MONASH

4/93

M O N A S H

U N I V E R S I T Y

ANCORA IMPARO

AUSTRALIA

MODELLING THE VICTORIAN ROAD TOLL

Tim R.L. Fry

Working Paper No. 4/93

June 1993

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 7326 0372 2

MODELLING THE VICTORIAN ROAD TOLL

Tim R.L. Fry

Working Paper No. 4/93

June 1993

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Modelling the Victorian Road Toll.

Tim R.L. Fry[1]
Department of Econometrics
Monash University
Clayton, Vic. 3168
Australia.

Abstract: In 1990, the Victorian road toll decreased 29% from the previous year to *548* fatalities at a time when the State's economy was (and still is) experiencing one of its worst recessions in history. With more government agencies assuming significant roles in the management of road safety in Victoria and investing greater amounts of resources to foster a safer road environment for its many users, it is timely and pertinent to investigate the impact of the various contributing factors on the road toll. In particular, 'what impact have the economic recession and the various road safety measures, such as speed cameras, had?' This paper reports the results of modelling movements in the Victorian road toll for particular road user groups and in total using both linear and Poisson regression models.

---

[1] I would like to thank Lorna Heiman (Monash University Accident Research Centre) for helping to compile the database, the Transport Accident Commission, VIC ROADS and the Victorian Police for supplying much of the data used in this study. I would also like to thank Max Cameron and Thorolf Thoresen for their helpful discussions. The views expressed in this paper, however, are solely those of the author.

I Introduction.

In 1990, the Victorian road toll decreased 29% from the previous year to *548* fatalities at a time when the State's economy was (and still is) experiencing one of its worst recessions in history. With more government agencies assuming significant roles in the management of road safety in Victoria and investing greater amounts of resources to foster a safer road environment for its many users, it is timely and pertinent to investigate the impact of the various contributing factors on the road toll. In particular, 'what impact have the economic recession and the various road safety measures, such as speed cameras, had?'

In this paper we use a recently compiled Victorian database to attempt to model the determinants of monthly fatalities on Victorian roads between January 1985 and December 1990. We study a disaggregation of the total fatalities (or toll), namely the three road user groups: Vehicle Occupants, Pedestrians and All Other Users. We also model the total toll by using the "adding-up" of the three components (user groups) modelled.

Two types of statistical model are utilised. In common with much of the previous literature, we use linear regression models based upon the Normal distribution. However, recognising that the variables to be modelled are, by their nature, counts of a given event in a fixed time period (i.e. number of fatalities in the month) we also consider the use of models based upon the Poisson distribution (see Cameron and Trivedi (1986) for further details of such models).

The structure of this paper is as follows: in section II we consider the background to the current paper; section III describes the database used in the paper; section IV discusses the models which are estimated; the results of the estimation are presented and discussed in section V and the final section contains some conclusions and recommendations for future research.

II Background.

For some time there appears to have been a belief that there is an association between the state of the economy and the number of road accident fatalities. However, attempts to quantify such a relationship, particularly in Australia, have only appeared in recent years. In this section we review the previous literature on modelling the road toll in Australia and also consider the lessons which can be learned from these previous modelling exercises.

To date there has been very little modelling of the determinants of the road toll in Australia. This is primarily due to a lack of appropriate data. Studies have typically either been carried out using annual data (Campbell and Filmer (1981), Pascoe (1988), Thomson (1982) and Thomson and Mavrolefterou (1984)) or quarterly data for a single state (Bhattacharyya and Layton (1979) - Queensland and Layton and Weigh (1983) - Victoria). Only very recently has any modelling been carried out using monthly data for a single state - Victoria (Haque (1991) and Thoresen *et al* (1992)). Further, due to inconsistencies in reporting injuries, all of the above cited papers use fatalities or fatality rates as dependent variables to be modelled.

In using annual data the aforementioned papers have found that, in order to obtain a sufficiently large sample, data from the states and territories of Australia needs to be ‘pooled’. The exception being the study by Thomson and Mavrolefterou (1984) who use data on New South Wales. Despite having small sample sizes these models have typically used a large number of explanatory variables in a linear regression framework.

Using data for 1958/59 to 1976/77 for the six states of Australia Campbell and Filmer (1981) estimate a five equation simultaneous equation model for the four fatality rates (per vehicle kilometre) - drivers, passengers, motorcyclists and pedestrians and pedal cyclists. The fifth equation is for motor vehicle accident involvement. However, since no data is available on accident involvement this equation is substituted out of the system leaving a four equation system which is estimated by full information maximum likelihood (FIML).

The results obtained by Campbell and Filmer (1981) indicate that alcohol consumption, driver/rider experience, proportion of kilometres travelled in urban areas and a downward time trend are all important determinants of the number of fatalities. Their results on police activity and seat belt usage are however inconclusive.

Thomson (1982) is more concerned with selling the contribution that econometrics can make in the field of ‘road loss’ research. He reports the results of some preliminary work using pooled Victoria/N.S.W. data for 1952/53 to 1979/80, in which regression models are estimated for the number of fatalities for seven road user groups (male drivers under 21 years old, male drivers 21 → 29 years old, male drivers over 29 years old, all female drivers, passengers under 17 years old, passengers aged 17 → 29 years and passengers aged over 29 years). The models are estimated by OLS as single equations using Box-Cox transformations to obtain the 'appropriate' functional form.

The preliminary results indicate that random breath test activity, seat belt usage, (reduced) speed limits, the state of the economy and vehicle density (vehicles registered per kilometre of road) all have both statistically significant and expected directions of impact on the number of fatalities. The results also indicate that a linear specification is to be used. However, it is envisaged that the modelling exercise should progress to consider the use of a simultaneous equation specification, estimated by two stage least squares (2SLS) or FIML, for 13 road user groups. Unfortunately, the results of such a modelling exercise seem never to have been published.

Thomson and Mavrolefterou (1984) use annual N.S.W. data to establish relationships between the state of the economy and random breath testing activity and the number of road fatalities. A theoretical argument is given in this paper as to why increases in economic activity will increase activity on the road and hence the toll. Pascoe (1988) uses data from 1949/50 to 1984-85 and 'causality' tests to attempt to verify the link between the Australian economy (real gross domestic product) and road fatality numbers for four road user groups (drivers, pedestrians, others and all road users). His results provide weak evidence of the state of the economy influencing the number of fatalities.

Pascoe then fits single equation log-linear regression models for each of his road user groups and finds significant impacts of the state of the economy, random breath testing activity and vehicle population on the number of fatalities. Re-specifying and re-estimating using 2SLS gives a similar picture and the paper concludes with the recommendation that similar work should be considered with, at least, quarterly data.

To date two papers have appeared using quarterly data. The first by Bhattacharyya and Layton (1979) uses data on fatality numbers for Queensland from 1950 to 1976 to study the impact of the introduction of seat belt legislation in Queensland. The paper uses both Box-Jenkins techniques and a regression ('causal') model to study the impact of the legislation. Both models show that compulsory seat belt wearing significantly reduced the road toll in Queensland.

Layton and Weigh (1983), drawing on the methodology of Bhattacharyya and Layton (1979), estimate a quarterly regression model for Victorian road fatalities using data for 1961 to 1982. Their results show fuel sales, seat belt usage and random breath testing activity all have an impact upon fatality numbers.

