Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Mobile safety cameras: estimating casualty reductions and the demand for secondary healthcare

Lee Fawcett [a] & Neil Thorpe [b]

[a] School of Mathematics & Statistics, Newcastle University, Newcastle, NE1 7RU, UK

[b] Transport Operations Research Group, Newcastle University, Newcastle, NE1 7RU, UK

Published online: 15 Jul 2013.

To cite this article: Journal of Applied Statistics (2013): Mobile safety cameras: estimating casualty reductions and the demand for secondary healthcare, Journal of Applied Statistics, DOI: 10.1080/02664763.2013.817547

To link to this article: http://dx.doi.org/10.1080/02664763.2013.817547

Journal of Applied Statistics, 2013
http://dx.doi.org/10.1080/02664763.2013.817547

Taylor & Francis
Taylor & Francis Group

Mobile safety cameras: estimating casualty reductions and the demand for secondary healthcare

Lee Fawcett[a]* and Neil Thorpe[b]

[a]*School of Mathematics & Statistics, Newcastle University, Newcastle NE1 7RU, UK;* [b]*Transport Operations Research Group, Newcastle University, Newcastle NE1 7RU, UK*

(*Received 13 February 2013; final version received 17 June 2013*)

We consider a fully Bayesian analysis of road casualty data at 56 designated mobile safety camera sites in the Northumbria Police Force area in the UK. It is well documented that regression to the mean (RTM) can exaggerate the effectiveness of road safety measures and, since the 1980s, an empirical Bayes (EB) estimation framework has become the gold standard for separating real treatment effects from those of RTM. In this paper we suggest some diagnostics to check the assumptions underpinning the standard estimation framework. We also show that, relative to a fully Bayesian treatment, the EB method is over-optimistic when quantifying the variability of estimates of casualty frequency. Implementing a fully Bayesian analysis via Markov chain Monte Carlo also provides a more flexible and complete inferential procedure. We assess the sensitivity of estimates of treatment effectiveness, as well as the expected monetary value of prevention owing to the implementation of the safety cameras, to different model specifications, which include the estimation of trend and the construction of informative priors for some parameters.

Keywords: Markov chain Monte Carlo; mobile safety cameras; negative binomial distribution; Northumbria Safety Camera Partnership; regression to the mean

1. Background

In 2011, official statistics revealed that 203,950 people were reported as injured as a result of road traffic accidents in Great Britain [8]. Of these, 1901 people were killed and 23,122 were seriously injured placing a huge economic and human cost on society. Road casualty reduction is therefore a key aim of government transport policy with new road safety measures continually being tested in an attempt to reduce the number and severity of casualties. Implementing a road safety measure, whether this is a new junction layout, education programmes for young children or new technology to assist in the enforcement of traffic laws, clearly comes at a financial cost. The ability to isolate the effects of the measure on changes in the pattern of casualties is therefore vital – for assessing the potential impacts of alternative accident remedial measures, selecting

*Corresponding author. Email: lee.fawcett@ncl.ac.uk

locations that might benefit from treatment and for evaluating the actual performance of measures once implemented. As we discuss, the effectiveness of such measures is usually assessed via observational studies and, as such, can be prone to numerous sources of error; most notably selection bias, often resulting in *regression to the mean* (RTM).

1.1 *Selection bias in observational studies*

The problem of selection bias in observational studies is well known and well documented. When attempting to assess the effectiveness of a new road safety scheme, for example, sites selected for treatment are often those that have observed, over some pre-determined baseline period, an unusually high number of accidents or casualties; in any subsequent time interval, the accident/casualty count at these sites would probably reduce anyway, even if no treatment was implemented, simply because baseline counts were abnormally high. In such investigations, both ethical and economic concerns are often cited as reasons against completely randomised studies; as a result, post-treatment separation of the true causal effect from any change that would have occurred anyway, without treatment – the RTM effect – has received much attention. The main consequence of ignoring RTM is often an exaggerated treatment effect and possibly unjustified financial investment.

