INVESTIGATION OF ROUTE CHOICE IN RESPONSE TO VARIABLE MESSAGE SIGNS

Peter HIDAS
Senior Lecturer
School of Civil and Environmental Engineering
University of New South Wales,
Sydney, 2052, Australia
Fax: +612-9385-6139
Email: P.Hidas@unsw.edu.au

Emad AWADALLA
Research Student
School of Civil and Environmental Engineering
University of New South Wales,
Sydney, 2052, Australia
Fax: +612-9385-6139
Email: emad@civeng.unsw.edu.au

Abstract: One of the most promising uses of Variable Message Signs (VMSs) is in the area of incident management. If drivers' response, in terms of flow diversion, to various VMS messages can be predicted, it may be possible to influence drivers' route choice behaviour by the use of appropriate messages in order to achieve an 'optimum diversion rate'. The aims of this study are to evaluate drivers' response to VMS and to develop a prediction model of the diversion rate as a result of the provided information. A personal interview survey was conducted in the Sydney Metropolitan Region. Four scenarios of VMS messages were displayed to each respondent, including combinations of different causes of delay and different lengths of expected delay. This paper outlines the survey design, the model development process and presents some preliminary results of the prediction model developed from the analyses of the first 200 interviews.

Key Words: route choice, Variable Message Signs, incident management, stated preference survey

1. INTRODUCTION

Traffic congestion in urban areas constitutes a major problem in large cities due to the increase in population, commerce, and development activities. Increase in travel delay, fuel consumption, air pollution, and vehicle operating costs attributed to congestion imposes huge costs on society. For example, the estimated congestion cost in Sydney is at least $2 billion per year (Commeignes, 1992). Several international studies (eg. Grenzeback and Woodle 1992) found that as much as 50 percent of this congestion results from traffic incidents, ie. events that create an unexpected and temporary reduction of road capacity, such as a stalled vehicle, a crash, inclement weather or even a planned roadwork. The traffic congestion that results from these incidents can lead to additional accidents, cause delayed response to emergency situations and reduces the quality of life in a community.

Recognition of the significance of traffic incidents has prompted road authorities to develop and implement Incident Management (IM) plans in many metropolitan areas around the world. The majority of these plans concentrate on early detection and clearance of the incident so as to minimise the effects of incidents. These plans also aim to coordinate the actions of the various institutions involved in the IM plan such as the traffic authorities, police, fire department, tow truck operators, etc. However as indicated by the literature the weakest link in the chain of IM plans is the traffic management during the incident (Judycki and Robinson, 1992).

One of the most promising traffic management measures during traffic incidents is the use of traveler information systems to inform drivers in real time about the traffic conditions, the presence of incidents, and the expected delay. It is believed that these systems help the drivers to avoid the congestion by diverting from the congested links in the network. One instrument of traveler information systems is Variable Message Signs (VMS), which can be installed beside or above the carriageway, and can be used for various purposes including safety warnings, capacity variation, parking guidance and information, and flow diversion.

Early field studies at VMS locations have found evidence of traffic diversion in the range of 5 to 80 % of the total driver population subjected to the message. This range is clearly too wide for prediction and modelling purposes. In the last decade several international research studies have been conducted to investigate the influence of various VMS messages on drivers' route choice behaviour and to establish quantitative models of diversion rates. While these studies provide a better understanding of the factors influencing route choice behaviour in response to real time traffic information, the proposed models have serious limitations in predicting diversion rates under different conditions and in other countries. There is a need to develop more general models for prediction purposes in Australia and to calibrate such models based on local data. No information of this kind is available in Australia.

The objectives of this study are firstly to collect information on drivers' route choice behavior in response to Variable Message Signs and to develop a route choice model for predicting the degree of diversion as a result of the provided information. A second objective is to incorporate the route choice model in an urban network simulation model to study the benefits of using VMS as an incident management tool.

This paper reports on the first part of the study. Section 2 describes the survey design and some other important characteristics of the survey. Section 3 presents the subsequent steps of the data analysis, including some basic characteristics, the model formulation, model development and testing of the model fit. The last section includes a summary of the findings and further work.

2. THE SURVEY

Different approaches can be used to understand driver behavior in the presence of information. These approaches include: field surveys, interactive route choice simulators, and stated preference surveys. Field surveys are expensive, difficult to organise and cause traffic disruption (Bonsall et. al. 1998). Route choice simulators are an ideal tool for understanding driver behavior in the presence information, but they are very expensive and not suited to every problem (Bonsall 1994a, 1994b and Bonsall et. al. 1997). The stated preference (SP) approach was used to assess the effects of information provided by VMSs on drivers route choice by Wardman et al. (1997), Chatterjee et al (1998), and Abdel-Aty et al. (1997). The advantage of the stated preference approach stems from its ability to control the choice context and the independent variables that will enter the demand model, and its ability to design the questions to relate to drivers' journey rather than to a hypothetical journey. Also, the stated preference survey is cheaper than field surveys and route choice simulators. For these reasons, the stated preference survey method was selected for this study.

