

Volume 14 Number 1

January-March 1994

✓

NOTICE: This material may be protected by
copyright law (Title 17 U.S. Code)

NOTICE: This material may be protected by
copyright law (Title 17 U.S. Code)

Transport Reviews

T A Transnational
Transdisciplinary Journal

ISSN 0144-1647



Taylor & Francis
Publishers since 1798

(A)

Hazard-based duration models and their application to transport analysis

By DAVID A. HENSHER

Institute of Transport Studies, Graduate School of Business,
University of Sydney, NSW 2006, Australia

and FRED L. MANNERING

Department of Civil Engineering, University of Washington,
Seattle, WA, U.S.A.

A number of transport-related phenomena deal with a time element that defines the duration until an event's occurrence. Examples include the time that transpires until a trip is made, the length of time a commuter delays a trip departure to avoid traffic congestion, and the time until a newly introduced mode is used. Hazard-based duration models, which have enjoyed widespread use in a number of non-transport fields (e.g. economics, biostatistics), are an obvious choice for modelling such transport phenomena. The objective of this paper is to present hazard-based models, in a general way, to individuals interested in transport problems. In so doing, every effort is made to avoid a jargon-laden approach that typifies current articles and texts on the subject. It is hoped that such a presentation, along with an overview of existing transport applications of such models, will lead to an increased use of hazard-based duration models in transport.

1. Introduction

In many fields of study, developing an understanding of the factors that determine the time that transpires until or between the occurrence of specific events is often an important analytic focus. Data that deal with time (or duration) until event occurrence are commonly referred to as duration data. Such data are frequently encountered in transport applications. Examples include the time that transpires until an individual tries using a new mode or route, the time between vehicle purchases, the time between vehicle accidents, the time until trying a new technology (e.g. airbags in cars or an electric vehicle), the time until an incident (a vehicular accident or disablement) is cleared from a highway, the length of time a commuter delays a trip departure to avoid traffic congestion, the time that transpires between individuals' decisions to make a trip, and even the length of time waiting in vehicular queues at toll booths and/or international border crossings.

Despite the large amount of duration-related phenomena encountered in the transport field, surprisingly little has been done to analyse such phenomena statistically. In many respects this lack of analysis can be traced back to the historical development of transport modelling, which has been characterized by the use of cross-sectional data and methods. This cross-sectional mindset, with relatively few exceptions emanating from the growing interest in panel data, has pervaded both transport research and practice and has acted as a barrier to the exploration of the many duration-related issues encountered in the field.

The study of duration data is commonly undertaken in a number of non-transport disciplines. For example, industrial engineering looks at the length of time until machine failures (Mann *et al.* 1974), medical sciences look at the length of time of patient survival after corrective operations or disease treatments (Kalbfleisch and Prentice 1980, Fleming and Harrington 1990), and economists consider the length of time individuals are unemployed (Kiefer 1988). These fields have developed and applied hazard-function methods to study duration data statistically. A hazard-function may be interpreted as the instantaneous probability that episodes in the interval $(t, t + \Delta t)$ are terminating provided that the event has not occurred before the beginning of the interval. An episode (or spell) is the period of time between successive events. Events are changes in the set of all distinct values that a variable may take.

Such methods focus on the probability of an end-of-duration occurrence (e.g. a previously unemployed individual getting a job) given that the duration has lasted (i.e. conditioned on the duration lasting) to some specified time. This conditional probability of a duration ending is an extremely important concept because, in many instances, the probability of ending a duration is clearly dependent on the length of time the duration has lasted. For example, consider the availability of a new transport mode. It would be expected that the probability of an individual trying the new mode would change from day to day as a result of marketing efforts and word-of-mouth feedback from others that have already tried the mode. Such changes in probabilities are well-suited to hazard-based duration modelling methods which provide a tight link between theory and the empirical approach. While hazard-based approaches do not offer any computational advantage over approaches that fit probability distributions to the duration data directly, they do allow one to formulate the problem in terms of the conditional probabilities of interest, and such a formulation can provide valuable insight into the empirical estimation of the model.

The intent of this paper is to introduce hazard-based duration modelling, demonstrate its applicability to the study of transport problems, to review existing transport applications of hazard-based models and to provide directions for future research. The paper begins by providing an intuitive overview of hazard-based models with a focus on data structure and the problems that incomplete data may introduce. Next, hazard-based models are mathematically introduced and proportional and accelerated-lifetime forms are discussed. This is followed by a presentation of possible assumptions regarding the distribution of durations and the implications associated with these assumptions. Issues of heterogeneity and state dependence are then discussed along with possible methods of addressing these important issues. Other possible statistical complications and alternative modelling approaches are presented as well. The paper then gives an overview of known applications of hazard-based models to transport problems and concludes with a summary and directions for future research.

2. Hazard-based models: data structure and intuitive overview

To illustrate the dimensions of the problems presented when one chooses to analyse duration data, the example of the introduction of a new mode of travel is used. In this case, the analyst would not necessarily be interested in the 'equilibrium' state of mode acceptance, which is the state that is presumably being captured when one undertakes the estimation of standard logit-based mode-choice models. That is, standard logit-based analyses of mode choice probabilities assume an instantaneous adjustment to price, performance and other factors upon which modes are compared.

How
initial
slow
viabi
where
mode
St
exam
inform
indivi
C. At
by pe
event
spells
mode.
An
new m
B, it m
expose
area w
determ
the ad
(e.g. in
Left-ce
mode i
lengths
In g
of cens

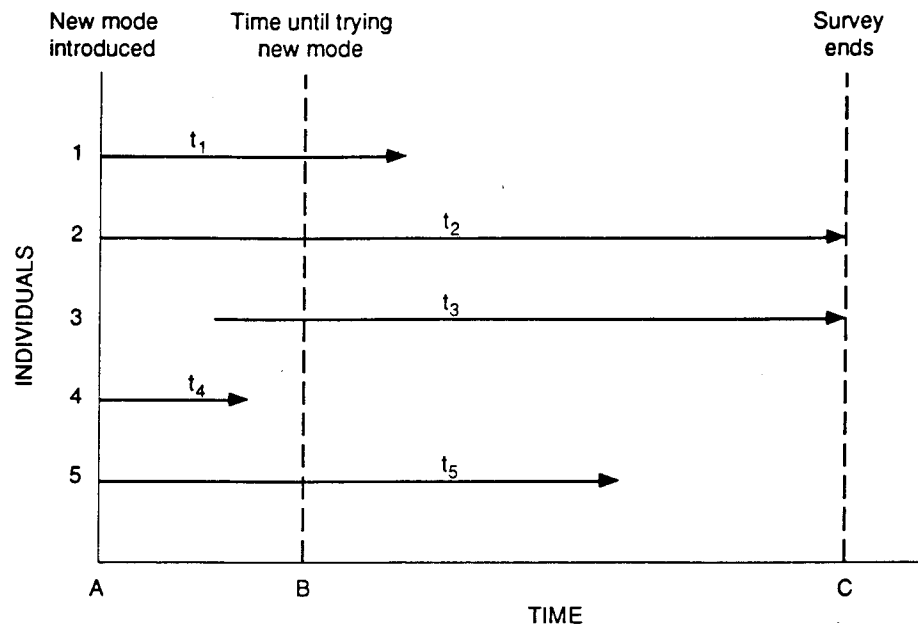


Figure 1. Example of duration data.