Recently, two papers have appeared which use monthly Victorian data (Haque (1991) and Thoresen *et al* (1992)). Haque models total fatalities with both annual data (1966 to 1990) and monthly data (January 1985 to December 1990). He fits a linear regression model and finds that the state of the economy (as measured by the number unemployed), fuel sales, a 'package' of road safety measures (modelled with a dummy variable) and a strong downward trend all influence the Victorian road toll.

The lack of a conceptual framework to give structure to the statistical modelling has meant that much of the research in this field can be accused of the heinous crime of 'data mining'. The paper by Thoresen *et al* (1992) is an exception in that in their paper they exploit a conceptual model - the "Road Trauma Chain" - originally developed by Cameron (1990) to define both the structure of the model to be fitted and to aid in variable selection. Using this modelling framework they specify and estimate a multiplicative model for fatalities for each of four road user groups (Vehicle Occupants, Pedestrians, Bicyclists/Motorcyclists and Total) in Victoria.

Their resultant models are estimated by ordinary least squares (OLS) as single equations using a double-log, or constant-elasticity, specification. The data used is monthly from January 1985 to December 1990. They find that road safety 'countermeasures', such as speed camera activity and the number of random breath tests, the state of the economy, social and seasonal factors all have impacts upon the road toll for the road user groups they study.

Both the existing Australian literature and a recent survey paper by Hakim *et al* (1991) on modelling the road toll indicate strongly that when considering modelling the road toll the analysis should be carried out at as disaggregated a level as possible using monthly data. Further, and perhaps most importantly, a conceptual model should underpin the specification of the functional form of the model and the selection of appropriate explanatory variables to incorporate in the model.

From a statistical viewpoint it is also desirable to consider the use of count data models, as the dependent variable (e.g. number of monthly fatalities for a particular road user group) is, by its nature, a count (see Fridstrøm and Ingebrigtsen (1991) for further details). In addition, at the level of disaggregation and high frequency the modelling is carried out at, the dependent variable may not be able to be well approximated by a normal distribution as is done when fitting linear regression models.

For the reasons outlined in this section, this paper considers the modelling of monthly fatalities by road user group in Victoria using a conceptual model from the road safety field (Cameron (1990)). It also specifies and estimates both linear regression and count data models. The next two sections of this paper describe in more detail the data and models used.

III The Data.

The data used in this study is largely drawn from Thoresen *et al* (1992) and the reader is referred to that paper for a more extensive description of the database. However, in this paper some additional data on hours of daylight has also been utilised to model seasonal factors for certain road user groups in a more appropriate way than using the use of dummy variables to represent such factors. The sample period studied is January 1985 to December 1990. This is the longest period for which consistent monthly data on all of the variables is available[2].

The Thoresen *et al* (1992) database is a particularly rich source which includes both published data on economic and social factors and previously unpublished data on the operation of major road safety measures and also on fatalities by disaggregated road user groups. The data on the road safety measures is of particular interest as it allows us to avoid the use of dummy variables to proxy such activity.

The variables used as explanatory variables in the models in this paper fall into three groups: road safety activities, economic and social factors and seasonal factors. The road safety variables used are the number of random breath tests performed, the number of traffic infringement notices issued from speed camera activities and driver seat belt usage (%). The social and economic variables are total number (000's) unemployed, retail turnover of alcohol (80/81 $million) and the number of new probationary licences issued and the seasonal variables are seasonal dummy variables and hours of daylight.

The choice of these variables and their incorporation into the appropriate part of the model specification is described in the next section.

---

[2]Data for 1991 on several of the variables of interest is, as yet, still either unavailable or provisional and thus 1991 is not included in the sample.

IV The Models.

As in Thoresen *et al* (1992), the conceptual framework used in this paper is the "Road Trauma Chain". The use of this chain leads naturally to a multiplicative or constant-elasticity model structure. Thus continuous variables should enter the linear models in a logarithmic form. This chain also helps in the selection of appropriate explanatory variables for use in models of fatalities for each of the road user groups.

The dependent variables which we choose to model are monthly numbers of fatalities by road user group and these are plotted in Figure 1. The reasons for the choice of the monthly number of fatalities are both methodological and practical. Andreasson (1991) argues that the use of fatality rates (per capita or per kilometre) does not provide a consistent measure across time when the relationship between fatalities and the exposure factor (population or travel) is non-linear. The "Road Trauma Chain" implies a multiplicative model and thus numbers rather than rates are modelled in this paper. Further, we model fatalities rather than 'serious casualities' (fatalities + serious injuries) as the fatalities data is not subject to the definitional changes which occur with the available data on 'serious injuries'. A final reason for the choice is that both the public and policy makers seem to prefer analysis of fatalities.

The explanatory variables used in the models are chosen through use of the "Road Trauma Chain" and with particular reference to the variable selections made by Thoresen *et al* (1992). Thus for vehicle occupants the road safety variables used are speed camera activity and the number of random breath tests. The random breath test variable is lagged one month as it is believed to act as a short term deterrent. That is, an increase in the number of random breath tests in a month is hypothesised to reduce the number of fatalities in the following month. The speed camera variable has an inbuilt lag as it refers to infringement notices issued and these relate to offences already committed. Again, an increase in speed camera activity is expected to decrease the toll.

Economic and social variables used are the number unemployed and retail turnover of alcohol, both of which are hypothesised to have positive impacts upon the toll. The number of new probationary licences issued is used as a measure of inexperienced road users and also is expected to have a positive impact. Seasonal factors in this toll are modelled using seasonal dummy variables, as such seasonality is caused by a large number of factors including road conditions, weather and school and public holidays and data is not available on many of these factors.

For pedestrians and all other road users the road safety and economic and social variables used are the same as for vehicle occupants with the same hypothesised impacts. Seasonal factors in the 'all other users' toll is also modelled using seasonal dummy variables. However, seasonality in the pedestrian toll is modelled using the variable hours of daylight as one of the main seasonal factors in such accidents is visibility, with less fatalities when there is better visibility.

The omission of variables relating to road usage is perhaps a little surprising. However, the available variable (fuel sales in megalitres) is not a good measure of road usage, as fuel sales constitute an intermediate variable lying between the causes of travel and the outcome of travel, namely fatalities, modelled in this paper. It is also highly collinear with unemployment in that both reflect the 'state' of the economy. Thus fuel sales are excluded from the models presented in this paper[3].

Two sets of statistical models are used to model fatalities. As noted in section II, the dominant model in previous Australian research has been the linear regression model. Thus, in common with previous studies, we use linear regression models to model fatalities for three road user groups in Victoria. We consider two variants of the linear regression model. Namely, single equation and multiple (seemingly unrelated) equation models. However, since the data on fatalities is count data we also investigate the use of count data models based upon the Poisson distribution.

