Evaluating the
Supplementary Road Safety Package:

Models That Count

Stephen Irvine

Supervisor: Dr. Murat Genç

A dissertation submitted in partial fulfilment of the requirements for the degree of

Bachelor of Commerce with Honours in Economics

October 2004

Department of Economics
University of Otago
ABSTRACT

In an attempt to lower the level of road trauma in New Zealand, the Land Transport Safety Authority introduced the Supplementary Road Safety Package (SRSP) in October 1995. The package consists of targeted speed and alcohol enforcement, and features graphic television advertising highlighting the consequences of unsafe driving. Over the first four years the campaign was allocated a budget of NZ$50.06 million and charged with reducing 80 fatalities, 450 serious injuries and 1600 minor injuries. Although a requirement of the package’s approval was that it be thoroughly evaluated, no consistent conclusion has been drawn. Recognising the discrete and strictly positive nature of road trauma measures, this dissertation adds to the body of literature by adopting statistical modelling techniques specifically designed for the analysis of such count variables: The Poisson and Negative Binomial regression models. While the Poisson model finds a significant level effect on the number of serious injuries from the SRSP’s introduction, no statistically significant effect is found using the more appropriate Negative Binomial model.
ACKNOWLEDGEMENTS

Though a daunting task at first, putting this dissertation together has proved to be a rewarding experience. Thank you to the following people for their various contributions to this work.

A massive thanks to my supervisor Dr Murat Genc. Not only did he initially suggest the methodological approach I have adopted, he made himself available in-between his teaching and research responsibilities, and the Bridge gaming ‘circuit’, to provide general guidance, read through my various drafts and correct my sometimes interesting grammar. This dissertation is definitely much better for your guidance.

For making the year all-the-more interesting, cheers to all the 4th year Economics students (and the lone Masters student). As the year progressed the sing-a-longs, random conservations, and bad Economics jokes increased without restraint; making it a truly ‘memorable’ year (perhaps sometimes in a slightly irritating sense!). In fact, thanks to the staff in the department too, for making it a bearable place to live out the majority of my life this year!

Cheers to my friends outside of university too – while most of you can’t be credited with helping get this dissertation done, I appreciate you guys for always suggesting an alternative activity to aid in my procrastination! Special mention to Lisa – thanks for being conned into being my final draft proof-reader.

Thanks to Joanne Leung, Paul Graham and Jagadish Guria of the NZ Land Transport Safety Authority, and Birabaharan Appapillai of the Police Infringement Bureau for the provision of much of the data required for the empirical part of my dissertation. Certainly without access to road trauma and enforcement data, statistical evaluation would not have been possible.

For discussing the intricacies of marketing and the measuring of advertising levels, thanks to Terry Macpherson of Massey University. As an economist I was a fish out of water in many respects, but hope I have been able to do the data justice.
## CONTENTS

1. Introduction 1

2. Literature Review 3
   2.1. Model Structure and Estimation 3
   2.2. Dependent Variable 5
   2.3. Independent Variables 7
      2.3.1. Compulsory Breath Testing and Speed Cameras 8
      2.3.2. Trend and Seasonality 9
      2.3.3. Socioeconomic Factors 11
      2.3.4. Advertising – Supplementary Road Safety Package 12
   2.4. Results 14

3. Analysis of Count Data 16
   3.1. The Poisson Regression Model 16
   3.2. The Negative Binomial Model 18
   3.3. Goodness of Fit Tests 18
   3.4. Misspecification Tests 20

4. Model Formulation 21
   4.1. Dependent Variable 21
   4.2. Independent Variables 23
   4.3. Model Structure 29

5. Results 30

6. Conclusions 36

7. Limitations and Future Research 37

8. References 38

9. Appendix 40
1. INTRODUCTION

In 1994 there were 580 fatalities and 3,268 serious injuries resulting from traffic accidents in NZ with the corresponding social productivity cost estimated at $3.3 billion in 1994 dollars. The true cost is expected to be markedly larger however, as it would also include third party trauma and the cost of traffic disruptions (Tay 2001).

The Supplementary Road Safety Package (SRSP) was introduced by the Land Transport Safety Authority (LTSA) in October 1995 in an effort to reduce New Zealand’s cost of road crashes. Based on an Australian campaign implemented in Victoria in 1989 which achieved reductions of 50% in road fatalities over three years of operation, the New Zealand campaign consisted of targeted speed and alcohol enforcement and was supported by graphic and shocking television advertising highlighting the consequences of unsafe driving. Initial funding was $12.2 million per year over four years, of which $7.1 million was allocated to the shock advertising campaign. The advertising budget was increased to $8.4 million in 1996 to cover seat belt publicity, and again in the following year to $9.8 million. Though the complete campaign included targeted use of Compulsory Breath Testing (CBT) and the speed cameras, the advertising campaign played a prominent role in the SRSP, at least in terms of budget allocation.

The SRSP was considered necessary if the target of reducing road trauma to 420 fatalities and 11,000 Police-reported serious injuries by 2001 was to be met. However, the ‘shock’ nature of the advertising campaign was met with opposition, with public outcries deeming their graphic nature unnecessary, ineffective, or both.

A requirement of the SRSP’s approval was that it be regularly and independently evaluated over time against its performance targets. While a number of studies have sought to evaluate the effectiveness of the advertising campaign to date, no consistent conclusion has been drawn. Tay (2001, p.14) argues that such disparities are characteristic of policy evaluation in social sciences in general:

> In modelling the relationship between the policy variables and the performance measure, the analysts have to make many assumptions and depending on their choice, the outcome may differ considerably.
Moreover, the 'fundamental problem' of such analyses put forward by authors is that of non-experimental data – the analyst has no control over the conditions under which the data are generated. As a result, there is no control group to isolate potential confounding factors, and evaluations are based on fitting historical observational data to statistical models, whose results can vary considerably depending on the analysts’ prior beliefs (Tay, 2003). As Gregg Easterbrook put it: “Torture numbers, and they'll confess to anything.”

This dissertation aims to extend the literature on the evaluation of road safety countermeasures by adopting regression techniques specifically designed for the analysis of count data. Although the Poisson regression model is considered the benchmark for such analysis, the assumptions underlying it prove too restrictive for most real world data, and we turn our attention to the more appropriate Negative Binomial model.

The structure of this dissertation is as follows. Section two provides an overview of the literature to date, highlighting several issues authors must consider when evaluating the effectiveness of road safety countermeasures. Section three lays out the theory of the 'count data' estimation techniques to be employed in the empirical component of this dissertation. Section four outlines the construction of the model and the choice of control variables, while section five contains the empirical results obtained and a discussion of the findings. Section six summarises the conclusions reached and section seven outlines the limitations of this work and suggests directions for future research.
2. LITERATURE REVIEW

As highlighted by Lewis (2001) in “Same Data, Different Conclusions: Analysis of the New Zealand Drink-Driving Campaign”, the results obtained from analysing non-experimental data can differ widely. A review of the literature surrounding the evaluation of Road Safety advertising highlights several issues where authors depart from previous evaluation: the model structure and estimation technique employed, the measurement used for enforcement levels, the treatment of trend and seasonality, controlling for socioeconomic influences, and of course the way in which the advertising campaigns effect is captured. These issues and their treatment in previous literature are outlined in this chapter.

2.1 Model Structure and Estimation:
Statistical evaluation of the SRSP began with Macpherson and Lewis (1998) and Cameron and Vulcan (1998) using monthly and quarterly road trauma data respectively, both estimating their models via Ordinary Least Squares (OLS). Concerns as to the suitability of OLS modelling of road trauma were first raised by Stroombergen (1998), when his re-investigation of Cameron and Vulcan’s model found evidence of first order autocorrelated error. Highlighting the possible distortionary effects on coefficient estimates for both model versions, Stroombergen re-estimated the models by Prais Winsten estimation and included lagged dependent variables in an attempt to account for the autocorrelation. In the re-estimated relationship, the dummy variable for 1995/1996 became insignificant, while the 1996/1997 dummy was 50-60% smaller in the model using unemployment as the economic variable. The intervention effect of the CBT programme was also found to be 30-50% and 50-70% smaller in the ‘new car’ and unemployment model versions respectively. As a consequence, Stroombergen suggested future evaluation use more advanced estimation techniques such as Maximum Likelihood Estimation to cope with the problem of lagged dependent variables and autocorrelated errors.

Tay (2001) also noted inconclusive Durbin Watson statistics in the previous evaluations, and re-estimated Cameron and Vulcan’s models via a first order autoregressive (AR) model using OLS and the Cochrane Orcutt two-step procedure.
The results obtained from the OLS and AR models were qualitatively and quantitatively very similar, which given the insignificant estimate of rho and inconclusive Durbin Watson statistics, was not wholly unexpected (Tay, 2001).