2.1 SURVEY DESIGN

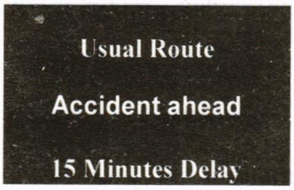
The data was collected using personal interviews in which an interviewer is present to record the responses provided by the respondent in answer to a series of questions posed by the interviewer.

The survey objective was to collect information on the following parameters:

- ***VMS content***: previous studies have found that the cause of an incident displayed on the VMS affects respondents' route choice behaviour. Three causes were used: accident, roadwork, and congestion.
- ***Extra Travel time***: the difference in travel time between the usual and alternative routes in the normal conditions also has an effect on the route diversion.
- ***Familiarity with the network***: to investigate the effect of drivers' familiarity with the network on the route diversion due to the provided information from VMS.
- ***Experience with VMS***: prior experience with VMS may affect driver response to VMS.
- ***Driver attributes***: the effects of drivers' personal characteristics, age and sex.
- ***Willingness to divert at later points:*** if a driver decided to stay on the usual route, his/her willingness to divert at the next potential diversion points.

The questionnaire was designed to investigate the impact of these parameters on route choice in the presence of VMS information. An important issue in the survey design was the selection of the trip to be considered in the interview. Most overseas studies used one or a few pre-selected trips in their surveys. While this type of survey may provide more accurate data pertaining to the conditions of the selected trip, it has several significant disadvantages. First, it makes it far more difficult to find respondents having experience with the pre-selected route. Second, and more importantly, the results are not well suited for the development of a general model. In particular, they provide very limited information on the relationship between the trip time and the 'Extra Travel Time", ETC (ie. the difference between the usual and alternative trip time), a crucial factor identified by most previous studies.

While you are on the way to your destination, imagine that before you approach the intersection at which you could divert to the other route, you see a VMS showing the message below:



Which route would you select?

Usual[1]

Alternative[2]

Figure 1 An example of the key questions used in the interview survey

For these reasons we have decided to use a more open-ended questionnaire survey method. Respondents were first asked to select a car trip in the Sydney Metropolitan Area which is longer than 15 minutes, and which may have at least one alternative route (beside the "usual route" used by the respondent). Details of these two routes were recorded in the first part of the questionnaire. Then, a number of hypothetical VMS messages were shown to each respondent, and they had to state which route they would select in the given situation. An example of these questions is shown in Figure 1.

This survey method was selected because it can provide information on a more continuous range of trip times and thus it enables us to develop an estimation model of the diversion rate which can generally be applied anywhere in Sydney at any location regardless of the trip origin, destination and trip length. From a survey organisational viewpoint, this method also made it easier to find the required number of respondents. A potential weakness of the method is that the variance of the answers collected from a wide range of trips included in the analysis may be larger than in the case of one specific trip between fixed origin and destination, and therefore the model developed from these results may be less accurate. Nevertheless, this data collection method is considered more appropriate for the intended purpose: the development of a general route choice model for incident management.

2.2 Survey Work

The required sample size was estimated as 400 questionnaires to satisfy a 90% confidence limit, which is considered acceptable in this type of survey. The survey started at the end of September 1999 and it was expected to finish by the end of the same year. However, only 200 questionnaires were completed by January 2000, and for various administrative reasons the survey could not proceed any further before July 2000. The second round of 200 questionnaires was obtained by mid-November 2000. A total of 22 interviewers participated in the two rounds of the survey.

The analysis presented in this paper has been carried out on the data collected from the first round of survey, 200 questionnaires. A repeat analysis of the full data set is currently in progress.

3. DATA ANALYSIS

3.1 General Characteristics

Table 1 shows the trip and respondent characteristics as revealed in the answers to the questionnaires. As shown in the table, most of the respondents are commuter drivers, 71%, and males, 68%. Also, the age of most of the respondents, 87%, ranges from 17 to 50 years. About half of the respondents drive on the usual route daily, while 28% use it few times a week.

Table 1 Trip and Respondent Characteristics

(a) Trip Purpose and Time Distribution

Trip Purpose	Observed	%	Trip Time (min.)	Observed	%
Commuting	142	71	10 - < 15	28	14
Shopping	15	7.5	15 - 20	31	15.5
Entertainment	18	9	20 - 25	45	22.5
Others	25	12.5	25 - 30	21	10.5
			30 - 35	27	13.5
			35 - 40	14	7
			40 - 45	14	7
			45 - 50	8	4
			50 - 55	5	2.5
			55 - 60	7	3.5

(b) Frequency of Travel on the Usual and Other Routes

Usual Route	Observed	%	Other Route	Observed	%
Every day	97	48.5			
Few times a week	56	28	Few times a week	23	48.5
Once a week	31	15.5	Few times a month	61	28
Less	16	8	Once a month	31	15.5
			Less	85	8