We turn now to the specification of the linear regression models. Our models can be written as follows:

$$y_{jt} = \mathbf{x}'_{jt}\beta_j + u_{jt}\ , \quad j = 1,\ 2,\ 3\ \ ; \quad t = 1,\ \ldots,\ 72\ ,$$

where j denotes road user group (vehicle occupants, pedestrians and all other users), t denotes time period (month), $y_{jt} = \log(F_{jt})$ with F_{jt} denoting the number of fatalities in month t for user

---

[3]Should an appropriate measure of road usage, say vehicle kilometres travelled, become available then an equation for this road usage should be added to the model specification. Currently, no such monthly data exists.

group j and u_{jt} is a random error term, which is assumed to follow a normal distribution. As the structure of the model implied by the chain is multiplicative the vector of explanatory variables, $\mathbf{x}_{jt}$ comprises of the logarithm of the appropriate continuous variables, such as number unemployed. The exception to this are the variables relating to speed camera activity, which has several zero observations, thus precluding the use of the logarithmic transformation and seasonal dummy variables. These variables are therefore included without transformation.

The difference between the single equation and multiple equation variants of the linear regression model fitted concerns the distributional assumption made about the u_{jt}. In the single equation variant of the model it is assumed that the u_{jt} are independently distributed as normal variables with mean zero and variance σ_j^2 (j = 1, 2, 3). In the multiple equation (or seemingly unrelated) model the $\mathbf{u}_t = (u_{1t}, u_{2t}, u_{3t})'$ follow a multivariate normal distribution with mean zero and variance-covariance matrix Σ.

As the specification of the $\mathbf{x}_{jt}$ is such that some variables are not included in all models there will be an efficiency gain in estimating the seemingly unrelated model if the disturbances, the u_{jt}, are <u>not</u> independent (i.e. $\Sigma \neq I$). This is an assumption which will be tested.

The dependent variables are the monthly *number* of fatalities. Such variables are count variables and any statistical modelling should take this into account. A natural distributional choice in modelling count data is the Poisson distribution. We therefore consider modelling the road toll for each road user group using a single equation Poisson regression model as follows:

$$P(F_j = f_{jt}) = \frac{\exp(-\lambda_{jt})\lambda_{jt}^{f_{jt}}}{f_{jt}!}, \quad f_{jt} = 0, 1, \ldots \; ; \; j = 1,2,3,$$

$t = 1, \ldots, 72$ and $\log(\lambda_{jt}) = \mathbf{x}'_{jt}\beta$. The conceptual model used tells us that the specification should be a constant-elasticity model. Thus continuous variables should enter $\mathbf{x}_{jt}$ in logarithmic form (see Gourieroux (1984)).

The Poisson distribution has two potentially restrictive assumptions underlying its use. The first of these is the assumption of a constant accident probability, independent of previous months, throughout each month. Fridstrøm and Ingebrigtsen (1991, p364) argue that this restriction may be satisfied for fatality data. The other restriction is that the expected value, conditional on the explanatory variables, equals the variance conditional on the explanatory variables. This restriction may not hold for fatality data and will be tested using regression based tests (see Cameron and Trivedi (1990) for details).

In all three model specifications (single equation linear and Poisson regression models and the seemingly unrelated regression model) we model the three road user groups indexed *j*. That is, vehicle occupants (j = 1), pedestrians (j = 2) and all other road users (j = 3). To model the total road toll for all road users we exploit the fact that these three user groups add to the total. However, unlike the economics literature on modelling demand systems, as the models for the component user groups are either log-linear or non-linear we are unable to estimate a separate model for the total category. Rather we produce predicted values and/or estimated impacts for the total through "adding-up" the component models' predicted values and/or impacts.

V Results.

The models have been estimated using the LIMDEP package (Greene (1991)). The seemingly unrelated regression (SUR) and Poisson regression specifications are estimated by maximum likelihood and the single equation linear models are estimated by OLS. The parameter estimates, standard errors, summary measures and diagnostics are presented in Tables 1 to 6.

The results are encouraging in that in all specifications (SUR, single equation and Poisson regression) the explanatory variables (probationary licences, speed cameras, seat belt usage, random breath tests, alcohol, unemployment and daylight) all have the correct sign for the estimated coefficients and, with the exception of random breath tests, similar magnitudes. Since the coefficients for all variables, except speed cameras and dummy variables, represent elasticities it is particularly reassuring that there is agreement from all three specifications on the likely impacts of these variables. We note that there is also a reasonable degree of agreement between the specifications on the statistical significance of the explanatory variables.

As noted in section IV, estimating the SUR specification allows us to test whether the single equation variant is appropriate. That is, we are able to test H_0: $\Sigma = I$ against H_1: $\Sigma \neq I$. The result of the likelihood ratio test of this hypothesis is that, at the 5% level of significance, we reject the null hypothesis of a diagonal variance-covariance matrix. Thus there is a gain in statistical efficiency in the use of the seemingly unrelated regression model.

An implication of this result is that we will not discuss our single equation models in this section. Rather, we present Tables 3 and 4 only for for reasons of completeness and because much of the previous work has used single equation models.

The SUR specification has a level of goodness of fit for each individual equation similar to that of the single equation models developed by Thoresen *et al* (1992) and an acceptable level of goodness of fit for the system as a whole. There is also no evidence of serial correlation in the model. However, the estimated skewness ($\sqrt{b_1}$) and excess kurtosis (b_2-3) in Table 2 suggest that the assumption of normality may not be appropriate.

To test whether the estimated values of skewness and excess kurtosis are significantly different from zero, the value implied by the normal distribution, we may use the asymptotic distributions given for these statistics in Harvey (1990). Unfortunately, convergence to the asymptotic distribution, especially for the kurtosis statistic, is slow. As the moments, but not the forms, of the exact distributions are known we therefore use the methodology of Evans and Fry (1992) to obtain 'approximate' critical values for the estimated skewness and kurtosis coefficients.

For a 5% level of significance these 'approximate' critical values are -0.55591 and 0.54391 for the skewness coefficient and -0.75856 and 1.21876 for the excess kurtosis coefficient. We therefore do not reject the null hypothesis of zero skewness for occupants and all other road user groups models but reject it for the pedestrian model. The null hypothesis of zero excess kurtosis is not rejected for the occupants and pedestrian models but rejected for the all other road user model. These hypothesis tests suggest that the normality assumption may be reasonable for the occupants model, that is is not reasonable for the all others model, and that it is not clear whether it is reasonable for the pedestrian model.

Using the test developed by Jarque and Bera (1987) (hereafter JB) we can test the joint hypothesis that skewness and excess kurtosis are both zero, which implies the normal distribution. This test has an asymptotic chi-squared distribution with two degrees of freedom. However, as convergence to this asymptotic distribution is also slow we use the small sample critical values given by Jarque and Bera for n = 75[4] (5% small sample critical value = 4.27). Carrying out this test, at the 5% significance level, we find that the normal distribution is not appropriate for the model for pedestrians and 'others'. These test results appear to cast doubt upon the validity of the multivariate normality assumption which underlies the SUR specification.

For reasons discussed earlier, we also consider the use of single equation Poisson regression specifications to model each of the

---

[4]The value n = 75 is the nearest to our sample size of 72.

dependent variables. The results of this analysis are contained in Tables 5 and 6. The overall fit of each of the models is assessed by a likelihood ratio test (LRT) of the hypothesis of zero cofficients (no impact) on all variables in the model excluding the constant. This hypothesis is rejected, at the 5% level, for all three user groups. Thus the models provide a statistically significant degree of explanatory power.