However, there may be a strong interest in looking at the rate at which individuals initially try the new mode on their way to establishing an equilibrium state, because slow acceptance may create political and financial pressures that could affect modal viability. Analysis of this acceptance rate is a classic application of duration data, where duration in this case is defined as the time between the introduction of the new mode and the time individuals first try the new mode.

Structurally, the data needed to model this duration problem is illustrated by the example provided in figure 1. In this figure, five individuals are sampled to obtain information on their trying a new mode. Information on the modal choices of these individuals is collected over some period of time until the survey is terminated at time C. At time C, there will likely be a group of individuals (e.g. individuals represented by person 2 in figure 1) that either; (a) will never try the new mode, or (b) will eventually try the new mode, but just have not done so up to time C. The duration spells of these individuals will be censored since they are not observed trying the new mode. This type of censoring is referred to as right-censoring.

Another type of censoring could arise if the survey was begun some time after the new mode was introduced. In figure 1, for example, if the survey was started at time B, it may be difficult to determine when an individual (such as individual 3) was first exposed to the new mode. Such an individual may have moved into the geographic area where the mode is available after the mode was first introduced. Being unable to determine when durations begin is referred to as left-censoring. Left-censoring poses the additional problem of not knowing the value of the determinants of duration (e.g. income, household size, attitudes) at the beginning of the duration period. Left-censoring can be avoided, in this case, by beginning the survey when the new mode is first introduced (time A in figure 1). This will ensure full knowledge of the lengths of durations as well as possible determinants of durations.

In gathering duration data, it is important to avoid left-censoring because this type of censoring is difficult to handle in hazard-based models of duration. In contrast,

right-censoring can be more easily handled. The modelling consequences of censoring will be discussed later in this paper.

3. Hazard-based models: mathematical approach

Mathematically, the hazard function can be expressed in terms of a cumulative distribution function, $F(t)$, and a corresponding density function, $f(t)$. The cumulative distribution is written as,

$$F(t) = \Pr [T < t] \quad (1)$$

where \Pr denotes the probability, T is a random time variable, and t is some specified time. In the case of the time until the acceptance of a new mode, (1) gives the probability of trying the new mode before some transpired time, t .

The corresponding density function (the first derivative of the cumulative distribution with respect to time) is,

$$f(t) = dF(t)/dt \quad (2)$$

and the hazard function is,

$$h(t) = f(t)/[1 - F(t)] \quad (3)$$

where $h(t)$ is the conditional probability that an event will occur between time t and $t + dt$ given that the event has not occurred up to time t . In words, the hazard, $h(t)$, gives the rate at which events (such as trying a new mode) are occurring at time t , given that the event has not occurred up to time t .

Another important construct in hazard-based models is the survivor function. The survivor function gives the probability that a duration will be greater than or equal to some specified time t . That is, the probability that an individual remains in the state ('survives') until time t . The survivor function is written as,

$$S(t) = \Pr [T \geq t] \quad (4)$$

and therefore is related to the cumulative distribution function by,

$$S(t) = 1 - F(t) \quad (5)$$

and to the hazard function by

$$h(t) = f(t)/S(t) \quad (6)$$

Graphically, hazard, density, cumulative distribution and survivor functions are illustrated in figure 2. This figure provides a visual perspective of the equations presented above.

Turning specifically to the hazard function, its slope has important implications. Recall that, in the introduction, we talked about the possibility that the probability of ending a duration may be dependent on the length of the duration. This is referred to as duration dependence and the first derivative of the hazard function with respect to time (i.e. the slope of the hazard function) provides this information.

To illustrate this, consider the four hazard functions shown in figure 3. In this figure, the first hazard function, $h_1(t)$, has $dh_1(t)/dt > 0$ for all t . This is a hazard that is monotonically increasing in duration implying that the longer individuals go without exiting a duration, the more likely they are to exit soon. The second hazard function has $dh_2(t)/dt < 0$ for all t and is monotonically decreasing in duration. This implies the longer individuals go without exiting a duration the less likely they are to exit soon.

Fig

The
inde
func
leng
I
the h
durat
durat
mode
hazar
accel

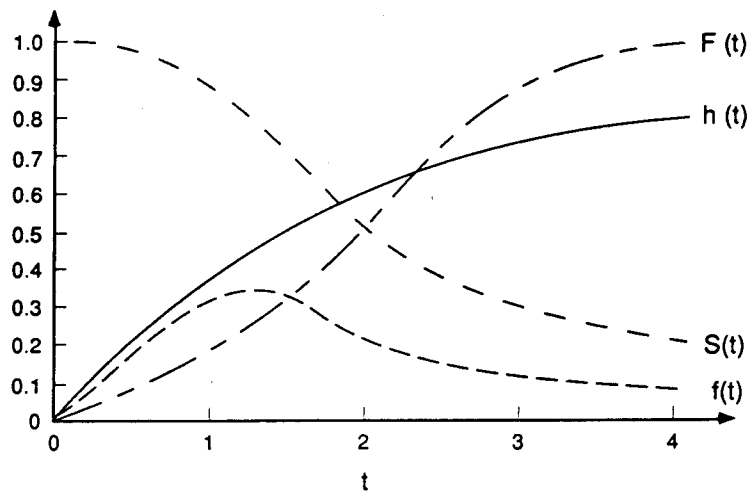


Figure 2. Illustration of hazard ($h(t)$), density ($f(t)$), cumulative distribution ($F(t)$) and survivor functions ($S(t)$).

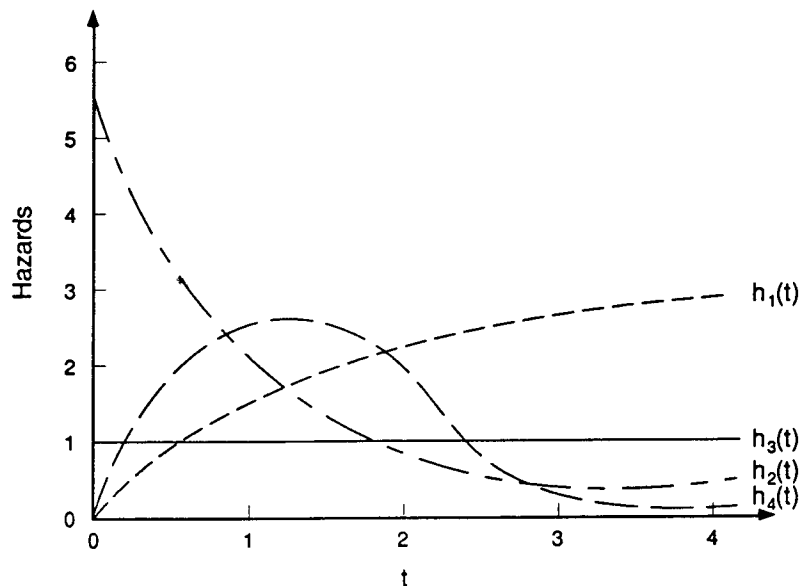


Figure 3. Illustration of four alternative hazard functions.

The third hazard function has $dh_3(t)/dt = 0$ which means that exit probabilities are independent of duration and no duration dependence exists. Finally, the fourth hazard function is non-monotonic and has $dh_4(t)/dt > 0$ and $dh_4(t)/dt < 0$ depending on the length of duration t . In this case the exit probabilities increase or decrease in duration.

