Pergamon

Accid. Anal. and Prev., Vol. 28, No. 1, pp. 73–87, 1996
Elsevier Science Ltd. Printed in Great Britain

0001-4575(95)00043-7

SPEED LIMITATION AND MOTORWAY CASUALTIES: A TIME SERIES COUNT DATA REGRESSION APPROACH

PER JOHANSSON

Department of Economics, University of Umeå, S-901 87 Umeå, Sweden

(Received 16 November 1993)

Abstract—The paper deals with the effect of a lowered speed limit on the number of accidents in which there are fatalities, injuries and vehicle damage on Swedish motorways. Two models extending the Poisson and negative binomial count data models are used for estimation. The extended models account for both overdispersion and potential dependence between successive counts. The inferences of the parameters depend on the assumed form of overdispersion. It is found that the speed limit reduction has decreased the number of accidents involving minor injuries and vehicle damage. Furthermore, the models allowing for serial correlation are shown to have the best *ex ante* forecasting performance.

1. INTRODUCTION

A structural form model is developed in order to provide an explanation of the mean level of the number of casualties* on highly trafficked stretches of Swedish motorways [see Hakim et al. (1991) for a review of macro models for road accidents]. The main interest is focused on the effect on casualties of a reduction in the speed limit. In the structural model, the level of exposure is endogenous and unobservable. Two structural equations for the number of casualties and the level of exposure are formalized. A reduced form is estimated under different assumptions regarding the error terms in these two equations.

The analysis of the effects of known events (e.g. of a political decision) is called intervention analysis (Box and Tiao 1975). Examples from the road accidents field include Michener and Tighe (1992) on the effect of increasing the speed limit on fatal highway accidents, Durbin and Harvey (1985) and Harvey and Fernandez (1989) on the effect of seat belt legislation on British road casualties and Bhattacharyya and Layton (1979) on the effectiveness of seat belt legislation. It should be noted that, even if a particular intervention has a negative effect on the number of casualties, it is not always obvious that it should have been imposed. With this in mind a social cost-benefit analysis could be conducted (see, e.g. Forester et al. 1984; Kamerud 1988). Intervention analysis with time series data is generally performed with ARIMA models†, but the structural time series model approach can be used (cf. Harvey and Durbin 1986). The inference in these models relies on the assumption that data are normally distributed. Obviously count data are not normally distributed‡. For data with small counts, the analysis should be conditional on a natural distribution for counts, such as the Poisson or negative binomial (NB) distributions. These two distributions are used in Michener and Tighe (1992) and Fridstrøm et al. (1992) in modelling road accidents in the U.S. and the Nordic countries excluding Island, respectively.

Michener and Tighe (1992) studied the number of fatal accidents and claimed that, whilst the number of fatalities may not be independent events, the number of fatal accidents are. Fridstrøm (1991) claimed that the Poisson distribution is the natural distribution for accident counts. When the number of victims is the dependent variable, overdispersion (i.e. the observed variance of the count variable exceeds the variance according to the assumed theoretical distribution) must always be expected, since the number of victims are not mutually independent events. Using the models presented below, these claims of Michener and Tighe (1992) and Fridstrøm (1991) can be tested.

For cross sectional data, it is often reasonable to

*Casualties is used as a generic term for accidents in which there are fatalities, injuries (severe or minor) or vehicle damage.

†With traffic accident data many authors test proportions before and after the intervention, see, for example, Baum et al. (1990). See, however, Lloyd (1990).

‡Count data with a large mean is frequently considered as approximately normal.

think of the frequencies or counts as independent observations with a mean level that is constant or a deterministic function of explanatory variables. The traditional Poisson regression model (e.g. Hausman et al. 1984; Cameron and Trivedi 1986) and the generalized linear model (GLM) approach (see McCullagh and Nelder 1983) are of this type. If the frequencies are observed over a time period, the observations may be dependent or serially correlated. If this dependence is neglected, inferences about the estimated model may be misleading.

The introduction of serial correlation in time series models can be made in different ways. Cox (1981) conceives of two approaches—observations and parameter driven processes. Examples of observation driven processes are ARIMA models for Gaussian data and the approach of Zeger and Qagish (1988) for count data. In parameter driven models the autocorrelation is introduced through a latent process. Examples with count data are the approaches introduced by Smith (1979) and Zeger (1988).

In Section 2 the literature on road casualties is reviewed and a structural model is developed. The data are presented in Section 3. In Section 4 the Poisson model, the negative binomial model as well as the models originally suggested by Zeger (1988) and Smith (1979) are presented. In Section 5, the estimation of parameters, overdispersion and autocorrelation test, model evaluation and prediction in these models are discussed. Section 6 presents the empirical results. The final section contains a discussion of the results.

2. A MODEL OF ROAD CASUALTIES

Earlier studies (see Hakim et al. 1991) have used a great variety of variables in order to explain the number of casualties, e.g. the exposure level, socio-economic variables (e.g. consumption of alcohol, disposable income and unemployment), demographic variables (e.g. age and sex), natural phenomena* such as weather conditions and political decisions. Accordingly, this section begins by considering which explanatory variables should be included and how they effect the number of casualties. This discussion is largely based on previous empirical work. Thereafter, a new structural economic model is presented.

Exposure. Based on earlier studies (see Hakim et al. 1991), we expect that the number of casualties will increase when the exposure level (the aggregate mileage driven) increases. Fridstrøm (1991), Michener and Tighe (1992) and others used the logarithm of the exposure to obtain constant elasticity of expected casualties. Since a real exposure level cannot be obtained, various proxies have been used, for example, delivered gas to gas stations, registered vehicles and licensed drivers (see Michener and Tighe 1992 for a discussion of different exposure proxies). Fridstrøm (1991) claimed that the traffic volume depends on the speed limit and road geometry.

Economic variables. Economic conditions may be regarded as affecting traffic casualties either directly, by having an effect on the risk level, and/or indirectly by influencing the exposure level.

Both demand and supply-oriented explanations suggest a negative relationship in the long run between economic growth and the number of casualties (Hakim et al. 1991). The demand-oriented hypothesis suggests that the demand for safer cars increases with income, leading to fewer casualties. From the supply side, the money spent on road-safety features increases. However, Forester et al. (1984) claimed that the effect of income on the number of accidents cannot be determined by sign, since the value of time can be represented by real income. An increase in income would lead to a reduction in time spent driving safely and thereby increasing the number of accidents. On the other hand, if safety is a normal good, an increase in real income will increase demand for safety and hence reduce the number of casualties.

The number of work trips seems to be inelastic with respect to income, whereas young drivers' travel frequency and the number of recreational trips are quite elastic (Hakim et al. 1991). Brännäs and Laitila (1991) estimated models for travel distance. They found business travel distance to be elastic with respect to income, whereas work, shopping, visit and recreation distance are inelastic. Similar results for travel frequencies are obtained by Brännäs (1987). Hakim et al. (1991) state that the price of gasoline affects the number of trips made by high-risk groups.

Hence, the net effect of economic growth on the number of casualties is not clear. The positive effect from an increase in exposure may be offset by a negative effect from the demand and supply of safety. The inclusion of a trend variable in the model may catch the effect of a gradual change in attitudes and of the construction of safer cars and roads induced by the economic growth, yielding a positive relation between income and the number of casualties.

Natural phenomena. Since we have monthly data, seasonal variation in casualty frequencies can be studied. Fridstrøm et al. (1992) introduced seasonal dummy variables in an analysis of casualties in Sweden and claimed that the dummies "tend to

*The effect of the moon on road accidents is studied in Laverty et al. (1992).

absorb a large part of the variation attributed to weather, daylight and—in particular—exposure". Higher fatality frequencies are noted during the summer and fall and lower ones during the winter and spring periods.

Political decisions. The speed limit reduction is exogenous since it was introduced as a means of reducing the carbon dioxide (CO_2) emissions and was not due to a dramatic increase in casualties. The effect of the reduction in speed can come from two sources. The effect from an accident is not as severe compared with that from an accident at a higher speed and the number of accidents may decrease as a result of a lower risk level. The effect of the speed limit reduction on the number of accidents should be an effect of a reduction in the risk level and a test of the risk-homeostatic theory of Wilde (1982). According to this theory, individual behaviour is adjusted in response to perceived risk in an attempt to balance perceived and desired risk. The risk-homeostatic theory leads to the conclusion that the only way to have a permanent reduction in accidents is by introducing measures that alter the attitude towards risk. The speed limit reduction will have an effect only if the speed limits are obeyed. The median speed increased by 1 km/h in 1980–1985 and thereafter does not seem to have increased (TSV 1987). In 1986, two changes occured which were expected to have a negative effect on the number of road casualties. There was a sharpening in the law with regard to speeding (among other things the driving licence was automatically withdrawn if the speed limit was exceeded by 30 km/h; SOU 1991a) and it became compulsory for adults to use seat belts in the back seat. Since the increase in speed seems to be a linear function of time, it can presumably be handled by the inclusion of a trend variable.

Sociological variables. In a number of studies (see Hakim et al. 1991) consumption of alcohol has been seen to increase the frequency of fatal accidents. In a recent paper by Michener and Tighe (1992) differences in the minimum beer-drinking age in different U.S. states were not shown to have any effect on fatal accident frequencies. The Swedish alcohol policy has small regional variations and there has been no major change in the attitudes and policy towards alcohol during the time period studied. Thus, no alcohol variable is included in the model. Young drivers form a high risk group (Hakim et al. 1991) and it is of interest to test whether they behave differently than other drivers. Unfortunately no "good" data on age is available so that no such variable is included in the model.

