A duration model for the TTB of Lyon

Iragaël Joly

Workshop on Activity Based Approach
May 28-31

A duration model for the TTB of Lyon

Iragaël JOLY
Ph.D. Student
Laboratoire d'Economie des Transports, ENTPE
14 avenue Berthelot
69363 Lyon Cedex 07
iragael.joly@let.ish-lyon.cnrs.fr

Abstract:
Escaping unidimensional analysis limits and linear regression irrelevancy, the duration model incorporates impacts of covariates on the duration variable and permits to test the dependence of TTB on elapsed time. We apply the duration model to the TTB of Lyons (France), in the perspective of a discussion of Zahavi's hypothesis. The duration dependence estimation illustrates covariates effect on TTB and suggests a non-monotone hazard for their distribution, which conflicts with the TTB stability hypothesis and more generally with the classic travel time minimisation problem.

1. Introduction

The interest in behavioural analysis has led the research in transportation to the time-use analysis. The activity-based approach raises the problem of the way individuals decide both the participation and the duration of one or another activity. The individual's activity pattern includes its timing, its duration, location of activity, mode of transport and activity sequencing. Some researchers have specifically studied duration of activities.

The interest of this paper is that of the time allocated to urban travel during a day. The individual travel time budget (TTB) is then computed as the sum of the durations of all the trips realised in one day. The TTB has been claimed by Zahavi (1980) as being a constant amount of time about 1 hour per day per capita. He also claimed that this amount of time is similar between different cities and different time periods. Then, the Zahavi's conjecture can be formulated as the spatial and temporal stability of the TTB. Since then, it has become a common conjecture in the transportation research field. The cities sprawl can easily be interpreted as a consequence of the increase in disposable speeds. Hence speeds and any policies favouring speeds become responsible of the increasing mobility. Recently, Schafer and Victor (2000) have used the constant TTB concept to construct a transport demand model and to predict the future mobility of the world population.

On one hand, the relative stability has been confirmed in some researches (Zahavi and Ryan, 1980 ; Zahavi and Talvitie, 1980 ; Hupkes, 1982 ; Bieber, et al., 1994 ; Vilhemson, 1999 ; Schafer and Victor, 2000). On the other hand, a lot of authors have adopted the opposite direction (Purvis, 1994 ; Levinson and Kumar, 1995 ; Kumar and Levinson, 1995 ; Godard, 1978 ; Van der Hoorn, 1979 ; Landrock, 1981 ; Gordon, et al., 1991 ; Kitamura, et al., 1992). The critics of Zahavi's conjecture have been concerned with the influence of some socio-economic, activity-related and area specific variables. For examples, variables such as income, car-ownership, age, timing of the trips or urban density are shown to influence the TTB. This multiple critiques are warnings to the abusive application of the constant TTB concept in a non-world level.

A key question of the TTB is its level of observation and application. But most of the critiques are at disaggregated level such as national, regional or urban level. The TTB needs to be explored at different levels of observation. Here, we propose the urban level and examine the TTB of the city of Lyon (France).

Furthermore, the TTB has been studied with respect to several variables, such as characteristics of individuals, transportation system, or activities. But most of the time the analysis has been unidimensional or limited to the linear analysis. Then, the study of TTB can be improved by a modelisation that incorporates a set of variables and that overtakes the limits of the traditional linear model. In this paper, it will be done with the duration models methodology (or survival analysis).

Unlike the classic estimation methodologies, such that linear or logistic regressions, the duration models are useful to study the duration allocated to the different activities. First, this method models the impact of variables on the duration variable under study: the TTB. Second, it is adapted to the duration data that are non-negative and that can be censured and time varying. Third, this kind of model permits the examination of the duration processes in which the temporal dynamic needs to be included. Here, the conditional probability of the ending of a process, given that it has lasted to some specified time, permits to discuss Zahavi's hypothesis and the minimisation of travel time. The hypothesis of 1 hour TTB would lead this probability to increase around 1 hour of elapsed time. More generally, the minimisation of travel time in the allocation of time process would imply an increasing probability of the end-of-duration with elapsed time.

The duration model concentrates on the modelling of the conditional probability of ending which integrates the principle of the temporal dynamics. It permits the likelihood of ending an activity to depend on the length of elapsed time since the start of the activity.

Used in biometrics and industrial engineering fields, the duration models have been applied in transportation fields in multiple ways: accident analysis (Jovanis and Chang, 1989 ; Mannering, 1993, Nam and Mannering, 2000), car ownership (Mannering and Winston, 1991 ; Gilbert, 1992 ; Hensher, 1998), traffic queuing (Paselk and Mannering, 1993), duration before acceptance of a new toll (Hensher and Raimond, 1992), and traveler's activity behaviour. The analysis of the activity behaviour focus on: the time spent at home between trip generating activities (Hamed and Mannering, 1993 ; Mannering et al., 1994 ; Misra and Bhat, 2000) ; the duration of out-of-home activities (Niemer and Morita, 1996 ; Kitamura et al., 1997 ; Bhat, 1996a,b ; Timmermans et al., 2002) ; the duration between two occurrences of an activity (Schonfelder and Axhausen, 2000 ; Bhat et al., 2002).

Hensher and Mannering (1994) and Bhat (2000) present detailed overviews of the existing applications of duration models in transportation field.

The purpose of the paper is to analyse the TTB of Lyon in order to discuss Zahavi's hypothesis. The stability of the TTB will be tested through the functional form resulting from the duration model. Furthemore, the duration model technique permits to examine how this TTB is dependant on some variables. The second part presents Zahavi's hypothesis and TTB studies. The robustness of stability is then discussed with respect to the level of observation. In the third part, data and duration model method are presented. Finally results of the non-parametric, semi-parametric and parametric estimations are presented and lead critics of the TTB stability and the allocation of time mechanism.

2. Zahavi's hypothesis – stability of travel budgets

Zahavi's hypothesis has been defined at two different levels. First, at aggregate level (world wide level), the means TTB of agglomerations are similar between cities and time of observation. Second, at the disaggregate level (local level), travel expenditures exhibit regularities that are supposed to be transferable between cities and time.

Zahavi (1973) and Zahavi (1974) show that TTB and TMB (transport monetary budget) are linked to the socio-economic characteristics of individuals, and to the attributes of transport system supply and urban structure. Furthermore, the regularity of such relationships observed between different cities, leads to their integration in a travel-demand forecast model. These regularities permit to consider the travel expenditure of one individual as a budget of which amount is rationally determined. Zahavi is one of the first to suggest the expenditures budgets concept and to incorporate time budget in the optimisation program of the individual travel choices.

Both TTB and TMB appear as constraints in Zahavi's (1979) model, named the Unified Mechanism of Transport (UMOT). Zahavi determines the TTB as an inverse function of mean travel speed. Then, he reduces the problem of resources allocated to transportation at a simple repartition of fixed amounts of time and money resources between different modes.

Far from this simple tool, numerous disaggregated models consider travel time as expenditure to be minimised. Still, minimisation of travel time is a way to escape the integration in the

allocation of time mechanisms of the derived demand concept that characterises travel demand.