Information relating to duration dependence, as derived from the first derivative of the hazard function with respect to time, can provide important insights into the duration process being modeled. However, there are clearly important determinants of duration (e.g. socioeconomic characteristics) that must be accounted for in the modelling approach as well. These determinants, or covariates, are included in hazard-based models using two alternate methods; proportional hazards and accelerated lifetime.

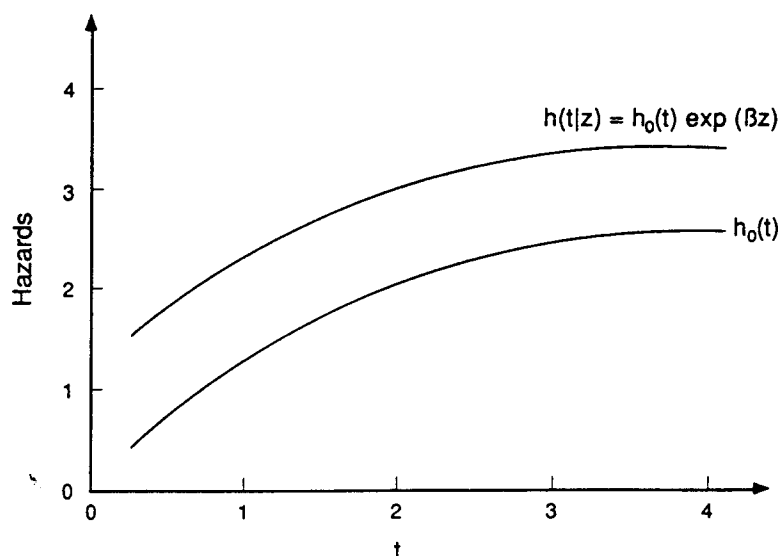


Figure 4. Illustration of the proportional hazards model.

Proportional hazards models operate on the assumption that covariates (i.e. factors that affect duration) act multiplicatively on some underlying hazard function. The proportionality is due to the decomposition of the hazard rate into one term dependent upon time, and another dependent only on the covariates. This concept is illustrated in figure 4.

In this figure, the underlying (or baseline) hazard function is denoted $h_0(t)$, and this is the hazard function assuming all elements of the covariate vector, Z , are zero. The manner in which covariates are assumed to act on the baseline hazard is usually specified as the function $\exp(\beta Z)$, where β is a vector of estimable parameters. Therefore the hazard rate with covariates, $h(t|z)$, is given by the equation (as shown in figure 4),

$$h(t|Z) = h_0(t) \exp(\beta Z) \quad (7)$$

Proportional hazards models have enjoyed considerable popularity in a variety of fields (see Fleming and Harrington 1990). These models can easily handle right-censored data and they provide a nice interpretation of estimated parameters (i.e. simple multiplicative effect on the underlying hazard). The assumption of proportionality, however, limits the application set. For example, if a covariate is car ownership (zero, non-zero), the quotient of the hazard rate of owners and non-owners should not vary over time. This restriction can be relaxed to an extent by introducing class-specific hazard rates: $h_c(t|z) = h_{0c}(t) \exp(\beta z)$ where $c = 1, \dots, C$ classes.

An alternative approach to incorporating covariates in hazard-based models is the accelerated lifetime model. This model assumes that the covariates rescale time directly (i.e. accelerate time) in a baseline survivor function which is the survivor function when all covariates are zero. Assuming that the covariates act in the form $\exp(\beta Z)$, as was the case for the proportional hazards model, the accelerated lifetime model can be written as,

$$S(t|Z) = S_0[t \exp(\beta Z)] \quad (8)$$

and it follows that this model can be written in terms of hazard functions as,

$$h(t|Z) = h_0[t \exp(\beta Z)] \exp(\beta Z) \quad (9)$$

Accelerated lifetime models have, along with proportional hazards models, enjoyed wide-spread use (see Kalbfleisch and Prentice 1980). The selection of accelerated lifetime or proportional hazards models is often determined on the basis of distributional assumptions (i.e. the assumed distribution of durations). Commonly used distribution assumptions are discussed below.

4. Distributional alternatives

Two general approaches to implementing hazard-based models are possible. One is to assume a distribution of duration (e.g. Weibull, exponential, etc.) and the other is to apply a generalized approach that does not require a distributional assumption. The former approach is called 'fully parametric' because a distributional assumption is being made for the hazard along with an assumption on the functional form specifying how covariates interact in the model (i.e. the $\exp(\beta Z)$ used in the previous section). The latter approach is semi-parametric because only the covariate functional form is specified.

Fully parametric models can be estimated in proportional hazards or accelerated lifetime forms, and a variety of duration-distribution alternatives are available including gamma, exponential, Weibull, log-logistic, and log-normal. The choice of any one of these alternatives can be justified on theoretical grounds, and each has important implications relating to the shape of their underlying hazard functions. Three common distributions; exponential, Weibull, and log-logistic, are summarized below.

The exponential distribution is the simplest to apply and interpret. With parameter $\lambda > 0$, the exponential density function is,

$$f(t) = \lambda \exp(-\lambda t) \quad (10)$$

with hazard,

$$h(t) = \lambda \quad (11)$$

Equation (11) implies that this distribution's hazard is constant and thus the probability of exiting a duration is independent of the length of time of the duration. This is a fairly restrictive assumption because the exponential distribution does not allow any sort of duration dependence to be captured.

The Weibull distribution is a more generalized form of the exponential in that it allows for positive duration dependence (hazard is monotonic increasing in duration), negative duration dependence (hazard is monotonic decreasing in duration) or no duration dependence (hazard is constant in duration). With parameters $\lambda > 0$ and $P > 0$, the Weibull distribution has density function,

$$f(t) = \lambda P (\lambda t)^{P-1} \exp[-(\lambda t)^P] \quad (12)$$

with hazard,

$$h(t) = \lambda P (\lambda t)^{P-1} \quad (13)$$

In (13), if the Weibull parameter P is greater than one, the hazard is monotone increasing in duration, if P is less than one it is monotone decreasing in duration, and if P equals one, the hazard is constant in duration and reduces to the exponential distribution's hazard (i.e. $h(t) = \lambda$). Since the Weibull distribution is a generalized form