(c) Previous Experience with VMS

Previous Experience	Observed	%	Quality of VMS Information	Observed	%
A Lot	35	17.5	Very Satisfactory	14	17.5
Little	119	59.5	Satisfactory	86	59.5
Nil	46	23	Uncertain	42	23
			Very Unsatisfactory	12	0.0

(d) Age and Gender Distribution

Age	Observed	%	Gender	Observed	%
17-35	88	44	Male	138	50.57
35-50	86	43	Female	63	49.43
>50	26	13			

About 60% of the respondents had little previous experience with VMS, while a quarter of them had no previous experience at all. Fifty six percent of the respondents with previous experience assessed the quality of the previous information, given by VMS, as satisfactory while, a quarter of them were uncertain. The trip times recorded from the survey show a fairly even and continuous distribution: 90% of the trip times ranged from 10 to 45 minutes, while 10% of the trip times were between 45 and 60 minutes.

3.2 Analysis of Extra Travel Time, ETT, and trip travel time

The extra travel time (ETT) is the travel time difference between the alternative and the usual route under normal conditions. Previous studies have shown that the ETT has a strong effect on the route choice behaviour in the presence of the VMS information. However, most previous studies were based on one or two fixed trips, therefore the results were relevant to the given fixed trip travel time. We believe that there is a strong relationship between the trip travel time and the ETT, which affects the probability of diversion. Thus, a separate analysis was carried out to investigate the combined effect of the ETT and the travel time on the diversion rate in the presence of the provided information.

When analysing the full data set regardless of the VMS content (Accident, Roadwork or Congestion), the following general characteristics were found:

- The diversion rate decreases as the extra travel time, ETT, increases;
- The diversion rate decreases with the increase of travel time on the usual route for travel times up to 45 minute. The travel time has no noticeable effect on the diversion rate when it becomes more than 45 min.

A more detailed analysis was carried out on a subset of data collected for the displayed cause of Accident, and statistical models were built to represent the relationship between the extra travel time, the usual travel time and the diversion rate. Separate models were created for the five VMS message contents (Accident – Long Delay, Accident – Delay, Accident – 10 Min.

Delay, Accident – 20 Min. Delay, Accident – 30 Min. Delay). An illustration of the results for the 'Accident – 20 Min. Delay' scenario is shown in Figure 2.

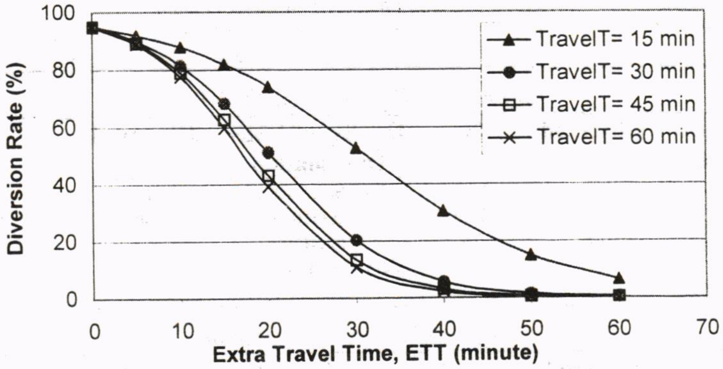


Figure 2 Effect of ETT on Diversion Rate (Accident delay 20 min)

In general, the diversion rate increases with the increase of the displayed delay time in each case. The diversion rate at ETT 20 minute for 10, 20 and 30 minute delay are about 15%, 50% and 90% for travel time of 30 minute. The diversion rate for the long delay is more than for delay. Also, the diversion rate for the long delay is approximately the same as the 20-minute delay.

Results of this analysis are used in the development of the prediction model.

3.3 Model Formulation

The model is required to predict the diversion rate, which is bounded between 0 and 1, based on the displayed VMS and the attributes of the drivers. The logit/logistic models are the suitable forms for the required model for the following reasons. The form of the binary logit models is (Cramer 1991):

$$\ln\left[\frac{P}{1-P}\right] = \beta_o + \Sigma \beta_i x_i \tag{1}$$

where

P	probability of event occurrence
1-P	probability of event non-occurrence
β_o	Constant
β_i	coefficients of the explanatory variables
X_i	the explanatory variables

The above equation of the logit models is suitable when the explanatory variables are demographic variables only. The demographic variables do not vary according to the response category (usual/alternative route) chosen by the individual; rather, they vary only across individuals. On the other hand, the conditional logistic models are used when the explanatory variables are choice specific variables. The choice specific variables take different values depending on the response category even for the same individual (Liao, 1994).