Marking a change of methodology for New Zealand SRSP evaluation Tay, Ozanne and Santiono (1999) explicitly allowed for the discrete nature of the number of road accidents by adopting a Poisson Regression approach. The authors highlight the consequences of failing to allow for the nature of discrete “count” variables, stating that estimates from standard linear regression techniques such as OLS are biased, inconsistent and inefficient (Tay et al., 1999). In support of this view, Tay (2003) found via the Jacque-Bera test that the normality assumption was violated in preliminary OLS estimation of his fatal crash model. Further, Povey, Keal and Frith (2003) considered a Negative Binomial error within their number of road related injuries model to be more appropriate, citing an inflated variance due to injuries being ‘clustered’ within crashes.

Guria and Leung (2002) adopted Principal Component Analysis (PCA) citing insufficient degrees of freedom and strong collinearities between explanatory variables. PCA transforms a number of possibly correlated variables into a set of uncorrelated “Principal Components” (PCs) which are linear combinations of the explanatory variables, but without much loss of much information. Degrees of freedom are conserved as a smaller number of PCs are included than the number of variables that would otherwise be included in OLS. To obtain parameter estimates the authors performed OLS based on the PCA results, avoiding the possibility of bias in the PCA estimates.

A General to Specific approach to modelling was adopted in general, in that variables found to be insignificant were excluded from final models. However, this was ultimately at the discretion of the author, for example Cameron and Vulcan (1998) retained the seasonal dummy variables, as well as dummies for the SRSP, CBT and speed camera programmes regardless of significance, to arrive at their final estimates.

---

1 The authors had as few as 10 observations, with 8 candidate explanatory variables.
In terms of variable transformation, both Macpherson and Lewis (1998) and Cameron and Vulcan (1998) adopted a double log model specification, offering only the explanation that multiplicative (rather than additive) structures had been previously used in evaluations of road trauma. However using data provided by Macpherson and Lewis (1998), Tay (1999) re-estimated the same relationship using linear, log-linear, semi-log and double-log models, concluding that the double log model actually produced the worst fit for the data.

2.2 Dependent Variable:
In choosing the performance measure to gauge the effect of an intervention, there is argument among researchers about whether we should be analysing driving behaviour, or the results of bad behaviour. In other words, measuring the effect by seeing if drivers drink less, or seeing if accident statistics improve (Lewis, 2001).

Macpherson and Lewis (1998) considered the number of Evidential Breath Tests (EBTs) a more direct behavioural measure of drink driving behaviour and consequently of the advertising campaign’s effect on the driving population. They argued that while serious casualties align to the publicly stated SRSP goals, there are a number of possible factors other than the alcohol levels of drivers that contribute to crashes such as vehicle age, road conditions and the time of day. In addition, later dubbed by Tay (2001) as the “non-encompassing and non-exclusivity” argument, Macpherson and Lewis (1998, p.43) reasoned:

...people watching an advertisement may drink and drive and not have a crash, and others who don’t see the advertisement may drink and drive and crash. Others still, may not drink and drive as a result of watching the advertisement, and also have a crash.

Given these complications, the number of EBTs was chosen by the authors as the performance measure\(^2\). Tay (2001, p.4) argued that a similar argument can also be applied to EBTs:

\(^2\)The authors do note however the frequent disparities between the number of drink-driving offenders on CBT records, and the number actually convicted.
...not all drink-drivers who watched the advertisements were caught by the police and among those who were caught, some did not watch the advertisements.

Cameron and Vulcan (1998) analysed serious casualties⁢, noting that previous research regarding the experience of the Victorian campaign had been successful in developing models for monthly variations in road trauma. The quarterly models of Guria and Leung (2003) also examined serious casualties, given concerns that the number of fatal crashes within a quarter may be too small relative to random variation, to detect the effects from road safety interventions. Following estimates used in social cost calculations that only about half of serious injuries are reported to the Police, the serious casualty and serious crash series were adjusted to account for such underreporting⁴.

Allowing for separate effects of the SRSP dependent on the time of day, Cameron and Vulcan (1998) also constructed a model for serious casualty crashes subdivided into high or low alcohol hours, and urban or rural areas⁵. Similarly, in their analysis of the effect of enforcement on the mean open road speed, Povey, Keal and Frith (2003) excluded reported injury crashes during high alcohol hours, citing that significant interventions related to drink driving were implemented over their study period, while speed enforcement is generally concentrated outside high alcohol hours.

While Cameron and Vulcan (1998) consider serious casualties as the definitive performance measure for the SRSP’s effect, they also constructed a model of fatalities based on their serious casualty models. The authors expected the fatality model to produce a worse fit for the data, and consequently to weaken statistical tests of significance due to a lower frequency of occurrence. Tay (2003) also analysed the total number of fatal crashes, noting that the SRSP outcomes had already been examined via drink-driving behaviour and serious crashes in previous evaluations. Further, while investigating the advertising campaign among different segments of society, Tay et al (1999) focussed only on fatal crashes related to alcohol, speed and

---

³ Serious casualties are defined as fatalities plus serious injuries.
⁴ Guria and Leung (2003) note that this adjustment implicitly applies a constant reporting rate over time, which we have no reason to assume.
⁵ High alcohol hours are defined by the LTSA as 10pm to 4am daily plus 4am to 6am on Fridays, Saturdays and Sundays. Urban areas are road with a speed limit up to 70km/h.
drugs – in line with the advertising’s targeted objectives. Perhaps the main reason fatal crash numbers have been used widely in road safety literature is due to their very high average cost per crash, extensive media attention, and high reliability in data collection (Tay, 2003).

Guria and Leung (2003) also used fatalities and fatal crash data within their annual models arguing that they are a good approximation of the overall road safety outcome from interventions, and their majority contribution to the total social cost of road trauma. In addition, one version of the model excluded motorcycle crashes, the authors reasoning that motorcyclists usually face a higher risk than other road users, and that part of the decline in total road trauma experienced since 1990 could be due only to the decline in their use.

Further controlling for changes in traffic over time, Guria and Leung (2003) normalised all dependent variables by the Transit New Zealand Transport Volume Index (TVI). They reasoned that the actual risk per kilometre could be decreasing over time, but due to traffic growth alone we may observe either no change, or even an increase in road trauma. While this implicitly assumed a linear relationship between the road utilisation and road trauma, inclusion of the explanatory variable alongside a linear time trend is infeasible due to high correlation between them. Rather than explicitly including a measure of traffic volume in their analyses, some authors have instead chosen to capture such factors by including a linear trend and/or an ‘economic’ variable.

2.3 Independent Variables:
Though Cameron and Vulcan (1996) were unable to incorporate a statistical model due to lack of data since the SRSP’s implementation, they suggested four possible contributing factors to the reduction in road trauma observed during 1995/1996. First was the impact of the CBT and speed camera programmes. Second, the pre-existing downward trend in road related accidents observed since the early 1990s. Third were socioeconomic factors, particularly accounting for the economic recovery during 1995/1996 which may have caused the SRSP’s effectiveness to be under-estimated. Their final candidate factor was the effect of the SRSP. The treatment of these factors and their incorporation in various studies are discussed below.
2.3.1 Compulsory Breath Testing and Speed Cameras

In previous literature, the use of dummy variables to account for the speed camera and CBT programmes has been the most prevalent technique, which serves to measure the average safety effect over the period examined (Chang & Yeh, 2003). Cameron and Vulcan (1998) included dummy variables for both the introduction of the CBT and speed camera programmes in the second and fourth quarters of 1993 respectively. While the CBT dummy was considered for inclusion in all four model versions, the speed camera dummy was only included in the model of low alcohol hour urban crashes, as previous evaluation had determined that the speed camera programme’s effect was significant only within these periods\(^6\). Consequential re-evaluations of the Cameron and Vulcan study by Stroombergen (1998), White (2000), Tay et al (2000) and Tay (2001) continued the dummy variable approach. However, Tay (2001) and Guria and Leung (2002) noted the difficulty in separating the SRSP effects from those of the CBT and speed camera programmes given their introduction in quick succession; choosing instead to include one dummy variable to control for the effects of both enforcement programmes. Tay et al (2003) also controlled for the CBT and speed camera programmes via dummies, but allowed the magnitudes of these variables are able to differ according to age bracket and sex.

Evaluation of Evidential Breath Tests by Macpherson and Lewis (1998) and a later re-evaluation by Tay (2003) used the number of CBTs performed to represent the level of general police enforcement. Povey, Keal and Frith (2003) included the number of speed camera and non-speed camera tickets to proxy for camera and non-camera enforcement respectively. A dummy variable for the implementation of the Highway Patrol was also considered for inclusion but was later omitted due to high collinearity with the number of non-camera tickets issued. Guria and Leung (2002) captured the level of police enforcement by constructing a Strategic Police Hours index. Included are the number of budgeted hours appropriated for alcohol, safety belt, speed control, and general visible patrol, as taken from published annual New Zealand Road Safety Programmes.