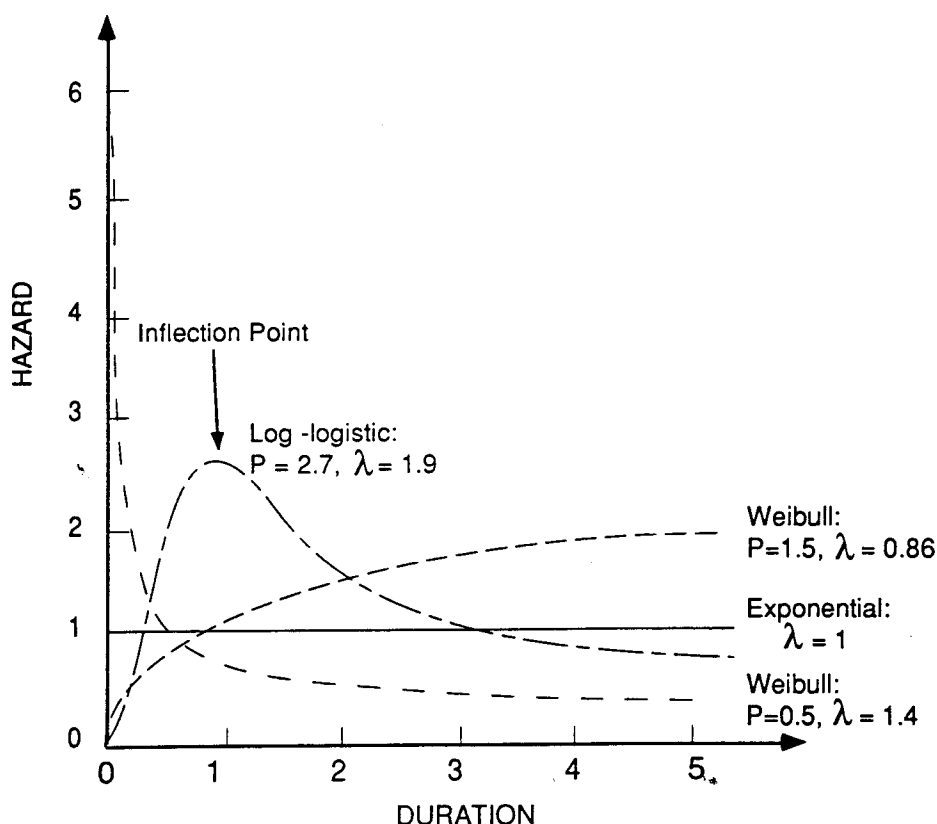


Figure 5. Hazard function distributions (Kiefer 1988).

of the exponential distribution it provides a more flexible means of capturing duration dependence, but it is still limited due to the monotonicity restriction that it places on the hazard. In many applications, a non-monotonic hazard may be theoretically justified.

The log-logistic distribution allows for non-monotonic hazard functions and is often used as an approximation of the more computationally cumbersome log-normal distribution. The log-logistic, with parameters $\lambda > 0$ and $P > 0$ has the density function,

$$f(t) = \lambda P (\lambda t)^{P-1} [1 + (\lambda t)^P]^{-2} \quad (14)$$

and hazard function,

$$h(t) = [\lambda P (\lambda t)^{P-1}] / [1 + (\lambda t)^P] \quad (15)$$

Note that the log-logistic's hazard is identical to the Weibull's except for the denominator. Equation (15) shows that if $P < 1$, the hazard is monotone decreasing, if $P = 1$, the hazard is monotone decreasing from parameter λ , and if $P > 1$, the hazard increases from zero to a maximum at time $t = [(P-1)^{1/P}] / \lambda$ and decreases toward zero thereafter.

Figure 5 shows a comparison of the hazards of the three distributions discussed. In this figure an exponential distribution is presented along with monotone increasing and decreasing Weibull distributions, and a non-monotonic log-logistic distribution. The selection of a distribution is in part guided by reasonable hypotheses on behavioural response over time. For example, in the case of a new mode, individuals who are eager to choose it but then lose interest might be represented by the

log-logistic; those unaffected by advertising and word-of-mouth might be represented by the exponential distribution.

The alternative to assuming a distribution of the hazard is to use a non-parametric approach for modelling the hazard. This is convenient when little or no knowledge of the functional form of the hazard is available. Such an approach was developed by Cox (1972) and is based on the proportional hazards approach. The Cox proportional hazards model is semi-parametric because $\exp(\beta Z)$ is used as the functional form of the covariates. The model is based on the ratio of hazards so that the probability of an individual, i , exiting a duration at time t_i , given that at least one traveller exits at time t_i , is given as,

$$\exp(\beta Z_i) / \sum_{j \in R_i} \exp(\beta Z_j) \quad (16)$$

where R_i denotes the set of individuals with durations greater than or equal to t_i .

The Cox proportional hazard model has been used in a number of fields (see Fleming and Harrington 1990, Breslow 1974, Elandt-Johnson and Johnson 1983). Some caution should be exercised when applying semi-parametric models. If the hazard is generated from a known distribution, and a Cox model is applied, statistical efficiency will be lost since information regarding the hazard's distribution is not being used. This could result in less precise coefficient estimates as reflected by their higher standard errors. Although this efficiency matter is of some concern, several studies (e.g. Efron 1977, Oaks 1977) have found the asymptotic variance matrix of Cox model estimators to be close to those generated from fully parametric hazards models. Thus, in most cases, Cox models can be applied without serious efficiency losses.

5. Heterogeneity

The assumption implicitly made in proportional hazards models is that the survivor function (4) is homogeneous over the population being studied. As such, all of the variation in durations is assumed to be captured by the covariate vector Z . A problem arises when some unobserved factors (i.e. not included in Z) influence durations. This is called unobserved heterogeneity and can result in a major specification error that can lead one to draw erroneous inferences on the shape of the hazard function and covariate coefficient estimates (Heckman and Singer 1984, Lancaster 1979, Gourieroux *et al.* 1984). Ignoring heterogeneity is the equivalent to leaving out an important covariate in the $\exp(\beta Z)$ function. Fortunately, a number of corrections have been developed to account for heterogeneity explicitly. The most common is to include a heterogeneity term that is designed to capture unobserved effects across the population, and work with the conditional duration density function. With a heterogeneity term, v , having a distribution over the population, $g(v)$, and with a conditional duration density function, $f(t|v)$, the unconditional duration density function can be determined from,

$$f(t) = \int f(t|v)g(v)dv \quad (17)$$

With this formulation, hazard models can be derived using procedures identical to those used in the derivation of the non-heterogeneity hazards models.

The problem in operating such an heterogeneity model is that a distribution of heterogeneity in the population must be selected. There is seldom any theoretical justification for selecting one distribution over another, and the economics and

marketing literature is strewn with papers that have used a wide-variety of heterogeneity distributions, the most popular of which is the gamma distribution (Hui 1990, Gupta 1991, Greene 1992). The selection of a heterogeneity distribution must not be taken lightly. The consequences of incorrectly specifying $g(v)$ are potentially severe and can result in inconsistent estimates as demonstrated both theoretically and empirically by Heckman and Singer (1984). Fortunately, from the perspective of choosing among many possible distributions, it has been shown (Kiefer 1988) that if a correctly specified duration distribution is used, the coefficient estimate results are not highly sensitive to alternate distributional assumptions of heterogeneity. To avoid concern about heterogeneity assumptions entirely, Heckman and Singer (1984) propose a non-parametric representation of heterogeneity that requires no prior parametric assumptions. Their method has been successfully applied and appropriate software is available (Vilcassim and Jain 1991).

6. State dependence

State dependence in duration models considers the effect that past duration experiences have on current durations. Such dependence can capture important habitual behaviour effects that can be strong indicators of the length of durations. Heckman and Borjas (1980) provide an extensive discussion of state dependence issues in hazard-based models.

In most models of duration, three types of state dependence can exist: duration dependence, occurrence dependence and lagged duration dependence. Duration dependence simply focuses on the conditional probability of a duration ending soon, given that it has lasted until some known time. This type of state dependence is captured in the shape of the hazard function (see figure 5). For example, a monotone increasing hazard (Weibull with $P = 1.5$ and $\lambda = 0.86$ as shown in figure 5) has positive duration dependence since the longer the individual's duration, the more likely the duration will end soon. Most hazard models (with the notable exception of the exponential distribution) implicitly embody some form of duration dependence.

Occurrence dependence captures the effect that the number of previous durations has on the current duration. For example, individuals that have delayed their departure from work to home to avoid traffic congestion four times during the past week may have different current-day departure-delay durations than individuals that have delayed only once in the past week. The four-delay individuals may have longer or shorter current-day delay durations because they are more experienced delaying and perhaps have a better notion of when to leave to optimize their avoidance of traffic congestion. Occurrence duration is accounted for by including the number of previous duration occurrences in the covariate vector Z .

