View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Sydney eScholarship

Institute of Transport Studies
Graduate School of Business
The University of Sydney
NSW 2006

Working Paper
ITS-WP-95-2

The Duration Between Traffic Accidents in the Taxi Sector: An Empirical Inquiry

Mohammad M. Hamed*
David A. Hensher
Hashem R. Al-Masaeid*

* Department of Civil Engineering,
Jordan University of Science and Technology Irbid P.O.Box 3030, Jordan.

February 10, 1995

NUMBER: Working Paper ITS-WP-95-2

TITLE: The Duration Between Traffic Accidents in the Taxi Sector: An Empirical Inquiry

ABSTRACT The taxi, as a supplier of urban passenger transport, is often sidelined in travel demand studies, yet its role in the overall transport task is far from marginal. Taxis are the most intensively used automobile transportation and consequently have a very high exposure rate in respect of potential accident. Very little empirical research has been undertaken into the risks associated with using a taxi. Drivers of taxis are usually assumed to be experienced drivers with a safety history appropriate for the responsibility of transporting the public. Yet this is not always the case. To provide some understanding of the exposure to risk which passengers place themselves in when hiring a taxi, we explore the safety record of drivers over time. Of particular interest is the frequency of accidents over a given time period, and in particular the elapsed time between a taxi drivers initial accident and subsequent accidents. To what extent is the duration between accidents a positive or negative function of experience, personal attributes, temporary impairment, fatigue etc? A set of proportional hazards models are developed to explain the time interval (duration) between traffic accidents of taxi drivers in Amman, Jordan. Three econometric models are developed to predict the duration between the date when a taxi driver begins to drive and the date when the first, second, and third accidents occur respectively. Estimation results show that driving-related capabilities, socioeconomic characteristics, and temporary impairments significantly influence the duration between traffic accidents. The duration to the first traffic accident is lower than the duration between the first and second traffic accidents or the second and third traffic accidents.

KEYWORDS: Taxis, duration modelling, traffic accidents

AUTHORS: Mohammad M. Hamed*
David A. Hensher
Hashem R. Al-Masaeid*
* Department of Civil Engineering,

Jordan University of Science and Technology Irbid P.O.Box 3030, Jordan.

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

Telephone: + 61 2 550 8631
Facsimile: + 61 2 550 4013
E-mail: davidh@gsb.su.oz.au

DATE: February 10, 1995

1.0 INTRODUCTION

The taxi, as a supplier of urban passenger transport, is often sidelined in travel demand studies, yet its role in the overall transport task is far from marginal. Taxis are the most intensively used automobile transportation and consequently have a very high exposure rate in respect of potential accident. Very little empirical research has been undertaken into the risks associated with using a taxi. Drivers of taxis are usually assumed to be experienced drivers with a safety history appropriate for the responsibility of transporting the public. Yet this is not always the case. To provide some understanding of the exposure to risk which passengers place themselves in when hiring a taxi, we explore the safety record of drivers over time. Of interest is the frequency of accidents over a given time period, and in particular the elapsed time between a taxi drivers initial accident and subsequent accidents. To what extent is the duration between accidents a positive or negative function of experience, personal attributes, temporary impairment, fatigue etc?

An accident is said to occur when one or more of the traffic system elements fails. This system can be efficiently simplified by a road-vehicle-driver system. The driver is the only instantaneous decision making component in this system. Therefore, the driver's role is to process information from the road, the traffic, and the vehicle, and to make the appropriate decision(s) or action(s), and observe and respond to the new situation that results.

Drivers have to process huge amounts of information and cope with random and complex variations of the traffic state. Sensorial and psychological functions are called on during driving, to varying degrees of intensities (Chich 1985). Cognitive impairments negatively influence the processing of information. As such, cognitively impaired drivers are expected to have high risk of being involved in a traffic accident when compared with unimpaired drivers (Cooper et al. 1993).