As mentioned in section IV, a potentially restrictive assumption of the Poisson regression model is the equality of the conditional mean and variance. To test this assumption we use the optimal regression based test of Cameron and Trivedi (1990), which is also a score test of the Poisson distribution against the alternative of the Katz system of distributions (see Lee (1986)).

Specifically, we test:

$$H_0: \mathrm{Var}(F_{jt}) = \lambda_{jt} \text{ against } H_1: \mathrm{Var}(F_{jt}) = \lambda_{jt} + \alpha.g(\lambda_{jt}),$$

with two choices (λ_{jt} and λ^2_{jt}) for the $g(.)$ function. The results of these tests, labelled CT1 and CT2 respectively, are found in Table 6. The test statistic has an asymptotic standard normal distribution and we do not reject the null hypothesis for the vehicle occupants and pedestrians models. However, we do reject the null hypothesis for the all other road users model. The results suggest that the Poisson specification may not be appropriate for the all others road user group.

Another way to assess the adequacy of the assumption of the Poisson distribution is to investigate the standardised, or 'Pearson', residuals $\hat{v}_{jt} = (f_{jt} - \hat{\lambda}_{jt})/\hat{\lambda}_{jt}$. Aitken *et al* (1989) point out that, if the Poisson regression model is well specified, these residuals will have an asymptotic standard normal distribution. We compute these residuals and test for their normality in the same way we tested for normality of the residuals in the SUR model.

First testing for zero skewness and zero excess kurtosis separately we find that zero skewness is not rejected for the pedestrian and 'others' models, but is rejected for the occupants model. The hypothesis of zero excess kurtosis is not rejected for any of the models. The JB test of the joint hypothesis of zero skewness and zero kurtosis rejects the null hypothesis for the occupants model but cannot reject the hypothesis for the other two models.

It is not apparent from the above which model specification is more appropriate and since, to our knowledge, no tests exist to discriminate between our SUR specification and our Poisson specification we can only draw the broad conclusion that neither is ideal. Thus, to attempt to distinguish between the two, we now consider how well they track the observed data and use them to estimate some impacts of interest.

Figures 2 to 5 depict the actual and predicted values from each model specification for each of the road user groups and for the aggregate of the user groups (i.e. the monthly total road toll). Overall both specifications seem to track the data fairly well. In addition, the tracking abilities of the SUR and Poisson specifications are remarkably similar and so these plots also do not enable us to discriminate between the models.

Our final tables (Tables 7 and 8) give the estimated impacts of the changes between 1989 and 1990 in the explanatory variables in the models. So, for example, the increase in speed camera activity in 1990 is estimated to reduce the vehicle occupant, and hence the total, toll by 53 fatalities with the SUR specification and by 59 fatalities with the Poisson specification. Again the models are strikingly similar in their results. The exception being random breath testing which has a very low impact in the Poisson model.

Comparing these impacts with Thoresen *et al* (1992, Table 3) we find that, with the exception of random breath tests, the impacts from our models are broadly in line with their estimated impacts. The estimated random breath test impact is surprising low and is thus of some cause for concern as it is believed that these tests are effective in reducing fatalities.

We conclude this section by re-iterating that we find the results from the two model specifications which we have fitted broadly similar, *but* neither is an ideal model specification nor can we discriminate between them. These are issues which we address in our final section.

VI Conclusions.

This paper may be viewed as representing a third attempt, following on from Haque (1991) and Thoresen *et al* (1992), to produce a model for monthly fatalities on Victorian roads. The results reported here are consistent with the earlier work in that road safety measures, social and economic factors and seasonal factors all have significant and expected impacts upon the road toll. Thus a clear pattern of results is emerging concerning factors which influence the Victorian road toll.

The models used in this paper are more sophisticated than the those used in the earlier work and may be viewed as improving the specification of an appropriate model for the Victorian road toll. Further, the SUR and Poisson specifications are easy to estimate and test in an existing econometrics package, LIMDEP, making their use in future road toll modelling work attractive.

What our results have not been able to give us is an indication of which model specification (SUR or Poisson) is more appropriate. From our results both specifications are similar but not ideal. However, given the count data nature of the dependent variables and our finding that there is a gain in estimating the regression model as an SUR we might argue that an appropriate specification is a seemingly unrelated Poisson regression model.

The issue which arises is whether further work ("part four") is warranted. We should note that any futher modelling of the Victorian road toll would also need to consider whether the use of a simultaneous equation model specification for road usage and road user group fatalities, similar to that used by Campbell and Filmer (1981), is warranted. Unfortunately, road usage data is not available at a monthly level for Victoria and seemingly unrelated Poisson specifications are not easy to estimate. Thus it is probable that this paper will represent the end of the modelling process for the monthly road toll in Victoria.

References.

Aitken, M., Anderson, D., Francis, B. and Hinde, J. (1989), *Statistical Modelling in GLIM*, Oxford University Press, Oxford.

Andreasson, D.C. (1991), "Population and Registered Vehicle Data vs. Road Deaths", *Accident Analysis and Prevention*, **23**, 343-351.

Bhattacharyya, M.N. and Layton, A.P. (1979), "Effectiveness of Seat Belt Legislation on the Queensland Road Toll - An Australian Case Study in Intervention Analysis", *Journal of the American Statistical Association*, **74**, 596-603.

Breusch, T. and Pagan, A. (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica*, **47**, 1287-1294.

Cameron, A.C. and Trivedi, P.K. (1986), "Econometric Models Based on Count Data: Comparisons and Applications of some Estimators and Tests", *Journal of Applied Econometrics*, **1**, 29-54.

Cameron, A.C. and Trivedi, P.K. (1990), "Regression-Based Tests for Overdispersion in the Poisson Model", *Journal of Econometrics*, **46**, 347-364.

Cameron, M.H. (1990), "Vehicle Crashworthiness Ratings", Preliminary Report, *Monash University Accident Research Centre.*

Campbell, R. and Filmer, R.J. (1981), "Deaths from Road Accidents in Australia", paper presented at the 10th Conference of Economists, Canberra.

Evans, M.A. and Fry, T.R.L. (1992) "Testing the Linear Regression Model Using Burr Critical Value and p-Value Approximations", in press *Journal of Quantitative Economics*, Volume **9**.

Fridstrøm, L. and Ingebrigtsen, S. (1991), "An Aggregate Accident Model based on Pooled, Regional Time-Series Data", *Accident Analysis and Prevention*, 23, 363-378.

Gourieroux, C. (1984), *Econometrie des Variables Qualitatives*, Economica, Paris.

Greene, W.H. (1991), *LIMDEP Version 6.0: User's Manual and Reference Guide*, Econometric Software, New York.

Hakim, S., Shefer, D., Hakkert, A.S., and Hocherman, I. (1991), "A Critical Review of Macro Models For Road Accidents", *Accident Analysis and Prevention*, **23**, 379-400.

Haque, O. (1991), "Unemployment and Road Fatalities", Report GR 91-10, VIC ROADS.

Harvey, A.C. (1990), *The Econometric Analysis of Time Series, 2nd. Edition*, Phillip Allan, London.

Jarque, C.M. and Bera, A.K. (1987), "A Test for Normality of Observations and Regression Residuals", *International Statistical Review*, **55**, 163-172.