The difficulty of this integration problem explains both the popularity of the Zahavi's hypothesis of a 1 hour TTB and its use at non-world levels. This constancy leads mechanically to systematic reinvestments of travel-time savings in transport. Then it gives the responsibility of the increasing mobility to the speeds and to any transport policies aiming at improving circulation conditions.

Schafer (2000) and Joly (2003)[1] confirm the stability observed by Zahavi at aggregate level. The mean TTB of these three studies are closed of 1 hour. Differences appear because of the divergent methods and definitions used. The mean TTB of Zahavi is defined on the mobile population and only for motorised modes of transport, while Schafer studies the entire population and all modes. Finally, in its precedent TTB analysis of 100 cities of the world, Joly (2003) obtains TTB for the urban population, but only motorised trips are observed. Nevertheless, the TTB distributions of the three works show similar attributes as, for example, closed interquartile ranges, the mean and similar dispersion around the mean.

The stability seems to be valid only at the world level. The disaggregation of the level of observation reduces the robustness of the TTB stability hypothesis. For example, at a continental level, Levinson, Wu and Rafferty (2003) conduct an analysis on US cities with regression models, using data from the United States (2000 Census). They show significant effects of congestion, income, population, population density and area. In the same way, Joly (2003) shows the opposition of two urban organisations characterised by distinct TTB dynamics. First, an extensive model composed of North American and Oceanic agglomerations, which develop by the extension of their space and time consumption. Second, an intensive model characterising European cities and Asiatic metropolis find stability in consumption of space and time. Hence, a "European" city with near stable TTB is opposed to a "North American" city with a TTB that appears to be sensitive to variables such as urban density, mean GDP per capita, mean road speeds and daily travel distance.

Numerous studies using finer scale of observation questioned the apparent stability. Zahavi and Talvitie (1980), Zahavi and Ryan (1980), Chumak and Braaksma (1981), Hupkes (1982) are the first to valid the stability. Since then, despite the difficulties of comparison, a large part of the different studies of TTB do not support the Zahavi's hypothesis. Mokhtarian and Chen (2002) present overview of the variables found to affect the TTB in numerous studies. Hence, TTB varies with socio-economic variables such as age, gender, employment status, car ownership, household size and income. Activity-related characteristics and area-specific characteristics are referred as influencing variables. For example, the out-of-home duration and the activity duration are studied, but most of the time, it requires the use of the travel time needed to access the activity and leave the definition of the TTB as a sum. Finally area-specific attributes are studied. For example, population and urban density are influencing variables (Landrock, 1981, Gordan et al., 1989). However, these studies can hardly be compared because of the divergent definitions of urban, sub-urban or rural attributes.

---

1 Joly I., (2003), Les rapports espace-temps de la mobilité quotidienne et les systèmes productifs des transports urbains - Une analyse de la base UITP sur les systèmes de transports urbains de 100 villes du monde, directed by Alain Bonnafous, research report, for the French Commissariat Général du Plan.

3. Data and methods

The data source used in the present study is a household mobility survey conducted between November 1994 and April 1995 by the CERTU (Centre d'Etudes sur les Réseaux, les Transports, l'Urbanisme et les constructions publiques) in the French agglomeration of Lyon. The survey collects data on socio-demographic and mobility characteristics of the 6000 households and of each individual in the household. The survey also includes information on a week day mobility of all members of the household above 5 years of age. Each trip is described by (a) the starting and stopping times, (b) the types of activities at origin and at destination, (c) the travel mode. Thus, the one-day out-of-home activity diary can be deduced, from the first trip to the last trip of the day. *Table 1* specifies the definition of the variables.

Table 1 : Variables definitions

Variable	Definition
Household responsibility variables	
Number of children	Number of children is reported exactly, and could be classified in: Number of children under 5 years of age and Number of children of age between 6 and 17, or presence of kids
Transport variables	
Licence holding	1 if holder
Number of private vehicles at free-disposal	Number of private vehicles at free-disposal: from 0 to 4 and more
Principal mode used	Coded from 1 to 7: walk, cycle, motorcycle, transit, private vehicle (driver), private vehicle (passenger) and "others". The mode is principal if it represents the greater part of the number of trips of the day
Socio-economic variables	
Monthly Household income	Monthly Household income is defined in steps : 0 ; 2500 ; 5000 ; 7500 ; 10000 ; 12500 ; 15000 ; 20000 ; 30000 ; 50000 ; 50000 and more (in F) High incomes are defined as over 20000F
Sex	Male = 1 / Female = 0
Age	Age is reported exactly, and could be used as class variables : from 0 to 19 years ; from 20 to 49 years; 50 years and more
Household size	Number of members in the household
Employment status	Coded from 1 to 9: full-time worker, part-time worker, student, scholar, unemployed, retired, stay at home, formation, and other
Worker	1 if part-time or full-time worker
Others variables	
Day of the trips	From 1 to 5: Monday to Friday
Household localisation	From 1 to 8: hyper-centre, Lyon-Villeurbanne, 1st ring East and West, 2nd ring East and West, 3rd ring East, and external zone

Table 2 describes the summary statistics. Despite the mean TTB is distant from 1 hour, it is included in the relatively close interval of TTB obtained by both Zahavi and Schafer. Here, the TTB of more than 6 hours (less than 1% of the sample) can not be assimilated to daily urban and frequent mobility and then are excluded from the analysis.

Table 2 : Summary statistics of TTB (in min)

Mean	76.5	Median	65	Interquartile Range	60	Quantile 75% (Q3)	100
Standard error	51.74	Mode	60	Range	353	Quantile 25% (Q1)	40

Frequently, the TTB is analysed from unidimensional point of view. Nevertheless, the study of the TTB could be improved by a modelisation that incorporates simultaneously a set of variables. In this paper, it will be done by the duration models methodology. First, this

technique is adapted to deal with duration data that are non-negative and that can be censured[2] and time-varying. The linear classical methods are irrelevant to model positive variables or partially observed or measured variables. Furthermore the qualitative models as the logistic regression integrate not easily variables that could change during the observation period. Second, the duration model introduces the duration dependence concept. It models the conditional probability of the end-of-duration of a process, given that it has lasted to a specified time, and permits the likelihood of ending to be depending on the length of elapsed time. Hence, this probability can vary during the process. Finally, this conditional probability can question the TTB stability hypothesis and the minimisation of the temporal component of travel costs. Indeed, the estimation of this conditional probability, called hazard rate, will inform us on the temporal dynamics of TTB. Then, increase of this probability in elapsed time will imply accelerated decrease of estimated TTB. Given TTB stability around 1 hour, the hazard should increase faster after 1 hour of elapsed time in transport. Hence, given the minimisation of travel time expenditure, we should observe, at least, a monotonically increasing hazard, with elapsed time.