The probability of an accident is dependent on several individual driver attributes. These attributes consist of a driver's socioeconomic characteristics, temporary impairments, and overall capabilities. Driver characteristics include accident proneness, social maladjustments, and personal maladjustments. Temporary impairments are fatigue, high blood alcohol level, and the effect of drugs. Driving-related capabilities, represented by measures of experience, recognise the role of the learning process which leads to the increase of driving skills. Accumulated experience is expected to make a driver aware of several automation procedures. This in turn, will give the driver a higher ability to gather information about the position of the vehicle relative to the road direction and other vehicles' position, and to use the information to safely guide the vehicle within the lane through the traffic.

Taxi drivers constitute a significant proportion of the daily driving population in many cities. Although there is an increase in the number of studies of the rate of occurrence of traffic accidents made over a specified period of time (e.g. Saccomanno and Buyco 1988, Jovanis and Deller 1983, Jovanis and Chang 1987, Khasnabis and Reddy 1983), there has been no investigation of the taxi market per se. Jones et al. (1991) have considered the frequency and duration of traffic accident, defining duration as the length of time between the time the police receive a report of an accident until they leave the accident scene. This specification of duration is important for planning to minimise delays in processing an accident (and in saving lives), however of equal importance is a better appreciation of the accident profile of major carriers of people - taxi drivers. Frequency models do not provide information on the duration between traffic accidents (i.e. time between accident occurrences).

By addressing the factors that are likely to affect the duration between traffic accidents, decision-makers can design improved safety oriented measures geared toward increasing this duration. An understanding of possible factors contributing to higher likelihood of accidents in the taxi driver set may provide support for reform in taxi safety regulations. If the number of working hours, number of rest hours, number of calls attempted during a specified period of time, vehicle condition (e.g regularity of vehicle maintenance), and the driver's socioeconomic characteristics are shown to have statistical links with undesirable and/or desirable accident histories, then the industry regulators have useful information to assist them in regulatory reform.

The objective of this paper is to develop a set of disaggregate models to predict the duration between traffic accidents. The first duration is defined as the duration between the date when a taxi driver began driving and the date when the first accident occurred. The second duration refers to the time until the occurrence of the second traffic accident (see Figure 1).

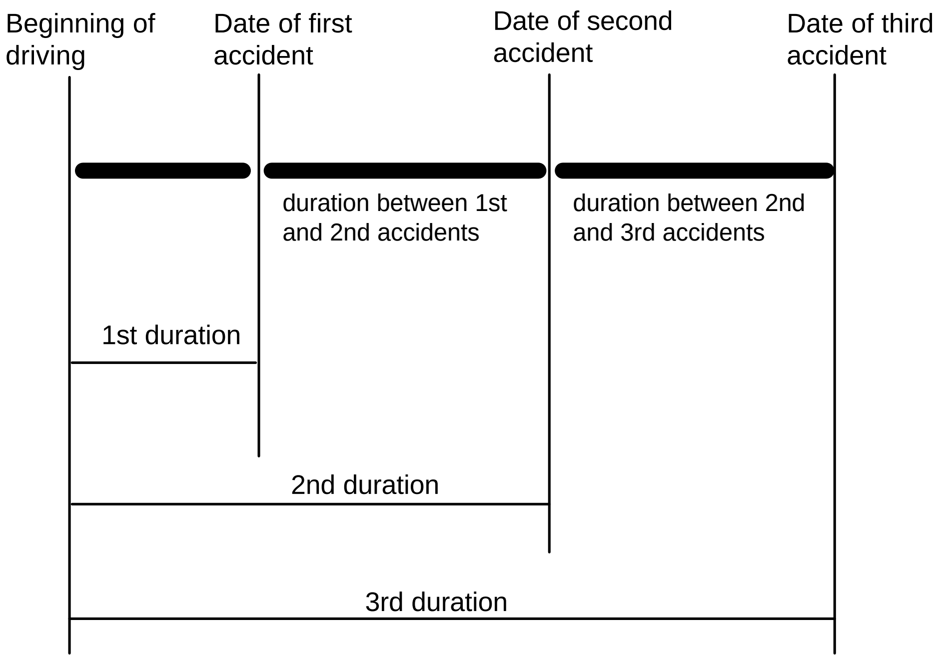

Figure 1. Duration between traffic accidents

2.0 MODEL DEVELOPMENT

The duration between the initial state and the first accident, and then between each subsequent pair of accidents is modelled as a set of independent events. A single state specification of a duration model defines the two states - no accident and accident. All durations are specified as durations of no traffic accidents. We are interested in determining the probability of a taxi driver ending the duration of no accident conditioned on zero traffic accidents over a predetermined elapsed time. This probability is likely to vary as the length or duration of no accidents increases, perhaps due to increased driver's experience, better vehicle or road maintenance, or reduced number of working hours.