Layton, A.P. and Weigh, J.C. (1983), "The Efficacy of Some Recent Australian Road Safety Initiatives", *The Logistics and Transportation Review*, **19**, 267-278.

Lee, L.F. (1986), "Specification Test for Poisson Regression Models", *International Economic Review*, **27**, 689-706.

Pascoe, M. (1988), "Road Safety and the Economy", Reference Paper 133, Bureau of Transport and Communications Economics.

Thomson, J. (1982), "Road Loss Research in Australia: The Potential Contribution of Econometrics", *Australian Road Research Board Proceedings*, **11**, 221-232.

Thomson, J. and Mavrolefterou, K. (1984), "Assessing the Effectiveness of Random Breath Testing", *Australian Road Research Board Proceedings*, **12**, 72-80.

Thoresen, T., Fry, T.R.L., Heiman, L., and Cameron, M.H. (1992), "Linking Economic Activity, Road Safety Countermeasures and Other Factors with the Victorian Road Toll", Report 29, *Monash University Accident Research Centre*.

TABLE 1: Seemingly Unrelated Regression Model Estimates.

Variable Name	Occupants	Pedestrians	Others
Constant	17.006 *(6.2004)*	10.203 *(2.1971)*	1.5654 *(2.5081)*
Probationary Licences	0.15842 *(0.10517)*		
Speed Cameras	-0.00000494 *(0.00000352)*		
Seat Belt Usage	-3.3929 *(1.4339)*		
Random Breath Tests		-0.11571 *(0.10715)*	
Alcohol	0.36223 *(0.36404)*	1.89460 *(0.55817)*	0.59518 *(0.46257)*
Unemployment	-0.20995 *(0.19478)*	-0.56499 *(0.31568)*	-0.40616 *(0.28993)*
Daylight Hours		-2.11390 *(0.49925)*	
February	-0.001239 *(0.093463)*		-0.047736 *(0.18360)*
March	0.064498 *(0.085293)*		-0.078075 *(0.17833)*
April	0.062288 *(0.093463)*		-0.510990 *(0.18670)*
May	0.059684 *(0.100760)*		-0.021088 *(0.19165)*
June	-0.085206 *(0.106120)*		-0.285170 *(0.20181)*
July	-0.167510 *(0.101010)*		-0.610710 *(0.19873)*
August	-0.139290 *(0.094074)*		-0.557540 *(0.19459)*
September	0.090270 *(0.087084)*		-0.333630 *(0.18401)*
October	0.056313 *(0.088635)*		-0.601780 *(0.18342)*
November	-0.165780 *(0.083929)*		-0.397940 *(0.18087)*
December	0.058644 *(0.115440)*		-0.416630 *(0.19809)*

Standard errors in *parentheses*.
Log-likelihood = -5.2237, system R^2 = 0.7997.

TABLE 2: S.U.R. Summary Measures and Diagnostics.

	Occupants	Pedestrians	Others
R^2	0.4874	0.3992	0.3513
s^2	0.0220	0.1309	0.0961
dw	1.8507	2.4089	1.9272
$\sqrt{b_1}$	0.4379	-0.7234	-1.0088
b_2-3	-0.4450	1.1900	3.1891
JB	2.8952	10.5280	42.7230

Estimated Variance-Covariance Matrix:

$$\hat{\Sigma} = \begin{bmatrix} 0.0220 & & \\ 0.0194 & 0.1309 & \\ -0.0021 & -0.0202 & 0.0961 \end{bmatrix},$$

Likelihood Ratio test of diagonal Variance-Variance Matrix = 9.5954 with 3 degrees of freedom.

Estimated Correlation Matrix:

$$\hat{R} = \begin{bmatrix} 1.0000 & & \\ 0.3615 & 1.0000 & \\ -0.0457 & -0.1801 & 1.0000 \end{bmatrix}.$$

TABLE 3: Single Equation Model Estimates.

Variable Name	Occupants	Pedestrians	Others
Constant	20.638 *(7.424)*	10.177 *(2.287)*	1.4728 *(2.796)*
Probationary Licences	0.12193 *(0.1269)*		
Speed Cameras	-0.00000584 *(0.00000425)*		
Seat Belt Usage	-4.0274 *(1.722)*		
Random Breath Tests		-0.11086 *(0.1154)*	
Alcohol	0.19260 *(0.4319)*	1.91190 *(0.5871)*	0.63214 *(0.5156)*
Unemployment	-0.13678 *(0.2281)*	-0.56560 *(0.3273)*	-0.42507 *(0.3227)*
Daylight Hours		-2.13070 *(0.5258)*	
February	-0.086474 *(0.11180)*		0.025633 *(0.20690)*
March	0.072207 *(0.10200)*		-0.090463 *(0.20090)*
April	0.014159 *(0.11090)*		-0.472770 *(0.20980)*
May	0.060617 *(0.11910)*		-0.026787 *(0.21500)*
June	-0.111560 *(0.12430)*		-0.268310 *(0.22580)*
July	-0.205380 *(0.11880)*		-0.575290 *(0.22270)*
August	-0.146980 *(0.11160)*		-0.549150 *(0.21870)*
September	0.095040 *(0.10370)*		-0.337000 *(0.20710)*
October	0.016683 *(0.10630)*		-0.551410 *(0.20680)*
November	-0.158610 *(0.10050)*		-0.397490 *(0.20390)*
December	0.086715 *(0.13890)*		-0.409130 *(0.22310)*

TABLE 4: Single Equation Summary Measures and Diagnostics.

	Occupants	Pedestrians	Others
R^2	0.5044	0.3993	0.3554
s^2	0.0278	0.1406	0.1185
dw	1.8579	2.4108	1.9207
BP	10.2783	25.5681	54.8177
$\sqrt{b_1}$	0.3753	-0.7294	-1.1230
b_2-3	-0.5992	1.2042	3.5003
JB	2.7673	10.7346	51.8898

Degrees of Freedom for the Breusch-Pagan (BP) test of heteroscedasticity (Breusch and Pagan (1979)) are 16, 4 and 13 respectively.

TABLE 5: Poisson Regression Model Estimates.

Variable Name	Occupants	Pedestrians	Others
Constant	21.506 *(7.087)*	9.9539 *(1.843)*	1.0530 *(2.948)*
Probationary Licences	0.12511 *(0.1223)*		
Speed Cameras	-0.00000551 *(0.00000432)*		
Seat Belt Usage	-4.1986 *(1.641)*		
Random Breath Tests		-0.00957 *(0.0928)*	
Alcohol	0.22369 *(0.4253)*	1.75290 *(0.4654)*	0.83902 *(0.5796)*
Unemployment	-0.18649 *(0.2222)*	-0.67926 *(0.2787)*	-0.53225 *(0.3619)*
Daylight Hours		-2.04680 *(0.4150)*	
February	-0.086144 *(0.11090)*		0.056615 *(0.19610)*
March	0.075016 *(0.09679)*		-0.079549 *(0.19070)*
April	0.004640 *(0.10750)*		-0.429890 *(0.22300)*
May	0.047093 *(0.11440)*		0.000532 *(0.20530)*
June	-0.121720 *(0.12240)*		-0.241150 *(0.22960)*
July	-0.225040 *(0.11860)*		-0.429890 *(0.23490)*
August	-0.156750 *(0.11130)*		-0.521990 *(0.23620)*
September	0.083834 *(0.09847)*		-0.279970 *(0.20810)*
October	-0.002561 *(0.10210)*		-0.436840 *(0.21500)*
November	-0.183100 *(0.10130)*		-0.303900 *(0.20270)*
December	0.075886 *(0.13430)*		-0.416190 *(0.22750)*

TABLE 6: Poisson Regression Diagnostics.