\(^6\) As suggested by the authors, see Mara et al (1996).
2.3.2 Trend and Seasonality

It is common practice in multiple regression models of time series data to include a trend variable. Besides mitigating the problem of a 'spurious' regression, the trend variable serves as a proxy for those variables whose data are difficult to obtain (Tay, 2001). Cameron and Vulcan (1996) suggested that while the CBT and speed camera programmes may have influenced the downward trend in road trauma, part of the reduction attributed to the programmes may have instead been due only to a pre-existing downward trend. They reasoned that if both a pre-existing downward trend and an economic recovery were in effect during the first year of the SRSP, the two factors may have balanced and the observed 11.5% reduction in fatalities could be attributed to the SRSP. Based on the 19% reduction (adjusted for economic effects) realised during the Victorian campaign in 1990, the authors concluded that this balancing assumption was reasonable in the New Zealand case – attributing the full 11.5% reduction to the SRSP.\(^7\)

Macpherson and Lewis (1998) explicitly included a linear time trend to represent the net effect of tightening safety standards, increased traffic growth, and road engineering improvements. Cameron and Vulcan (1998) also suggested the inclusion of a trend component would capture the effects of factors smoothly increasing (rather than decreasing) over time, such as traffic volume. They reasoned that small and gradually increasing factors such as road engineering improvements may have contributed to the downward trend in road trauma, but could not be explicitly included in the model. Expecting a positive trend, but initial analysis revealing a negative coefficient, the authors considered the trend to be capturing effects of the CBT and SRSP and excluded it from their final models. In addition, they reasoned that removal of the trend did not deteriorate the models' explanatory power for quarterly fluctuations, concluding as in their 1996 study that the systematic improvements in roading offset the increases in kilometres travelled.

Re-estimating Cameron and Vulcan's model with a trend term, White (2000) highlighted the significant impact the inclusion of a trend term had on the model. The SRSP dummy variables for the first two years were found to be statistically

\(^7\) Recall that Cameron and Vulcan (1996) did not employ a statistical model – their conclusions are based on assumption.
insignificant and only marginally significant in the models with the unemployment rate and number of new cars as the economic variable respectively. The Durbin Watson statistics of the models also improved, from rejection to indeterminate in the unemployment model.

Tay (2001) also re-evaluated the Cameron and Vulcan model, suggesting his most important addition being the inclusion of a trend term. Firstly, given the time ordered nature of the residuals, low Durbin Watson statistics observed were suggestive of an important time dependent variable being omitted. In addition, a simple F-test of linear restriction indicated that inclusion of a trend significantly increased the models' explanatory power. Secondly, observing that estimates for the policy variables were sensitive to inclusion / exclusion of the trend warranted investigation into whether the negative trend is affected by the introduction of the CBT and SRSP campaigns. Tay (2001) included two interaction terms in his alternative model to test for such structural change in the trend: CBT-Trend and Advert-Trend. The CBT programme alone did not significantly impact the trend, though the combined effect of both the CBT and SRSP campaigns (i.e., periods when both in effect) was found to have a significant negative impact on the trend during the first two years of the SRSP.

Dummy variables were included to capture either monthly or quarterly variations in road trauma in most models, in accordance with standard econometric practice. However, while Tay et al (1999) accounted for alcohol consumption and weather conditions via inclusion of the quarterly amount of alcohol available for consumption, and the number of wet days within a month respectively, no explicit account for a linear trend or monthly seasonal variation was allowed for within their segmentation analysis of the SRSP.

Noting that two of the quarterly dummies were consistently insignificant in previous models, Tay (2001) included the December quarter dummy only, allowing for the distinct peaks in road trauma series referred to in the media as the "Christmas Road Toll". Aslop and Langley (2000) also highlighted the increased exposure to risk of injury or death during the Christmas period, which has often been explained by greater traffic volumes and driver fatigue.
2.3.3 Socioeconomic Factors

A recent study of the New Zealand experience by Scuffham and Langley (2002) highlighted the importance of controlling for economic factors. They reasoned that economic activity has at least a partial explanation for changes in road trauma, suggesting a possible two-fold influence: either directly by affecting the exposure to a crash via distance travelled or the number of discretionary trips, and/or indirectly by affecting the risk of a crash for a given level of exposure, perhaps via an income effect on vehicle safety. Cameron and Vulcan (1996) also highlighted the importance of accounting for socioeconomic factors in analysis of road trauma countermeasures, suggesting that the economic recovery experienced during 1995/1996 could have bought about a higher level of road deaths due only to growth in traffic volume.

The unemployment rate has been used extensively as a proxy for such economic conditions, capturing any short term fluctuations in traffic and economic activities. Macpherson and Lewis (1998) included the unemployment rate to account for the income effect on alcohol consumption. Similarly, Cameron and Vulcan (1998) estimated separate models using either the unemployment rate or the number of new cars registered, accounting for the amount of discretionary travel and increases in the size of the vehicle fleet respectively. Guria and Leung (2002) took a more detailed view of the function of unemployment in their models, adding that it also affects the quality, in terms of drink-driving, and quantity of travel. On the other hand, the authors suggest an income effect on vehicle safety and roading of changes in income; concluding that the inclusion of the unemployment rate captures the net effect of these two opposing forces.

Tay (2001) chose to include the number of new cars in his re-evaluation, suggesting it had performed consistently better as the 'economic variable' in previous models by Cameron and Vulcan (1998) and White (2000). Two statistical approaches were used for choosing between competing models: the Discriminating approach, using goodness-of-fit measures; and the Discerning approach, using a non-nested F test. Both approaches favoured the models with the number of new cars included as the economic variable (Tay, 2001). Within their 29-year annual model, Guria and Leung

---

8 See Scuffham and Langley (2002) for a summary of previous research.
(2002) expected the number of new cars registered to account for the removal of excise duties on imported used cars in 1990, and the resulting higher level of interaction between vehicles and the overall quality advances of the fleet.

Research by Scuffham and Langley (2002) provided complimentary evidence for the preference expressed by Tay (2001), finding that the unemployment rate is not significant in explaining the number of fatal crashes within a period. However, they did find the change in unemployment lagged by 9-12 months significant, suggesting that unemployment is important in explaining the short run dynamics of fatal crashes (Scuffham & Langley, 2002).

2.3.4 Advertising – Supplementary Road Safety Package

Macpherson and Lewis (1998) and Tay et al (2003) gauged exposure to advertising via an “Adstock” variable, attempting to measure the retained awareness of current and past levels of advertising. The Adstock variable is derived from Targeted Audience Rating Points (TARPs), the percentage of the target audience watching the advertisement relative to the advertising expenditure. The SRSP advertising campaign planned to achieve a total of 800 television TARPs per month following its introduction (Guria and Leung, 2002).

Cameron and Vulcan (1998) and the subsequent re-evaluations modelled the effect of the SRSP via dummy variables. They included two dummy variables, one each for the 1995/1996 and 1996/1997 years the campaign had been in effect, allowing for the campaign’s effect to differ across the years. Tay et al (2003) also adopted the dummy approach, but allowed their magnitude to differ by sex and across four age segments.

Highlighting the traditional dummy variable approach measuring the average safety effect over the period of analysis only, various authors have allowed for a temporal safety effect of road safety interventions9. In their analysis of the ‘Criminal Sanction for Drunk Driving’ (CSFDD) policy implemented in Taipei City in 1999, Chang and Yeh (2003) suggested time variation in such policy is due either to changes in the policy itself, or people’s changing perception of the effect over time. The CSFDD

---

9 See Fry (1996) for further work on advertising ‘wear-out’, and an application to the Victorian road safety campaign.
policy was similar to that of the SRSP in that it was designed to complement increased enforcement initiatives and was introduced amidst mass media promotion and targeted enforcement. Controlling for Police manpower and socioeconomic factors, various functions of 'time elapsed since implementation' were included to capture the variation in the policy's safety effect. The time model specifications were found to out-perform the dummy approach and effectively capture the Life Cycle of the policy on accident rates.

Guria and Leung (2002) used a Real Expenditure on Safety Information and Promotion index, encapsulating a broader definition of 'advertising' than just television campaigns; which includes road safety publications and pamphlets, displays, and educational resources. They argued that the use of different advertising mediums all improve the public's awareness of the road safety message. Given the existence of road safety advertising before the imposition of the SRSP, Tay (2001) also noted the importance of controlling for the level of advertising before the campaign's introduction, in order to differentiate between the general advertising effect, and that of the 'core fear' SRSP campaign (Tay, 2001). However, citing that the appropriate data were not supplied, the author assumed that the effect of advertising in general would be captured partially by the trend variable, and the Advert-Trend interaction term would be testing the specific effect of the SRSP campaign.

Given the SRSP campaign featured a distinct change in the style of advertising to appeal to fear and emotion, Guria and Leung (2002) included two interaction terms, SRSP*SHindex and SRSP*ADindex to test for complementary impacts of the changes in advertising style and targeted enforcement on road trauma. Contrary to earlier studies where an SRSP dummy had been included to represent a 'level effect', the SRSP interaction terms now meant the authors were considering changes in the slope of the trend in road trauma. Given that the strong link between advertising and enforcement in the Victorian road safety campaign was considered essential to its success, Macpherson and Lewis (1998) tested for such a link in the New Zealand experience but found a much weaker correlation.

2.4 Results:
While some authors attempted to estimate the road trauma saved in relation to the publicly stated goals of the SRSP, other authors were content with testing a significant impact, regardless of magnitude.