In the first duration, if the driver has not been involved in any traffic accident (up to the time when the empirical evidence was observed), the observation is a censored observation; if an accident has occurred the observation is complete. However, in the second duration two states arise, in the first state the driver has not been involved in any accident (no accident state); in the second state the driver has been involved previously in one traffic accident only. The third duration represents the duration to the third traffic accident from the date of obtaining a driving license. Again two states arise. The first state deals with drivers who have never been involved in a traffic accident.

The second state deals with those drivers who have had either one or two accidents previously.

Accident histories can be characterised as a time-sequenced set of events. For each unit of analysis, accident histories provide information about the exact duration until a state transition as well as the occurrence and sequence of accidents. Accident information can be identified for each taxi driver in a given period of time. A duration model in its statistical form is referred to as a hazard function. Formally, 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,

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

where Prob denotes the probability, T is a random continuous time variable, and t is some specified time. Equation (1) for example, identifies the probability of a taxi accident before some transpired time (assuming no left censoring). The corresponding density function 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 accident will occur between time t and t+dt given that an accident has not occurred up to time t:

$$Prob\ (T_o \geq t+1 | T_o \geq t) \quad (4)$$

Information relating to duration dependence, as derived from the first derivative of the hazard function with respect to time (i.e. its slope) provides insights into the duration process being modelled. Plotting the hazard function against time gives important empirical information for the parameterisation of the baseline hazard (Hensher and Mannering 1994). The probability of ending a duration or spell in a particular state may be dependent on the length of the duration. There may also be important determinants of duration (e.g. work practices) that should be included in the modelling approach. These covariates are included in hazard-based models using *proportional hazards* (Cox 1972).

Proportional hazards models operate on the assumption that covariates act multiplicatively on some underlying or baseline 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 (Prentice and Gloeckler 1978). To accommodate time varying covariates we assume that they are well approximated by their mean over the interval. This gives a clue to the interval size given the particular application (Hensher and Raimond 1992). A relatively general form of the hazard is specified as:

$$h_o(t) = \lambda_b(t) \exp (z_o(t)\beta) \quad (5)$$

where $\lambda_b(t)$ is an arbitrary baseline hazard and $\exp (z_o(t)\beta)$ is the parametric component including time varying covariates associated with an origin state *o*. A discrete set of time intervals are observed. The conditional probability rule in (4) translates into the following function, given (5):

$$\text{Prob } (T_o \geq t+1 | T_o \geq t) = \exp (-\exp (\gamma (t) + z_o(t) \beta)) \quad (6)$$

where $\gamma (t) = 1n (\int_t^{t+1} \lambda_b(u)du)$. u is any function in terms of time. The model allows for a continuous 'failure' time T_o and (right) censoring c_o, but with observation taking place only at t_o, t = 0,1,2,..., *J*-1, or in the final interval (*J*, ∞). If the baseline hazard is assumed to be well approximated by its mean over the time interval, it is completely captured by the single term $\gamma (t)$. Left censoring may exist if an event was well under way when the observation period commenced. Right censoring exists since the endpoint of the last episode of an individual cannot be observed.

The specification of the hazard function addresses explicitly the conditional probabilities of ending the duration of no traffic accident. Thus the hazard rate is the rate with which the durations of no accidents are ending at time t, conditioned on the taxi driver having survived (not involved in a traffic accident) until time t. This approach has been adopted by Jovanis and Chang (1989) and Jones et al. (1991) to study general traffic accidents. Hensher and Mannering (1994) review the limited but growing transport literature on duration modelling.

The effect of the covariates is to multiply the base-line hazard function by a scale factor, which does not depend on the duration of no accident. The sign of the parameter estimates indicates the direction of the effect of the exogenous variable(s) on the conditional probability. For instance, if *number of driving hours* has a positive coefficient, then this variable is expected to increase the risk of traffic accident occurrence or decrease the duration of no accident.