In the road safety literature, an empirical Bayes (EB) approach (see Section 2.1) is usually employed to quantify the RTM effect. Although the effect is variable, studies typically show a reduction in casualty frequency owing to RTM of between 20% and 30% [13]; in other words, estimates which do *not* account for RTM would typically be biased by 20–30%. In fact, so widespread is the acceptance of the selection bias phenomenon and the resulting RTM effect that most studies do very little – if anything at all – to actually check for the presence of selection bias and the appropriateness of the standard procedure used to quantify the RTM effect.

Figure 1 serves to illustrate the discussion on the effects of selection bias thus far. The four scenarios show the possible contributions of RTM to any reduction in casualty frequency as a result of a road safety measure, applied in each scenario at the timepoint shown. In each case, the *counterfactual outcome* shown is obtained as a result of assuming the trend in casualty frequency is the same for locations treated with the safety measure and those that are not (horizontal line at the bottom of the plot), and this trend is assumed constant across the entire timeframe depicted. In scenarios 1, 2 and 3, we can see that treatment has been applied at the peak of a possible ‘blip’ in casualty counts; in these situations, we might expect counts to regress to some underlying average anyway, even without any treatment intervention. In scenario 2, casualty counts after treatment have reverted to their underlying mean level, and no further, suggesting a benefit illusion and no genuine treatment effect. Scenario 1 shows RTM with further reduction, suggesting at least some genuine treatment effect. In scenario 3, the post-treatment casualty count has not even reduced to its underlying mean level, possibly suggesting that the intervention has resulted in an overall *increase* in casualties. Scenario 4 shows no RTM effect; only here might we expect simple before/after comparisons to give unbiased estimates of treatment effectiveness, with the counterfactual outcome being exactly the same as the pre-treatment casualty count.

Of course, RTM is not the only non-treatment effect that can distort the apparent effectiveness of road safety schemes. For example, trends in risk, and changes in exposure to risk, might also contribute to any observed change in casualty frequency, as could improved car safety [3]. We will return to the issue of trend in Section 4.3.

1.2 *Road safety camera policy in the UK*

In 1996, a government report concluded that road safety cameras (notably speed cameras) could be an effective weapon in reducing casualty frequencies [14]. However, the relatively high

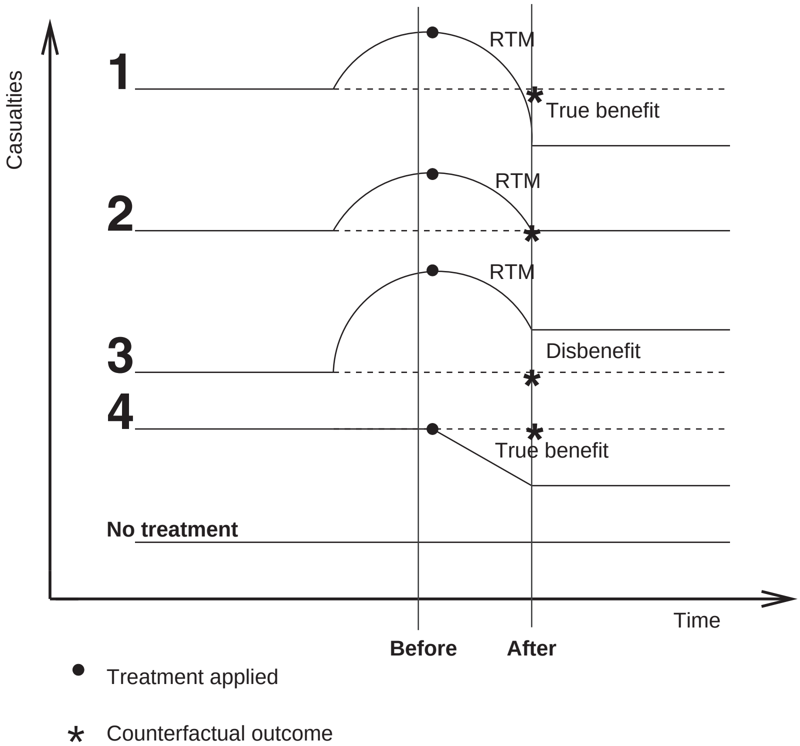

Figure 1. Hypothetical scenarios showing the possible effects of RTM, and a road safety scheme intervention, on casualty frequency.