In the absence of statistical tests, by comparing the 1995/1996 road trauma statistics with that for 1994/1995 Cameron and Vulcan (1996) concluded that the SRSP had a significant effect in its first year. The 12 month toll had fallen by a significant 11.5% as at December 1996, though the authors did not consider the 0.6% reduction in serious injuries to be significant.

In contrast, Macpherson and Lewis (1998) questioned the attribution of the 1995/1996 reduction, citing the pre-existing downward trend in the road toll prior to the SRSP’s implementation in 1995. Acknowledging that the road toll had fallen, the authors cited the 14% increase in the number of accidents, and the stable number of EBTs as proof of the campaigns ineffectiveness.

Results from Cameron and Vulcan (1998) were consistent with a road safety advertising campaign requiring time to develop and become effective. The first year of the SRSP was most effective in urban areas, while in the second year the strongest effect was found during high alcohol hours in both urban and rural areas. The authors estimated that a total of 281 and 771 serious casualty crashes were saved in the first and second years respectively, corresponding to 352 and 996 serious casualties.\footnote{This calculation is based on the overall average of 1.253 serious casualties per serious casualty crash during 1990-1997.}

Tay (2001) concluded that the SRSP was effective during the first two years. While the CBT programme alone did not have a significant impact on the trend in road trauma, when combined with the presence of the SRSP campaign there was a significant negative effect.

The Poisson marginal effects found by Tay (2003) implied that a one unit increase in ‘Adstock’ results in a decrease of 0.02 fatal crashes per month; i.e., that one crash per month is saved for every 50 Adstock. Similarly, one fatal crash per month is saved
for every 6,600 breath tests. Contrary to the expectation of advertisings’ supportive role in road safety campaigns, the advertising campaign was found to have a significant positive direct effect.

Guria and Leung (2003) found the SRSP to be effective in reducing road trauma, estimating 285-360 fatalities and 1,700 serious injuries saved during the five years since its implementation. However, the authors concluded that no definite conclusion could be drawn on the individual contributions of enforcement and advertising during the campaign.

Results from Tay et al (2000) indicated that the SRSP had a significant effect for all male age groups 15-55 years of age. Except for females ages 25-34, the SRSP was not found to have an impact on female drivers (or males aged 55 years plus). These results are in line with the authors’ hypothesis, as the SRSP television campaign was targeted at young male drivers.
3. ANALYSIS OF COUNT DATA

3.1 The Poisson Regression Model:
The Poisson regression model is considered the benchmark for count data in the same way Ordinary Least Squares is for real-valued continuous data. Data on count variables are of a discrete, strictly non-negative nature, and have inherent heteroskedastic error.

Due to these characteristics a linear specification for the conditional mean $E[y_i | x_i]$ is not appropriate. Further, we are generally unable to take the logarithmic transformation of count data, due to observations being able to take the value zero. Consequently, an exponential form is adopted, also serving to ensure predicted values of the conditional mean are always positive:

$$E[y_i | x_i] = \exp(x_i \beta), \quad (1)$$

where $y_i$ is count data taking the values $i = 1, 2, \ldots, n$, $\beta$ is the vector of regression coefficients, and $x_i$ is a matrix of covariates for subject $i$.

As the above specification is nonlinear in parameters, we are unable to use Ordinary Least Squares estimation, and Nonlinear Least Squares, while possible, is inappropriate due to the inherent heteroskedastic error of count data. Estimation of (1) requires Maximum Likelihood Estimation (MLE), under the assumption of a Poisson distributed error. The central assumptions underlying the Poisson distribution are that of independent events, but more characteristically equality of the mean and variance:

$$\text{Var}(y_i | x_i) = \lambda = E(y_i | x_i), \quad (2)$$

where the parameter $\lambda$ is equal to both the mean and variance of the distribution. All the probabilities and higher moments of the distribution are determined entirely by the mean. This assumption is certainly restrictive and is commonly violated by real world data. Such violations of this condition result in either underdispersion or overdispersion, where the variance is less than or exceeds the mean respectively.

12 Adapted from Cameron and Trivedi (1998) and Winkelmann (2003).
If we choose to maintain the Poisson model given a violation of this central assumption, the standard errors require adjustment. Suppose instead of mean and variance equality we have:

$$\text{Var}(y_i | x_i) = \alpha \times \lambda = \alpha \times \exp(x_i' \beta),$$  \hspace{1cm} (3)

with the data being underdispersed if $\alpha < 1$ or overdispersed if $\alpha > 1$. After estimation of $\beta$, an estimate of $\alpha$ is obtained and the standard errors are corrected by division by $\sqrt{\alpha}$. As count data commonly exhibit overdispersion, failure to correct standard errors results in erroneously large t-statistics and hence invalid hypothesis testing.

While this Poisson assumption is unlikely to hold in practice the Poisson estimator of $\beta$ is robust, assuming the mean function $\lambda$ has been correctly specified, even if the Poisson assumption is not appropriate. Hence assuming an independent sample and correct specification of the conditional mean, parameters within a Poisson regression can be estimated via Maximum Likelihood Estimation (MLE) with the same desirable properties of other MLE estimators – consistency, asymptotically efficiency and normality.

Due to exponentiation of the model, interpretation of the coefficients differs from that in OLS since:

$$\frac{\partial E[y_i | x_i]}{\partial x_{ji}} = \exp(\beta_i + \beta_2 x_{j2} + ... + \beta_k x_{jk}) \times \beta_j = E[y_i | x_i] \times \beta_j,$$  \hspace{1cm} (4)

$$\beta_j = \frac{\partial E[y_i | x_i]}{\partial x_{ji}} \times \frac{1}{E[y_i | x_i]} = \frac{\partial \ln E[y_i | x_i]}{\partial x_{ji}},$$  \hspace{1cm} (5)

therefore, $\beta_j$ is the semi elasticity of $E[y_i | x_i]$ with respect to $x_{ji}$. A one unit change in variable $j$ will result in a change of the conditional mean by $E[y_i | x_i] \times \beta_j$ percent. If a regressor has been transformed by the natural logarithm, $\beta_j$ is then an elasticity.

Clearly, unlike the standard regression model, the marginal effects (partial derivatives) of the explanatory variables have to be computed. It is common practice to calculate these at the predicted count (conditional mean) of the dependent variable, using equation (4) above where the predicted count is given by

$$r = e^{\beta_0 + \beta_1 x_1 + ... + \beta_k x_k}.$$  \hspace{1cm} (6)
3.2 The Negative Binomial Model:
An extension of the Poisson model, The Negative Binomial (NB) model, is the most commonly used alternative when the strict requirements of independence of the underlying process and inclusion of all relevant regressors are in doubt. The principal advantage of the NB model over Poisson is the account taken for ‘non-Poissonness’ from overdispersed variables, though it is also more efficient than Poisson.

While there are several derivations of the NB model, Stata 8 includes two of these: mean dispersion (the default) and constant dispersion. In both models, the conditional variance is necessarily greater than the conditional mean\(^{13}\).

The relationship between the variance and mean within the mean dispersion model is given by

\[
\text{Var}(y_i | x_i) = \exp(x_i'\beta) + \sigma^2[\exp(x_i'\beta)]^2,
\]

while the constant dispersion model is given by

\[
\text{Var}(y_i | x_i) = (1 + \sigma^2)\exp(x_i'\beta).
\]

Utilising a dispersion function \(\phi_i\), such that \(\text{Var}(y_i | x_i) = \Phi_i \text{E}(y_i | x_i)\):

\[
\phi_i = (1 + \sigma^2),
\]

\[
\phi_i = 1 + \sigma^2 \exp x_i'\beta,
\]

equations (9) and (10) demonstrate the dispersion functions for the constant and mean forms respectively. In practice however, the choice between the two parameterisations is trivial as they yield similar results.

3.3 Goodness-of-Fit Tests
A misspecification is a violation of any of the three Poisson assumptions: the Poisson distributed error, independence of events, or equality of mean and variance assumptions. In practice, the variance function is a natural starting point for tests of misspecification as the equality of mean and variance assumption is the most commonly violated. For this reason, Goodness-of-Fit tests are also useful in gauging

\(^{13}\) This is in fact true for Negative Binomial models in general – under or equi-dispersed data cannot be estimated by a Negative Binomial model.
the appropriateness of the Poisson specification versus alternative count models, such as the NB model.

The Deviance test is based on the difference of two log-likelihood ratio values, equivalent to the log of the ratio and is the default within Stata 8. Under $H_0$ (if the restriction is correct) the test statistic is

$$X^2_D = -2(L_r - L_u) \sim \chi^2_{n-k}$$

where $L_r$ is the log-likelihood achieved if the Poisson model gave a perfect fit to the data, and $L_u$ is the log-likelihood achieved by the model under scrutiny. The Deviance statistic is approximately chi-squared distributed with $n-k$ degrees of freedom. A value of the deviance greatly in excess of $(n-k)$ or more practically when the Deviance statistic divided by $(n-k)$ degrees of freedom is significantly greater than one, is suggestive of overdispersion either due to omitted variables and/or 'non-Poisson’ form.