	Occupants	Pedestrians	Others
LRT	58.740	49.540	24.090
CT1	-1.167	0.094	-4.512
CT2	-0.915	-0.021	-4.563
$\sqrt{b_1}$	0.599	0.352	0.036
b_2-3	-0.373	0.190	1.105
JB	4.723	1.595	3.679

Degrees of freedom for LRT are 16, 4 and 13 respectively.

TABLE 7: S.U.R. Estimated Impacts of Explanatory Variables in 1990.

Variable Name	Occ's	Ped's	Others	Total
Probationary Licences	6			6
Speed Cameras	-53			-53
Seat Belt Usage	-34			-34
Random Breath Tests		-14		-14
Alcohol	-18	-28	- 6	-52
Unemployment	-29	-24	-11	-64
Daylight Hours		1		1
Total Estimated	-128	-65	-17	-210
Actual Change	-141	-66	-21	-228
Actual 1989 Level	517	159	100	776

TABLE 8: Poisson Estimated Impacts of Explanatory Variables in 1990.

Variable Name	Occ's	Ped's	Others	Total
Probationary Licences	4			4
Speed Cameras	-59			-59
Seat Belt Usage	-42			-42
Random Breath Tests		- 2		- 2
Alcohol	-11	-25	- 8	-44
Unemployment	-26	-29	-14	-69
Daylight Hours		1		1
Total Estimated	-134	-55	-22	-201
Actual Change	-141	-66	-21	-228
Actual 1989 Level	517	159	100	776