The Cox model specification implies strong interdependence among the risk components and assumes that the underlying distribution for duration between traffic accidents is unknown (Jovanis and Chang (1989)). This latter advantage of the Cox model is important since little if anything may be known about the underlaying distribution of duration between traffic accidents (Cox and Oakes 1988). With the above specification, the survival function can be written as

$$S(t) = e^{(-L(t) \exp(\gamma z_0))} \quad (7)$$

where L(t) is the cumulative hazard function.

Another appealing advantage with the above approach is that it accounts explicitly for censoring. Cox (1972) suggested a partial-likelihood approach to estimate the vector γ without specifying the form of $h_0(t)$. Let the observed durations be ordered as below, $d_1 < d_2 < d_3 ... < d_k$. γ_i represents the covariate associated with d_i. Then the probability that a taxi driver i ending the duration of no traffic accident at time d_i, given that the driver ended duration at time d_i, is given as,

$$P_i = e^{(\gamma z_{0i})} \Big/ \sum_{r \in C_i} e^{(\gamma z_{0i})} \quad (8)$$

The partial log-likelihood function can be written as:

$$L(\gamma) = \sum_{i=1}^{K} (\gamma z_{0i}) - \mathrm{Log}\Big(\sum_{j \in C_i} e^{(\gamma z_{0i})}\Big) \quad (9)$$

3.0 DATE DESCRIPTION AND ANALYSIS

A taxi driver survey was conducted in July 1993 in the City of Amman, Jordan's capital. A sample of 370 taxi drivers were interviewed and details of their accident history. All reported accidents were checked through police records to secure cross-checking data on driving license type, number of traffic accidents since holding the driving license along with their date of occurrence.

The survey collected a wide range of information relating to the driver's socioeconomic characteristics including marital status, number of children in the household, number of employed persons, total household monthly income, number of household members along with their age and sex, and household location. Information was also collected on the vehicle use and condition as well as the working conditions and schedules for drivers including the number of driving hours per day, number of calls attempted per day, number of rest hours (off-duty) per day, vehicle regular maintenance, years of experience as a taxi driver, and average number of miles travelled per day.

273 of the surveyed taxi cab drivers have been involved in at least one traffic accident. 187 and 60 of those drivers have been involved in two and three traffic accidents respectively. The data also showed that the average duration to the occurrence of the first, second, and third traffic accidents were 2.04, 4.46, and 10.45 years, respectively. Figure 2 shows the observed distribution of taxi traffic accident durations for each of the three accidents and the duration between each adjacent accident.