An alternative Goodness-of-fit measure, the Pearson chi-squared statistic, is defined by:

$$X^2_P = \sum_{i=1}^{n} \left( \frac{y_i - \hat{\lambda}_i}{\hat{\lambda}_i} \right)^2 \sim \chi^2_{n-k}$$

where $y_i$'s are actual values and $\hat{\lambda}_i$'s are model-predicted values, with the statistic being chi-squared distributed with $(n-k)$ degrees of freedom. Similarly to the Deviance statistic, in practice the calculated statistic is compared to $(n-k)$, and is indicative of overdispersion when the Pearson statistic divided by $(n-k)$ degrees of freedom is significantly greater than one.

Finally, the somewhat controversial group of Pseudo-R-squared Goodness-of-Fit measures are used to estimate the percentage of variation explained by the model. In contrast to the R-square within linear models, different pseudo R-squares may yield substantially different values, or even values greater than one. The specific Pseudo R-square statistic used within Stata 8 is the McFadden form, utilising log-likelihood ratios:

$$R^2_p = 1 - \frac{L_P}{L_C}$$
where \( L_F \) is the log-likelihood of the full model and \( L_C \) is the log-likelihood value of the constant-only model. However, as with other forms of the statistic, this should be interpreted with caution. In the words of Conroy (2003), on the topic of the R-square measure in general:

As a sole criterion for model selection, it should only be used when there is no-one in the office capable of formulating a theory (and the cleaners have gone home).

### 3.4 Misspecification Tests

To more formally test the appropriateness of the Poisson model against the NB alternative, we can use one of three tests: Likelihood Ratio, Wald, and Lagrange Multiplier\(^{14}\). For simplicity, we use the Likelihood Ratio test to test the equality of mean and variance assumption imposed, against the alternative that overdispersion is present. Given that the variance of the NB model alternative is equal to \((\mu + k\mu^2)\), the NB model reduces to the Poisson model when \(k = 0\). i.e.,

\[
H_0 : k = 0 \\
H_1 : k > 0
\]

With the test statistic given by:

\[
LR = -2(L_F - L_u) \sim \chi^2_{(1)}
\]  

where \( L_F \) is the log likelihood gained from the Poisson model, and \( L_u \) is the log likelihood from the Negative Binomial model. The statistic is chi-square distributed with one degree of freedom; the null hypothesis of the Poisson form being rejected if \( LR \) exceeds the critical value.

Unlike other estimation techniques however, the Poisson and NB models lack any specific diagnostic testing. While techniques such as OLS have residual tests for heteroskedasticity, and misspecification tests such as RESET, modelling with the count data models employed here relies on a robust logical framework for the model specification process. As such, the following section outlines the choice of dependent and explanatory variables to be used in estimation.

\(^{14}\) All three are asymptotic and equivalent tests.
4. MODEL FORMULATION

In the absence of an economic theory of road trauma or model diagnostic tests to prescribe a ‘correct’ empirical model, this section provides the logical arguments for the choice of variables to be included within the empirical analysis in section five.

4.1 Dependent Variable:
Macpherson and Lewis (1998) used the number of evidential breath tests as the measure of performance in their analysis, arguing that it was a more direct behavioural measure of driver behaviour. However, their concern over the “non-encompassing and non-exclusivity” argument when examining road trauma statistics was acceptably countered by Tay (1999).

While fatal crashes are subject to major annual fluctuations within smaller countries and represent only a small number of total crashes, they are considered a good approximation of the SRSP benefits given that they impose the highest economic costs on communities (Toomath, 1998). Because all fatal crashes are investigated and recorded by the Police, investigating fatal crash numbers avoids known problems of underreporting and measurement error of serious injuries and serious injury crashes. Adjusting the series for such underreporting by using the above reporting ratio estimates would implicitly assume a constant reporting rate over time (Guria & Leung, 2002).

Scuffham and Langley (2002) contended that while the public relates more to the number of fatalities, as more than one fatality may result from a single crash, the series’ variance is artificially inflated. Hospitalisations might also be used but they also suffer from this “non-independence” of events, as well as changing admission practices over time (Scuffham & Langley, 2002). Cameron and Vulcan (1998) side-stepped this issue by calculating the number of serious casualties based on the overall average of 1.253 serious casualties per serious casualty crash during 1990-1997. However, the NB model, an extension of the Poisson regression, is able to allow for

---

15 Only 47.6% of serious injury crashes and 51% of serious injuries are Police reported. Reporting ratios are for rural areas sourced from Guria and Leung (2002), as used over the decade to 1999/2000 in social cost calculations.
such overdispersion in the data and will enable more direct evaluation of the road trauma saved in respect to the SRSP's publicly stated goals: 80 fatalities and 450 serious injuries during the 4 years 1995/1996 to 1998/1999.

Hence the performance measures adopted here are the number of fatalities and the number of serious injuries within a quarter on New Zealand roads. As count variables these are initially assumed to be Poisson distributed, though a better fit is expected under the assumption of a Negative Binomial distributed error.
4.2 Independent Variables:

4.2.1 A Linear Trend

Including a trend variable is common practice in multiple regression models involving time series data, mitigating spurious correlation and serving as a surrogate for relevant variables or intervention effects on which data are either unavailable or immeasurable (Tay, 2001). Within the realm of road safety evaluation such omitted variables include improvements in road engineering and vehicle safety. However, this term will also capture the net effect of the recent ‘pre-existing’ downward trend in worldwide road trauma and motorcycle use, and the rise in overall traffic growth (Cameron & Vulcan, 1998; Guria & Leung, 2002).

While Cameron and Vulcan (1998) highlighted the value of including a time trend, they chose to exclude it from their model. Expecting a positive trend due to “the apparent socio-economic system in New Zealand during the 1990’s and the attendant growth in road use”, but instead finding a significant negative trend, led the authors to conclude that the negative effect found was an artefact of the introduction of Compulsory Breath Testing and the SRSP. They also argue that the inclusion of the trend offers no further explanatory power, yet the results from replication of their models by White (2000) with a trend term shows a significant impact on the results drawn.

Clearly, if all possible contributing factors are known and all required data can be sourced, the model is fully specified and there is no reason to include a trend variable. As in most areas of social sciences there is no firm economic theory to dictate which explanatory variables should appear in a causal model of fatal crashes, hence a linear time trend is included within the analysis to capture factors contributing to the number of crashes not otherwise explicitly controlled for, the net of which is expected to be negative (Tay, 2001).
4.2.2 Seasonal Dummy Variables
Inspection of Figures 1 and 2 above reveals a distinct seasonal pattern in road trauma. Three quarterly dummy variables are included to control for the average seasonal effect of each quarter, with quarter one being the 'control'. While past evaluations have found only the fourth quarterly dummy significant, we have no reason to assume this will be the case when adopting estimation techniques other than OLS.

4.2.3 An Economic Indicator
The quarterly number of new cars registered has been included in previous evaluations to represent the injection of new cars into the fleet and to proxy for overall economic conditions – accounting for variations in discretionary travel and fluctuations in road trauma levels over and above those due to increases in the total size of the vehicle fleet (Cameron and Vulcan, 1998). This variable also captures the sharp increase in imported used cars following the abolition of excise duties in July 1990 which lowered the price of domestic used cars. This increased the exposure level (and hence crash risk) and the overall size of the vehicle fleet on the roads. On the other hand, the increase in imported used cars may have resulted in the replacement of older vehicles, serving to improve the overall safety level of the vehicle fleet. Consequently, including the number of new vehicle registrations would account for the overall effect of the higher level of interaction between vehicles and the overall quality of the vehicle fleet (Guria & Leung, 2002).

Figure 3: The number of new cars registered in NZ 1990-2003
The unemployment rate is also a candidate proxy for economic conditions. Increased unemployment is thought to reduce both the ability to pay for travel, and the demand for employment related travel and discretionary trips (Cameron & Vulcan, 1998; Scuffham & Langley, 2002). While suggestive of a positive relation between the unemployment rate and road trauma, there is also an income effect on roading and vehicle quality. As a result, inclusion of the unemployment rate would represent the net effect of the increase in travel and the income effect on roading (Guria & Leung, 2002).

However, Scuffham and Langley (2002) found that the actual level of unemployment has no significant effect of the number of crashes. Changes in unemployment lagged by 9-12 months were found to have a significant effect on crashes, suggesting that unemployment contributes to short run dynamics, with a much less significant long run role. In line with these findings, Tay (2001) found that models using the number of new cars as the economic variable performed consistently better in terms of explanatory power than those using unemployment.

Guria and Leung (2003) suggested that analysing the raw number of serious casualties or deaths over time is not informative due to increases in traffic growth. The level of road risk could be decreasing over time due to road safety interventions, yet increased traffic volume could outweigh any safety effect realised resulting in an increase in road deaths or accidents. While normalisation of the dependent variable by a measure of traffic volume can account for this, we implicitly assume a constant risk per
kilometre travelled\textsuperscript{16}. Further, we are not able to include the TVI as an explanatory variable due to its high collinearity with time.

On the contrary, earlier authors have argued that since the TVI has an almost perfect linear trend and does not display any seasonal variation, it does not offer any substantial additional explanation of serious injuries fluctuations, and need not be included in models of road trauma (Cameron & Vulcan, 1998).