The conditional logit model estimates the effects of a set of choice specific independent explanatory variables, z_i, on a dependent variable with response categories. The form of the binary model according to the general model given by Liao (1994) is:

$$\ln\left[\frac{Prob\ (y=1)}{Prob\ (y=2)}\right]=\Sigma\alpha_i\left(z_{1i}-z_{2i}\right) \tag{2}$$

where

Prob (y=1) probability of the occurrence of event 1 (using alternative route)
Prob (y=2) 1-Prob(y=1) = probability of the occurrence of event 2 (using usual route)

When the explanatory variables include both choice specific and demographic variables, a mixed model will be utilized which combines the conditional logistic model and the logistic model. The form of the binary mixed model according to the general mixed model given by Liao, 1994 is:

$$\ln\left[\frac{Prob(y=1)}{Prob(y=2)}\right]=\Sigma\alpha_j\left(z_{1j}-z_{2j}\right)+\Sigma\beta_i x_i+\beta_o \tag{3}$$

where both choice-specific explanatory variables z_j, and demographic variables x_i are included in the same model and the subscripts i and j represent the two types of explanatory variables.

3.4 MODEL DEVELOPMENT

The binary mixed model presented in Eqn. (3) can be also written in the form:

$$P=\frac{1}{1+e^{U}} \tag{4}$$

where P is the probability of diverting to the alternative route, and U is the relative utility of the alternative route compared with the usual route. The factors expected to affect the relative utility (and thus, the diversion rate) in the presence of the VMS information are: VMS message content, Extra travel time, Familiarity with the network, Willingness to divert at later points, Trip purpose, Experience with VMS, and Driver attributes.

The VMS message content includes different combinations of the incident cause and its severity. The incident cause used in the model is one of: Accident (Acc), Roadwork (Rdwk) and congestion (Cong). The value of severity is one of: delay time in minutes (Minutes), Long delay (Long Delay), and Delay (Delay). Also, the travel time on both the usual and the alternative routes is entered the model as an independent variable.

The demographic variables considered in the study are: person's age, sex, familiarity with the road network, previous experience and assessment of the VMS information, willingness to divert at the later diversion points, and the trip purpose. The choice-specific variables are the travel time and the displayed delay with its cause on the VMS. Thus, the relative utility of the alternative route is calculated as a linear combination of the choice-specific and demographic variables:

$$U = f(\textit{Message Content, Usual and Extra Travel Time, driver characteristics}) \tag{5}$$

In the logit regression, the method used for estimating the coefficients of the independent variables is the maximum likelihood. This method yields coefficient values that maximize the probability of obtaining the observed set of data (Hosmer & Lemeshow 1989).

3.5 MODEL BUILDING

The goal of the model building is to develop the best and most reasonable prediction model fit to the data. The model building procedure consists of three stages: (1) selection of the variables and verification of the importance of each independent variable included in the model, (2) checking the linearity in the logit and suggesting modifications if required, and (3) checking if interaction terms need to be included among the independent variables in the model.

3.5.1 Variable Selection

The variable selection stage was executed using the stepwise method. A critical aspect of using the stepwise logistic regression is the choice of alpha level, the cut off values, for the statistical significance to judge the importance of variables. Alpha values are recommended to be set in the range 0.15 to 0.20 rather than .05 since the use of 0.05 often result in excluding important variables from the model (Hosmer & Lemeshow 1989).

The SPSS package was used to carry out the backward/forward stepwise logistic regression for the collected data. Most of the independent variables included in this study are nominal scaled and discrete. Thus it is inappropriate to treat them as continuous scaled variables. The values of these variables should be recorded by creating a new set of variables corresponding to the original categories. The number of the new variables is less than the number of variable categories by one. The omitted category is the reference category. The SPSS logistic regression procedure creates automatically new variables for the independent variables declared as categorical. The independent variables and the coding system are shown in Table 2.

Seven runs of the SPSS logistic regression procedure were executed to check the effect of the cut-off values (alpha level), the stepwise approach (forward, backward), and to get the best model fit the data. Four runs out of these were executed using the backward stepwise logistic regression with 0.05, 0.10, 0.15, and 0.20 cut off values for both entry and removal. Another two runs were executed using the forward stepwise logistic regression with 0.10 and 0.15 cut off values for both entry and removal. The last run was carried out using the logistic regression with all of the independent variables in the model. For the purpose of assessing these models statistically, some important statistic measures are presented in Table 3.

Table 2 Independent Variables, Coding System and description

Variable	Codes used	Abbreviation
Accident Long Delays [1]	1= present 0 = absent	A1
Accident Delays [1]		A2
Congestion Long Delays [1]		C1
Congestion Delays [1]		C2
Roadworks Long Delays [1]		R1
Roadworks Delays [1]		R2
Accident # Minutes Delay [1]	Continuous variables	DelayAcc
Congestion # Minutes Delay [1]		DelayCong
Roadworks # Minutes Delay [1]		DelayRw
Extra Travel Time = (Travel time at alternative – Travel time at usual route)		ETT
ETT/ Travel time at usual route		ETTR
Age	1=17-35, 2=35-50, 3>50	Age
Sex	1= male, 2= female	Sex
Familiarity with the usual route	1=every day,2=few times aweek, 3= once a week, 4=less	Travlusu
Familiarity with the alternative routes	1= few times aweek,2= few times a month, 3= once a month, 4=less	Travlalt
Previous experience with VMS	1=a lot, 2=little, 3=nil	Experien
Assessment of previous VMS information	1= very satisfactory, 2=satisfactory,3=uncertain, 4=very unsarisfactory	VMS inform
Willingness to diver at later diversion points	0= no later points, 1= divert, 2= not divert	Divornot
Trip purpose	1=commuting,2=Shopping, 3=Entertainment, 4=others	Trippurp