FIGURE 1: Victorian Road Toll
Monthly Fatalities by Road User Group

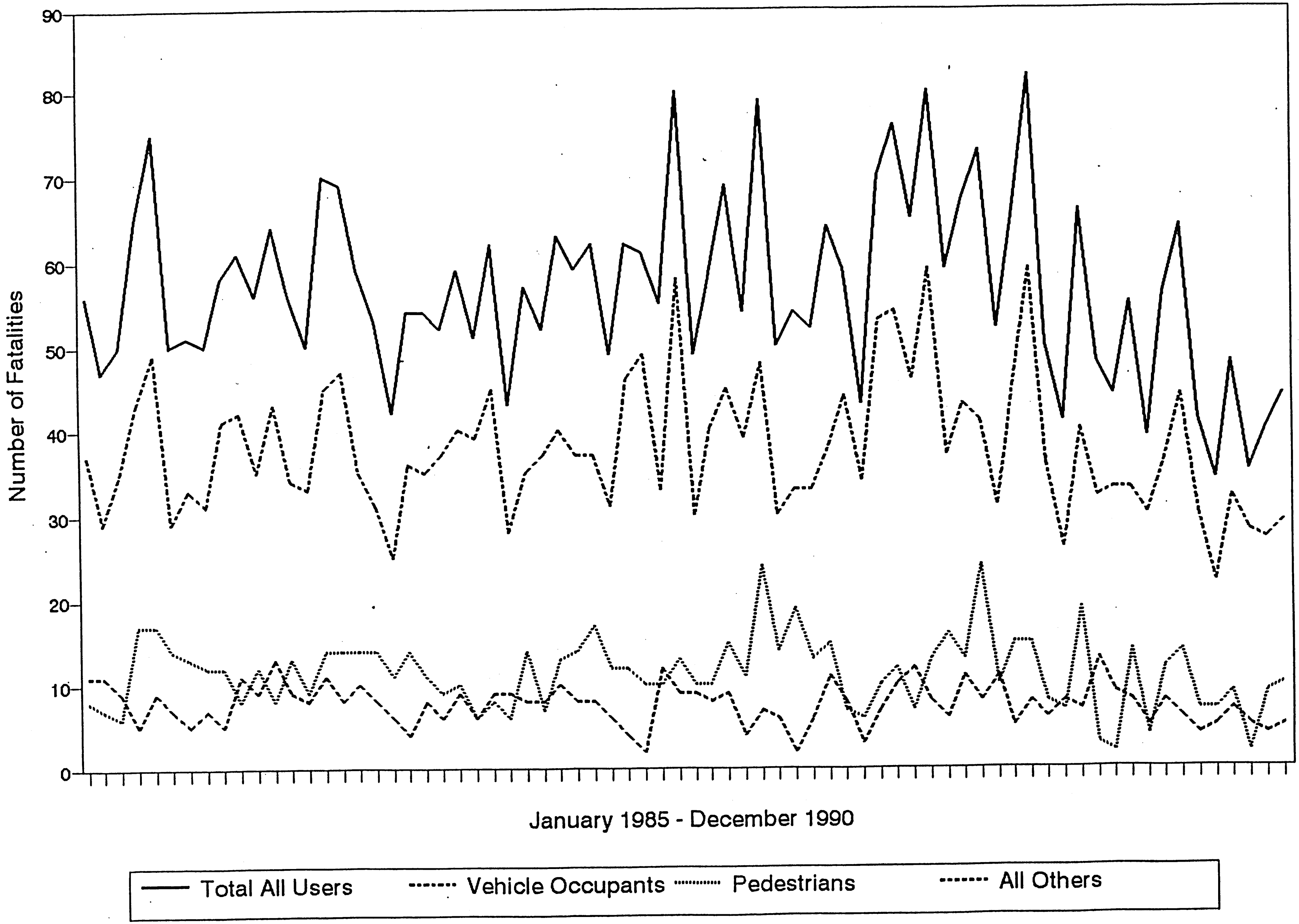


FIGURE 2: Actual v Predicted Toll

Total All User Groups

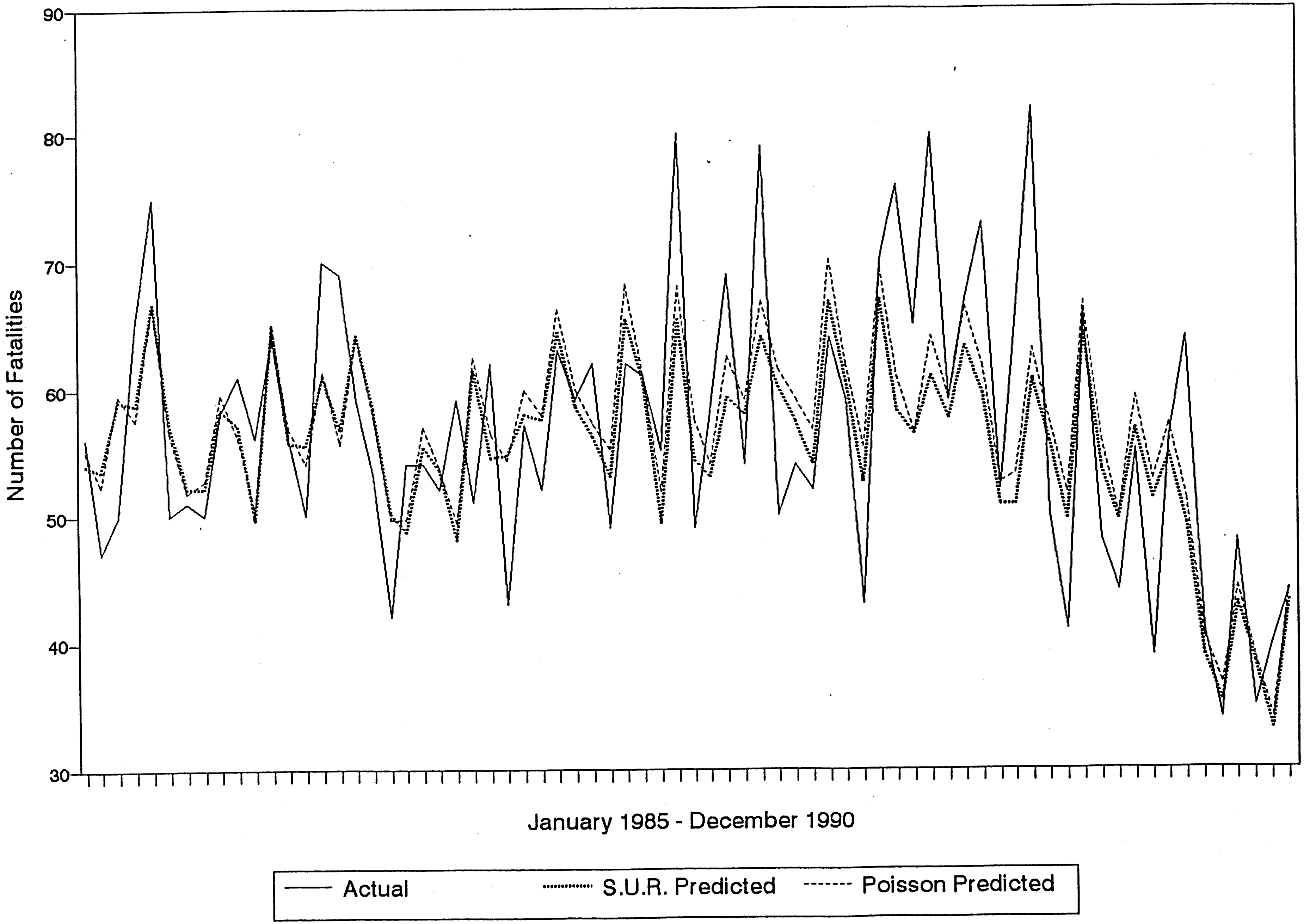


FIGURE 3: Actual v Predicted Toll