![Figure 5: Total vehicle kilometres travelled 1979-2000](image)

The number of new cars registered will be employed as the economic variable, given that previous studies found it to out-perform the unemployment rate as the proxy for economic conditions (Tay, 2001). A positive coefficient is expected. Also, while the influence of traffic volume on the level of road trauma is accepted, it is not considered for inclusion within the models for three reasons: no consistent quarterly measure is available over the sample period, the near-perfect linearity offers no explanatory power of quarterly fluctuations, and inclusion of a linear time trend is expected to acceptably capture the net effect of this and other immeasurable influences varying systematically over time.

4.2.4 Alcohol Available for Consumption
The influence of alcohol on the frequency of road trauma is a well documented and accepted fact\textsuperscript{17}. The amount of 'pure' alcohol available for consumption (in million litres) is included to control for the seasonal effect of alcohol on road related trauma.

\textsuperscript{16} Guria and Leung (2003) use Transit New Zealand's annual TVI index within their annual models, and construct a quarterly index using data from Annual Average Daily Traffic (AADT) counts from telemetry sites.

\textsuperscript{17} See Cameron and Vulcan (1996).
over and above the seasonal fluctuations captured by the quarterly dummy variables. A positive effect on road trauma is expected.

![Alcohol Available for Consumption](image)

**Figure 6:** Total alcohol available for consumption 1990-2003

4.2.5 The Speed Camera Programme

Capturing the effect of the speed camera programme via a dummy variable measures only the average effect over the period examined. Further, given the introduction of the CBT and speed camera programmes in quick succession, separating out the respective intervention effects would be statistically difficult, as found in previous evaluations. Rather than including a single dummy to represent the onset of both, the quarterly number of speed camera tickets issued will be used to control for variations in speed camera enforcement, with the expectation of a negative coefficient.

![Speed Camera Tickets Issued](image)

**Figure 7:** Speed camera tickets issued 1990-2003

4.2.6 The Compulsory Breath Testing Programme

Like the speed camera Programme, controlling for the effect of the CBT programme via a dummy variable is not the preferred approach when there is enforcement data available. Following the lead of Tay (2003), the number of quarterly CBTs will be
used to control for the effect of the CBT programme, with the expectation of a negative coefficient.

![Compulsory Breath Tests Performed](image)

**Figure 8: Compulsory breath tests performed 1990-2003**

### 4.2.7 Advertising – Supplementary Road Safety Package

As outlined by Tay (2003), testing the impact of the “shock” advertising style of the SRSP requires controlling for the level of advertising both before and during the campaign’s implementation. Previous authors have achieved this via an Adstock variable derived from TARPs, or the Real Expenditure on Safety and Promotion. Since data on TARPs has only been regularly collected since the imposition of the SRSP, they are unable to control for the level of advertising over the whole sample period. Real Expenditure on Safety and Promotion, while only available as an annual measure, includes a broader definition of advertising than television advertisements alone and is available over the entire period examined. Hence, the Real Expenditure on Safety and Promotion is included as the measure of advertising intensity as a yearly covariate, with the expectation of a negative effect on road trauma. While requiring the assumption that expenditure is spent evenly over a year, it serves to capture the effect of all types of advertising included within the SRSP campaign.

![Real Expenditure on Safety and Promotion](image)

**Figure 9: Real expenditure on safety and promotion 1990-2003 (1999 dollars)**
Having controlled for the level of advertising via Real Expenditure on Safety and Promotion, the average ‘level’ effect of the SRSP’s change in advertising style is captured via a dummy variable. A negative coefficient is expected on both advertising expenditure and the SRSP dummy variable.

4.3 Model Structure

The dataset employed is quarterly and covers the period of September 1990 to March 2003 – 51 periods. The Poisson Regression model will be adopted (via Maximum Likelihood Estimation) as a starting point for evaluation. Estimations are performed using Stata 8, with the models taking the following form:

\[ E[y_i | x_i] = \exp(x'_i \beta) \]  

(15)

where:

- \( E[y_i | x_i] \) is the conditional mean of the serious injury or fatality series,
- \( \beta \) is a vector of coefficients on the independent variables \( x_i \):
  - \( Time \) – a linear time trend
  - \( Q2, Q3, Q4 \) – dummy variables taking the value 1 for the second, third and fourth quarter respectively
  - \( Alcavail \) – the total volume of alcohol available for consumption (million litres)
  - \( Newcar \) – the number of new cars registered
  - \( Sctick \) – the number of speed camera tickets issued
  - \( Cbts \) – the number of Compulsory Breath Tests performed
  - \( Realad \) – annual Real Expenditure on Safety and Promotion in 1999 dollars
  - \( SRSP \) – a dummy variable taking the value one for a period where the SRSP campaign is implemented.

However, given the restrictive nature of the assumptions underlying the Poisson model, the Negative Binomial model is expected to be more appropriate in practice.
5. RESULTS

The fatality and serious injury histograms above indicate that both variables are skewed to the left. Consequently, estimation via OLS will result in a non-normal error term and misleading inference, as standard F and t tests are based on the normality assumption (D. Owen, personal communication, October 7, 2004)

The Poisson model is accepted as the starting point for analysis of count variables, characterised by the assumption of equal mean and variance of the dependent
variable. To informally gauge the suitability the Poisson model, we first calculate the summary statistics of the variables.

Table 1: Summary statistics of serious injuries

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Smallest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>520</td>
</tr>
<tr>
<td>5%</td>
<td>552</td>
</tr>
<tr>
<td>10%</td>
<td>576</td>
</tr>
<tr>
<td>25%</td>
<td>625</td>
</tr>
<tr>
<td>50%</td>
<td>690</td>
</tr>
<tr>
<td>75%</td>
<td>820</td>
</tr>
<tr>
<td>90%</td>
<td>970</td>
</tr>
<tr>
<td>95%</td>
<td>1021</td>
</tr>
<tr>
<td>99%</td>
<td>1118</td>
</tr>
</tbody>
</table>

The fatality summary statistics show very strong evidence for overdispersion in the serious injuries series, with the variance being 29 times larger than the mean.

Table 2: Summary statistics of fatalities

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Smallest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>86</td>
</tr>
<tr>
<td>5%</td>
<td>100</td>
</tr>
<tr>
<td>10%</td>
<td>108</td>
</tr>
<tr>
<td>25%</td>
<td>117</td>
</tr>
<tr>
<td>50%</td>
<td>136</td>
</tr>
<tr>
<td>75%</td>
<td>151</td>
</tr>
<tr>
<td>90%</td>
<td>165</td>
</tr>
<tr>
<td>95%</td>
<td>171</td>
</tr>
<tr>
<td>99%</td>
<td>199</td>
</tr>
</tbody>
</table>

The variance of fatalities is nearly four times larger than its mean, which indicates the series may be overdispersed. From this informal evidence, we do not expect the Poisson model to fit either dependent variable very well, but will estimate it for comparative purposes. For more insightful interpretation of some of the coefficients, in addition to estimating the coefficients (semi-elasticities), we also calculate the marginal effects at the predicted count of the dependent variable.

Initial estimation revealed that the Real Expenditure on Safety and Promotion was not significant at any reasonable level, and was consequently dropped from the final model. While it may be conceivable that the level of advertising expenditure has no significant effect on road trauma outcomes, the insignificance may be more plausibly explained by its inclusion in a quarterly model as a yearly covariate. However even if quarterly expenditure data could be sourced, advertising expenditure will vary for reasons other than intensity: by year, month, programme, contract and competitive
pressures. The Real Expenditure on Safety and Promotion does not take these factors into account and will flatten the real levels of advertising exposure (T. Macpherson, personal communication, September 28, 2004). In addition, the theoretically ‘better’ Adstock measure employed by Macpherson and Lewis (1998), is not able to control for advertising intensity over the full period of analysis, as efforts to measure TARPs (on which the Adstock measure is based) at regular intervals were only made following the implementation of the SRSP.

Table 3: Results for serious injury model

<table>
<thead>
<tr>
<th></th>
<th>Poisson</th>
<th>Negative Binomial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Marg. Eff.</td>
</tr>
<tr>
<td>Time</td>
<td>-0.013348</td>
<td>-9.661853</td>
</tr>
<tr>
<td>Q2*</td>
<td>-0.101975</td>
<td>-71.87575</td>
</tr>
<tr>
<td>Q3*</td>
<td>-0.171927</td>
<td>-119.4545</td>
</tr>
<tr>
<td>Q4*</td>
<td>-0.504237</td>
<td>-325.969</td>
</tr>
<tr>
<td>Alcavai</td>
<td>0.180685</td>
<td>130.7818</td>
</tr>
<tr>
<td>Newcar</td>
<td>0.000000467</td>
<td>0.0033801</td>
</tr>
<tr>
<td>Cbs</td>
<td>0.000000036</td>
<td>0.0002643</td>
</tr>
<tr>
<td>Sctick</td>
<td>-0.0000012</td>
<td>-0.0008635</td>
</tr>
<tr>
<td>SRSP*</td>
<td>-0.0555728</td>
<td>-40.4710</td>
</tr>
</tbody>
</table>

Log Likelihood = -286.4884
Pseudo R-square = 0.7056
Goodness-of-Fit Chi-square = 143.54
Prob > Chi-square (41) = 0.000

Log Likelihood = -266.6223
Pseudo R-square = 0.1821
LR test of a = 0:
Chi-square (1) = 39.73
Prob > Chi-square = 0.000

Note:
Estimated with constant, but not reported.
1%, 5%, 10% significance indicated by a, b, c respectively.
* Marginal effect is for a change of dummy from 0 to 1.