Finally, lagged duration dependence captures the effect that the lengths of previous durations have on current duration. Returning to the example of delaying departure from work, an individual who has delayed a specified amount of time on a preceding day may have developed a habitual pattern that would make previous-day delay duration a good predictor of current-day delay duration. Again, this type of state dependence is accounted for by including lagged durations in the covariate vector Z .

Great caution must be exercised when including and interpreting state dependence. The common problem is that unobserved effects (heterogeneity) remain in the model and are 'picked up' in the coefficients of the state variables included in the covariate vector Z . For example, suppose that income is an important determinant of the length

of time that individuals delay their departure from work, but a duration model is estimated without income (i.e. income becomes the equivalent of an unobserved effect). If a lagged duration variable is included in the model, its estimated coefficient will be capturing lagged duration effects as well as residual income effects because income also determines, and is therefore correlated with, lagged duration dependence. In the presence of such heterogeneity, inferences drawn on state dependence could be erroneous because non-state effects are being captured.

Elbers and Ridder (1982) have shown that if heterogeneity is properly accounted for, duration dependence (i.e. the dependence captured by the shape of the hazard function) can be captured accurately. However, there are really two types of heterogeneity. One is 'pure' heterogeneity which refers to unobserved factors that are not influenced by previous duration involvement (as discussed above in the heterogeneity portion of the paper). The second is 'state dependent' heterogeneity and refers to unobserved factors that are influenced by an individual's previous duration involvement. This second type of heterogeneity is extremely difficult to distinguish from occurrence and lagged duration dependence even if heterogeneity is explicitly accounted for as shown in (17), because such corrective methods typically capture 'pure' but not 'state dependent' heterogeneity (see Heckman and Borjas 1980). One relatively simple solution to this problem is to instrument state variables by regressing them against exogenous covariates and using regression-predicted values as variables in the duration model.

To summarize, state dependence must be treated with considerable caution because the potential for serious misspecification is always present. The analyst must use a carefully thought out statistical approach to incorporate state effects.

7. Other modelling issues

Censoring, as discussed earlier and illustrated in figure 1, is an important concern in hazard-based model estimation. Right-censoring can be handled in both proportional hazards and accelerated lifetime models of duration. All that is required is a relatively minor modification to the likelihood function, and then estimation can proceed using standard maximum likelihood methods. However, when correcting for right-censoring, the assumption that is typically made is that individuals censored at any given time are a representative sample of the individuals continuing their durations up to the given time. This assumption usually holds, but unusual duration termination patterns could invalidate the standard right-censoring correction procedure and require further modification to the likelihood function (see Kalbfleisch and Prentice 1980).

Having data that is left-censored (see figure 1) presents a serious modelling problem. With left-censoring, the likelihood function soon becomes unwieldy. The problem becomes one of determining the distribution of duration 'start times', from which the contribution of left-censored observations to the model's likelihood function can be determined. In the presence of state dependence and heterogeneity, accounting for left-censoring is extremely difficult. For further information on the left censoring problem, the reader is referred to Heckman and Singer (1984) and Fleming and Harrington (1990).

As a final point, it should be noted that accounting for heterogeneity in a Cox semi-parametric duration model is conceptually straight forward but computationally cumbersome. The reason for this is that heterogeneity in the Cox model structure involves multiple integration over all observations. The numerical integration required can be prohibitive in large data sets. As a result, most studies that address

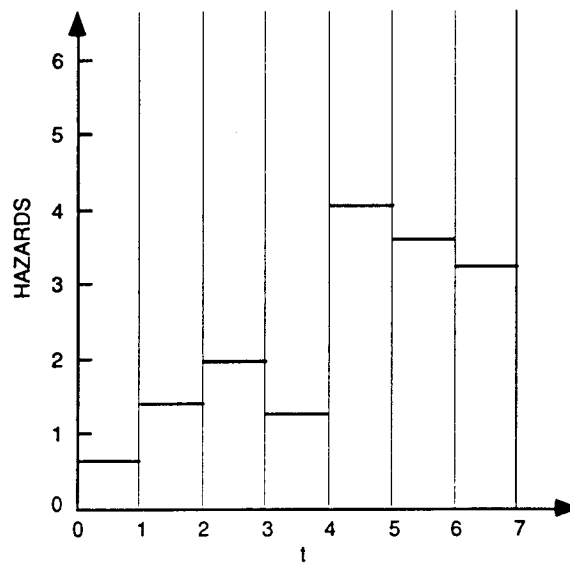


Figure 6. Illustration of a non-monotonic hazard function resulting from logistic regression analysis.

heterogeneity assume a parametric form of the underlying hazard. A discussion of the Cox model with heterogeneity is presented in Han and Hausman (1990).

A final modelling concern relates to time-varying covariates, which are covariates that change during individual durations. Empirically, time-varying covariates can be incorporated into hazard models by allowing the covariate vector to be a function of time (i.e. $Z(t)$ instead of Z) and re-writing the hazard and likelihood functions accordingly. The likelihood function understandably becomes more complex, but estimation is still possible and simplified by the fact that time-varying covariates usually do not change continuously over time (i.e. a few discrete changes can be more easily handled in the likelihood function). The problem with including time-varying covariates is that it becomes difficult to interpret coefficients and to separate out duration dependence (i.e. the shape of the hazard over time). For further information on time-varying covariates, the reader is referred to Peterson (1976) and Greene (1992).

8. Alternative modelling methods

Before the development of continuous-time hazard-based models (such as those described in previous sections), logistic regressions were used extensively in fields such as biostatistics to analyse duration data. This approach segments time into discrete intervals and then applies standard logistic regression methods to predict the probability of individual durations ending in these discrete time intervals. As illustrated in figure 6, such an approach allows for a quite general form of the hazard, because it can change, either increasing or decreasing, from one time interval to the next. However, the hazard is assumed to be constant within each time interval.

When compared to continuous-time methods, the logistic regression approach does not fare well. This is due to possible statistical efficiency losses resulting from the use of discrete time. Both proportional hazards and accelerated lifetime approaches have the capability of utilizing more time-related information than discrete-time techniques. This is because they consider exact failure times in a continuous-time context as

opposed to failure times that are defined by a more vague discrete-time interval. However, some studies have shown that under certain conditions, for example, logistic regression and Cox, proportional hazards models can produce very similar results (see Green and Symons 1983, Abbot 1985, Ingram and Kleinman 1989). These conditions are that the discrete time intervals chosen for the logistic regression must be sufficiently short and that the exit probabilities (i.e. the probability of durations ending) during the discrete-time intervals must be small.

Although logistic regression is theoretically inferior in terms of statistical efficiency, it offers at least two advantages over proportional hazards and accelerated lifetime approaches. First, handling time-varying covariates is comparatively easy because changes in covariates can be readily made from one discrete-time period to the next. Second, tied data, which can be problematic in proportional hazards and accelerated lifetime models, is not a problem in logistic regression approaches. Tied data occurs when a number of observations end their durations at the same time. This can result when data collection is not precise enough to determine the exact duration-ending times. Thus, duration exits tend to be grouped at specific times. In the presence of tied data, the likelihood function for proportional hazards and accelerated lifetime models becomes increasingly complex. Kalbfleisch and Prentice (1980) and Fleming and Harrington (1990) discuss tied data in the context of continuous-time models.