1 Note that only one of these variables is present in the model at a time.

Table 3 shows that the regression models from the backward and forward stepwise approaches are identical. The models are the same when the cut-off values are 0.10, 0.15 and 0.20. The change in Chi-square, χ^2, from the full model, containing all the independent variables, to the second model, alpha α= 0.10, is not statistically significant (p=0.99) indicating that the independent variables which are not included in the second model do not make any significant contribution to the prediction of the dependent variable. Also, the change in χ^2 from the second model, α =0.10, to the third model, α=0.05, is statistically significant (p=0.046) indicating that the independent variables that omitted from the third have a significant contribution in the prediction of the dependent variable. The value of –2LL for the second model is smaller than for both first and the third model indicating that the second model is better than the first and the third. Thus, the second model, α= 0.10 is selected as the basic model to indicate the independent variables that contribute significantly to the prediction of the dependent variable. These variables are: all the variables related to the displayed delay and its cause, age, willingness to divert at next diversion points, previous experience with VMS, familiarity with the alternative routes and the trip purpose.

Table 3 The Effect of cut off values (alpha level) of the stepwise regression

Model	Model χ^2 (D.F.)	Significance of χ^2 (D.F.)	Change of χ^2 from the previous model	Significance of the change of χ^2	-2LL
Containing all the independent variables	187.51 (31)	0	____	____	588.5
Using backward elimination α=0.10	188.25 (24)	0	0.74 (7)	0.99	587.7
Using backward elimination α=0.05	180.17 (21)	0	8.08 (3)	0.046	595.8
Using backward elimination α=0.15	The same results as the obtained model by using the backward elimination α=0.10				
Using backward elimination α=0.20					
Using forward selection α=0.10					
Using forward selection α=0.15					

Table 4 shows the basic model results which include the Wald test and its significance for each independent variable in the model. The Wald test shows that, all independent variables are statistically significant. Thus, all independent variables in the selected model should be kept in the model as they significantly affect the dependent variable.

The signs of the estimated variable coefficients shown in column B of Table 4 are in general accordance with *a priori* expectations. A positive coefficient indicates that the probability of diversion increases with the presence and/or increase of the variable concerned; variables with positive coefficients include: all incident cause and severity variables, the ratio of the extra travel time to the usual travel time, experience with VMS, and familiarity with the alternative routes. Negative coefficients, indicating an inverse relationship between the variable and the probability of diversion, were found for the extra travel time (ETT), the willingness to divert at a later point, less familiarity with the alternative routes and shopping trips as opposed to commuter trips. Some inconsistent signs were found for the familiarity with the alternative routes and trip purpose variables; these will be further investigated in the final analysis of the full data set.

3.5.2 Checking the Linearity in the Logit

At the variable selection stage, the logit regression model was assumed to have a linear form, i.e. it was assumed that each independent variable has a linear relationship with the logit. This assumption is common and acceptable at this stage as indicated by Hosmer & Lemeshow (1989). But after finishing the variable selection stage and identifying the important independent variables to be included in the model, this assumption should be checked and the appropriate modifications should be implemented. The linearity assumption in the logit should be checked for all the continuous independent variables included in the model.

The Box-Tidwell technique (Hosmer amd Lemeshow, 1989) was applied to check the linearity for the continuous independent variables in the basic model. For the ETT, the term

"ETT*Ln(ETT)" was added to the basic model and resulted in a new model with χ^2 of 188.9 and 25 degree of freedom, d.f. The difference of χ^2 between the new model and the basic model, 0.8 with 1 d.f., is not statistically significant, p=0.4. Thus, the coefficient of the added term is not statistically significant. Accordingly, it is accepted that the relationship between the ETT and the logit is linear.

Table 4 The basic model (backward stepwise regression with alpha=0.10)

Initial Log Likelihood Function -2 Log Likelihood 775.96264
* Constant is included in the model.