Of the Poisson estimates, eight are significant at the 1% level of significance, with the SRSP dummy being significant at the 1.9% level. By themselves these results would suggest the Poisson model explains the serious injury series well. However, both the Goodness-of-fit Chi-square (Deviance) statistic and the Likelihood Ratio test of the significance of the overdispersion parameter α in the Negative Binomial model conclude in favour of the Negative Binomial alternative being more appropriate.
The significance of the SRSP dummy variable is the most notable difference between the Poisson and Negative Binomial estimations. The p-value for the SRSP dummy variable is 0.019 in the Poisson model. However, when the effect is estimated via the Negative Binomial model, taking the overdispersion into account; the SRSP dummy variable is significant only at levels of significance greater than 12.3%. This illustrates the consequences of failing to correct the Poisson standard errors in the presence of overdispersion – while the Poisson estimates are consistent, the p-values are erroneously small (i.e., large t-statistics) and hypothesis testing is invalid. Adopting a Poisson model, such as above, to model the overdispersed serious injury series, the researcher would wrongly conclude that the SRSP dummy variable is statistically significant.

More formally testing the appropriateness of the Poisson model against the Negative Binomial model, we employ the Likelihood Ratio-test as outlined in section three:

\[ LR = -2(-286.4884 + 266.6223) = 39.73 \]

which given the 5% chi-square critical value of 3.84, rejects the null hypothesis concluding in favour of the Negative Binomial model. Hence even though the Poisson estimates are consistent given the overdispersion, our interpretations will focus on the Negative Binomial estimates to facilitate valid hypothesis testing.

As mentioned in section three, the lack of specific diagnostic tests for the Poisson and Negative Binomial model means we are unable to test for ‘model misspecification’ as such. Compounded by the lack of a formal casual theory of road trauma, the above model is constructed via a robust logical framework.

The trend and quarterly dummies are all plausibly negative and significant at the 1% level. The significant time trend confirms the influence of immeasurable and/or omitted effects on the level of serious injuries, such as road engineering improvements. In contrast to previous studies the fourth quarter dummy is negative, implying that ceteris paribus and on average, the December quarter has 50% less serious injuries in comparison to the first quarter. While this may be picking up the effects of the pre-existing trend in the serious injury series or the safety effect from increased CBT and speed camera enforcement during the Christmas holiday season,
perhaps after controlling for the other factors, the first quarter is indeed the riskiest time of the year on New Zealand roads. On reflection this result is not wholly counterintuitive given the traditional New Zealand Christmas summer holiday season continues over much of January and February. Ceteris paribus and on average, serious injuries are 10% and 17% lower than the first quarter for quarters two and three respectively.

Both the amount of alcohol available for consumption and the number of new cars registered are significant and positive as expected. An increase in the volume of total alcohol available by one million litres increases the number of serious injuries by 18%. From the calculated marginal effects of speed camera tickets, we can infer that one additional speed camera ticket decreases the predicted number of serious injuries in a quarter by 0.0008763; i.e., one serious injury is saved each quarter for every 1,141 speed camera tickets issued.

Contrary to expectation, the number of CBTs performed has a positive and significant marginal effect. The marginal effect of an increase in one CBT per quarter increases the number of serious injuries 0.0002676; i.e., one serious injury is related to every 3,736 CBTs performed. While this may at first seem counter-intuitive, the positive correlation can be explained by the fact that CBT enforcement by the Police is typically higher in periods where road trauma outcomes are expected to be poor such as holiday weekends or the Christmas period. Thus, while the coefficient reveals a positive correlation, this does not imply causality; i.e., the ineffectiveness of the CBT programme.

As mentioned previously, the SRSP dummy is significant only at the 12.3% level of significance. While statistically insignificant at any reasonable level, interpretation of the coefficient and marginal effect respectively would imply that the introduction of the shock advertising campaign resulted in an average ‘level effect’ decrease of 6.22% (or 45.33) in the predicted count of serious injuries. In comparison to the publicly stated goal of saving 450 serious injuries during the 4 years 1995/1996 to1998/1999, the estimated effects imply that the SRSP easily exceeded its goal, saving 725 serious injuries. However, it is important to stress that the above
interpretation has no statistical basis as the coefficient is insignificant at any reasonable level.

Attempting to gauge the effectiveness of the SRSP in saving fatalities, we estimate a model with the number of fatalities as the dependent variable\textsuperscript{18}. However, the results from the fatality model are much less illuminating as only the time trend has a statistically significant effect. While the Goodness-of-fit chi-square (Deviance) statistic fails to reject the null hypothesis of the Poisson form at any significance level less than 38%, the coefficients of the model are largely insignificant. Given the non-rejection of the Poisson form, no statistical evidence of overdispersion is found and estimating the Negative Binomial model is not appropriate.

As highlighted by Guria and Leung (2002), the poor explanatory power of the quarterly fatality model may be due to the stochastic nature of the fatality series: i.e., too few fatalities occurring within a small observation period (i.e., a quarter) relative to random variation to statistically detect the intervention effects. While Guria and Leung (2002) adopted the number of serious casualties (fatalities plus serious injuries) as their quarterly dependent variable to mitigate this problem, that will not be pursued here. We have already successfully modelled the serious injury series, and the marginal effects from such an aggregated model (a ‘serious casualty’ saved) are not informative with respect to the SRSP’s goals.

\textsuperscript{18} See the Appendix for the Stata 8 output
6. CONCLUSIONS:

Though the Supplementary Road Safety Package consisted of a move to targeted speed and alcohol enforcement, and graphic television advertising, this dissertation has not found any statistical evidence of a significant impact on the level of trauma on New Zealand roads.

The consequences of using inappropriate modelling techniques in the evaluation of road safety countermeasures have been explored. Through a review of previous literature, Ordinary Least Squares has been found inappropriate for estimation involving count data, due to heteroskedastic the error and the discrete, non-negative nature of the data. Further, while the Poisson model is considered the benchmark for analysis of count data, overdispersion commonly found in ‘real world’ data positively biases t-ratios rendering hypothesis testing invalid. Having found overwhelming evidence for overdispersion in the serious injuries series, the Negative Binomial model was deemed more appropriate for gauging the SRSP’s effect.

As expected given the consistency of Poisson in light of overdispersion, the estimates from the two models were quantitatively similar. However, while the Poisson model concluded that the SRSP campaign had a statistically significant negative impact on the level of serious injuries at the 1.9% level, the empirically more appropriate Negative Binomial model found no such effect at any reasonable level. Hence, the conclusion of this empirical investigation is that the SRSP produced no statistically significant impact on the level of serious injuries in New Zealand.
7. LIMITATIONS AND FUTURE RESEARCH:
Due to the time and space restraints of this dissertation, there are several areas for future research. Though this analysis found the Real Expenditure on Safety and Promotion to be insignificant, it does not necessarily follow that road safety advertising in general is ineffective. Firstly, its inclusion as a yearly covariate required the assumption that expenditure is spread equally throughout the year. Also, real expenditure may be a poor indicator of advertising intensity as the amount spent varies for a number of factors: year, month, programme, contract and competitive pressures. However the preferred ‘Adstock’ measure could not be employed in this investigation due to data availability issues.

Future effort into modelling the fatality series within a count regression framework is recommended. While the Poisson and Negative Binomial models estimated satisfactorily explained the serious injury series, the modelling framework performed poorly when transferred to the fatality series. As suggested by previous authors, the fatality series with such small observational periods may be too stochastic in nature for statistical techniques to distinguish the intervention effects.

While macro-level data was used in this analysis to enable comparison to previous evaluations, more micro-level evaluation is certainly possible. Crash, injury and fatality data is available in monthly aggregations by age and sex of driver, region, time of day, and severity of injury. Such micro analysis may require the use of Hurdle or Zero Inflated Poisson modelling techniques to deal with the possible excess of zero-count periods. However, such micro-level data is expected to enable investigation into the temporal safety effect of the campaign, alleviating the problem of the SRSP functions of time being unable to perform alongside a linear time trend.