The structural model

The functional form for the mean level of casualties can not be deduced from theory. A common representation (cf. Fridstrøm and Ingebrigtsen 1991; Michener and Tighe 1992) is to let the mean level, μ_t, at each time point t to be modelled as $\mu_t = \exp(\beta_0 + \mathbf{x}_t\beta)$, where $\mathbf{x}_t$ is a $(1 \times k)$ vector of the explanatory variables described above, β is the corresponding $(k \times 1)$ parameter vector and β_0 is an intercept.

Let the first element in $\mathbf{x}_t$ be equal to the logarithm of the exposure level, $\ln e_t$, and let the vector $\mathbf{x}_{2t}$ be the remaining variables in $\mathbf{x}_t$. The model can be written

$$\ln \mu_t = \beta_0 + \beta_1 \ln e_t + \mathbf{x}_{2t}\beta_2 \quad (1)$$

The level of exposure, e_t is an unobservable variable, and as has been mentioned above it is most likely not to be exogenous. Thus it is only possible to use proxy variables for e_t. The possible proxy variables have several limitations, for instance, delivered gas to gas stations suffers from a time lag between delivery to gas station and consumption (i.e. before a price increase on gas, gas stations bunker gas while this increase in delivered gas is not proportional to the vehicle mileage travelled). Instead of using a proxy variable for the exposure level we model it as a latent variable.

If the accident data is based on private car transportation, the exposure is equal to demand for travel. Let ordinary demand for travel be

$$\ln e_t = \gamma_0 + \gamma_p \ln p_t + \gamma_{DI} \ln DI_t \quad (2)$$

where p_t is the price for gasoline*, DI is disposable income, γ_0 is an unknown parameter and γ_p, and γ_{DI} are the unknown parameters that correspond to the gas price and disposable income, respectively†.

Variables in $\mathbf{x}_{2t}$. To capture a gradual change in the risk and the increase in speed a "linear"‡ trend, t $(t = 1, \ldots, 116)$ is included. No expected direction between the trend variable and the number of casualties can be given, since a negative and positive relation between the number of casualties and the gradual change in risk and the increase in speed, respectively, are expected. The seasonal effect is handled by using dummy variables, with December as a base. The speed limit decrease is modelled as a step intervention variable, $I_t = 0$, $t = 1, \ldots, 90$, and $I_t = 1$, $t = 91, \ldots, 116$, taking full effect directly after the speed limit decrease. The effect from the sharpening of the

*Gasoline amounts to 80% of the variable costs, see Hultkrantz (1992).

†The indirect utility function for two commodities corresponding to this demand function is given in Varian (1992, Ch. 8). It would be easy to include other variables such as weather if we think that they affect exposure, see, e.g. Hultkrantz (1992).

‡In fact, this is an exponential trend since the trend is linearly included in the exponent.

law regarding speeding and that it became compulsory for adults to wear seat belts in the back seat is also modelled as a step intervention variable, $law_t = 0$, $t = 1, \ldots, 59$ and $law_t = 1$, $t = 60, \ldots, 116$.

The model for exposure (2) is included in the mean level model (1) for accidents, so that the reduced form, with the expected positive/negative effect of the explanatory variable on the number of casualties below the parameter, can be written

$$\mu_t = \exp\Bigl(\alpha + \underset{-}{\varphi_1} \ln p_t + \underset{+}{\varphi_2} \ln DI_t + \underset{-}{\beta_2} I_t + \underset{-}{\beta_4} t + \beta_5 \, Law_t + \sum_{j=1}^{11} \delta_j S_{jt}\Bigr) \tag{3}$$

where

$$S_{jt} = \begin{cases} 1, & t = j, \quad j+12, \quad j+2\cdot 12, \ldots \\ 0, & t \neq j, \quad j+12, \quad j+2\cdot 12, \ldots \\ -1, & t = 12, \quad 2\cdot 12, \quad 3\cdot 12, \ldots \end{cases}$$

and $\alpha = \beta_0 + \beta_1 \gamma_0$, $\varphi_1 = \beta_1 \gamma_p$, $\varphi_2 = \beta_1 \gamma_{DI}$. The effect for December is

$$\delta_{12} = -\sum_{j=1}^{11} \delta_j$$

The problem with the estimable form (3) is that only the net effect of the economic variables can be obtained. It is not possible to separately identify β_1, the elasticity of expected casualties, with respect to exposure. Michener and Tighe (1992) assumes that the elasticity of expected accidents with respect to the exposure variable is one, so that the estimates may be interpreted as the elasticities of expected accidents with respect to the economic variables.

As we have omitted some variables, for instance, road geometry, speed in the demand function (2) and since a proxy for disposable income has to be used, we may include an error component ζ_t in (2), i.e.

$$\ln e_t = \gamma_0 + \gamma_p \ln p_t + \gamma_{DI} \ln DI_t + \zeta_t \tag{4}$$

The conditional mean function takes the form

$$\lambda_t = \mu_t \omega_t \tag{5}$$

where $\omega_t = \exp(\zeta_t \beta_1)$.

Further, as we have omitted variables that potentially affect the number of casualties, such as alcohol and speed variability we may introduce an additional random variable η_t. The conditional mean function then takes the form

$$\lambda_t = \mu_t \omega_t \eta_t \tag{6}$$

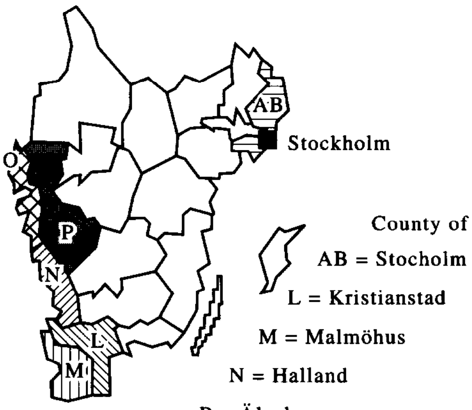

Fig. 1. The counties in the southern part of Sweden that are affected by the speed limit reduction.

3. DATA

The data set consists of monthly car accidents on motorways, January 1982 to December 1991, in the Swedish counties Stockholm, Kristianstad, Malmöhus, Halland, Gothenburg and Bohus and Älvsborg (see Fig. 1). On June 22 1989, a speed limit reduction was introduced, from 110 km/h to 90 km/h, on Swedish motorways and other major highways, yielding 90 observations before the speed limit reduction and 30 after. Since the maximum speed was 90 km/h for buses and trucks before the speed limit reduction and since cyclists, moped riders, pedestrians, etc. are not allowed on the motorways the reduction/increase in the number of car casualties should be the net effect of the speed limit decrease.

The variables are the number of fatal accidents* (Fatal), severe injury accidents (S-injury), minor injury accidents (L-injury), vehicle damage accidents (Vehicle), disposable income (DI) (quarterly data at 1985 prices) and price indices for petrol (PPI). The price index (with 1982 as base) is deflated using the consumer price index. Different investigations of the reliability of the data have shown that 100% of the total number of fatalities, 60% of the total number of severe injuries and 30% of the total number of minor injuries are reported to the authorities (SOU 1991b). The low coverage of the number of minor injury accidents can be attributed to cyclist and walker accidents in cities while there is a much better degree of coverage of motorway car accidents (see Nilsson 1993). Descriptive statistics for the variables are given in Table 1. The means of all variables are

*An accident is registered according to the most severe injury, i.e. if one person is killed while the others involved are only injured, the accident is reported as fatal.

