasticity of start time." Analogous quantities are defined for arrival time.

Suppose congestion delay time increases uniformly by a small fraction ε, that is, from $(y - y_0)$ to $(y - y_0)(l + \varepsilon)$. One can then compute the mean change in arrival time over the sample and divide by ε: this gives the "rate of shift" of arrival time. This can be expressed as

$$(4) \qquad r_A = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} T_{ij} \frac{d}{d\lambda} p_{ij} \left[\lambda(y - y_0)\right] \Big|_{(\lambda=1)} ,$$

where T_{ij} is the arrival time corresponding to time period j for subject i. N is the number of subjects, and M is the number of time intervals in the choice set used in the estimation procedure. Note that probabilities $P_{ij}(y)$, j = l,...,M are unchanged if the travel time y is increased by a constant independent of j. The rate-of-shift parameter r_A is expressed in minutes or other time units.

The effective "elasticity" of arrival time is defined similarly, except that the change in arrival time is expressed as a fraction of the corresponding off-peak travel time y_0 before taking the mean. This parameter, η_A is dimensionless, i.e., it has no units.

The rate of shift and "elasticity" for start time (i.e., time of leaving home), r_s and η_s, are defined analogously. Note that these are largely determined just by the rate of increase in travel time, rather than by the effects of rescheduling.

TABLE 34

	Cumulative Distribution (percent)	
Time (a.m.)	With Existing Congestion	Congestion Times Increased 25%
7.00	14.98	15.06
7.10	17.56	17.67
7.20	21.92	22.16
7.30	28.74	29.13
7.40	35.96	36.41
7.50	48.13	48.41
8.00	64.20	64.17
8.10	70.97	70.86
8.20	76.01	75.89
8.30	81.94	81.83

Aggregate probability distribution of arrival times at work for the drive-alone subsample.

In the case of the logit model for the choice probabilities, the expressions for these quantities may be simplified; for example, r_A in equation (4) is the sample mean of the covariance between T and $(y - y_0)\left(\alpha + \gamma\frac{\partial L}{\partial y}\right)$.

These summary statistics are presented in Table 35 for the drive-alone and shared-ride subgroups with work start times between six a.m. and ten a.m., corresponding to the estimated parameters given in Table 32. The greater degree of congestion experienced by the second subgroup may be due in part to the greater likelihood of arranging shared rides between areas of high residential and employment density. The difference between the elasticities of arrival time for the drive-alone and shared-ride subsamples appears particularly noteworthy.

TABLE 35

	Drive Alone	Shared Ride
Rate of shift of start time (minutes)	-2.91	-4.45
Rate of shift of arrival (minutes)	-0.46	-0.088
“Elasticity” of start time	-0.177	-0.243
“Elasticity” of arrival time	-0.024	-0.0059
Mean increase in on-vehicle time due to congestion	16.2%	24.3%

Summary statistics for changes in trip timing in response to increased congestion (sample with work start times between 6 a.m. and 10 a.m.).

Queue Formation at Bottlenecks

Because the effects of rescheduling are small when averaged over the entire sample, we consider, briefly, a case where its effects are expected to be significant. This occurs when congestion time may be viewed as due to the buildup of queues at bottlenecks on congested highways when demand exceeds capacity. Growth of the queue can result in a rapid rate of increase in congestion time, and thus make rescheduling worthwhile.

Traffic flow has been analyzed in terms of deterministic queue formation by May and Keller (1967) (the first in a series of investigations of increasingly complex traffic flow models) and by Small (1976). Demand was taken as given in the May-Keller model, and was used to determine the delay caused by queuing when demand exceeds the capacity of the bottleneck.

Here we consider also the feedback effect of delay on demand, i.e., on how the rate at which travelers arrive at the bottleneck is influenced by the existing length of the queue. With appropriate simplifications, the shape of the travel time curve can then be expressed in terms of a parameter representing the tradeoff between congestion time and "schedule delay." Further details of these results will be presented elsewhere.

In the simplest case, suppose all peak-time travelers have the same deadline T_1 for exit from the bottleneck, and all have the same tradeoff ρ between time spent waiting in the queue and time "wasted" through early arrival at the destination. We assume the simple utility model of equation (1), with $\rho \equiv \alpha/\beta$. [There is no difficulty, in principle, in using the model of equation (2) with an explicit term for the probability of being late; it then becomes necessary, however, to compute the resulting graphs numerically.] Suppose that the queue is stable , where a stable queue is defined as one for which no traveler could increase his utility by changing his arrival time at the queue. (Where utility is constant over a range of arrival times, the position of individual travelers in the queue will be determined by the unobserved stochastic component of the utility function.)

One then finds that the queuing delay has the triangular form shown in Figure 8, corresponding to a uniform arrival rate of travelers between times T_A and T_B . The parameters are:

$$T_A = T_1 - \frac{\rho n}{\rho C - q(\rho - 1)} \quad ;$$

$$T_B = T_1 - \frac{n}{\rho C - q(\rho - 1)} \quad ;$$

$$T_C = \frac{q}{C - q} (T_1 - T_B) + T_1 \quad ;$$

where n is the total number of peak travelers, C is the capacity of the bottleneck, and q is a constant "background" traffic flow $(q < C)$. [If the background flow q is zero, the formulae remain well-defined but the queue may become unstable, i.e., there may be no configuration in which no traveler could gain by further rescheduling.]