4. Overview of duration models

A duration model is based on the conditional probability of the end-of-duration of a process, given that this process has lasted to some specified time. This conditional probability is defined by the hazard function. Let *T* be a non-negative random variable representing the duration of a process. The hazard function, *h(t)*, is the instantaneous probability that the process ends in an infinitesimal interval *Δ* after time *t*, given that this process has lasted to the time *t*. The hazard function is given by:

$$h(t) = \lim_{\Delta \to 0^+} \frac{P(t \leq T < t + \Delta \,/\, T > t)}{\Delta}$$

This conditional probability can be expressed in terms of the density function, *f(t)* and cumulative density function, *F(t)*, of *T*.

$$f(t) = \lim_{\Delta \to 0^+} \frac{P(t \leq T < t + \Delta)}{\Delta}$$

$$F(t) = P(T < t) = \int_0^t f(u)du$$

The probability of end-of-duration in an infinitesimal interval after *t* is given by: *f(t)*Δ*. The probability that the process lasts to time *t* is *1-F(t)*. Hence, the hazard function can be written as:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} = \frac{dF(t)/dt}{S(t)} = \frac{-dS(t)/dt}{S(t)} = \frac{-d\ln S(t)}{dt},$$

where the complementary probability of *F(t)* is *S(t)*, the survival distribution (probability to survive until *t* or the endurance probability, Bhat, 2000):

$$S(t) = \Pr[T \geq t] = 1 - F(t)$$

The hazard and the survival functions are describing the duration process. So the shape of the hazard function has important implications for the duration dynamics.

To study this shape, one may use three approaches: parametric, non-parametric and semi-parametric estimations.

---

[2] The observation of an process is censured when its beginning (left censured) or its end (right censured) are excluded from the observation period.

Non-parametric approach
The non-parametric approach is similar to an exploratory data analysis. The survivor function is estimated using the Kaplan-Meier product limit estimator (Kaplan and Meier, 1958). The KM estimator of survival at time t_j is computed as the product of the conditional survival proportions:

$$S_{KM}(t_j) = \prod_{k=1}^{j} \frac{r(t_k) - d(t_k)}{r(t_k)},$$

where $r(t_k)$ is the total population at risk for ending at time t_k. $d(t_k)$ is the number of individuals stopping at t_k. The corresponding survival curve is a step function with a drop at each discrete end-of-duration time. The definition of these steps is of special importance in presence of discrete times, i.e. many unique event times. This discretisation may arise when the reported duration times are rounded off. Here, for the TTB of Lyon, we can show the rounding to the nearest 5 minutes in reporting the travel time duration. In presence of discrete times, event times are grouped into intervals. Then, the steps are defined by arbitrary determined intervals. Assuming a constant hazard within each discrete period, one can then estimate the shape of hazard by a continuous-time step-function. This method is known as the life-table method. In our case of rounded times, a width of 5 minutes is believed to be the suitable interval. The estimation of the hazard and the survivor functions characterising the distribution of the duration variable, T will be given at the midpoint of the interval.
This approach produces an empirical approximation of survival and hazard, but does not model effect of covariates. Then, only tests of classification effects of covariates on survival functions can be conducted.

Parametric approach
The incorporation of the effect of covariates can be done through two parametric forms: the proportional hazards form and the accelerated lifetime form. The first form assumes a multiplicative effect of covariates on a baseline hazard function. In the second form, a direct effect on duration is assumed.

Proportional hazard model
The proportional hazard model (PH model) assumes that the hazard function is decomposed as:

$$h(t/X) = h_0(t) \, g_0(X) = h_0(t) \, exp(-\beta X),$$

where $h_0(t)$ is the baseline hazard. $h_0(t)$ is a function of survival time and represents the duration dependence. $g_0(.)$ is a function of the covariates and gives the change of the hazard function caused by the covariates. The separation of the time effect and the covariates effects leads the PH model to assume the proportionality between the hazard rates of two individuals, i and j, with different attributes. Given that the covariates effects are not time dependent, the hazard ratio is given by:

$$\frac{h_i(t)}{h_j(t)} = exp\{\beta_1(x_{j1} - x_{i1}) + ... + \beta_k(x_{jk} - x_{ik})\},$$

The distributional assumptions for the baseline hazard $h_0(t)$, impose specific forms to the shape of the hazard function: constant, monotone or U-form.
The estimation will conduct to the distributional parameters and covariates estimators. Coefficient estimators can be interpreted either in terms of its effect on the hazard ratio or defined as derivative of the log-hazard with respect to the associated covariates:

$$\beta_k = \frac{\partial \ln h(t)}{\partial X_k}$$

Subsequently, positive coefficient implies that an increase in the corresponding covariate decreases the hazard rate and increases the expected duration. Hence, if the covariate j increases by 1 unit, the hazard changes by $100(e^{-\beta j} - 1)\%$. The hazard ratio interpretation is easier to be applied in case of binary covariate.

The accelerated lifetime model
The second parametric form for accommodating the effect of covariates with the duration dependence assumes that the covariates act directly on time. Then, the survival function in the ALT model is:

$$S(t/X) = S_0[t\exp(-\beta X)],$$

where $S_0(t)$ is baseline survivor function. Furthermore, corresponding hazard function is:

$$h(t/X) = \frac{-\partial S(t/X)/\partial t}{S(t/X)} = h_0[t\exp(-\beta'^* X)] \cdot \exp(-\beta'^* X)$$

The ALT model can be expressed as a log-linear model, such that $\ln t = \beta' X + \varepsilon$, with density function of the error term $f(\varepsilon)$, that differs according to the type of estimated model. Then, the coefficients can be interpreted after exponential transformation with respect to the following derivative:

$$\beta_k = \frac{\partial \ln T}{\partial X_k}$$

In case of binary covariate, e^{β} gives the expected survival time ratio. For quantitative covariates, $100(e^{\beta} - 1)$ gives variation in percent of the expected survivor time for each 1 unit increase of the covariate.

In the two parametric approaches, there exist a need to specify the used distribution function. The classically used distributions for duration distributions are the exponential, Weibull, log-logistic, Gompertz, log-normal, gamma, and generalised gamma distributions. Validity of the exponential and Weibull distributions can be graphically tested in the non-parametric approach[3].

The parametric approaches permit simultaneous estimation of covariates effects and of duration dependence. However, the distributional assumption for the baseline hazard is risky. Meyer (1990) has shown that the parametric approach inconsistently estimates the baseline hazard when the assumed parametric form is incorrect.

Semi-parametric approach
Finally the semi-parametric approach focuses solely on the covariates coefficient estimates. This estimation technique estimates the PH model using the partial likelihood framework suggested by Cox (1972), which do not need the specification of the baseline hazard function, $h_0(t)$. One avoids then the risk of a mis-specified baseline function. The quality of the estimation of the covariates coefficients is considered to be more robust than the fully-

---

[3] If the hazard is constant ($h(t) = \lambda$) then : $-\log S(t) = \int_0^t h(u)du = \lambda t$.

This implies that a plot of $-\log \hat{S}(t)$ against t should be a straight-line through the origin.