-2 Log Likelihood	587.711
Goodness of Fit	2018.845

	Chi-Square	df	Significance
Model Chi-Square	188.251	24	.0000
Improvement	-1.716	1	.1903

Classification Table for RESPONSE

		Predicted .00 (0)	Predicted 1.00 (1)	Percent Correct
Observed .00	0	77	96	44.51%
Observed 1.00	1	24	491	95.34%
			Overall	82.56%

Variables in the Equation

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
A1(1)	3.9985	.6310	40.1513	1	.0000	.2217	54.5184
A2(1)	3.0092	.5683	28.0379	1	.0000	.1832	20.2716
C1(1)	2.8321	.5834	23.5684	1	.0000	.1667	16.9817
C2(1)	2.3913	.5652	17.8982	1	.0000	.1431	10.9278
R1(1)	3.8915	.6682	33.9204	1	.0000	.2028	48.9850
R2(1)	1.5801	.5347	8.7340	1	.0031	.0932	4.8555
DELAYCO	.1772	.0311	32.4472	1	.0000	.1981	1.1939
DELAYACC	.2044	.0338	36.5707	1	.0000	.2111	1.2268
DELAYRW	.1840	.0323	32.4682	1	.0000	.1982	1.2020
ETT	-.2058	.0360	32.7821	1	.0000	-.1992	.8140
ETTR	1.6012	.7370	4.7205	1	.0298	.0592	4.9591
AGE			14.3777	2	.0008	.1156	
AGE(1)	.9078	.2406	14.2341	1	.0002	.1256	2.4789
AGE(2)	.5609	.3603	2.4239	1	.1195	.0234	1.7523
DIVORNOT			8.7802	2	.0124	.0785	
DIVORNOT(1)	-.1293	.3400	.1445	1	.7039	.0000	.8788
DIVORNOT(2)	-1.2794	.4926	6.7450	1	.0094	-.0782	.2782
EXPERIEN			7.2885	2	.0261	.0651	
EXPERIEN(1)	.6535	.2864	5.2056	1	.0225	.0643	1.9222
EXPERIEN(2)	.1021	.3158	.1044	1	.7466	.0000	1.1074
TRAVLALT			13.0495	4	.0110	.0807	
TRAVLALT(1)	.0046	.8417	.0000	1	.9956	.0000	1.0046
TRAVLALT(2)	-.3711	.7972	.2167	1	.6415	.0000	.6899
TRAVLALT(3)	.8588	.8465	1.0293	1	.3103	.0000	2.3603
TRAVLALT(4)	.3069	.7950	.1490	1	.6995	.0000	1.3592
TRIPPURP			7.4348	3	.0593	.0430	
TRIPPURP(1)	-.4974	.3883	1.6411	1	.2002	.0000	.6081
TRIPPURP(2)	.0264	.4373	.0037	1	.9518	.0000	1.0268
TRIPPURP(3)	.8846	.3937	5.0482	1	.0247	.0627	2.4221
Constant	-1.8148	.9585	3.5845	1	.0583		

The procedure was repeated for the variables ' delayco', 'delayacc' and 'delayrdwk', and for each of these the test showed that the relationship between these independent variables and the dependent variable is linear.

3.5.3 Variable Interactions

The interaction between any two independent variables indicates that the effect of one of them on the dependent variable varies with the values of the other independent variable.

In the basic model the variables: ETT, age, willingness to divert at next diversion points, the previous experience with the VMS, the familiarity with the alternative routes and the trip purpose are expected to have a potential to interact with the other variables included in the model. Accordingly, a total of 69 interaction terms can be included in the model.

To assess the contribution of the interaction terms to the models, the effect of including each interaction term was examined by adding this term, ie. the product of the variables, to the basic model and comparing χ^2 for the resulting model and the basic model. If the difference between χ^2 is significant, then the interaction term has a significant effect on the prediction of the dependent variable. A total of 69 runs of the SPSS logistic regression procedure were carried out to examine the effect of the interaction terms.

The results show that the interaction terms that can be added to the basic model based on a significance level of 0.10 or less are: Age*ETT, Age*travelalt, divornot*C1, divornot*experience, experience*trippurpose, travelalt*C2, trippurpose*A2, trippurpose*C2, and trippurpose* delayco. But, the interaction terms divornot*C1, and trippurpose*C2 will be excluded from the model based on 0.05 or less significance level. To check the effect of those two interaction terms, two models were fitted, one including the interaction terms in addition to the basic model based on 0.10 significance level, and the other based on 0.05 significance level. The difference in χ^2 between the two models, 8.4 with 5 d.f., is not statistically significant, p=0.15, thus the two interaction terms divornot*C1, and trippurpose*C2 will not be included in the model.

The Wald test for the model which includes the significant interaction terms indicates that the two interaction terms: travelalt*C2, trippurpose*A2 are not significant. Thus, the effect of each term was examined by removing it from the model and assessing the significance of the difference in χ^2. Removing the interaction term travelalt*C2 from the selected mode resulted in a new model with χ^2 of 278.2 and 49 d.f., and a difference of χ^2 between the two models of 8.6 with 4 d.f. which is not statistically significant, p=.08. Thus, the term travelalt*C2 will be removed from the selected model.. Removing the interaction term trippurpose*A2 from the selected mode yielded a model with χ^2 of 272.6 and 50 degree of freedom. The difference of χ^2,14.6 with 3 d.f. is statistically significant, p<.005 and thus, this interaction term has a significant effect on the prediction of the dependent variable and it will be kept in the model.

The final model consists of two parts: the basic model plus the following interaction terms: Age*ETT, Age*travelalt, divornot*experience, experience *trippurpose, trippurpose*A2, and trippurpose* delayco.