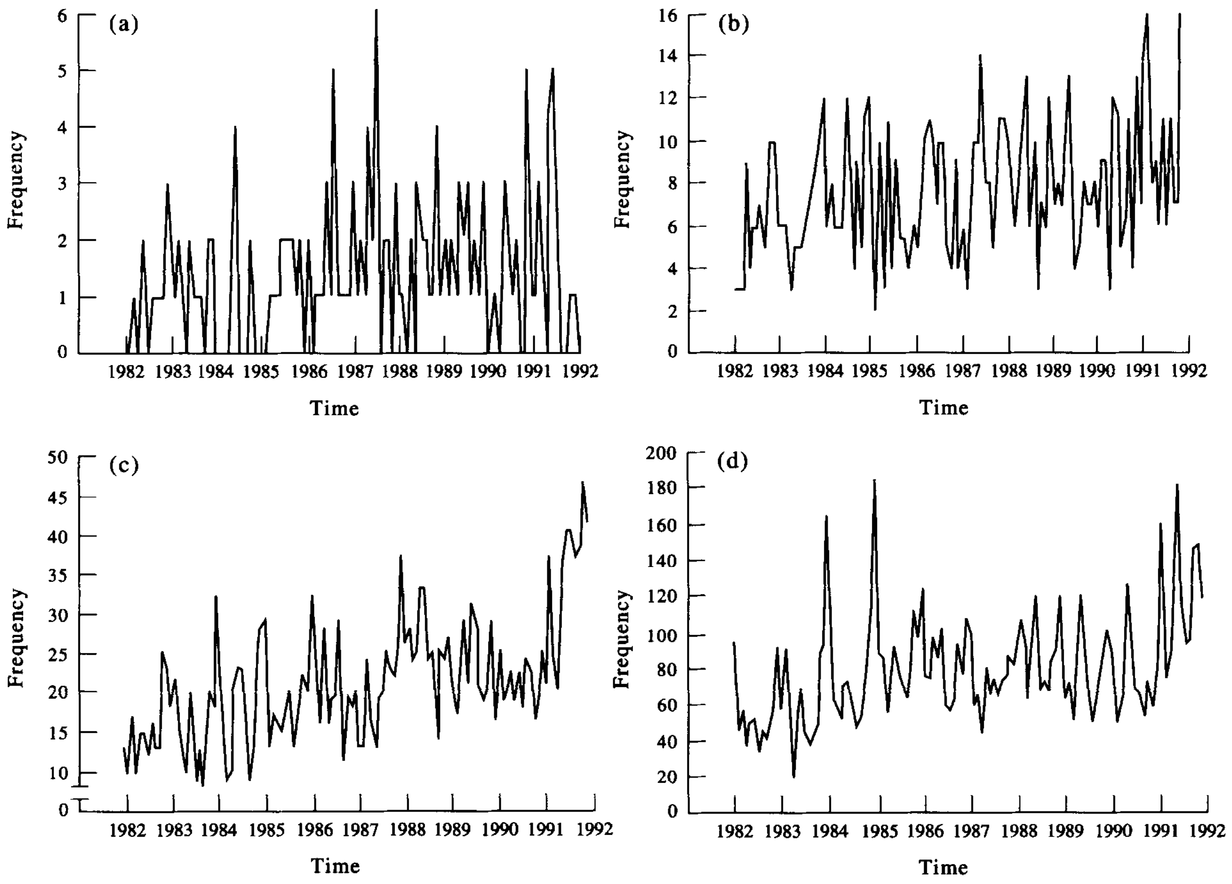

Fig. 2. Time series plot of the frequency of (a) fatal accidents (b) severe injury accidents (c) minor injury accidents and (d) vehicle damage accidents. (The dashed and dotted line indicates the time speed limit reduction in 1989.)

higher after the speed limit reduction. Overdispersion (with respect to the Poisson distribution, see Section 4) does not seem overly strong for the number of minor injury and fatal accidents. Time series graphs of the casualty data are given in Fig. 2. A positive trend and a seasonal pattern seem to be present. 116 observations are used for the estimation, leaving the final four observations for post-sample predictive comparisons.

Table 1. Descriptive statistics (disposable income, DI, measured in billions of SEK)

	Whole period		Before speed reduction		After speed reduction	
Variable	Mean	Variance	Mean	Variance	Mean	Variance
Fatal	1.41	1.64	1.34	1.46	1.60	2.19
S-injury	7.66	9.00	7.38	7.95	8.50	11.49
L-injury	21.22	59.82	19.48	43.29	26.43	74.99
Vehicle	79.59	848.33	75.43	691.16	92.07	1145.82
DI	118.06	124.95	114.71	112.58	128.13	27.39
PPI	91.96	91.81	90.57	83.03	96.14	97.84

4. TIME SERIES COUNT DATA REGRESSION MODELS

The Poisson distribution is a natural first candidate for describing count data such as accident frequency. The Poisson regression model is defined in terms of its density function, i.e.

$$P(y_t) = e^{-\mu_t}\mu_t^{y_t}/y_t! \quad (t = 1, \ldots, T). \tag{7}$$

In (7), y_t is the count or frequency variable at time t. The expected value, $E(y_t) = \mu_t$, equals the variance. The mean function is specified as $\mu_t = \exp(\mathbf{x}_t\beta)$*. In our model (3), $\mathbf{x}_t$ is the (1×17) vector of explanatory variables and β is the corresponding (17×1) parameter vector.

The models with an error term in the demand eqn (4) and in (6) are not consistent with the Poisson

*The function is the natural link in the GLM literature. Generally, it can be any function $h(\mathbf{x}_t)$. The advantage of using the exponential form is that μ_t is restricted to be positive. Gurmu and Trivedi (1993) suggest a test of the adequacy of the exponential specification against a more general specification.

assumption. We have a model that introduces overdispersion. The overdispersion is generated by η_t and/or ω_t, which may be serially correlated (this is likely to be the case if we have a common omitted variable, e.g. the speed level in both equations). The two functions may be written $\lambda_t = \mu_t v_t$, where v_t is either ω_t, η_t or $\omega_t \eta_t$. If normalised (then because of the normalisation, the intercept α in $\mathbf{x}_t \beta$ is changed) so that $\mathrm{E}(v_t) = 1$ and $\mathrm{V}(v_t) = \sigma^2$ the first two central moments of y_t are

$$\mathrm{E}(y_t) = \mu_t \quad \text{and} \quad \mathrm{V}(y_t) = \mu_t(1 + \sigma^2 \mu_t). \tag{8}$$

Further, if v_t is uncorrelated and gamma distributed, y_t is negative binomial (NB) distributed* with the mean and variance given in (8). Where there is both overdispersion and serial correlation, we may use the approach of Zeger (1988) or the one of Smith (1979) and Harvey and Fernandez (1989). Smith (1979) introduced a natural conjugate distribution to model local level models (i.e. the level is stochastic but not a function of exogenous variables) for counts and qualitative variables. No estimation of hyperparameters is performed. West et al. (1985) used the Kalman filter and various approximations in the estimation of unknown parameters. They also extended the local level model allowing for, e.g. a stochastic seasonal effect. Harvey and Fernandez (1989) extend Smith's (1979) approach by using maximum likelihood (ML) estimation of hyperparameters as well as by including exogenous variables. They label the model a structural approach (SA) for count data. While Smith's process is fully known distributionally, the correlation structure of the counts are restricted to be positive (see Brännäs and Johansson 1992). In Zeger's (1988) model (hereafter denoted the Zeger model) on the other hand the correlation structure is more flexible, but only the first two moments are known.

Zeger (1988) assumes that v_t is unobservable and stationary with $\mathrm{E}(v_t) = 1$ and $\mathrm{Cov}(v_t, v_{t+\tau}) = \sigma^2 \rho_v(\tau)$, where $\rho_v(\tau)$ is the autocorrelation function of $\{v_t\}$. We get the same expectation and variance as in (8) and

$$\begin{aligned} \rho_y(\tau) &= \mathrm{Cor}(y_t, y_{t+\tau}) \\ &= \rho_v(\tau)[\{1 + \sigma^2 \mu_t)^{-1}\} \\ &\quad \times \{1 + (\sigma^2 \mu_{t+\tau})^{-1}\}]^{-1/2} \quad (\tau \neq 0) \end{aligned} \tag{9}$$

To define the SA model Smith (1979) and Harvey and Fernandez (1989) used a recursive updating algorithm for the stochastic component. Assume that the count, y_t given v_t is Poisson distributed with the mean $\lambda_t = \mu_t v_t$. Let the distribution of v_{t-1} at time $t-1$ given the information set, $\mathbf{y}_{t-1} = (y_0, \ldots, y_{t-1})$, at time $t-1$ be gamma distributed with parameters a_{t-1} and b_{t-1}, i.e.

$$\begin{aligned} &f(v_{t-1} : a_{t-1} b_{t-1} \mid \mathbf{y}_{t-1}) \\ &\quad = e^{b_{t-1} v_{t-1}} v_{t-1}^{-a_{t-1}} / (\Gamma(a_{t-1}) b_{t-1}^{-a_{t-1}}) \\ &\quad = G(a_{t-1}, b_{t-1}) \end{aligned}$$

and let $v_t \mid \mathbf{y}_{t-1}$ be distributed as $G(a_{t|t-1}, b_{t|t-1})$, with $a_{t|t-1} = \omega a_{t-1}$, $b_{t|t-1} = \omega b_{t-1}$ and with $0 < \omega \leqslant 1$ an unknown parameter. From the properties of the gamma distribution, it follows that $y_t \mid \mathbf{y}_{t-1}$ has a $G(a^*_{t|t-1}, b^*_{t|t-1})$ distribution, where

$$a^*_{t|t-1} = \omega a_{t-1} \quad \text{and} \quad b^*_{t|t-1} = \omega b_{t-1} \mu_t^{-1} \tag{10}$$

The conditional distribution of y_t given the information set $\mathbf{y}_t$ is the NB distribution

$$\begin{aligned} &P(y_t \mid \mathbf{y}_{t-1}) \\ &= \int_{-\infty}^{\infty} P(y_t \mid \lambda_t) f(\lambda_t \mid \mathbf{y}_{t-1}) \, d\lambda_t \\ &= \binom{a^*_{t|t-1} + y_t}{y_t} b^*_{t|t-1}{}^{a^*_{t|t-1}} (1 + b^*_{t|t-1}) \\ &\quad {}^{-(a^*_{t|t-1} + y_t)} \end{aligned} \tag{11}$$

The posterior distribution of v_t is $G(a_t, b_t)$, with $a_t = \omega a_{t-1} + y_t$ and $b_t = \omega b_{t-1} + \mu_t$. The stochastic term v_t can be interpreted as a stochastic level. The recursion is defined implicitly (in terms of a_t and b_t) rather than explicitly as, e.g. in the local level model with Gaussian distributions.