And the plot of $\log[-\log \hat{S}(t)]$ against $\log(t)$ tests the Weibull distribution. In this case, the hazard is $\log h(t) = \alpha + \beta \log t$. Hence, a plot of $\log[-\log \hat{S}(t)]$ against $\log(t)$ should be a straight-line with β slope.

parametric approach (Oakes, 1977). But the Cox model excludes the baseline hazard and does not allow for consideration of the duration dependence.


5. *Estimation and Results*
Non-parametric estimation
Life table method constitutes the non-parametric approach and shows the first intuitions on the covariates effects and on the distribution to be used in the parametric approach. The graphical and statistical tests permit to identify influential classification variables.
Lifetable analysis offers first insight into the TTB temporal dynamic. The resulting survival and hazard functions are presented in figure 1. The survival curve presents two inflexion points. The first, near 20 minutes, seem to indicate the existence of minimum TTB level of 20 minutes, that is, almost accepted by all travellers. The second point, near 110 minutes correspond to a diminishing probability of the ending after 2 hours of travel. The survival decreases at a decreasing rate.

The hazard curve is characterised by peaks for 1, 2 and 3 hours that result from the rounding of declared travel times. The hazard curve presents clearly a point where the slope is reversed. The hazard is increasing until near 110 minutes, and then decreasing.

*Fig. 1* Survival and hazard curves for TTB in min
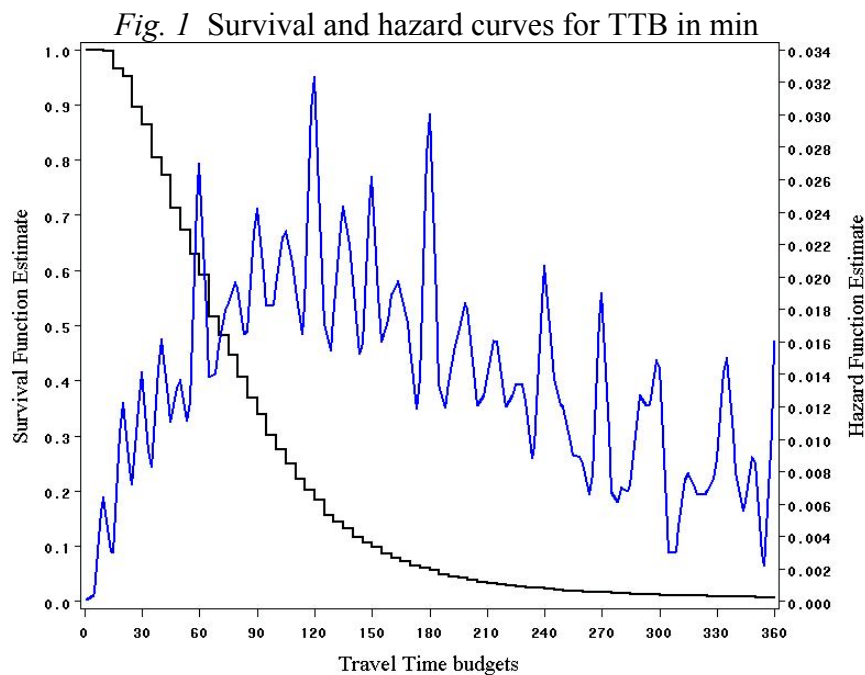
The non-monotonic form of the hazard suggests that log-logistic or log-normal will be appropriate distributions in a fully-parametric model. Furthermore, by the graphical test of linearity of the transformations ($-Log(\hat{S})$ and $Log[-Log(\hat{S})]$), the hypothesis of exponential and Weibull distributions are rejected. If confirmed by the parametric approach, the non-monotonic hazard will permit to discuss the TTB stability hypothesis and travel time minimisation.

The median survival times are presented in figure 2. For each time *t*, it approaches the expected survival time given that the process has lasted to *t*. For the initial TTB (TTB=0), the median survival time is 65 minutes, near the Zahavi's TTB level. The decreasing part suggests that travellers reduce the travel times during the first hour. But from 60 to 150

minutes, the median survival time is stable. Then, individuals that have already a 1 hour TTB, are expected to pass 30 minutes more in travel. And finally the median survival time is increasing after 150 minutes.

Fig. 2 Median survival lifetime

Median survival Lifetime
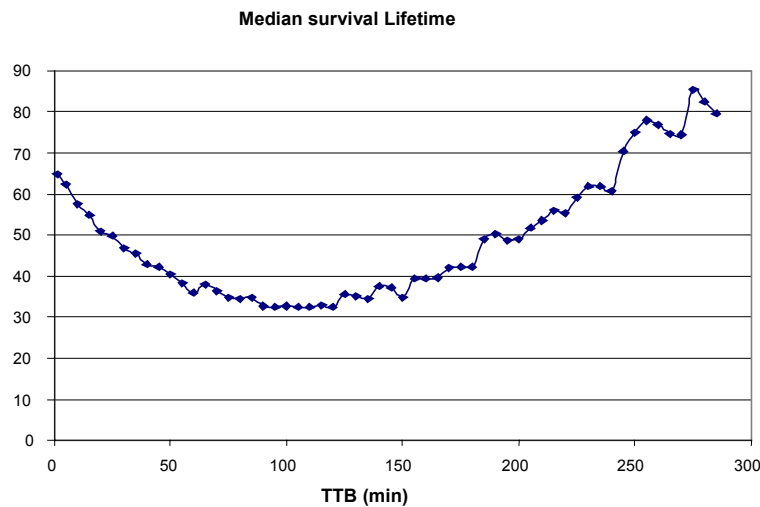
The hazard and the median survival time suggest a transition in the allocation of time to transportation, near the 110 minutes level. Everything happens as if, after this level, the travellers failed to diminish their travel times. Therefore, one can think of a segmentation of population. First, a group of individuals who minimises travel times and that is characterised by a near 1 hour TTB. Second, a group of travellers that abandon, or can not pursue the minimisation of travel times.

Finally the non-parametric estimation produces graphical and statistical tests of classification variables effect on survival. *Figures 3a, 3b, 3c and 3d* illustrate examples of the corresponding survival curves for the different classes of the variables: employment status, licensed driver, age and sex.

The form of the estimated survival curves for these classes are near the general survival curve. All the classification variables used produce distinct survival. For example, in figure 3a, workers are characterised by upper survival curve, then a worker will have higher TTB. A licensed driver will have higher TTB (*fig 3b*). The difference in survival curves between male and female appears after 60 min. Niemer and Morita (1996) show than women spend more time in activity linked to their household responsibilities. Here, it can explain that women have lower TTB than men (fig 3c). Finally, segmentation with respect to the class of age shows that young (under 20 of age) have the lowest TTB. Individuals between 21-49 years of age present the highest TTB (fig 3d).

Table 3 presents tests of equivalent survival for classification variables and surveys the observed effects. All the presented variables show graphically a classification power on survival, which is statistically validated.