3.6 ASSESSMENT OF THE MODEL FIT

The purpose of assessing the model fit is to test whether the predicted values give a reasonable description of the observed values. Assessing the model fit should include evaluating the fit of the overall model in addition to the accuracy of prediction. Evaluation of the overall model can be carried out using G_M/model χ^2, -2LL, and R^2_L. On the other hand, the evaluation of prediction accuracy can be carried out using τ and λ parameters (Menard 1995). Assessment of the model fit was carried out for the final model as well as the basic model. Table 5 shows the values of model χ^2, -2LL, R^2_L, τ and λ for both models.

Table 5 Assessment of fit measures for the final and the basic model

	Final Model	Basic Model
χ^2 (d.f.)	278.2 (49)	188.3 (24)
Significance of χ^2	< 0.001	< 0.001
-2LL	497.8	587.7
λ	0.68	0.44
τ	0.60	0.54
d (for λ)	6.15	4.66
d (for τ)	12.3	10.9
Significance of λ	< 0.001	< 0.001
Significance of τ	< 0.001	< 0.001
R^2_L	0.36	0.24

Table 5 shows that the model χ^2 is significant for both models, but the χ^2 is larger for the final model indicating a better fit than the basic model. Also, the –2LL value is smaller for the final model indicating the same as the χ^2. The values of λ and τ for the final model indicate good prediction efficiency for the final model and are slightly higher than for the basic model. The significance of λ and τ indicates a strong relationship between the observed and predicted classification of cases for both models. The value of R^2_L for the final model, 0.36, indicates a moderate strong relationship between the dependent and independent variables and it is higher than for the basic model.

3.7 THE VALUES OF DISPLAYED DELAYS

The model was used to estimate the values of the displayed delays in units of travel time (in minutes) as the absolute ratio of the coefficient of the incident severity variable (ie. the displayed delay) and the coefficient of the travel time on the usual route. Table 6 shows the model coefficients and the estimated values of delay. As expected, all the incident cause and severity variables have a positive coefficient, indicating an increased probability of diversion, while the negative coefficient of the usual travel time variable indicates that the diversion probability decreases when the trip travel time on the usual route increases.

Table 6 shows that the values of quantitative delay displayed by the VMS (eg. "15 Minutes Delay" or "30 Minutes Delay") vary between 1.03 to 1.22 minute of the usual route travel time depending on the stated cause of delay. There is no significant difference between the causes of delay. This result may be explained by the assumption that the drivers may not pay attention to the delay cause when the value of delay time is displayed on the VMS.

When the VMS message displays a qualitative indication of the incident severity (ie. "Long Delay" or "Delay"), this message is interpreted by drivers as equivalent to a certain amount of delay time. The values of "Long Delay" vary between 18.21 and 25.34 minutes of the usual route travel time according to the stated cause of delay. The values of "Delay" vary between 8.85 and 19.04 minutes of the usual route travel time according to the stated cause of delay. Table 6 shows also that there is no significant difference between the coefficients of "Long

Delay" and "Delay", neither for the "Accident", nor for the "Congestion" causes. On the other hand, there is significant difference between the coefficients of "Long Delay" and "Delay" for "Roadwork".

Table 6 Coefficient estimates and the value of delay

	Coefficient	SE	t-ratio	value in minutes*
Acc Minutes	0.1617	0.0275	5.88	1.22
Rdwk Minutes	0.1468	0.0275	5.33818182	1.11
Cong. Minutes.	0.1366	0.0261	5.23371648	1.03
Acc. Long Delay	3.3656	0.5816	5.78679505	25.34
Rdwk. Long Delay	3.1387	0.5988	5.24164997	23.63
Cong. Long Delay	2.4183	0.5375	4.49916279	18.21
Acc. Delay	2.5287	0.5126	4.93308623	19.04
Rdwk. Delay	1:1754	0.4816	2.44061462	8.85
Cong.Delay	1.9140	0.5089	3.76105325	14.41
TTIMEUSU	-.13280	0.0163	-8.147239264	--
Constant	-0.4496	0.3651	-1.23144344	--

*Expressed in units of travel time

4. CONCLUSIONS AND FURTHER WORK

This paper presented results of the effects of VMS messages on drivers' route choice behaviour in Sydney, based on the analysis of the first 200 interviews collected in 1999. A statistical analysis of the data was carried out and a general route choice model was developed which can be used to predict the diversion rate resulting from various VMS message contents. The full set of 400 interviews is now available and the statistical analysis of the full data set is currently in progress. All steps of the model development will be repeated, however, we expect that the final model will not be significantly different from the one presented here.

The following conclusions can be drawn from this study. The probability of diversion increases as the displayed delay time on the usual route increase. The probability of diversion in case of "Long Delays" is higher than for "Delays". The cause of the incident (Accident, Roadwork or Congestion) does not have an obvious effect on the probability of diversion when the displayed severity is "Long Delays" and delay in minutes, while it has a clear effect when the severity is "Delays". The probability of diversion in case of Accident is higher than that for Congestion and Roadworks when the severity is "Delays".