In many instances, problems with data ties and the need for time-varying covariates are considered more severe than possible efficiency losses. Consequently, discrete-time approaches continue to be developed. A good example is the recent work of Han and Hausman (1990). They developed a generalized discrete-time hazard approach that also accounts for possible heterogeneity. Their model provides for a non-parametric baseline hazard but assumes that heterogeneity is gamma distributed. The Han and Hausman work clearly shows that discrete-time techniques have their place in duration analysis.

9. Competing risks

Traditional duration analyses assume that durations end as a result of a single event. For example, the length of time a traveller stays at home before making a trip (i.e. home-stay duration) can be assumed to end when a trip is made. However, multiple duration-ending outcomes may be worthy of consideration because different outcomes could produce different durations. For the duration of travellers' home-stays, the type of trip ending the duration (e.g. shopping, social, work) could affect the length of duration. This possibility of multiple duration-ending outcomes is referred to as competing risks.

In the past, many researchers have assumed that a competing risks model with n possible outcomes had a likelihood function that could be separated into n distinct pieces. Under such an assumption, estimation could proceed by estimating separate hazard models for each of n possible outcomes. Unfortunately, separately estimating competing risks hazards inherently assumes independence among risks. This is frequently done (e.g. Katz 1986, Gilbert 1992) but may not always be appropriate because it ignores potentially important interdependence among risks. Treating competing risks independently is analogous to assuming recursivity in more traditional simultaneous equations problems (i.e. those problems that can be solved using three-stage-least squares and similar methods).

Accounting for interdependence among competing risks is not an easy task, but has been done by Diamond and Hausman (1984) and Han and Hausman (1990). Diamond and Hausman develop a model with strict parametric assumptions on the nature of interdependence. Han and Hausman extend this work by providing a flexible parametric form of interdependence. Their approach also allows one to test statistically whether the more common assumption of independence among competing risks is valid.

10. Review of transport applications

The application of hazard-based duration models in the transport field is comparatively new, with most work beginning in earnest in the late 1980s. This is surprising since the number of possible applications of hazard-based models in transport is quite large.

In applying hazard-based modelling methods to transport problems, it is important to view them as a reduced form of some underlying behavioural choice process. Such a view can help guide the selection of covariates, functional form of the hazard distribution, heterogeneity treatments and state dependence alternatives. Consideration of reduced forms is an important point and a departure from recent transport modelling which has tended to focus almost exclusively on complex behavioural choice processes with correspondingly complex and convoluted modelling methods. Although this complex behavioural focus centred on dynamic structural equations will continue as a research theme with a discrete-time focus, the transport researcher can gain additional insights into the underlying behavioural processes in continuous time by considering reduced form approaches.

A list of known transport applications of hazard-based studies is presented in table 1. This table shows the distribution assumptions made and whether or not heterogeneity and/or state dependence were considered. The studies listed in this table are discussed briefly below.

Some of the earliest applications of hazard-based models in transport dealt with accident analysis. Jovanis and Chang (1989) used a Cox proportional hazards model to look at the probability of accident occurrence on individual trips. They define duration as the length of trip-time before accident occurrence, with a non-accident trip being right-censored. Their study of accident reports from a less-than-truckload freight carrier yielded important results relating to the effect of driver characteristics and fatigue.

In other work, Chang and Jovanis (1990) provide a general structure for studying accident occurrence with hazard-based methods. Their paper addresses the many important theoretical and conceptual concerns involved in such studies. Lin *et al.* (1992) apply hazard-based methods to study the safety impacts of existing driving-hour regulations on less-than-truckload carriers. Their analysis extended and considerably expanded the earlier work of Jovanis and Chang (1989). Further work by Yang *et al.* (1992) applied a Cox proportional hazards model to study multiple-stop effects on truckers' driving risk. This paper provides an excellent demonstration of the flexibility of hazard-based approaches, and their empirical analysis uncovered many important relationships that would have been difficult if not impossible to capture using non-hazard-based analytic methods.

Jones *et al.* (1991) applied a fully parametric log-logistic accelerated-lifetime model to study the time required to restore capacity on Seattle freeways after the occurrence of a capacity-reducing traffic accident. They found that the log-logistic

Table 1. Review of transport applications of hazard-based duration models.

Study	Distribution assumptions					
	Cox proportional	Exponential	Weibull	Log-logistic	Heterogeneity	State dependence
Jovanis and Chang (1989)	•					
Chang and Jovanis (1990)	•					
Lin <i>et al.</i> (1992)	•					
Yang <i>et al.</i> (1992)	•					
Jones <i>et al.</i> (1991)		•	•	•		
Mannering (1991)			•			
Mannering (1993)			•			
Mannering and Hamed (1990)			•			
Hamed and Mannering (1993)			•			
Hamed <i>et al.</i> (1993)	•		•			•
Mannering <i>et al.</i> (1992)	•		•			•
Mannering and Winston (1991)			•			
Hensher (1992)	•			•		
Paselk and Mannering (1993)			•		•	
Hensher and Raimond (1992)			•		•	
Kim and Mannering (1992)			•		•	
Gilbert (1992)			•			

hazard was monotone decreasing, indicating that the longer the capacity reduction lasted, the less likely it was to end soon. This suggested that the Seattle area's accident management programme had a problem with severe accidents, and numerous recommendations were made as to how the shape of the hazard function could be changed through improved accident management procedures.

Mannering (1991, 1993) applied hazard models to study the time between individuals' traffic accidents. In fitting Weibull distributions (Mannering 1993), it was found that male drivers had a decreasing hazard (i.e. the longer they go without having an accident, the less likely they were to have an accident soon) whereas females had a constant hazard, indicating that accident probabilities were independent of the time that transpired without having an accident. These results suggested fundamental differences in gender and many interesting behavioural possibilities for these differences were proposed.

Hazard-based models have also been applied to study the dynamic effects of travel demand. For example, Mannering and Hamed (1990) applied a Weibull model to determine the length of time travellers delay their departure from work to avoid traffic congestion. This model was integrated with a logit-based choice model, thus demonstrating compatibility with more traditional transport modelling approaches. Also, Hamed and Mannering (1993) applied a Weibull model to study the time travellers spend at home between trip-generating activities (home-stay duration) and Mannering *et al.* (1992) and Hamed *et al.* (1992) applied a Cox proportional hazards model to study the same problem. These studies demonstrate another potentially important application of hazard-based models.

Hazard-based models have also been used to study automobile ownership. Mannering and Winston (1991) fitted a Weibull model to study the time between households' vehicle purchases. In other work, Hensher (1992) applied a Cox proportional hazards model to study the duration of automobile ownership in a household fleet, recognizing that many of the exogenous variables affecting the amount of time a vehicle is in a household change over time. Gilbert (1992) used a fully parametric Weibull duration model, specifying separate hazard functions for three different events that may end an ownership spell—replacement with a new vehicle, replacement with a used vehicle and disposal without replacement.

The work of Paselk and Mannering (1993) used hazard models to study vehicular delay at international border crossings. They used a number of fully parametric models but found that the log-logistic, with non-monotonic hazard, provided the best fit. Using data from the US/Canadian highway border crossing in Blaine, WA, their study of the hazard function revealed an increasing hazard until a vehicular delay of 21 minutes was reached and a decreasing hazard thereafter. This hazard inflection point indicated a deterioration in system operation at around 21 minutes of delay, and knowledge of this allowed corrective recommendations to be made.