Fig. 3a TTB survival curves
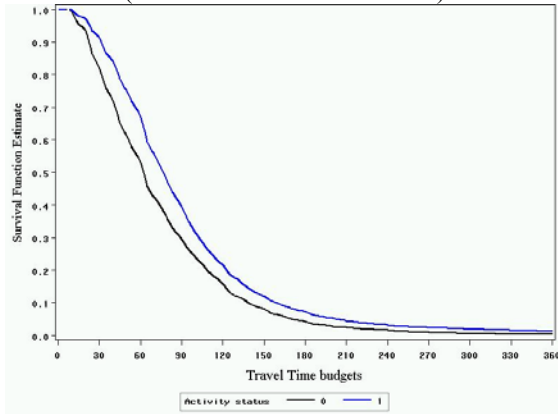(worker = 1 / non-worker = 0)

Fig. 3c TTB survival curves
(male=1 / female=2)
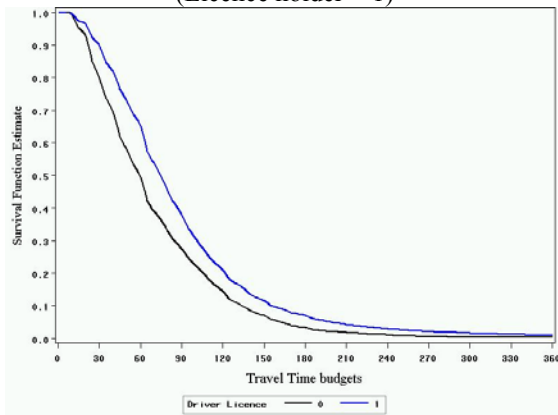

Fig. 3b TTB survival curves
(Licence holder = 1)

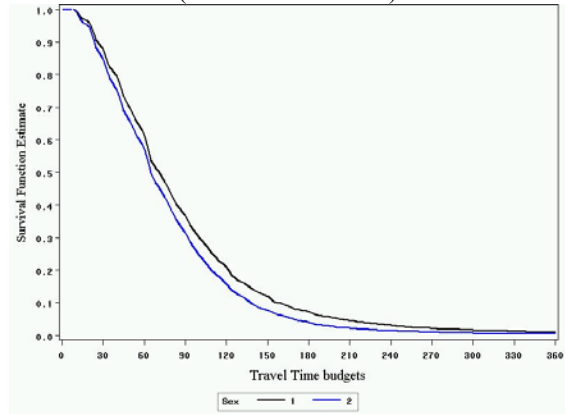*Fig. 3d* TTB survival curves
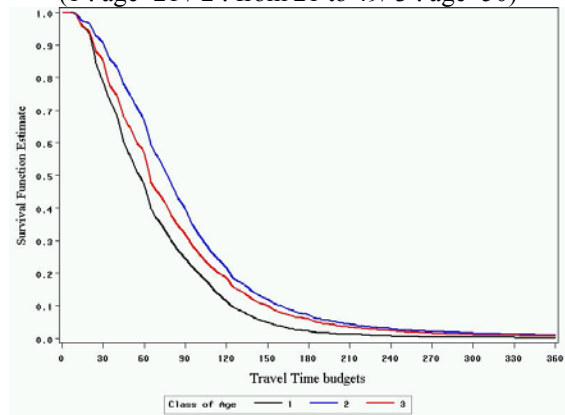(1 : age<21 / 2 : from 21 to 49/ 3 : age>50)

Table 3 : Chi-squared tests for differences in survival over stratification

	Log rank (p-level)	Wilcoxon (p-level)	-2log(LR) (p-level)	Observed effect
Day of the trips	61.1356 (<0.0001)	74.1796 (<0.0001)	30.1252 (<0.0001)	From Monday to Friday: increasing survival and increasing TTB
Worker	213.6524 (<0.0001)	297.9043 (<0.0001)	152.9904 (<0.0001)	Longer TTB for workers
Employment status	560.0853 (<0.0001)	672.5742 (<0.0001)	323.9391 (<0.0001)	By order of increasing survival: scholar, stay at home, other, retired, partial time worker, full-time worker, unemployed, student, formation.
Licence holding	240.6896 (<0.0001)	321.4084 (<0.0001)	147.7335 (<0.0001)	Longer TTB for licence holder
Age class	363.4827 (<0.0001)	438.6494 (<0.0001)	214.6413 (<0.0001)	By order of increasing survival: 0 to 20; older than 50; 20 to 50 years
Sex	84.9996 (<0.0001)	57.4785 (<0.0001)	53.2191 (<0.0001)	Longer TTB for male
Number of private vehicles at free-disposal	32.8779 (<0.0001)	36.8848 (<0.0001)	23.1019 (0.0031)	Longer TTB for 3 cars and more. Closed survivals for 0, 1 and 2 cars
Income Class	55.7909 (<0.0001)	73.3633 (<0.0001)	36.9774 (0.0136)	Longer TTB for classes of income greater than 20000F
High Income	44.5189 (<0.0001)	55.0466 (<0.0001)	29.4089 (<0.0001)	Longer TTB for high income household member
Household Size	46.4324	38.4409	28.4577	Decreasing TTB with household size

	(<0.0001)	(<0.0001)	0.0004	
Number of children under 5 years of age	13.1797 (0.0014)	6.4970 (0.0388)	7.3517 (0.0253)	Shorter TTB for individuals with 2 and more children under 5 years of age
Number of children of age between 6 and 17	102.8578 (<0.0001)	100.6159 (<0.0001)	56.9043 (0.0236)	Decreasing survival with number of children between 6 to 17 years of age
Number of children	117.7232 (<0.0001)	110.0801 (<0.0001)	64.0930 (<0.0001)	Decreasing survival with number of children
Presence of kids	43.5829 (<0.0001)	35.7499 (<0.0001)	25.5583 (<0.0001)	Shorter TTB with presence of kids
Household localisation)	43.8228 (<0.0001)	51.4101 (<0.0001)	25.4067 0.0006	Longer TTB for central localisation. Shorter TTB for 1st ring
Principal mode used	1469.2518 (<0.0001)	2091.1187 (<0.0001)	646.1205 (<0.0001)	By order of increasing survival : walk, motorcycle, cycle, private vehicle as passenger and driver, transit.

These non-parametric tests of survival equivalence inform us about the relation between TTB and the considered variables. But these tests are only unidimensional. The intuition given by these tests needs to be examined by considering the whole set of variables. Then, we estimate the semi-parametric Cox model, which is multidimensional and does not need to specify a distribution *a priori*.

Semi-parametric approach - Cox approach

Cox estimators are presented in table 4. The covariates selection method used is a stepwise process. In the PH model, estimates can be interpreted with their hazard ratios. It is defined as the ratio of hazards for different values of the considered covariate. Then, for example, the hazard ratio of the binary variable high household income is 0.887. The hazard rate of high household income individuals is 88% of the hazard of individuals that are not in this class. The covariates with hazard ratio less than 1 ($\beta<0$) will reduce hazard rate and as a consequence increase survival and TTB. This is the case of the following covariates: male, high household income, full-time worker, unemployed, central localisation, Friday.