The empirical model presented here is static in nature. However, modelling the time series element of road trauma makes a substantial difference to the count modelling techniques, requiring the use of specialised techniques beyond those adopted here. As a general overview, such “cross-sectional dependence”, can either be investigated via an explicit lag structure in the endogenous count variable, or a multiplicative error term that follows an autoregressive process (Winkelmann, 2003).
8. REFERENCES:


### APPENDIX:

*Output from Estimation of Serious Injury Series:*

```plaintext
. poisson sinjuries t q2 q3 q4 alcavail newcar cbts sctick srsp
Iteration 0: log likelihood = -286.4902
Iteration 1: log likelihood = -286.48843
Iteration 2: log likelihood = -286.48843

Poisson regression
Number of obs = 51
LR chi2(9) = 1371.28
Prob > chi2 = 0.0000
Pseudo R2 = 0.7056

Log likelihood = -286.48843

|      | Coef.    | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|------|----------|-----------|------|-----|----------------------|
| t    | -0.0133486 | 0.0008794 | -15.18 | 0.000 | -0.0150723 - 0.011625 |
| q2   | -0.1019751 | 0.0160616 | -6.35  | 0.000 | -0.1334552 - 0.0704955 |
| q3   | -0.1719275 | 0.0167377 | -10.27 | 0.000 | -0.2047329 - 0.1392222 |
| q4   | -0.5042376 | 0.059331  | -8.50  | 0.000 | -0.6205242 - 0.3879511 |
| alcavail | 0.1806858 | 0.0214364 | 8.43   | 0.000 | 0.1386712  0.2227004 |
| newcar | 0.467e-06  | 1.04e-06  | 4.47   | 0.000 | 2.62e-06 - 6.72e-06  |
| cbts  | 3.65e-07   | 7.39e-08  | 4.94   | 0.000 | 2.20e-07  5.30e-07  |
| sctick | -1.19e-06  | 2.91e-07  | -4.11  | 0.000 | -1.76e-06 - 6.23e-07 |
| srsp  | -0.0555728 | 0.0236001 | -2.35  | 0.019 | -0.1018281 -0.0093175 |
| _cons | 5.814306   | 0.1103083 | 52.71  | 0.000 | 5.598106 - 6.030506  |

. poisgof
Goodness-of-fit chi2 = 143.5394
Prob > chi2(41) = 0.0000

. mfx compute
Marginal effects after poisson
y = predicted number of events (predict) = 723.80779

| variable   | dy/dx     | Std. Err. | z     | P>|z|  | [ 95% C.I. ] | X  |
|------------|-----------|-----------|------|-----|-------------|----|
| t          | -9.661853 | 0.63481   | -15.22 | 0.000 | -10.9061 - 8.41765 | 26  |
| q2*        | -71.87575 | 11.02     | -6.52  | 0.000 | -93.475 - 50.2765 | 2.35294 |
| q3*        | -119.4564 | 13.153    | -9.071 | 0.000 | -141.315 - 97.5945 | 2.254902 |
| q4*        | -325.969  | 34.397    | -9.48  | 0.000 | -393.387 - 258.551 | 0.254902 |
| alcavail   | 130.7818  | 15.504    | 8.44   | 0.000 | 100.395 161.168    | 6.50098 |
| newcar     | 0.0033801 | 0.00076   | 4.47   | 0.000 | 0.001899 0.004861  | 38025.1  |
| cbts       | 0.0002643 | 0.00005   | 4.94   | 0.000 | 0.000259 0.000369  | 225719   |
| sctick     | -0.008635 | 0.0022    | -4.11  | 0.000 | -0.01276 - 0.00451 | 73109.4  |
| srsp*      | -40.47102 | 17.293    | -2.34  | 0.019 | -76.3652 - 6.57681 | 607843   |

(*) dy/dx is for discrete change of dummy variable from 0 to 1.
```
. nbreg sinjuries t q2 q3 q4 alcavail newcar cbts sctick srsrp

Fitting Poisson model:
Iteration 0:  log likelihood = -286.4902
Iteration 1:  log likelihood = -286.48843
Iteration 2:  log likelihood = -286.48843

Fitting constant-only model:
Iteration 0:  log likelihood = -387.75436
Iteration 1:  log likelihood = -326.10758
Iteration 2:  log likelihood = -325.96826
Iteration 3:  log likelihood = -325.96793
Iteration 4:  log likelihood = -325.96793

Fitting full model:
Iteration 0:  log likelihood = -303.07563
Iteration 1:  log likelihood = -272.37302
Iteration 2:  log likelihood = -269.271
Iteration 3:  log likelihood = -266.67189
Iteration 4:  log likelihood = -266.6223
Iteration 5:  log likelihood = -266.62227

Negative binomial regression
Number of obs = 51
LR chi2(9) = 118.69
Prob > chibar2 = 0.0000
Pseudo R2 = 0.1821
Log likelihood = -266.62227

|          | Coef.   | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|----------|---------|-----------|-------|-------|---------------------|
| sinjuries | -0.0132809 | 0.0014926 | -8.90 | 0.000 | -0.0162063 -0.0103555 |
| t        | -1.06292 | 0.0272547 | -3.90 | 0.000 | -1.1597102 -0.0652873 |
| q2       | -0.1772519 | 0.0280604 | -6.32 | 0.000 | -0.2323492 -0.1223546 |
| q3       | -0.5135711 | 0.100485 | -5.16 | 0.000 | -0.7121004 -0.3214523 |
| q4       | 0.1842405 | 0.0359973 | 5.12  | 0.000 | 0.113678 -0.2547945 |
| alcavail | 4.95e-06  | 0.00134  | 2.69  | 0.000 | 1.34e-06 8.58e-06  |
| newcar   | 3.70e-07  | 1.25e-07 | 2.96  | 0.003 | 1.25e-07 6.15e-07  |
| cbts     | -1.22e-06 | 4.6e-07 | -2.50 | 0.013 | -2.16e-06 -2.60e-07 |
| sctick   | -0.0622063 | 0.0401015 | -1.55 | 0.121 | -1.408037 -0.0163911 |
| srsrp    | 5.78859 | 0.1859395 | 31.13 | 0.000 | 5.424155 6.153024  |
| _cons    | -0.5996207 | 0.3088077 | -3.11 | 0.002 | -0.930759 -0.268482 |
|      /lnalpha | 3.088077 | 0.3088077 | -3.11 | 0.002 | -0.930759 -0.268482 |
| alpha    | 0.0024882 | 0.0007684 | 0.003 | 0.999 | 0.000074 0.004917 |

Likelihood-ratio test of alpha=0:  chibar2(01) = 39.73 Prob>=chibar2 = 0.000

. mfx compute
Marginal effects after nbreg
y = predicted number of events (predict)
= 723.78604

| Variable | dy/dx | Std. Err. | z     | P>|z|   | 95% C.I. | X |
|----------|-------|-----------|-------|-------|---------|---|
| t        | -9.612528 | 1.08136 | -8.89 | 0.000 | -11.732 -7.4931 |
| q2*      | -7.835346 | 18.671 | -0.41 | 0.683 | -111.427 -38.2397 |
| q3*      | -13.05657 | 18.686 | -0.69 | 0.492 | -35.646 19.534 |
| q4*      | -33.129 | 58.016 | -0.57 | 0.573 | -147.839 70.651 |
| alcavail | 13.3507 | 26.063 | 0.52 | 0.603 | 82.2677 186.434 |
| newcar   | 0.0035914 | 0.00134 | 2.69 | 0.007 | 0.000971 0.006212 |
| cbts     | 0.002676 | 0.0009 | 2.96 | 0.003 | 0.0009 0.00445 |
| sctick   | -0.009763 | 0.0035 | -2.80 | 0.005 | -0.01634 -0.00318 |
| srsrp*   | 45.33443 | 29.431 | 1.54 | 0.123 | 103.019 12.3501 |

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Appendix
Output from Estimation of Fatality Series:

```
.poisson fatalities t q2 q3 q4 alcavail newcar cbts sctick srsp
Iteration 0:  log likelihood = -192.97296
Iteration 1:  log likelihood = -192.97295

Poisson regression
Number of obs = 51
LR chi2(9) = 165.31
Prob > chi2 = 0.0000
Pseudo R2 = 0.2999
Log likelihood = -192.97295

| fatalities | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|------------|-------|-----------|------|------|----------------------|
| t          | -0.011422 | 0.0020677 | -5.52 | 0.000 | -.0154745 -.0073695 |
| q2         | 0.0014265  | 0.0378075  | 0.04  | 0.970 | -.0726748 .0755278  |
| q3         | -0.0608577  | 0.0390364  | -1.56 | 0.119 | -.1373676 .0156523  |
| q4         | 0.1447487   | 0.137652   | 1.05  | 0.293 | -.1250442 .4145416  |
| alcavail   | -0.0174481  | 0.0496104  | -0.35 | 0.725 | -.1146826 .0797865  |
| newcar     | 1.16e-06   | 2.46e-06   | 0.47  | 0.636 | -3.65e-06 5.98e-06 |
| cbts       | 4.33e-08   | 1.71e-07   | 0.25  | 0.800 | -2.91e-07 3.78e-07 |
| sctick     | 1.53e-07   | 6.70e-07   | 0.23  | 0.819 | -1.15e-06 1.47e-06 |
| srsp       | 0.014388   | 0.0549018  | 0.26  | 0.793 | -.0932175 .1219936 |
| _cons      | 5.212924   | 0.2568408  | 20.30 | 0.000 | 4.709526  5.716323 |
```

```
.poisngof
Goodness-of-fit chi2 = 42.50841
Prob > chi2(41) = 0.4059
```