5. ESTIMATION

For a Poisson regression model, a number of consistent estimators can be used to give estimates of β, e.g. ordinary least squares (OLS) (in the regression of the log frequencies on explanatory variables, thus only possible if no zeros are present), pseudo maximum likelihood (PML) and maximum likelihood (ML). The ML estimator based on (7) has gradient vector $\mathbf{X}'(\mathbf{y} - \mu)$ and Hessian matrix $-\mathbf{X}'\mathbf{M}\mathbf{X}$, where $\mathbf{X}$ is the $(T \times k)$ matrix of observations on exogenous variables, $\mathbf{y}$ is the T-vector of time series observations on the count variable, $\mathbf{M} = diag(\mu)$ and μ has tth element μt. The covariance matrix of $\hat{\beta}$, $\hat{\mathbf{A}}_T^{-1} = (\mathbf{X}'\hat{\mathbf{M}}\mathbf{X})^{-1}$, is evaluated at the estimate $\hat{\beta}$.

The estimation of σ^2 and β in (8) can be performed with various estimators. Such estimators are the ML, the quasi-likelihood (QL), the PML and the quasi-generalized pseudo-maximum-likelihood (QGPML) of Gourieroux et al. (1984). If the mean function is correctly specified, the PML is a consistent estimator of β and σ^2. If both the mean and variance

*Cameron and Trivedi's (1986) negbin II model.

functions are correctly specified, the QGPML estimator is consistent.

Consistent estimates of β in the Zeger model described in eqns (8) and (9) can, e.g. be obtained with Poisson ML, see Brännäs and Johansson (1994). A consistent estimator of the covariance matrix of the Poisson ML estimate, $\hat{\beta}$ is of the form

$$\mathrm{Cov}(\hat{\beta}) = \hat{\mathbf{A}}_T^{-1}\hat{\mathbf{B}}_T\hat{\mathbf{A}}_T^{-1} \tag{12}$$

where $\hat{\mathbf{B}}_T = \mathbf{X}'\hat{\mathbf{V}}\mathbf{X}$ and $\hat{\mathbf{V}}$ is an estimator of $\mathbf{V} = \mathrm{Cov}(\mathbf{y}) = \mathbf{M} + \sigma^2\mathbf{M}\mathbf{R}_v\mathbf{M}$, with $\mathbf{R}_v^{ij} = \rho_v(|i-j|)$. This covariance matrix estimator extends the one of Gourieroux et al. (1984) for overdispersion.

Let the unknown parameters be collected into β and $\theta' = (\sigma^2, \theta_p')$ where θ_p completely specifies $\rho_v(\tau)$. For a consistent estimator of θ, the β vector can also be consistently estimated from the QL estimating equation.

$$\frac{\partial\mu'}{\partial\beta}\mathbf{V}^{-1}(\beta, \hat{\theta})(\mathbf{y} - \mu) = 0 \tag{13}$$

Zeger (1988) shows that under mild regularity conditions $\sqrt{T}(\tilde{\beta} - \beta)$ is asymptotically multivariate normal with zero mean and covariance matrix

$$\mathbf{V}_{\tilde{\beta}} = \lim_{T\to\infty}\left(\frac{\partial\mu'}{\partial\beta}\mathbf{V}^{-1}\frac{\partial\mu}{\partial\beta}\Big/T\right)^{-1} \tag{14}$$

To estimate the elements of θ, Zeger (1988) suggests two approaches. In one, the parameters are estimated by moment estimators. The estimator of σ^2 is

$$\hat{\sigma}^2 = \sum_{t=1}^{T}[\tilde{y}_t^2 - \hat{\mu}_t]\Big/\sum_{t=1}^{T}\hat{\mu}_t^2 \tag{15}$$

and that of $\rho_v(\tau)$ is

$$r_\tau = \sum_{t=\tau+1}^{T}\tilde{y}_t\tilde{y}_{t-\tau}\Big/\left[\hat{\sigma}^2\sum_{t=\tau+1}^{T}\hat{\mu}_t\hat{\mu}_{t-\tau}\right] \tag{16}$$

where $\tilde{y}_t = y_t - \hat{\mu}_t$ is a residual. In the second approach, an approximating autoregression is utilized to reduce the computational burden arising from the inversion of $\mathbf{V}$ in each iterative step. In the following only the former approach is used. However, for the covariance matrix estimator of the Poisson ML (12), the $\mathbf{V}$ does not need to be inverted. Further Brännäs and Johansson (1994) have suggested alternative estimators of θ.

There is a decision problem for both the estimation of the PML-type covariance matrix (12) and the QL estimation at what value on τ and onwards should r_τ be set equal to zero. For the estimation of the PML covariance matrix, Johansson (1993) indicates that τ should be large since no extra bias of the standard errors is introduced. Brännäs and Johansson (1994) indicate that the bias in the estimation of $\rho_v(\tau)$ is not severe when using eqn (16) if $\sigma^2 > 0$. In the case of $\sigma^2 = 0$, the r_τ estimates can be greater than one in absolute value as well as of the wrong sign. The small sample performance of different estimators of the covariance matrix (12) is evaluated indirectly via t-tests in Brännäs and Johansson (1994). The conclusion to be drawn is that the Poisson ML estimator has favourable large, as well as small, sample properties even when there is serially correlated overdispersion. The size properties of the associated corrected t-test statistics are reliable when we have a reasonably large sample. Modest, if any, efficiency gain is obtained by using the QL estimator defined through (13).

Brännäs and Johansson (1992) examine the small sample performance of the SA model. A general conclusion is that the t-test statistic power functions are well-behaved, both when the correct model is the SA and when it is of the Poisson type.

A Newton–Raphson ML algorithm is used for the Poisson and NB models and the estimated covariance matrices are calculated using the inverse of the analytical Hessian matrices. The corrected covariance matrix for the Poisson estimator is estimated according to eqns (12), (15) and (16). The τ is set to 30. QL estimation of the Zeger model is performed with the six first (to ensure that the estimate of $\mathbf{V}$ is positive definite (pd))* estimated autocorrelations (16) and the covariance matrix is estimated by evaluating the QL covariance matrix (14) at the QL estimates. The estimation of the SA model parameters in (11) is performed by maximizing the log-likelihood with respect to ω and β (the GAUSS BFGS algorithm using numerical gradients is used).

5.1. *Tests for overdispersion and serial correlation*

In the above models it can be shown that serial correlation can be present only when there is overdispersion. Hence, the first thing to do is to test for overdispersion. If this is detected, then serial correlation should be tested for. Under the null hypothesis of a Poisson distribution, the local Lagrange multiplier test given by Cox (1981) can be used. In order to improve on the small sample performance of the test, Dean and Lawless (1989) give the approximation

$$s = \sum_{t=1}^{T}[\tilde{y}_t^2 - y_t + \hat{h}_{tt}\hat{\mu}_t]\Big/\sqrt{2\sum_{t=1}^{T}\hat{\mu}_t^2} \tag{17}$$

where $\hat{h}_{tt}$ is the tth diagonal element of $\hat{\mathbf{M}}^{1/2}\mathbf{X}(\mathbf{X}'\hat{\mathbf{M}}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{M}}^{1/2}$. The statistic s is one sided

*The estimated matrix $\mathbf{V}$ is not restricted to be pd. In the simulation study carried out by Brännäs and Johansson (1994) the estimated matrix $\mathbf{V}$ was not pd in more than 50% of the replications.

and asymptotically distributed as an N(0, 1) variate. In several Monte Carlo studies, the s statistic has been found to have good size properties in small samples (e.g. Dean and Lawless 1989; Brännäs 1992). Brännäs and Johansson (1992, 1994) indicated that the detection of overdispersion is reliable if the Zeger and SA models are being used. They also indicate that, if ω is close to one, the SA model will be very close to an ordinary Poisson process with no overdispersion and no serial correlation. Thus, a test of $H_0: \omega = 1$ against $H_A: \omega < 1$ is a test of the presence of overdispersion and serial correlation. It is important to note that under H_0, ω is on the edge of the parameter space. This implies that the statistic is distributed as $\frac{1}{2}(\chi_0^2 + \chi_1^2)$, where χ_0^2 is a degenerate distribution with all its mass at the origin. A desired size α is obtained by setting the size of 2α in a χ_1^2 distribution (cf. Harvey 1989b, Ch. 5).

A number of plausible tests of serial correlation are possible. Harvey and Fernandez (1989) made some suggestions from model selection and tests of misspecification. As in the GLM literature they recommend using the standardized (Pearson) residual for misspecification tests. However, they did not suggest any tests for serial correlation. The small sample performance of the portmanteau test statistics of Box and Pierce (1970) and Ljung and Box (1978),

$$Q_{BP} = T \sum_{k=1}^{K} r_k^2$$

and

$$Q_{LB} = T(T+2) \sum_{k=1}^{K} r_k^2 \Big/ (T-k)$$

respectively, are examined in Brännäs and Johansson (1994). Both statistics are powerful, but their sizes are too large.

5.2. Prediction

The predictors of the mean level in the Poisson and NB models are straightforward, since in the first case no error term implying overdispersion is present and in the latter the error term is time homogenous. The predictors for the Zeger and the SA models are described below. We assume that for any time period h, the future values of $\mathbf{x}_{T+h}$ are known.