Modes of transport have the highest hazard ratio. It can be ordered by increasing TTB effect: walk, cycle, motorcycle, car as driver, car as passenger and transit. Walk, cycle and motorcycle hardly decrease the expected TTB with a hazard ratio near 200%.

Table 4 : Results of Cox model estimation – Stepwise selection.

Variable	DF	Estimates	Standard Error	Chi-Squared	P-value Pr > ChiSq	Hazard Ratio
Number of children under 5 years of age	1	0.04195	0.02010	4.3553	0.0369	1.043
Sex	1	-0.17793	0.01954	82.8864	<.0001	0.837
High household income	1	-0.11983	0.02503	22.9241	<.0001	0.887
Monday	1	0.13216	0.02440	29.3423	<.0001	1.141
Tuesday	1	0.07924	0.02422	10.7027	0.0011	1.082
Friday	1	-0.07370	0.02790	6.9753	0.0083	0.929
Age > 50 years	1	0.08848	0.02408	13.4976	0.0002	1.093
Central localisation	1	-0.15168	0.03405	19.8467	<.0001	0.859
1st ring East	1	0.06568	0.02700	5.9187	0.0150	1.068
3rd ring East	1	0.10015	0.02980	11.2917	0.0008	1.105
Full-time worker	1	-0.04902	0.02426	4.0829	0.0433	0.952
Scholar	1	0.40004	0.03017	175.8488	<.0001	1.492
Unemployed	1	-0.13900	0.04513	9.4871	0.0021	0.870

Variable	DF	Estimates	Standard Error	Chi-Squared	P-value Pr > ChiSq	Hazard Ratio
Principal mode : motorcycle	1	1.32032	0.15256	74.9003	<.0001	3.745
Principal mode: walk	1	1.35821	0.06434	445.5931	<.0001	3.889
Principal mode: transit	1	0.33295	0.06410	26.9832	<.0001	1.395
Principal mode: cycle	1	1.18782	0.12593	88.9632	<.0001	3.280
Principal mode: private car as driver	1	0.70001	0.06139	130.0142	<.0001	2.014
Principal mode private car as passenger	1	0.68918	0.06248	121.6516	<.0001	1.992

This semi-parametric approach confirms the non-parametric intuitions on covariates effects and selects the most influential covariates to be included in the model. But the hazard function is not estimated with this method and also it gives no information on the duration dependence. In the final part of the estimation, the full parametric model allows to estimate covariates coefficients and the duration dependence simultaneously.


Parametric approach
Classically, applied duration models to duration activity have used Weibull distribution function (Mannering et al., 1994; Kitamura et al., 1997). This distribution corresponds to a monotone hazard, which in this application to the TTB is not observed. The non-parametric approach concludes to a non-monotonic hazard function and rejects exponential and Weibull distribution functions. Then, the accelerated lifetime models with the log-normal and log-logistic distributions are estimated. The log-logistic model produces the best likelihood and residuals[4]. Then, only the log-logistic model is presented. For this estimation, the set of covariates is identical to the set resulting from the stepwise process of the Cox model.
The plots of Cox-Snell residuals evaluate the goodness-of-fit (figures 4a and b). Figure 4a presents Cox-Snell residuals of the log-logistic model, which is the best model compared to the exponential, Weibull (figure 4b) and log-normal models.

---

[4] In general, likelihood-ratio statistics can be used to compare models, those are nested within another. The exponential, Weibull and log-normal models are special cases of the generalised gamma model, and can then be compared. But the log-logistic is not nested within the generalised gamma distribution. Then, we compare the goodness-of-fit of the log-normal and log-logistic model with likelihood level and residuals of Cox-Snell.

Cox-Snell residuals plot for log-
logistic model
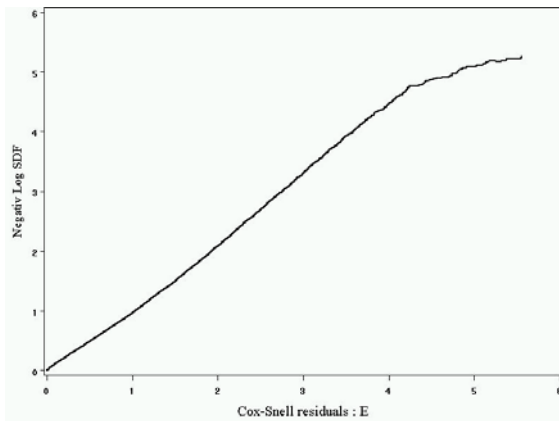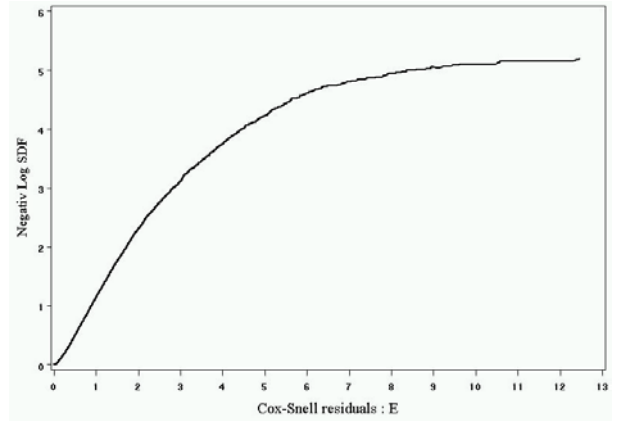
Fig. 4b Cox-Snell residuals plot for log-
normal model

Table estimates (table 5) shows the used covariates. In a first step, the model is estimated with the set of covariates produced by the stepwise Cox model. The class of ages is found to be non-significant. Then, the age is substituted to the class of ages and is significant. The covariates are all significant at 5%, except the number of children of age less than 5 years and residents of the 1st ring East, which are significant at 10% and the unemployed at 15%.

In an accelerated lifetime model, estimates can be interpreted in terms of expected time ratio. For example, the expected TTB of men is 8% greater than the expected TTB of women. For the quantitative covariate: age, the expected TTB will decrease by 2% for each year.

The estimates are closed to those produced by the Cox model. Transport modes are still the most influential covariates. They can be ordered in the same way. Walk, cycle and motorcycle have the most decreasing effect on expected TTB. Scholar has shorter TTB. Full-time worker and unemployed have longer TTB. High-income household member shows longer TTB. Finally, Monday and Tuesday present shorter TTB and Friday longer TTB.