The rates of growth and decline of the delay time in Figure 8 are $1/(\rho - 1)$ and $(1 - q/C)$, respectively.

The assumption that all travelers have the same exit deadline T_1 is not essential, and Figure 9 illustrates the case of a continuous distribution of exit deadlines, represented by the curve DD' (giving the cumulative distribution as a function of t). The queue has the same form as in Figure 8. The point Z, where the queue finally disperses, occurs when the slope of the curve DD' equals $C - q$. The straight line XZ, of slope $C - q(\rho - 1)/\rho$, gives the cumulative distribution of peak travelers leaving the queue, and its intersection with the demand curve DD' at X determines the time at which the queue begins to form. The line XY has slope $C\rho/(\rho - 1) - q$; XYZ then gives the cumulative distribution of peak travelers entering the queue.

A typical value for the maximum rate of increase in travel time for trips involving potential bottlenecks (e.g., the Caldecott Tunnel and the Bay Bridge) is about 0.3; setting this equal to $1/(\rho - 1)$ gives $\rho \approx 4.3$. The similarity of this value to that obtained from the estimates in Table 33 (0.582/0.132 ≈ 4.4) suggests that the utility model is not entirely unrealistic.

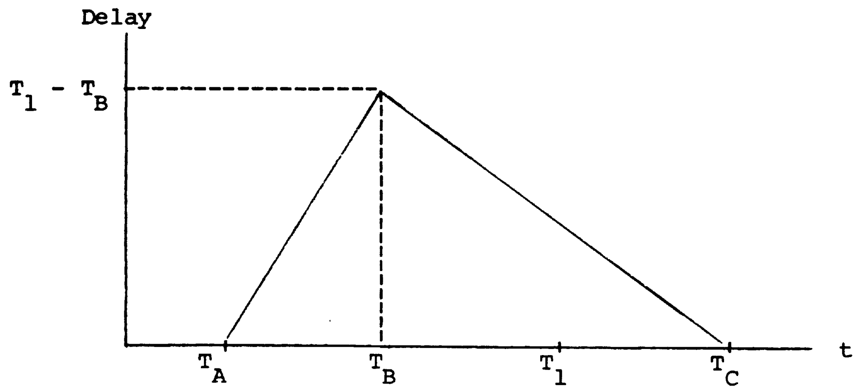

Figure 8: Delay as a Function of Arrival Time for a Simple Stable Queue

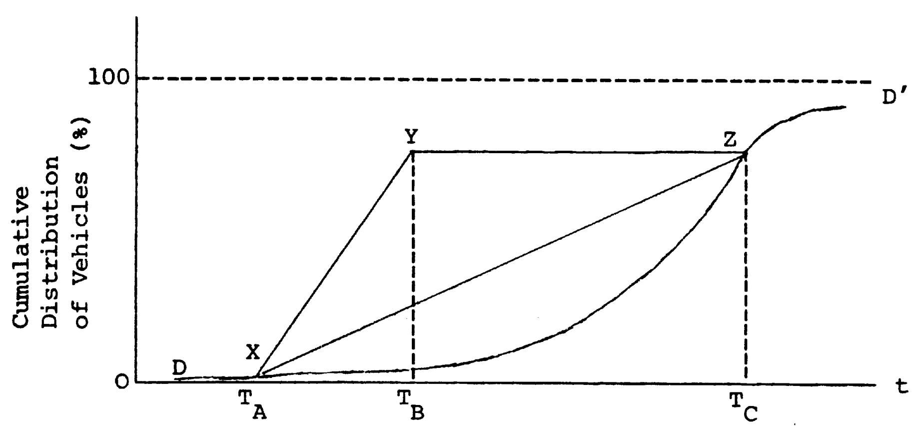

Figure 9: Queue Formation in Relation to the Cumulative Distribution of Exit Time "Deadlines"

As a final exercise in this simplified queuing model, we consider the change in utility (benefit to peak travelers) of an increase in the capacity of the bottleneck. Utility will be expressed in units of the "value" of schedule delay in minutes. As an example, suppose T_1 is fixed, $C = 10{,}000$ vehicles per hour, $n = 5000$, and $\rho = 2$. The queue parameters (see Figure 8) are, then,

$$T_1 - T_B = 15 \text{ minutes} ,$$

$$T_1 - T_A = 30 \text{ minutes} ,$$

and the average utility is -30 minutes (in schedule delay equivalent units). Suppose capacity is increased by twenty percent. The average reduction in travel time would be 2.5 minutes if the incoming flow remained unchanged (as in the May-Keller model), but is only 1.25 minutes when we allow for rescheduling. Thus, if we consider only the reduction in travel time, the effect of rescheduling into the peak <u>halves</u> the benefit obtained from the increase in capacity.

If, however, the change in utility is calculated, we find $\Delta U = 2.5$ minutes (in schedule delay units) with no rescheduling, but $\Delta U = 5$ minutes when rescheduling is taken into account. The reduction in schedule delay achieved by rescheduling, therefore, more than offsets the corresponding increase in average travel time. We see that the gain for peak travelers is (at least in this special case) <u>greater</u> than would have been estimated for a fixed incoming flow.