The best h-step ahead linear minimum mean square error predictor for the Zeger model, defined in Brännäs (1995), is of the form

$$\begin{aligned}\hat{y}_{T+h|T} &= \mathrm{E}(y_{T+h} \mid \mathbf{y}_T) \\ &= \mu_{T+h} + \mathrm{Cov}(y_{T+h}, \mathbf{y}_T)\mathbf{V}^{-1}(\mathbf{y}_T - \mu) \qquad (18)\end{aligned}$$

where $\mathrm{Cov}(y_{T+h}, \mathbf{y}_T)$ is the $(1 \times T)$ vector of the first T elements in the final row of the extended covariance matrix $\mathbf{V}_{T+h} = \mathrm{Cov}(\mathbf{y}_{T+h})$. Given $\mathbf{x}_{T+h}$ and the estimated parameters, an estimate of $\mathbf{V}_{T+h}$ can be obtained. If no autocorrelation is present, $\mathrm{Cov}(y_{t+h}, \mathbf{y}_T) = 0$ and the best predictor is simply the unconditional mean μ_{T+h}. The variance of the h-step ahead prediction is (see Brännäs 1995)

$$\begin{aligned}\mathrm{Var}(\hat{y}_{T+h}) = \mu_{T+h}&(1 + \sigma^2\mu_{T+h}) \\ &- \mathrm{Cov}(y_{T+h}, \mathbf{y}_T)\mathbf{V}^{-1} \qquad (19) \\ &\times \mathrm{Cov}'(y_{T+h}, \mathbf{y}_T)\end{aligned}$$

It is easy to see that the predictor is unbiased and that the variance of the prediction decreases if the serial correlation is properly accounted for. If there is no serial correlation, the expression (19) reduces to the NB model prediction variance and, if in addition $\sigma^2 = 0$ to the Poisson model prediction variance.

For a given value of μ_{T+1} and the information set $\mathbf{y}_T$, the mean and variance of the one-step predictive distribution of $\mathbf{y}_{T+1}$ in the SA model are given from the properties of the NB distribution

$$\hat{y}_{T+1|T} = \mathrm{E}(y_{T+1} \mid \mathbf{y}_T) = \mu_{T+1}(a_T/b_T) \qquad (20)$$

and

$$\mathrm{Var}(y_{T+1} \mid \mathbf{y}_T) = \mathrm{Var}(\lambda_{T+1} \mid \mathbf{y}_T) + \mathrm{E}(\lambda_{T+1} \mid \mathbf{y}_T) \qquad (21)$$

where $\mathrm{Var}(\lambda_{T+1} \mid \mathbf{y}_T) = [\mu_T^2/\omega][a_{T+1}/b_{T-1}^2]$ and $\mathrm{E}(\lambda_{T+1} \mid \mathbf{y}_T) = \mathrm{E}(y_{T+1} \mid \mathbf{y}_T)$. By repeatedly inserting $a_t = \omega a_{t-1} + y_{t-1}$ and $b_t = \omega b_{t-1} + \mu_t$ into eqn (20) it follows that

$$\hat{y}_{T+1|T} = \mu_{T+1}\left[\sum_{j=0}^{T-1} \omega^j y_{T-j} \Big/ \sum_{j=0}^{T-1} \omega^j \mu_{T-j}\right] \qquad (22)$$

This is an exponential weighting moving average scheme. The h-step ahead predictive distribution for y_{T+h} is difficult to derive (Harvey 1989b, Ch. 6). What can be shown is that for a given value on μ_{T+h}, the mean of the predictive distribution of y_{T+h} is of the form

$$\hat{y}_{T+h|T} = \mu_{T+h}\left[\sum_{j=0}^{T-1} \omega^j y_{T-j} \Big/ \sum_{j=0}^{T-1} \omega^j \mu_{T-j}\right]$$

5.3. Model evaluation

In addition to the test of overdispersion and serial correlation discussed in Section 5.1, we now consider informal criteria for diagnostically checking other aspects of the model specification. The residuals of the Poisson, NB model and the Zeger model are of the form $\hat{\varepsilon}_t = y_t - \hat{\mu}_t$. For the SA model, the residual takes the form of a one step prediction error $\hat{v}_t = y_t - \mathrm{E}(y_t \mid \mathbf{y}_{t-1})$. The standardised (Pearson) residuals are defined by

$$\hat{\varepsilon}_{pt} = \varepsilon_t/\sqrt{\mathrm{Var}(y_t)} \quad \text{and} \quad \hat{v}_{pt} = \hat{v}_t/\sqrt{\mathrm{Var}(y_t \mid \mathbf{y}_{t-1})}$$

If the parameters in the Poisson, NB and SA models are known, it is easy to show that the residuals are independently distributed with expectation zero and unit variance (for the SA model, see Harvey 1989b, Ch. 6). The residuals in the Zeger model are not independent but have expectation zero and unit variance.

In the Poisson model the standardized residuals are bounded below by $-\hat{\mu}_t^{1/2}$. Thus a residual plot of the residuals should not be interpreted in the same way as in a normal regression model. McCullagh and Nelder (1983) suggested that the residuals, $\hat{\varepsilon}_{pt}$, should be plotted against $2\hat{\mu}_t^{1/2}$ and $2\sin^{-1}(\hat{\mu}_t)^{1/2}$ for the Poisson and NB models, respectively. In such plots, the property from the normal regression model that the slopes for constant y are minus one is preserved. A plot of $\hat{v}_{pt}$ and $\hat{\varepsilon}_{pt}$ against $2\hat{\mu}_t^{1/2}$ may reveal isolated large residuals, a general curvature and a trend or spread in the residuals. A general curvature in the residuals is an indication of an unsatisfactory scale for the explanatory variable(s) or an unsatisfactory functional form for the mean. A trend in the spread indicates an unsatisfactory variance function.

As measures of goodness of fit, the sum of the squared standardized residuals and the conventional coefficient of determination, R^2, are used. The sums of squared standardized residuals, $E=\Sigma_{t=1}^{T}\hat{\varepsilon}_{pt}^2$ and $V=\Sigma_{t=1}^{T}\hat{v}_{pt}^2$ are approximately distributed as $\chi^2(T-p)$, if we have independent observations, and as $\chi^2(T-t_0-p)$, respectively. Here, p is the number of estimated parameters and t_0 is the time period for the first non zero observation.

The statistic used as a measure of post-sample goodness of fit is

$$S_p^2=(1/h)\sum_{j=1}^{h}\hat{\zeta}_{T+j|T}^2 \tag{23}$$

where $\hat{\zeta}$ is either $\hat{\varepsilon}$ or $\hat{v}$. According to Harvey (1989a, Ch. 5), this statistic is more meaningful than the within sample comparison based on R^2. The h-step ahead residuals for the SA and Zeger models are computed using eqns (22) and (18), respectively.

6. RESULTS

First, a model with the explanatory variables introduced in Section 2 is estimated for all dependent variables. Next, a likelihood ratio (LR) test of seasonality (H_0: no seasonal dummies against H_A: seasonal dummies) is performed.

Fatal accidents

A model with no seasonal dummies can be rejected (LR = 22.25 with 11 degrees of freedom). No overdispersion is indicated ($s=-0.037$). Since no overdispersion is detected no test for serial correlation is performed and no model taking care of overdispersion and serial correlation needs to be estimated. Despite this, the SA model is estimated for comparison. We can see in Table 2 that the within sample goodness of fit as well as the post sample predictive statistic are very similar for the two models. The predictions and post sample predictions (with a 95% confidence interval) for the Poisson ML estimates are shown in Fig. 3.

The post sample prediction overestimates the number of fatal accidents, but the seasonal pattern is acceptably captured. Plots of the standardized residuals plotted against $2\hat{\mu}_t^{1/2}$ are presented in Fig. 4. No unexpected pattern is seen (the pattern in the residuals is the expected one, since y only takes discrete values, the residuals for constant y should have a slope of about minus one). However, four residuals are greater

Table 2. Goodness of fit statistics

	$\chi^2(E/V)$	R^2	S_p^2
Number of fatal accidents			
Poisson	95.51	.27	2.00
SA	96.02	.23	1.79
Number of severe injury accidents			
Poisson	101.10	.22	13.20
SA	106.70	.19	14.37
Number of minor injury accidents			
Poisson	146.30	.44	220.87
NB	112.70	.44	220.87
Zeger	111.77	.47	190.20
SA	156.40	.42	72.90
Number of vehicle damage accidents			
Poisson	507.46	.50	1269.00
NB	111.93	.55	1208.98
Zeger	89.11	.55	931.30
SA	519.60	.42	956.79

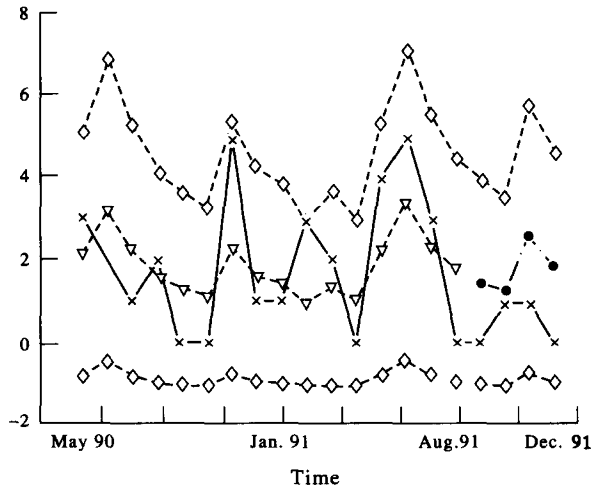

Fig. 3. Reported number of fatal accidents (solid line), post sample predictions (dashes and dots), predictions (short dashes) from the Poisson ML estimates (May 1990–August 1991). A 95% confidence interval is given by the long dashes.