The log-logistic scale is smaller than 1, corresponding to a non-monotonic hazard with inverted U-shape. Hazard is then increasing until 76min and decreasing afterwards. The median residual TTB is then decreasing and increasing.
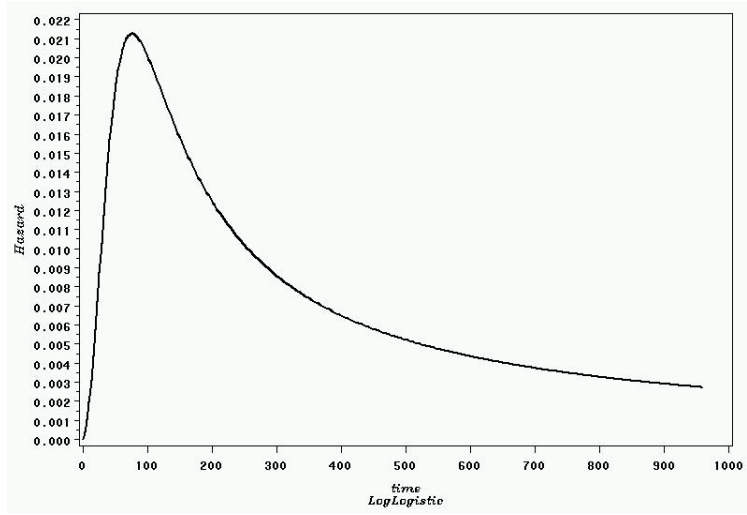
Fig. 5 Estimated hazard for log-logistic model

LogLogistic

<p style="text-align:center;">*Table 5* : Log-logistic parametric model</p>

Variable	DF	Estimate	Standard Error	Chi-Squared	P-value Pr > ChiSq
Intercept	1	4.82600	0.04182	13314.9011	<.0001
Number of children under 5 years of age	1	-0.02309	0.01350	2.9243	0.0873
Sex	1	0.08295	0.01225	45.8459	<.0001
High income household	1	0.07726	0.01582	23.8599	<.0001
Monday	1	-0.09897	0.01563	40.0949	<.0001
Tuesday	1	-0.05386	0.01550	12.0782	0.0005
Friday	1	0.05154	0.01767	8.5048	0.0035
Central localisation	1	0.09469	0.02141	19.5589	<.0001
1st ring East	1	-0.02904	0.01712	2.8781	0.0898
3rd ring East	1	-0.07263	0.01903	14.5728	0.0001
Age	1	-0.0021850	0.0004271	26.1761	<.0001
Full-time worker	1	0.04855	0.01513	10.2925	0.0013
Scholar	1	-0.27025	0.02316	136.1100	<.0001
Unemployed	1	0.04662	0.03000	2.4149	0.1202
Principal mode : motorcycle	1	-0.80281	0.09230	75.6598	<.0001
Principal mode: walk	1	-1.05826	0.03852	754.9015	<.0001
Principal mode: transit	1	-0.23885	0.03816	39.1771	<.0001
Principal mode: cycle	1	-0.98843	0.08095	149.1079	<.0001
Principal mode: private car as driver	1	-0.54105	0.03672	217.1432	<.0001
Principal mode private car as passenger	1	-0.60411	0.03749	259.7182	<.0001
Logistic Scale	1	0.37796	0.0028350		
Log Likelihood		-12776.297			

6. Conclusion

The survival analysis presented in this paper is applied to the travel time budgets (TTB) of Lyon (France). The sum of daily travel times is analysed with respect to the non-parametric lifetable approach, the semi-parametric Cox approach and the full-parametric approach. The first method gives incentives to use a non-monotonic *a priori* distribution in the parametric model, but it can not estimate the covariates effects. The stepwise selection in the Cox model permits a selection of covariates to be included in the parametric approach. But it can not estimate the duration dependence. Finally, the parametric model is constructed using the resulting set of covariates of the Cox model and non-monotonic distributions. Transport modes hardly impact on daily travel duration. Classical covariates, such as presence of children, gender, age, household income, household localisation, employment status, day of trips, are shown to be significant. The log-logistic model gives best goodness-of-fit. The estimated log-logistic scale implies a non-monotonic inverted U-shaped hazard, with inflexion point near 75 min.

The Zahavi's hypothesis can then be discussed. The TTB stability is observed by Zahavi at the world level, and has been recently validated by Schafer and Victor (2000). It can explain a part of the observed reinvestment in transport of travel-time savings due to increased speeds. Furthermore, it gives responsibilities of increasing mobility to speeds and speed-policies.

However, the application of the stability hypothesis to a finer level of observation, is irrelevant. Zahavi's studies and numerous followers on the analyses of TTB in different cities,

have shown many relationships existing between TTB and certain socio-economic, urban and transport variables. Major part of these analyses of TTB is unidimensional or limited to the linear model. The duration data can be examined with the appropriate method of survival analysis. The application of duration model to the TTB of Lyon shows effects of a set of covariates. This result shows the irrelevancy of the TTB stability hypothesis in the city of Lyon. The stability will mask the multiple mechanisms acting in the time allocation process. Furthermore, the non-monotonic hazard implies that the probability of ending daily transport, given it has lasted to a specified time, is not stable. Under TTB stability hypothesis, this conditional probability is expected to be monotonically increasing. The monotonic hazard will characterise a duration that is generated by a minimisation process. The estimated log-logistic hazard seems to show that everything happens as if 2 groups of travellers exist. The behaviour of a first group of individuals can be represented by the minimisation mechanism. A second group is composed of individuals that can not or do not want to minimise their TTB.

To gain robustness, the eventuality of heterogeneity between individuals needs to be included in this study. Furthermore, the application of duration model to the TTB failed to consider transport as a derived demand. Here, the duration of daily transport is disconnected from the pursued activities. Hence, the activity duration should be included in the covariates set. Duration models may offer an appropriate framework to reach the integration of derived demand concept into the allocation of time modelling. The competing hazards model allows modelisation of multiple duration processes in competition.

Références

Allison P.D., 1995, « Survival analysis using SAS – A practical guide », SAS Publishing.

Bhat C.R., T. Frusti, H. Zhao, S. Schönfelder, K.W. Axhausen, (2002), « Intershopping duration : an analysis using multiweek data », paper presented at the 81st Annual Meeting of Transportation Research Board, Washington, D.C..

Bhat C.R., (2000), « Duration modelling, in Handbook of transport modelling », (D.A. Hensher and K.J. Button, eds), Elsevier Science, pp.91-111.

Bhat C.R., (1996a), « A hazard-based duration model of shopping activity with non-parametric baseline specification and non-parametric control for unobserved heterogeneity ». Transportation Research part B, n°30, pp.189-207.

Bhat C.R., (1996b), « A generalised multiple durations proportional hazard model with an application to activity behaviour during the evening work-to-home commute », Transportation Research part B, n°30, pp.465-480.

Bieber A., M.-H. Massot and J.P. Orfeuil, (1994), « Prospects for Daily Urban Mobility », Transport Reviews, n°14(4), pp.321-339.

Chumak A. and Braaksma, (1981), « Implications of the travel-time budget for urban transportation modelling in Canada », Transportation Research Record, n°794, pp.19-27.

Cox D. (1972), « Regression models and life tables », Journal of the Royal Statistical Society Series, B34, pp.187-202.

Downes J.D., Morrell D., 1981, « Variation of travel time budgets and trips rates in Reading », Transportation Research, A n°15, pp.47-54.

Ettema D.F., Borgers AWJ, Timmermans HJP, (1995), « Competing risk hazard model of activity choice, timing, sequencing and duration ». Transportation Research Record, n°1493, pp.101-109.