Vehicle Occupants

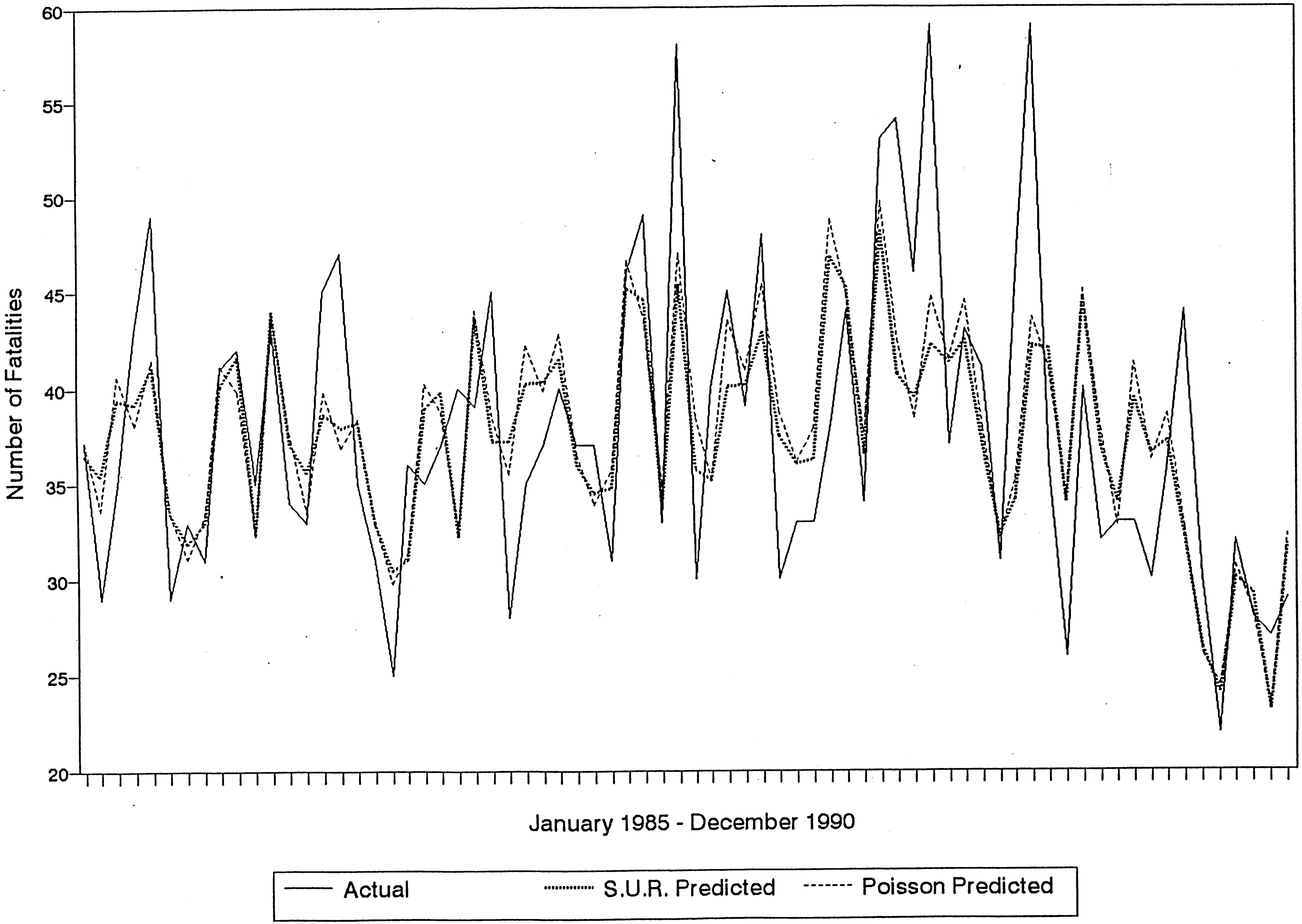


FIGURE 4: Actual v Predicted Toll

Pedestrians

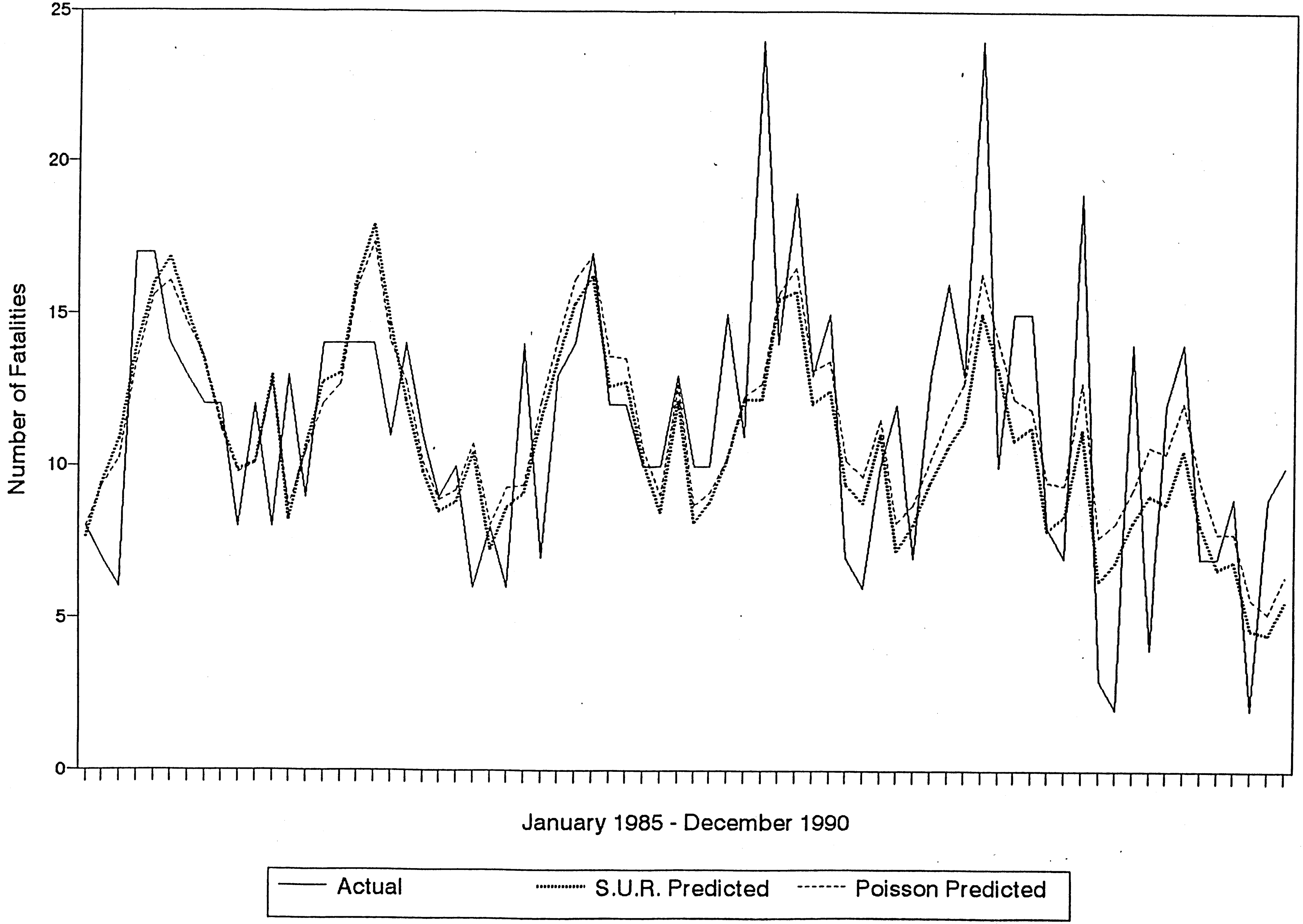
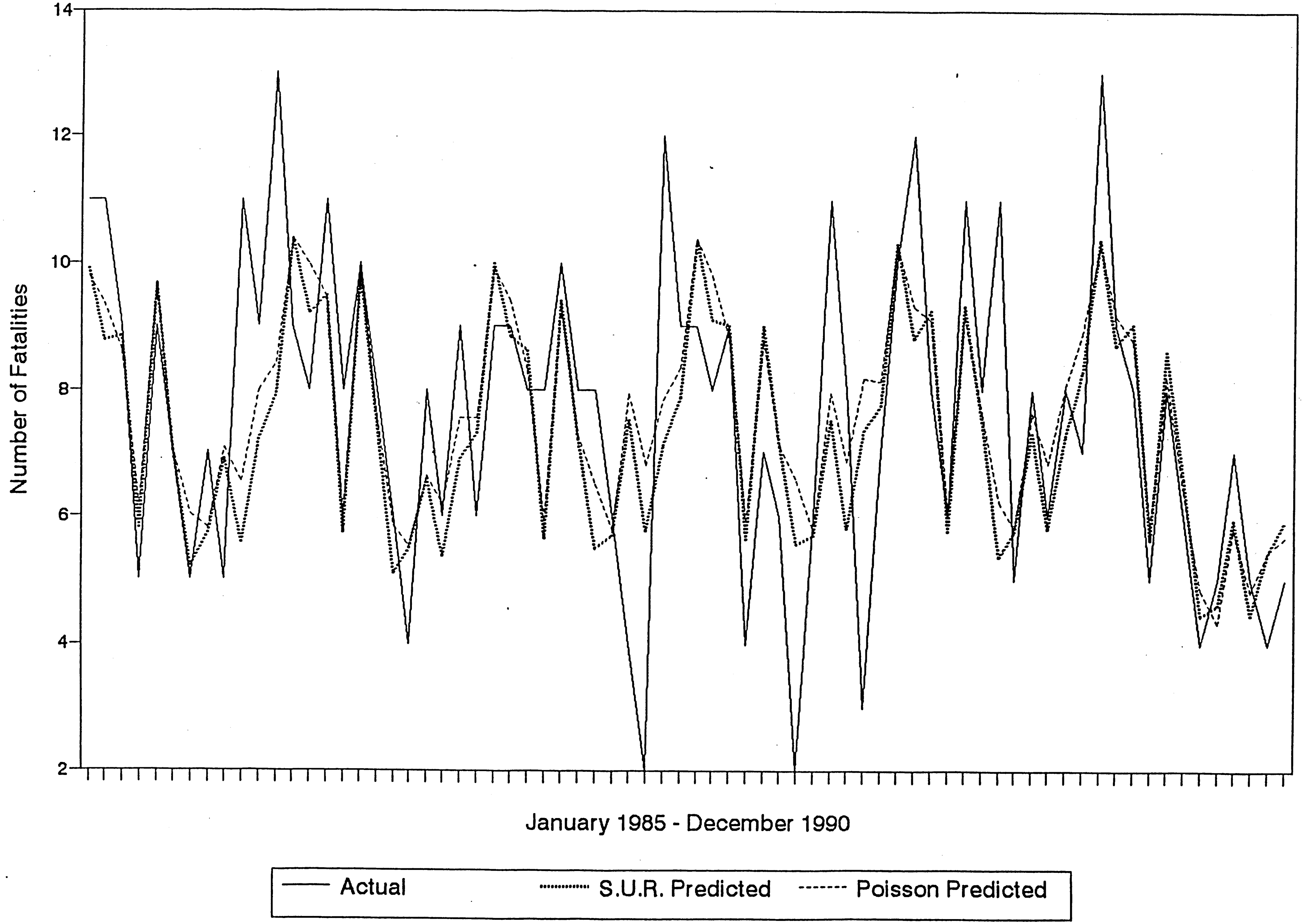


FIGURE 5: Actual v Predicted Toll

All Other Road Users



35.