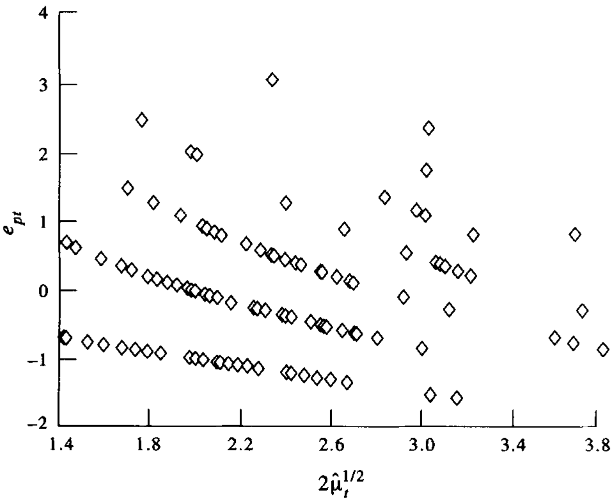

Fig. 4. Standardized Pearson residuals of the number of fatal accidents against $2\hat{\mu}_t^{1/2}$.

than two with one above three (six fatal accidents in July 1987). Still the functional form and the scale for the independent variables seem to be acceptable.

The estimation results are given in Table 3. The only significant effect is the positive seasonal effect for June (the mean number of fatal accidents is about 1.4 as many in June as in February). The explanation can be that there are more inexperienced drivers on the roads during the vacation periods, in combination with a fatigue effect from the heat. The effect can also be the result of more drunken drivers in the traffic during this period (especially in June when midsummer is celebrated). The seasonal effects can be summarized as increasing fatal accidents during the summer months (May–August) and November. This pattern does not conflict with the findings of Fridstrøm et al. (1992). They give three possible explanations for the reduced risk during the winter months (risk compensation behaviour, reduced exposure during the winter that is not controlled for through the gasoline sales proxy, and that visibility is increased when the roadside is covered by snow). The effect of the speed limit reduction is negative as hypothesized, but it is insignificant [Michener and Tighe (1992) obtained the same result]. The trend parameter is estimated to be negative and the parameter for the law dummy is estimated to be positive, however, insignificantly. The economic variables are theory consistent with the parameter estimates 1.01 and −0.42 for ln *DI* and gas price, respectively.

Severe injury accidents

We reject the null hypothesis of no seasonal dummies ($p = 0.10$). The overdispersion test is insignificant ($s = 0.152$). Hence, there is no evidence of overdispersion and no test for serial correlation needs to be performed. For comparison, the goodness of fit statistics for the SA model are also presented in Table 2. Both the post and the within sample goodness of fit statistics are in favour of the Poisson model. The predictions and residuals using the Poisson model are shown in Figs 5 and 6. The Poisson model seems acceptable since no unexpected pattern can be seen in the residuals and only two standardized residuals are greater than two.

The results for the Poisson ML estimator are presented in Table 3. The significant parameter is the positive effect of ln *DI* (one sided test of size 0.05) and the parameters for July. A 1% increase in income leads to a 2.76% increase in average reported severe injury accidents. The trend and the law dummy have small negative effects. The gas price parameter estimate has an unanticipated sign, however, insignificant. The effect of the speed limit reduction is negative and insignificant. The seasonal pattern is almost the same (increasing during the summer months, July–August) as for fatal accidents with the exception of a negative and almost significant effect for October and November.

Minor injury accidents

The model without seasonal dummies is rejected (LR = 36.88, $p = 0.0001$). In the overdispersion test, $s = 3.39$ indicates overdispersion. Further, a test of $H_0: \omega = 1$ against $H_A: \omega < 1$ yields $\chi^2 = 4.98$, which also indicates overdispersion. The estimate of σ^2 in the NB model is, however, insignificant. Hence the overdispersion is not severe. The estimated autocorrelations are large and positive for the six first lags and the portmanteau statistics ($\tau = 12$) Q_{BP} and Q_{LB} are significant at all reasonable levels of significance. The goodness of fit statistics are presented in Table 2. The NB and Zeger models perform best in terms of the within sample goodness of fit test. The approaches that take serial correlation into account perform better in the post sample predictions with the SA model emerging as the best one. In Fig. 7, the predictions and post sample predictions are shown for the QL estimator using the *h*-step predictor (18). The level is underestimated but the pattern is well captured. The residuals of the model can be judged as satisfactory. No unexpected pattern in the residuals is seen. However, there are five residuals greater than two, with one close to three (the peak in January 1984).

The results of the different estimators are presented in Table 3. All estimation methods give the same signs on the parameters. Generally, the estimates for the four estimators are close, with the smallest estimates and *t*-values being achieved for the Zeger

Table 3. Parameter estimates

Estimator		α/ω	1	Law	Trend	ln PPI	ln DI	Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Number of fatal accidents																			
Poisson	Estimate	−9.90	−.09	.16	.00	−.42	1.01	−.09	−.51	−.18	−.53	.28	.68	.40	.09	−.11	−.35	.34	−.02
	t-value	−.23	−.25	.46	.33	−.40	.28	−.31	−1.50	−.62	−1.59	1.19	3.44	1.34	.28	−.30	−.89	1.05	
Number of severe injuries																			
Poisson	Estimate	−32.02	−.15	−.15	.00	.41	2.76	.13	−.07	.02	−.06	−.11	.15	.35	.09	.07	−.27	−.31	.10
	t-value	−1.84	−.95	−1.04	.10	.91	1.90	1.22	−.59	.14	−.50	−.99	1.50	2.75	.65	.47	−1.74	−1.95	
Number of minor injuries																			
Poisson	Estimate	−9.03	−.21	−.12	.01	.19	.93	.16	−.00	−.03	−.09	−.10	.12	.10	.12	−.24	−.19	.06	.10
	t-value	−.83	−2.22	−1.30	2.71	.69	1.03	2.44	−.01	−.49	−1.24	−1.42	1.86	1.23	1.46	−2.51	−2.01	.70	
PML*	*t*-value	−.69	−1.33	−.88	1.94	.44	.82	2.55	−.01	−.50	−1.27	−1.46	1.93	1.10	1.33	−2.38	−1.83	.61	
NB	Estimate	−9.03	−.21	−.12	.01	.19	.93	.16	−.00	−.03	−.09	−.10	.12	.10	.12	−.24	−.19	.06	.10
	t-value	−.69	−1.51	−1.04	2.02	.51	.83	2.49	−.01	−.29	−.90	−1.19	1.32	1.07	1.16	−2.13	−1.30	.49	
Zeger	Estimate	−7.64	−.19	−.12	.01	.26	.78	.16	.00	−.05	−.09	−.10	.11	.09	.09	−.24	−.18	.08	.12
	t-value	−.60	−1.40	−.88	2.14	.70	.74	2.35	.02	−.60	−1.23	−1.34	1.65	.96	1.02	−2.35	−1.72	.84	
SA	Estimate	.79	−.16	−.23	.01	.08	1.17	.17	.00	−.03	−.09	−.10	.11	.10	.12	−.21	−.20	.05	.09
	t-value	4.98†	−2.99	−4.13	1.86	1.52	21.86	3.35	.09	−.53	−1.75	−2.02	2.07	1.85	2.19	−3.99	−3.97	.94	
Number of vehicle damage accidents																			
Poisson	Estimate	−4.21	−.24	−.36	.01	.30	.59	.37	.11	−.09	−.25	.12	.18	−.11	−.21	−.25	−.11	.04	.194
	t-value	−.78	−4.87	−7.72	7.38	2.07	1.30	11.93	3.09	−2.37	−6.33	3.74	5.75	−2.53	−4.57	−5.14	−2.36	.92	
PML*	*t*-value	−.29	−1.88	−2.90	2.93	.81	.48	4.23	1.19	−.97	−2.74	1.41	2.12	−.97	−1.85	−2.18	−.89	.33	
NB	Estimate	−4.11	−.25	−.36	.01	.25	.59	.39	.11	−.07	−.27	.11	.17	−.12	−.21	−.24	−.12	.05	.21
	t-value	−.30	−1.95	−2.83	3.00	.57	.54	7.60	1.64	−.86	−3.16	1.14	2.10	−1.13	−1.57	−1.88	−.67	.49	
Zeger	Estimate	3.53	−.34	−.32	.01	.39	−.12	.38	.10	−.09	−.26	.11	.16	−.16	−.25	−.27	−.07	.10	.25
	t-value	.25	−2.75	−2.78	3.72	1.09	−.11	4.87	1.19	−1.07	−3.22	1.41	2.14	−1.59	−2.49	−2.61	−.60	.92	
SA	Estimate	.22	−.11	.13	.01	1.09	−.37	.33	.11	−.11	−.24	.12	.19	−.18	−.28	−.29	−.02	.14	.23
	t-value	255.8†	−2.54	2.92	.20	24.66	−8.28	7.85	2.56	−2.64	−5.65	2.90	4.44	−4.16	−6.50	−6.73	−.44	3.30	
Total number of casualties (the whole sample period)																			
Zeger	Estimate	3.38	−.28	−.23	.01	.38	−.08	.31	.04	−.09	−.23	.04	.15	−.10	−.18	−.25	−.04	.13	.24
	t-value	.29	−2.26	−1.87	3.64	1.12	−.08	4.80	.62	−1.35	−3.39	.70	2.39	−1.20	−2.18	−2.94	−.49	1.39	

*Variance estimate according to eqn (12).
†$(\hat{\omega}-1)^2/v(\hat{\omega})$.