Gilbert C., (1992), « A duration model of automobile ownership », Transportation Research part B, n°26, (2), pp.97-114.

Goodwin P.B., 1981, « The usefulness of travel budgets », Transportation Research part A n°15, pp.97-106.

Godard X., 1981, « Les budgets temps de transport – Analyse de quelques agglomérations françaises », Rapport IRT, n.31.

Gordon P., A. Kumar and H.W. Richardson, (1989), « The influence of metropolitan spatial structure on commuting time », Journal of Urban Economics, n°26, pp.138-151.

Gordon P., H.W. Richardson and M.J. Jun, (1991), « The commuting paradox – Evidence from the top twenty », Journal of the American Planning Association, n°54(4), pp.416-420.

Gunn H.F., 1981, « Travel budgets – a review of evidence and modelling implications », Transportation Research part A, n°15, pp.7-24.

Hamed M.K., F. Mannering, (1993), « Modelling travellers post-work activity involvement : Toward a new methodology », Transportation science, n°27(4), pp.381-394.

Hensher D., (1998), « The timing of change for automobile transactions : competing risk multi-spell specification », in Travel Behaviour research : updating the state of play, J.D. Ortuzar, D. Hensher, S. Jara-Diaz, Elsevier.

Hensher D., F. Mannering, (1994), « Hazard-based duration models and their application to transport analysis », Transportation Reviews, N°14 (1), pp.63-82.

Hensher D. and T. Raimond, (1992), « The timing of change : discrete and continuous time panels in transportation », Paper presented for the First US Conference on Panels for Transportation planning, Lake Arrowhead, California.

Hupkes G., 1982, « The law of constant travel time and trip-rates », Futures, February, pp.38-46.

Jovanis P. and H.L. Chang, (1989), « Disaggregate model of highway accident occurrence using survival theory », Accident Analysis and Prevention, n°21, pp.445-458.

Kaplan E. and P. Meier, (1958), « Non-parametric estimation from incomplete observations », Journal of the American Statistical Association, n°53, pp.457-481.

Katiyar R., Ohta K., 1993, « Concept of « Daily Travel Time » (DTT) and its applicability to travel demand analysis », Journal of the Faculty of Engineering, The University of Tokyo (B), Vol. XLII (2), pp.109-121.

Kiefer N.M., (1984), « Econometric duration data and hazard functions », Journal of Econometric Literature, n°27, pp.646-679.

Kitamura R., J. Robinson, T. F. Golob, M. Bradley, J. Leonard and T. van der Hoorn, (1992), « A comparative analysis of time use data in the Netherlands and California », Research report number UCD-ITS-RR-92-9, Institute of Transportation Studies, University of California, Davis, June.

Kitamura R., C. Chen, R. Pendyala, (1997), « Generation of synthetic daily activity-travel patterns », Transportation Research Board, n°1607, pp.154-162.

Kumar A. and Levinson D., (1995), « Temporal variations on allocation of time », Transportation Research Record, n°1493, pp.118-127.

Levinson D. and A. Kumar, (1995), « Activity, travel and the allocation of time », APA Journal, Autumn, pp.458-470.

Levinson D., Y. Wu and P. Rafferty, (2003), « Rational Locator Re-examined: Are travel times still stale ? », paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne.

Landrock J.N., 1981, « Spatial stability of average daily travel times and trip rates within Great Britain », Transportation Research part A, n°15, pp.55-62.

Mannering F., (1993), « Male/Female characteristics and accident risk: some new evidence », Accident Analysis and Prevision, n°25(1), pp.77-84.

Mannering F., E. Murakami, S.G. Kim, (1994) « Temporal stability of traveller's' activity choice and home-stay duration : some empirical evidence ». Transportation, n°21, pp.371-392.

Mannering F. and C. Winston, (1991), « Brand loyalty and the decline of American automobile firms », Bookings Papers on Economic Activity, Microeconomics, pp.67-114.

Meyer B.D., (1990), « Unemployment insurance and unemployment spells », Econometrica, n°58, pp.775-782.

Misra R., C.R. Bhat, (2000), « Activity travel patterns of non-workers in the San Francisco bay area : explanatory analysis », Transportation Research Record, n°1718, pp.43-51.

Mokhtarian P. L., C. Chen, 2002, « TTB or not TTB, that is the question: A review and analysis of the empirical literature on travel time (and money) budgets », available from the authors

Nam D.and F.Mannering, (2000), « An exploratory hazard-based analysis of highway incident duration », Transportation Research part A, n°34, pp.85-102.

Niemeier D.A. and J.G. Morita, (1996), « Duration of trip-making activities by men and Women », Transportation, n°23, pp.353-371.

Oakes, D., (1977), « The asymptotic information in censored survival data », Biometrika, n°64, pp.441-448

Paselk T. and F. Mannering, (1993), « Use of duration models for predicting vehicular delay at US/Canadian border crossings », paper presented at the 1993 Annual Meeting of the Transportation Research Board, Washington, DC.

Prendergast L.S., Williams R.D., 1981, « Individual travel time budgets », Transportation Research part A, n°15, pp.39-46.

Purvis C.L., 1994, « Changes in regional travel characteristics and travel time expenditures in the San Francisco Bay area : 1960-1990 », Transportation Research Record, n°1466, pp.99-109.

Schafer A., D.G. Victor, 2000, « The Future mobility of the world population », Transportation Research part A, n°34, pp.171-205.

Schönfelder S., K.W. Axhausen, (2000), « Analysing the rhythms of travel using survival analysis », working paper, Paper submitted to TRB Washington DC - January 2001

Timmermans H., P. Waerden, M. Alves, J. Polak, S. Ellis, A.S. Harvey, S. Kurose, R. Zandee, (2002), « Time allocation in urban and transport settings: an international, inter-urban perspective », Transport Policy, n°9, pp.79-93.

Van der Hoorn T., (1979), « Travel behaviour and the total activity pattern », Transportation, n°8, pp.308-328.

Vilhelmson B., (1999), « Daily mobility and the use of time for different activities: the case of Sweden », Geojournal, n°48, pp.177-185.

Zahavi Y., 1973, « The TT-relationship : a unified approach to transportation planning », Traffic engineering and control, pp.205-212.

Zahavi Y., 1974, « Travel time budgets and mobility in urban areas », Report prepared for the U.S Department of Transportation, Washington, D.C. and Ministry of Transport, Federal Republic of Germany, Bonn.

Zahavi Y., 1979, « The 'UMOT' Project », Report prepared for the U.S. Department of Transportation and the Ministry of Transport of Federal Republic Of Germany.

Zahavi Y., 1982, « Discussion », Transportation Research Record, n°879, pp.25-27.

Zahavi Y., Ryan J. M., 1980, « Stability of travel components over time », Transportation Research Record, n°750, pp.19-26.

Zahavi Y., Talvitie A., 1980, « Regularities in travel time and money expenditures », Transportation Research Record, n°750, pp.13-19.