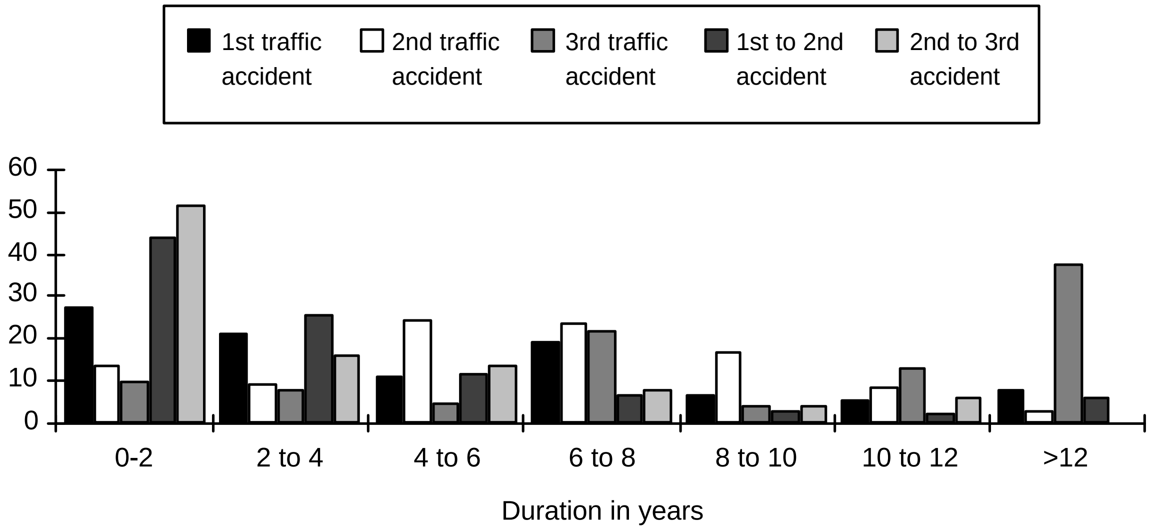

Figure 2. The Profile of Durations of Each Sequenced Traffic Accident

4.0 DISCUSSION OF EMPIRICAL RESULTS

The estimation results for the set of exogenous influences on the duration between traffic accidents are summarised in Table 1. Seven variables are included in the exogenous set, although not all are statistically significant across all three duration models. Estimation results suggest that the duration to the first traffic accident is shorter than the duration between the first and second or the second and third accidents. This finding is intuitively plausible, since the experience of taxi drivers usually increases over time. As elapsed time increases, a driver becomes more familiar with the transportation network as well as their vehicle.

Table 1. Duration Models for Traffic Accidents (t-statistics in brackets)
Dependent Variable: duration before a traffic accident

Exogenous Effects	1st traffic acc.	2nd traffic acc.	3rd traffic acc.
Age of driver (=1 if age < 30, 0 otherwise)	0.385 (1.94)	1.094 (3.56)	0.044 (0.17)
Driving experience (years)	-.031 (2.82)	-0.047 (2.24)	-0.032 (2.46)
Natural log of no. of driving hrs in day of acc.	0.456 (2.39)	0.510 (2.04)	0.985 (2.31)
Marital status (=1 if married, 0 otherwise)	-.602 (2.70)	-0.209 (1.58)	-0.357 (1.01)
Regular veh. maintenance dummy (1,0)	-.374 (0.62)	0.673 (1.58)	-0.138 (1.21)
No. of taxi calls in day of accident	0.005 (0.06)	0.007 (0.26)	0.017 (2.13)
No. of rest hours in day of accident	-.215 (1.87)	-0.274 (1.45)	-0.019 (1.27)
Log-likelihood at convergence	-723.16	-298.22	-371.49
Sample size	273	187	60
Log-rank test with 7 degrees of freedom	23.26	19.12	29.54

The first and second models show that younger (less than 30 years old) taxi drivers have a higher risk of being involved in a traffic accident when compared to older drivers. The age effect is not statistically significant for the third accident. The duration between traffic accidents for this category of drivers is shorter than that for older and more experienced drivers, supporting similar results by Jones et al. (1991). This is an expected finding, since the driving habits of younger taxi drivers are sometimes associated with risk-taking. Young drivers also lack the experience which usually provides the driving skills needed to avoid a traffic accident.

Marital status shows a considerable effect on the accident probability of a first accident, but is marginally significant for the second accident and not statistically influential at all for the third accident. Married drivers have lower risk of being involved in an initial traffic accident when compared with single drivers, a result *explained* by the greater responsibilities toward family and home. The driving habits of married drivers are consequently likely to be associated with less risk-taking.

Driving-related capabilities represented by experience indicate for all models that taxi drivers with more years of driving experience have a lower probability of ending their duration of no accident. Taxi drivers with vast experience not only have improved driving skills, but have also acquired more information about the transport network. As such, they are more likely to avoid highly congested roadways, avoid streets with bad driving surfaces, and have greater information about the temporal and spatial distribution of the potential sources of patronage. All these attributes combine to lower the rate of working hours which in itself reduces the exposure to risk and hence accident.

The driver's temporary impairments are represented by the number of driving hours, the number of rest hours and the daily number of calls responded to. The number of rest hours is statistically significant in the first accident only; as the number of rest hours per day increases the risk of being involved in a traffic accident decreases. The daily number of calls responded to is statistically significant in influencing the third accident, increasing the risk of traffic accident. The number of driving hours on the day of an accident is consistently significant across all models, suggesting that as the number of driving hours increases, ceteris paribus, the higher the probability of being involved in a traffic accident - the classic exposure observation.