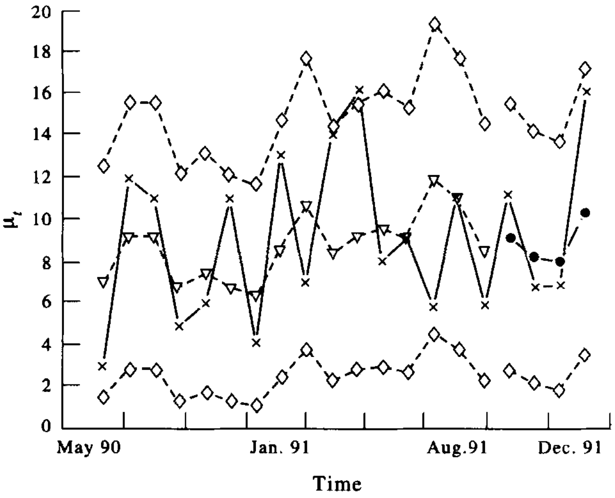

Fig. 5. Reported number of severe injury accidents (solid line), post sample predictions (dashes and dots), predictions (short dashes) from the Poisson ML estimates (May 1990–August 1991). A 95% confidence interval is given by the long dashes.

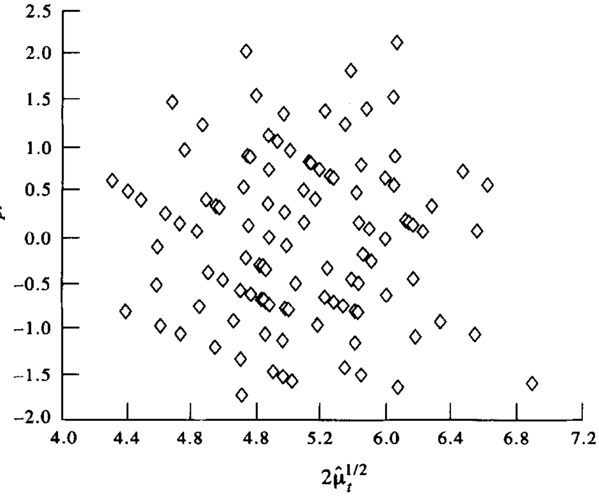

Fig. 6. Standardized Pearson residuals of the number of severe injury accidents against $2\hat{\mu}_t^{1/2}$.

model. The NB model gives the same estimates as the Poisson ML estimates, with standard errors close to those of the PML-type covariance matrix. The effect of the speed limit reduction is negative and significant for all models (one sided test of size 10%). The Zeger model gives a 17% reduction in reported accidents or an average monthly reduction of 3.5 accidents. The economic variables are still insignificant, the income effect remains positive but with the wrong sign for the gas price parameter. The trend parameter is now estimated to be small and positive and the parameter for the law dummy is estimated to be negative, however still insignificant. As before

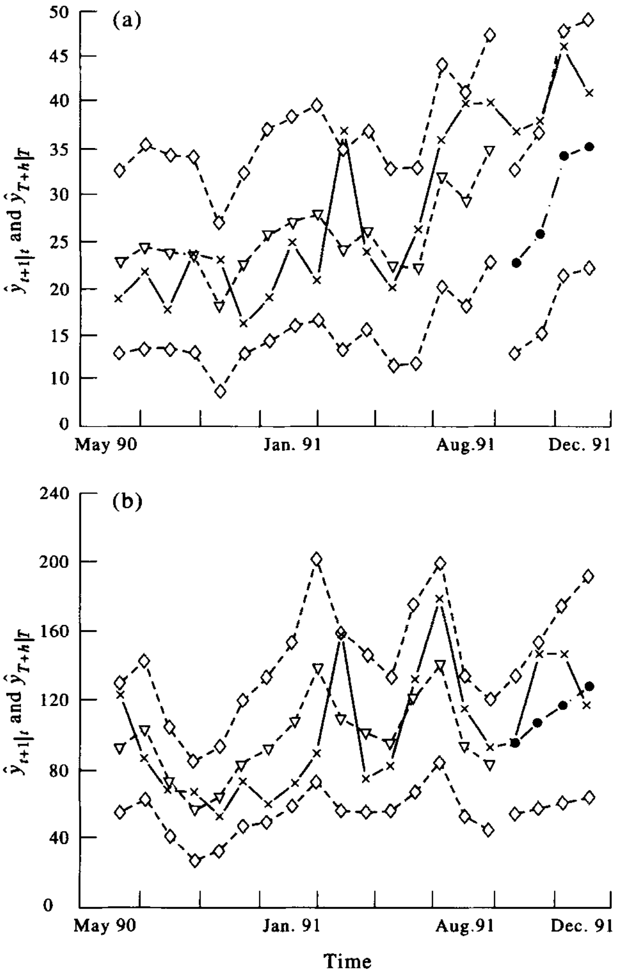

Fig. 7. Fitted values using h-step predictor (18) (short dashes), post sample predictions (dashes and dots) and reported number (solid line) of (a) minor injury accidents and (b) vehicle damage accidents from Zeger model (May 1990–August 1991). A 95% confidence interval is given by the long dashed lines.

we have an increase in accidents during the summer months (July–August) and, contrary to severe injuries, a negative significant effect in September.

Vehicle damage accidents

The model with seasonalities cannot be rejected at any reasonable level of significance (LR = 362.86). Overdispersion is present $s = 32.21$, the χ^2-test of $H_0: \omega = 1$ equals 255.9 and the estimate of σ^2 in the NB model is significant. The two first autocorrelation are estimated to be large and positive, while the autocorrelations for lag 3–6 are estimated to be large and negative. The portmanteau statistics ($\tau = 12$) Q_{BP} and Q_{LB} are significant at all reasonable levels of significance.

Both the within and the post sample goodness of fit statistics are in favour of the Zeger model (see

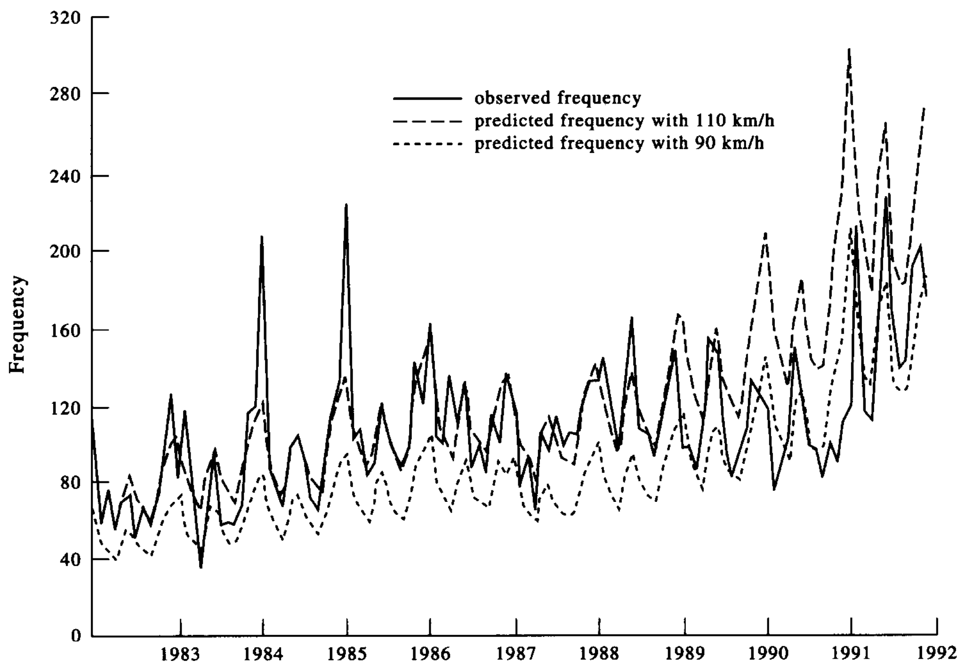

Fig. 8. Total number of casualties and estimated (using the whole sample period) mean predictions under 90 and 110 km/h with the Zeger model.

Table 1). In Fig. 7, predictions and post sample predictions are plotted for the QL estimator using the h-step predictor (18). The scale for the explanatory variables using the Zeger model seems to be good since no undesirable pattern is seen in the residuals. Five residuals are, however, greater than two in absolute value with one greater than three (the peak in January 1984). The reason that the SA model does not perform as well as the Zeger model may be due to the estimated large negative autocorrelations, for lags 3–6.

The parameter estimates are presented in Table 3. The four estimation techniques give the same seasonal pattern with significant positive peaks in January and June and negative ones in April, August and September. The intervention is negative and significant for all models. The Zeger model estimates give a 29% reduction in reported accidents or an average monthly reduction of 23 accidents. The parameter for $\ln DI$ is positive and smaller than previously for the static models and negative for the dynamic models (significant for the SA model). The estimated parameters for the gas price variable is still positive, insignificant for all models except for the SA model. The effect of the change in the law is now estimated to be negative and significant except for the SA model. The trend parameter is estimated to be small and positive, however, significant.

The total number of reported casualties for the whole sample period is estimated using the Zeger model. From Table 3 we can see that the estimates are close to those obtained with the number of vehicle damage accidents as independent variable. The speed limit reduction has reduced the number of casualties and the change in the law decreased the number of accidents. The total number of casualties and the predicted mean number of accidents with and without the speed limit reduction are shown in Fig. 8. By reducing the speed level from 110 km/h to 90 km/h the mean number of reported casualties each month is reduced by 29.