Both Hensher and Raimond (1992) and Kim and Mannering (1992) estimated Weibull models while accounting for possible heterogeneity (using a gamma distribution). Hensher and Raimond studied the time until acceptance of a new tolled roadway facility and found significant heterogeneity effects. They also proposed a way of transforming panel data collected and observed in discrete time into a continuous time data set, so that duration models can be used with the rich set of panel data now accumulating in transport. Kim and Mannering included state dependence effects in their study of the length of individuals' activity duration (time between successive vehicular trips) and found heterogeneity effects to be insignificant. Both of these

studies provide interesting demonstrations of how heterogeneity can be handled in transport applications.

11. Summary and conclusions

This paper provides an overview of the structure of hazard-based duration models along with a discussion of application issues. Transport studies that have used hazard-based models have been briefly discussed with a focus on methods used and findings made.

It is clear from the material presented in this paper that hazard-based modelling methods have great potential as a tool to be used in the study of a wide-range of transport phenomena. It is also clear that transport modellers have not exploited to the full the potential that hazard-based models offer the profession. This is evidenced by the fact that comparatively few researchers (see table 1) are involved in the application of such methods. This is particularly disturbing in the light of the need for understanding the dynamics of traveller behaviour (i.e. the timing of trip-related decisions) and the increasing availability of transport panel data which will, theoretically, allow the study of changes in travel behaviour over time. These types of problems are ideally suited to the application of hazard-based methods, as other fields (e.g. economics, biostatistics) have already demonstrated. It is hoped that this paper will serve as a catalyst for change, and that transport modellers will embrace hazard-based models as one of the more important modelling tools available to the profession.

Acknowledgments

This research was partially funded by grant A79131696 from the Australian Research Council whose support is gratefully acknowledged. The comments of Bill Waters II and Tim Raimond are appreciated.

Foreign summaries

Dans un certain nombre de problèmes de transport, on rencontre des variables caractérisées par le moment de leur occurrence dans un intervalle de temps. Ce sera par exemple le temps qui s'écoule avant que quelqu'un n'entreprenne un déplacement, le temps qu'un usager laissera passer avant de se mettre en route vers son travail ou son domicile de façon à éviter les pointes de trafic, ou encore la durée qui s'écoule avant qu'un mode de transport nouvellement mis en service soit effectivement utilisé. Les modèles à durée stochastique sont déjà utilisés largement dans d'autres domaines tels que l'économie et la statistique biologique; leur application aux phénomènes évoqués ci-dessus paraît aller de soi. Le but de l'article est de présenter ces modèles stochastiques, dans leurs caractéristiques générales, aux utilisateurs potentiels dans le domaine du transport. A cette fin, les auteurs se sont efforcés d'éviter tout le jargon technique qui caractérise trop d'articles et de textes courant sur ce sujet. Ils en attendent que cette présentation, couplée à l'évocation d'exemples de l'intérêt de leur application à des problèmes de transport, contribue à un recours plus fréquent à ces modèles stochastiques pour les problèmes de transport.

Eine Reihe von verkehrsbezogenen Phänomenen hat mit einem Zeitelement zu tun, das durch die Dauer bis zum Eintritt eines Ereignisses definiert ist. Die Beispiele schließen die Zeit ein, welche vergeht, bevor eine Fahrt unternommen wird, die Zeitdauer, um die der Pendler seinen Fahrtantritt verschiebt, um einem Verkehrsstau zu entgehen sowie auch diejenige Zeit, bis ein neu eingeführtes Verkehrsmittel genutzt wird. Ereignisorientierte Zeitmodelle die sich einer breiten Anwendung in nicht verkehrlichen Bereichen erfreuen (z.B. Ökonomie, Bio-Statistik), stellen eine einleuchtende Wahl zur Modellierung solcher Verkehrsphänomene dar. Ziel dieses Artikels ist es, ereignisorientierte Modelle allgemein verständlich für Personen darzustellen, die an Verkehrsproblemen interessiert sind. Deshalb wird jede Anstrengung

unternommen eine überzogene Fachsprache zu vermeiden, die aktuelle Artikel und Texte über diesen Gegenstand auszeichnen. Wir erhoffen uns von einer solchen Darstellung, zusammen mit einer Übersicht über bereits existierende Anwendungen solcher Modelle, eine zunehmende Nutzung von ereignisorientierten Zeitmodellen im Verkehr.

Varios fenómenos asociados al transporte deben tratar con el tiempo que transcurre hasta la ocurrencia de un evento. Ejemplos incluyen el tiempo transcurrido hasta que se realiza un viaje, la longitud de tiempo que un oficinista demora su partida del hogar para evitar la congestión de la mañana, y el tiempo que transcurre hasta que un medio recién introducido se comienza a utilizar. Los modelos de duración basados en obstáculos, que han gozado de amplio uso en varios campos distintos al transporte (ej. economía, bio-estadística) constituyen una obvia elección para modelar este tipo de fenómenos. El objetivo de este trabajo es presentar en general estos modelos a individuos interesados en problemas de transporte. Con este fin se han hecho ingentes esfuerzos para evitar el enfoque (casi coloquial) para iniciados que caracteriza los artículos y libros sobre este tema. Se espera que esta presentación, junto a una revisión de las aplicaciones existentes en transporte, lleve a una mayor utilización de este tipo de modelos en el área.

References

- ABBOT, R., 1985, Logistic regression in survival analysis. *American Journal of Epidemiology*, **121** (3), 465–471.
- BRESLOW, N. E., 1974, Covariance analysis of censored survival data. *Biometrics*, **30**, 89–99.
- CHANG, H.-L., and JOVANIS, P., 1990, Formulating accident occurrence as a survival process. *Accident Analysis and Prevention*, **22** (5), 407–419.
- COX, D. R., 1972, Regression models and life tables. *Journal of the Royal Statistical Society*, **B34**, 187–220.
- DIAMOND, P., and HAUSMAN, J., 1984, The retirement and unemployment behavior of older men. In: *Retirement and Economic Behavior*, edited by H. Aaron and G. Burtless (Washington, DC: Brookings Institute).
- EFRON, B., 1977, Efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, **72**, 557–565.
- ELANDT-JOHNSON, R., and JOHNSON, N. L., 1982, *Survival Models and Data Analysis* (New York: Wiley).
- ELBERS, C., and RIDDER, G., 1982, True and spurious duration dependence: the identifiability of the proportional hazard model. *Review of Economic Studies*, **44**, 403–409.
- FLEMING, T., and HARRINGTON, D., 1990, *Counting Processes and Survival Analysis* (New York: Wiley).
- GILBERT, C., 1992, A duration model of automobile ownership. *Transportation Research*, **26B** (2), 97–114.
- GREEN, M., and SYMONS, M., 1983, A comparison of the logistic risk function and the proportional hazards model in prospective epidemiologic studies. *Journal of Chronic Diseases*, **36** (10), 715–724.
- GREENE, W., 1992, *LIMDEP 6.0* (New York, NY and Sydney, NSW: Economic Software).
- GOURIEROUX, C., MONFORT, A., and TROGNON, A., 1984, Pseudo maximum likelihood methods: theory. *Econometrica*, **52** (3), 681–700.
- GUPTA, S., 1991, Stochastic models of interpurchase time with time-dependent covariates. *Journal of Marketing Research*, **18**, 1–15.
- HAMED, M., KIM, S.-G., and MANNERING, F., 1992, A note on travelers' home-stay duration and the efficiency of proportional hazards models. Working paper, University of Washington, Seattle.
- HAMED, M., and MANNERING, F., 1993, Modeling travelers' post-work activity involvement: toward a new methodology. *Transportation Science*, to be published.
- HAN, A., and HAUSMAN, J., 1990, Flexible parametric estimation of duration and competing risk models. *Journal of Applied Econometrics*, **1** (1), 36–64.
- HECKMAN, J., and BORJAS, G., 1980, Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Econometrica*, **47**, 247–283.