Previous studies have shown that the extra travel time (ETT, the travel time difference between the alternative and the usual route under normal conditions) has a strong effect on route choice behaviour in the presence of the VMS information. However, most previous studies were based on one or two fixed trips, therefore the results were relevant to the given fixed trip travel time only. This study has confirmed our assumption that there is a strong relationship between the trip travel time and the ETT, which affects the probability of diversion. Drivers are less likely to divert with the increase of ETT under normal conditions. The probability of diversion also decreases with the increase of trip travel time. Our model can be used with a range of trip times from 15 to 60 minutes in the Sydney Metropolitan region.

Drivers' age and sex, do not affect significantly the probability of diversion. Also, the trip purpose, drivers' assessment of previous VMS information, and drivers' familiarity with the usual route do not have a significant effect on the probability of diversion. These findings are consistent with the results obtained by Brocken (1991) and Benson (1996).

The findings of this study support the hypothesis that there is a predictable relationship between displayed VMS information and driver response, in the form of route selection, during non-recurrent congestion. If the diversion rate for various message contents can be predicted to a reasonable accuracy, this can be used as a valuable traffic management tool during incidents: a simulation model can estimate the 'optimum' diversion rate for any incident scenario, then the appropriate message content can be selected which would be most likely to generate the required diversion response from motorists.

The developed route choice model has the ability to predict the probability of diversion resulting from various VMS message contents for any car trip in Sydney Metropolitan Region. In the next phase of the study the prediction model will be incorporated into a microscopic urban transport network simulation model (SITRAS, see Hidas and Behbahanizadeh, 1998) for testing and validation. The model will be used to develop and evaluate alternative incident management plan options in the Sydney Metropolitan Region. Results of the second phase of the study will be reported in a subsequent paper.

REFERENCES

Abdel-aty, M. A. , Kitamura, R. ,and Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transpn. Res.*, **5C**, 39-50.

Benson, B. G. (1996). 'Motorist Attitudes About Content of Variable-Message Signs', *Transportation Research Record* 1550, Transportation Research Board, Washington, D. C., pp 48-57.

Bonsall, P. (1994a) *Driver response to VMS- QUO VADIS investigations.* First World Congress on Applications of Transport Telematics and Intelligent Vehicle-Highway Systems, Paris, December 1994.

Bonsall, P. (1994b) *Simulator experiments and modelling to determine the influence of route guidance on drivers' route choice.* First World Congress on Applications of Transport Telematics and Intelligent Vehicle-Highway Systems, Paris, December 1994.

Bonsall, P. , Firmin, P. , Anderson, M. , Palmer, I. ,and Balmforth, P. (1997). Validating the results of a route choice simulator. *Transpn. Res.*, **5C**, 371-387.

Bonsall, P., and Palmer, I. (1998). *Route choice in Response to Variable Message Signs- Factors Affecting Compliance.* Paper presented at the 8th World Conference on Transport Research, July 1998.

Brocken, M. G., and Van der Vlist, M. J. (1991). 'Traffic control with variable message signs', *Proceedings of Vehicle Navigation and Information Systems Conference*, pp. 27-45

Chatterjee, K., and Hounsell, N. (1998). *Modelling Driver Response to Variable Message Signs With RGCONTRAM.* Paper 373, Transportation Research Group, University of Southampton, U.K.

Commeignes, H., (1992). *The Cost of Congestion in Sydney*, a report prepared for the Roads and Traffic Authority of NSW, Pricing Strategy Branch.

Cramer, J. S. (1991). **The Logit Model: an introduction for economists**. Edward Arnold, London U.K.

Grenzeback, L. R. and Woodle C. E., (1992). The True Costs of Highway Congestion, *ITE Journal*, March 1992, pp. 16-20.

Hidas, P.,and Behbahanizadeh, K. (1998). SITRAS: A Simulation Model for ITS applications. *Proceedings of the 5th World Congress on ITS*. Seoul, Korea.

Hosmer, D W. and Lemeshow, S. (1989). **Applied Logistic Regression**. Wiley, New York, USA.

Judycki, D. C. and Robinson J. R. (1992). Managing Traffic during Nonrecurring Congestion, *ITE Journal*, March 1992, pp. 21-26.

Liao, T. F. 1994. Interpreting Probability Models: Logit, Probit and other Generalized Linear Models. Sage University Paper series on Quantitative applications in the Social Science, 07-101. Thousand Oaks, CA: Sage

Menard, S W (1995). Applied Logistic Regression Analysis. Sage University Paper series on Quantitative applications in the Social Science, 07- 106. Thousand Oaks, CA: Sage.

Wardman, M. , Bonsall, P. W. ,and Shires, J. D. (1997). Driver response to variable message signs: a stated preference investigation. *Transpn. Res.*, **5C**, 389-405.