7. DISCUSSION

When no overdispersion is present, very small differences between estimators are seen in the estimates and, hence, in the prediction performance. The Poisson model appears to perform satisfactorily. In the case of overdispersion and serial correlation, the SA and Zeger models perform better than the Poisson and the NB models when comparing their post sample predictions. The h-step predictor (18) seems, as expected, to be a better predictor than the mean function for the Zeger model.

The estimation results using the Poisson and Zeger models may be summarized as follows. The speed limit reduction has reduced the number of reported casualties. The intervention variable has a negative effect for all the dependent variables, but it is only significant for minor injury ($p = 0.081$) and vehicle damage accidents ($p = 0.003$). For severe injury accidents $p = 0.17$. The parameter for the law dummy variable is, as hypothesized, estimated to be negative for all dependent variables except for fatal accidents, while only significant for vehicle damage accidents. The economic variables do not contribute much in explaining the reported number of casualties. The gas price is insignificant throughout (except for in the SA model). The effect of the income variable is positive (except for the number of vehicle damage accidents under the SA and Zeger model) with parameter estimated ranging from 0.38 to 2.76 with the highest value significant for severe accidents.

The trend parameter estimates are small and significant for the observed number of minor injury and vehicle damage accidents. A small trend effect is almost expected, since a negative effect on the number of casualties from changes in attitudes and a positive effect on the number of casualties from a gradual increase in the speed on the roads was expected. One thing that is clear is that we have a peak in the number of casualties in June, but no firm explanation can be given as to why this should be the case. It may be an effect of the midsummer celebrations taking place in June every year, when drinking and driving increase.

To get some policy implications of the speed limit reduction, a cost-benefit analysis can, as was said earlier, be performed. Forester et al. (1984) estimate the number of lives saved by using the 55 mph as compared with the 65 mph speed limit on the U.S. highways to be 7466 per year. They concluded (the cost benefit analysis is based on the value society places on life saved compared with the value on the additional time required to travel) that the 55 mph is not cost-effective. The purpose of this study was not to perform a cost benefit analysis, but if one is to be carried out the number of casualties have to be divided into different groups since different injuries give rise to different costs. This study can be thought of as a first step towards such a cost-benefit analysis. On the income side, from a speed limit reduction, one has to include the reduction in death, injuries, vehicle damages, in fuel costs and the effect of less air pollution. On the cost side one has to include the increased cost of enforcement and travel time.

Acknowledgements—I thank Kurt Brännäs for generous and competent guidance through all stages of the work on this study. Runar Brännlund and Lars Hultkrantz have commented on one or several versions of the paper and thereby significantly improved the results. I also want to thank participants at ISF93 in Pittsburg and in the seminar at the Economic Department, University of Umeå for comments. Any errors are my own. Partial financial support from the Swedish Transport Research Board is gratefully acknowledged. I am also grateful to the Swedish Road Office and the Swedish Road and Traffic Research Institute for help with the data.

REFERENCES

Baum. H. M.; Wells, J. A.; Lund, A. K. Motor vehicle crash fatalities in the second year of the 65 mph speed limits. J. Safety Research 21:1–8; 1990.

Bhattacharyya, M. N.; Layton, A. P. Effectiveness of seat belt legislation on the Queensland road toll—an Australian case study in intervention analysis. J. Am. Stat. Assoc. 74:596–603; 1979.

Box, G. E. P.; Pierce, D. A. Distribution of residual autocorrelations in auto-regressive integrated moving average time series model. J. Stat. Comput. Simul. 41:229–241; 1970.

Box, G. E. P.; Tiao, G. C. Intervention analysis with applications to economic and environmental problems. J. Am. Stat. Assoc. 70:70–79; 1975.

Brännäs, K. Swedish travel frequencies. Statistical Research Report 1987–5, University of Umeå; 1987.

Brännäs, K. Flexible binary regression and panel data. Umeå Economic Studies. 274: University of Umea; 1992.

Brännäs, K. Prediction and control for a time series count data model. Int. J. Forcasting 11: 263–270, 1995.

Brännäs, K; Johansson, P. Estimation and testing in a model for serially correlated count data. In E. von Collani; R. Göb, eds. Proceedings of the Second Würzburg–Umeå Conference in Statistics. University of Wurzburg; 1992.

Brännäs, K.; Johansson, P. Time series count data regression. Communications in Statistics: Theory and Methods. 23: 2907–2925, 1994.

Brännäs, K.; Laitila, T. Modelling and prediction of travel distance by car. Transportation Planning and Technology 16:129–143; 1991.

Cameron, A. C.; Trivedi, P. K. Econometric models based on count data: comparison and applications and some estimators and tests. J. Appl. Econometrics 1:29–53; 1986.

Cox, D. R. Statistical analysis of time series, some recent developments. Scandinavian J. Statistics 8:93–115; 1981.

Dean, C.; Lawless, J. F. Tests for detecting overdispersion in Poisson regression models. J. Am. Stat. Assoc. 84:467–472; 1989.

Durbin, J.; Harvey, A. C. The effect of seat belt legislation on road casualties in Great Britain: report on assessment of statistical evidence. Annexe to Compulsory Seat Belt Wearing Report by the Department of Transport. (HMSO, London); 1985.

Forester, T. M.; McNown, R. F.; Singell, L. D. A cost-benefit analysis of the 55 mph speed limit. Southern Economic J. 50:631–641; 1984.

Fridstrøm, L. In favour of aggregate econometric accident models. Paper presented at the 6th International Conference on Travel Behaviour, Quebec; 1991.

Fridstrøm, L.; Ingebrigtsen, S. An aggregate accident model based on pooled, regional time-series data. Accid. Anal. and Prev. 23:363–378; 1991.

Fridstrøm, L.; Ifver, J.; Ingebrigtsen, S.; Krogard Thomsen, L. The contribution of exposure, weather, daylight, and randomness to the variation in accident counts: a four-country analysis. Paper presented at the 6th World Conference on Transportation Research, Lyon; 1992.

Gourieroux, C.; Montfor, A.; Trognon, A. Pseudo maximum likelihood methods: applications to Poisson models. Econometrica 52:701–720; 1984.

Gurmu, S.; Trivedi, P. K. Recent developments in models of event counts: a survey. Mimeo, Indiana University, Bloomington; 1993.

Hakim, S.; Shefer, D.; Hakkert, A. S.; Hocherman, I. A critical review of macro models for road accidents. Accid. Anal. Pre. 23:379–440; 1991.

Harvey, A. C. The econometric analysis of time series. Cambridge MA, The MIT Press; 1989a.

Harvey, A. C. Forcasting structural time series models and the kalman filter. Cambridge University Press: Cambridge; 1989b.

Harvey, A. C.; Durbin, J. The effects of seat belt legislation on British road casualties: a case study in structural time series modelling. J. Royal Stat. Soc. A149:187–227; 1986.

Harvey, A. C.; Fernandez, C. Time series for count data or qualitative observations. J. Bus. Econ. Stat. 7:407–417; 1989.

Hausman, J. A.; Hall, B. H.; Griliches, Z. Econometric models for count data with an application to the patents—R&D relationship. Econometrica 52:909–938; 1984.

Hultkrantz, L. An econometric study of the demand for recreational trips in Sweden. Umeå Economic Studies 281: University of Umeå; 1992.

Johansson, P. Truncation effect on a pseudo maximum likelihood covariance matrix estimator. Umeå Economic Studies 312: University of Umeå; 1993.

Kamerud, D. B. Benefits and costs of the 55 mph speed limit: new estimates and their implications. J. Policy Anal. Manag. 7:341–352; 1988.

Laverty, W. H.; Kelly, I. W.; Flynn, M. Geophysical variables and behavior: LXVIII, distal and lunar variables and traffic accidents in Saskatchewan 1984 to 1989. Preceptual and Motor Skills 74:483–488; 1992.

Ljung, G. M.; Box, G. E. P. On measure of lack of fit in time series models. Biometrika 66:67–72; 1978.

Lloyd, C. J. The comparison of accident counts up to the present date. J. Safety Res. 21:75–78; 1990.

McCullagh, P.; Nelder, J. A. Generalized linear models. London: Chapman and Hall; 1983.

Michener, R.; Tighe, C. A Poisson regression model of highway fatalities. Am. Econ. Rev. (Papers and Proceedings) 82:452–456; 1992.

Nilsson, G. Faktorer Bakom Trafiksäkerhetsförändringar. VTI Rapport 383: Linköping; 1993.

SOU Säkrare Förare. Slutbetänkande av Kommitén Körkort 2000, SOU, 1991:39; 1991a.

SOU Det Framtida Trafiksäkerhetsarbetet. Betänkande av Trafiksäkerhetsutredningen, SOU, 1991:79; 1991b.

Smith, J. Q. A generalization of the Baysian forcasting model. J. Royal Stat. Soc. B41:375–387; 1979.

TSV Hastighetsändringars Effekt på Skadefölj mm. Trafiksäkerhetsverket, Trafikbyrån, PM 1987:3; 1987.

West, M.; Harrison P. J.; Migon, H. S. Dynamic generalized linear models and Baysian forcasting (with discussion). J. Am. Stat. Assoc. 80:73–97; 1985.

Wilde, G. J. S. The theory of risk-homeostatic: implications for safety and health. Risk Anal. 2:209–255; 1982.

Varian, H. R. Microeconomic Analysis (London: Norton); 1992.

Zeger, S. L. A regression model for time series of counts. Biometrika 75:621–629; 1988.

Zeger, S. L.; Qagish, B. Markov regression models for time series: a quasi-likelihood approach. Biometrics 44:1019–1031; 1988.