- HECKMAN, J., and SINGER, B., 1984, A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, **52** (2), 271–320.
- HENSHER, D., 1992, Semi-parametric duration models of automobile ownership in multi-vehicle households. Working paper, Institute of Transport Studies, University of Sydney.
- HENSHER, D., and RAIMOND, T., 1992, The timing of change: discrete and continuous time panels in transportation. Paper prepared for the First US Conference on Panels for Transportation planning, Lake Arrowhead, California.
- HUI, W. T., 1990, Proportional hazard Weibull mixtures. Working paper, Department of Economics, Australian National University, Canberra.
- INGRAM, D., and KLEINMAN, J., 1989, Empirical comparisons of proportional hazards and logistic regression models. *Statistics in Medicine*, **8**, 525–538.
- JONES, B., JANSSEN, L., and MANNERING, F., 1991, Analysis of the frequency and duration of freeway accidents in Seattle. *Accident Analysis and Prevention*, **23** (2), 239–255.
- JOVANIS, P., and CHANG, H. L., 1989, Disaggregate model of highway accident occurrence using survival theory. *Accident Analysis and Prevention*, **21**, 445–458.
- KALBFLEISCH, J., and PRENTICE, R., 1980, *The Statistical Analysis of Failure Time Data* (New York: Wiley).
- KATZ, L., 1986, Layoffs, recall and the duration of unemployment. Working paper No. 1825, National Bureau of Economic Research, Cambridge, Massachusetts.
- KIEFER, N., 1988, Economic duration data and hazard functions. *Journal of Economic Literature*, **26**, 646–679.
- KIM, S.-G., and MANNERING, F., 1992, Panel data and activity duration models: Econometric alternatives and applications. Paper prepared for the First US Conference on Panels for Transportation Planning, Lake Arrowhead, California.
- LIN, T.-D., JOVANIS, P., and YANG, C.-Z., 1992, Modeling the effect of driver service hours on motor carrier accident risk using time dependent logistic regression. Working paper, University of California, Davis, California.
- LANCASTER, T., 1979, Econometric methods for the duration of unemployment. *Econometrica*, **47** (4), 939–956.
- LANCASTER, T., 1985, Generalized residuals and heterogeneous duration models with application to the Weibull model. *Journal of Econometrics*, **28**, 155–169.
- MANN, N., SCHAFER, R., and SINGPURWALLA, N., 1974, *Methods for Statistical Analysis of Reliability and Life Data* (New York: Wiley).
- MANNERING, F., 1991, Duration models and highway accident risk: a study of alternate parametric forms. *Proceedings of the Twenty-Second Annual Modeling and Simulation Conference* (Pittsburgh, Penn: University of Pittsburgh Press).
- MANNERING, F., 1993, Male/female driver characteristics and accident risk: some new evidence. *Accident Analysis and Prevention*, **25** (1), 77–84.
- MANNERING, F., and HAMED, M., 1990, Occurrence, frequency, and duration of commuters' work-to-home departure delay. *Transportation Research*, **24B** (2), 99–109.
- MANNERING, F., MURAKAMI, E., and KIM, S. G., 1992, Models of travelers' activity choice and home-stay duration: analysis of functional form and temporal stability. Working Paper, University of Washington, Seattle.
- MANNERING, F., and WINSTON, C., 1991, Brand loyalty and the decline of American automobile firms. *Brookings Papers on Economic Activity, Microeconomics*, 67–114.
- OAKS, D., 1977, The asymptotic information in censored survival data. *Biometrika*, **64**, 441–448.
- PASELK, T., and MANNERING, F., 1993, Use of duration models for predicting vehicular delay at US/Canadian border crossings. paper prepared for presentation at the 1993 Annual Meeting of the Transportation Research Board, Washington, DC.
- PETERSON, A., 1976, Bounds for a joint distribution function with fixed sub-distribution functions: applications to competing risks. *Proceedings of the National Academy of Sciences U.S.*, **73**, 11–13.
- VILCASSIM, N. J., and JAIN, D. C., 1991, Modeling purchase-timing and brand-switching behavior incorporating explanatory variables and unobserved heterogeneity. *Journal of Marketing Research*, **18**, 29–41.
- YANG, C.-Z., JOVANIS, P., and LIN, T.-D., 1992, Assessing motor carrier driving risk using Cox's semi-parametric model with multiple stop effects. Working paper, University of California, Davis, California.

Editorial suggestions for further reading

CHANG, H.-L., and JOVANIS, P. P., 1990, Formulating accident occurrence as a survival process. *Accident Analysis and Prevention*, **22** (5), pp. 407–419.

A conceptual framework for accident occurrence is developed based on the principle of the driver as an information processor. The framework underlies the development of a modelling approach that is consistent with the definition of exposure to risk as a repeated trial. Survival theory is proposed as a statistical technique that is consistent with the conceptual structure and allows the exploration of a wide range of factors that contribute to highway operating risk. This survival model of accident occurrence is developed at a disaggregate level, allowing safety researchers to broaden the scope of studies which may be limited by the use of traditional aggregate approaches. An application of the approach to motor carrier safety is discussed as are potential applications to a variety of transportation industries. Lastly, a typology of highway safety research methodologies is developed to compare the properties of four safety methodologies: laboratory experiments, on-the-road studies, multidisciplinary accident investigations, and correlational studies. The survival theory formulation has a mathematical structure that is compatible with each safety methodology, so it may facilitate the integration of findings across methodologies.

(Authors)

JOVANIS, P. P., and CHANG, H.-L., 1989, Disaggregate model of highway accident occurrence using survival theory. *Accident Analysis and Prevention*, **21** (5), pp. 445–458.

The analysis of discrete accident data and aggregate exposure data frequently necessitates compromises that can obscure the relationship between accident occurrence and potential causal risk components. One way to overcome these difficulties is to develop a model of accident occurrence that includes accident and exposure data at a mathematically consistent disaggregate level. This paper describes the conceptual and mathematical development of such a model using principals of survival theory. The model predicts the probability of being involved in an accident at time t given that a vehicle has survived until that time. Several alternative functional forms are discussed including additive, proportional hazards and accelerated failure time models. Model estimation is discussed for the case in which both accident and non-accident trips are included and for the case with only accident data. As formulated, the model has the distinct advantage of being able to consider accident and exposure data at a disaggregate level in an entirely consistent analytic framework. A conditional accident analysis is undertaken using truck accident data obtained from a major national carrier in the United States. Model results are interpretable and generally reasonable. Of particular interest is the fact that segmenting accidents into several categories yields very different sets of significant parameters. Driver service hours seemed most strongly to affect accident risk: regularly scheduled drivers who take frequent trips are likely to have a reduced risk of an accident, particularly if they have a longer (greater than eight) number of hours off-duty just prior to a trip.

(Authors)

TRAN

As
integr
made
Exper

(a)

(b)

In
owner:
enable
born d
previo
inform
analys
1987).

The
used fo

† IN
the Dep