5.0 CONCLUSIONS AND POLICY IMPLICATIONS

The study of the behaviour of taxi drivers in terms of exposure to risk and consequent accidents is a neglected but important topic, given their role in transporting passengers annually over substantial passenger kilometres. This study is an initial contribution. A knowledge of the temporal profile of initial and subsequent accidents and identification of the exogenous influences on variations in the duration to each accident (or no accident history) is useful in a review of safety regulations in the taxi industry.

The results show that the duration to the occurrence of the first traffic accident is lower than the duration between the first and second traffic accident. Furthermore, the duration between the first and second traffic accident is lower than the duration between the second and third traffic accident. Increasing exposure in the driving environment continues to exert a strong positive influence on the risk of being involved in an initial, second or third accident. After controlling for exposure, a number of other effects contribute, driving experience being the most important.

A number of taxi traffic accident reduction measures can be suggested from this preliminary empirical inquiry. Younger drivers might be required to attend mandatory education programs designed to improve their driving skills, promote safe driving, and develop safe and efficient techniques to respond to various critical conditions. This training is additional to that normally required to secure a taxi driving licence. To reduce driver fatigue, the profile of taxi driver daily working and rest hours should be studied in more detail.

A number of methodological extensions should be considered in further research. The specification of the hazard functions associated with each accident state in the current study assumed independence between the duration to an initial, a second and a third accident. Allowance for interdependencies using a multi-state or competing risks approach (Hensher 1994) is useful in establishing the likely influence of unobserved heterogeneity in the probability of a traffic accident which is correlated across the sequence of accidents (i.e. the full accident event history).

REFERENCES

Chich, Y. (1985). "Behavior of the driver," Paper presented at the Fifteenth International Traffic Engineering and Safety Study Week, Venice , Italy.

Cooper, P.J. (1990). "Differences in accident characteristic among elderly drivers and between elderly and middle aged drivers. *Accident Analysis and Prevention*, 22, pp.499-508.

Cooper, P.J. et al. (1993). "Vehicle crash involvement and Cognitive deficit in older drivers," *Journal of Safety Research*. 24 (1), pp.9-17.

Cox, D. (1972). "Regression models and life-tables," *Journal of Statistical Society*. 2, pp. 187-220.

Cox, D. and Oakes, D. (1988). *Analysis of survival data*. New York: Chapman and Hall.

Hensher, D.A. (1994) "The Timing of Change for Automobile Transactions: A Competing Risk Multispell Specification" presented at the *7th International Conference on Travel Behaviour,* Santiago, Chile, June 13-16.

Hensher, D.A. and Raimond, T. (1992) 'The Timing of Change: Discrete and Continuous Time Panels in Transportation", Presented at the *First U.S.A. Invitational Conference on Panels in Transport Planning,* UCLA Conference Center, Lake Arrowhead, California, October 25-27 (to appear in the Proceedings by Kluwer Publishers, Amsterdam).

Hensher, D.A. and Mannering, F.L. (1994) "Hazard-Based Duration Models and Their Application to Transportation Analysis" *Transport Reviews*, 14 (1), 63-82.

Jovanis, P. P. and Chang H. (1987). "Modelling the relationship of accidents to miles travelled," *Transportation Research Record*. 1068, pp. 42-51.

Jovanis P. P. and Chang H. (1989). "Disaggregate model of highway accident occurrence using survival theory," *Accident Analysis and Prevention*, 21 (5), pp. 445-458.

Jovanis, P. P. and Deller, J. (1983). "Exposure-based analysis of motor vehicle accidents," *Transportation Research Record*, 910, pp. 1-7.

Jones, B., Janssen, L., and Mannering, F. (1991). "Analysis of the frequency and duration of freeway accidents in Seattle," *Accident Analysis and Prevention*. 23 (4), pp. 239-255.

Khasnabis, S. and Reddy, T. (1983). "Systematic procedure for incorporating exposure factors in truck accident analysis," *Transportation Research Record*, 910, pp. 16-43.

Prentice, R.L. and Gloeckler, L.A. (1978)Regression analysis of grouped survival data with applications to breast cancer data, *Biometrics*, 34, 57-67.

Saccomanno, F. and Buyco, C. (1988). "Generalised loglinear models of truck accident rates," *Transportation Research Record*, 1172, pp. 23-31.