implementation and running costs were felt to prohibit their widespread deployment under prevailing funding mechanisms. In 1998, the government took the decision to allow traffic authorities to recover the cost of installing and operating speed cameras from the revenues generated from speeding offences detected by the cameras. As a result of the 1996 report and the introduction of the cost recovery approach, the government viewed speed cameras as an important part of its strategy to achieve its casualty reduction targets for 2010 [6]. In April 2000, a two year pilot programme commenced involving eight road safety camera partnerships. Results at the end of the first year prompted the government to take an earlier-than-expected decision to introduce legislation in 2001 to enable national roll-out of safety camera partnerships across Great Britain. By 2004, almost the entire area of England, Scotland and Wales was covered by 42 safety camera partnerships operating under the rules of the cost recovery programme introduced in 2001.

The rapid growth in speed camera activity, and subsequent increase in members of the public being punished for speeding offences, prompted a vigorous and detailed debate over the value of speed cameras in the national media. Opponents trying to discredit the operation of speed cameras (e.g. *SafeSpeed* and *Association of British Drivers*) focussed on a range of issues in an attempt to have the scheme abandoned, in particular disputing the claimed effectiveness of speed cameras as a casualty reduction measure. This tactic brought RTM effects to the forefront of the public debate to the extent that the impact of RTM was discussed in at least one daily national newspaper and eventually was incorporated (in 2005) in the official calculations of casualty reductions at speed camera sites. Cuts in local authority spending have refocussed attention on the claimed effectiveness – and value for money – of speed cameras, as road safety initiatives generally come under close scrutiny *vis-a-vis* other public spending priorities. In August 2010, various organisations in favour of safety cameras, including the *Royal Society for the Prevention of Accidents* and the *AA*, issued a speed camera communiqué claiming that speed cameras save 100 lives each year in the UK and actually 'pay for themselves'; the debate rages on.

1.3 *The Northumbria Safety Camera Partnership*

The Northumbria Safety Camera Partnership (NSCP) joined the national programme in April 2003. In February 2004, the Partnership commissioned a team of researchers to investigate specifically the impact of operating mobile road safety cameras on the demand for secondary health care at the region's hospitals. A mobile camera, as opposed to a fixed speed camera, is a portable unit that is operated from a Partnership vehicle at one of a number of designated sites. The study group collected data from 67 such sites in the region from a before period (April 2001–March 2003) and an after period (April 2004–March 2006).

Following other published studies whose aims were to evaluate the effectiveness of road safety schemes, the NSCP adopted a standard EB procedure (as outlined in Section 2.1 and demonstrated in Section 3.2) to separate real treatment effects at each of the 67 sites from the effects of RTM. The team then attempted to link police accident records to local hospital databases to quantify the cost savings to local National Health Service (NHS) secondary healthcare providers as a result of the implementation of the safety camera scheme. In tariff terms, they estimated the total cost of not having to treat the casualties 'saved' by the introduction of the safety cameras at about £30,000, over the two year treatment period; of course, grossing up to the national level would significantly increase this figure. For full details, see [4].

1.4 *Aims of this paper*

The primary objectives of the current paper are: (1) to consider the appropriateness of the current methodology for estimating the contribution of RTM to casualty reduction, and (2) to find improvements over the standard EB procedure for modelling casualty frequencies. In Section 2 we describe, generally, the standard EB approach for quantifying RTM, and outline the limitations of this approach as it is usually applied. The remainder of this paper then focuses on the safety camera data used in the NSCP-commissioned study. In Section 3, we consider some exploratory checks of the assumptions implicit in the standard approach for quantifying RTM. We then apply the standard EB procedure to the safety camera data and compare this to a fully Bayesian treatment, paying particular attention to the estimates of variability of expected casualty frequencies at each site. We also consider the resulting estimates of cost savings to the NHS, and the wider society, as a result of implementing the safety camera scheme, again comparing both empirical and fully Bayesian approaches. In Section 4, we consider some alternative model structures. In particular, we: assess the sensitivity of our estimates to different prior distributions for the mean casualty rate at each safety camera site; attempt to construct more realistic prior distributions for the regression coefficients used to predict the mean casualty rate; identify the contribution of trend to any change in casualty frequencies at safety camera sites.

2. Statistical modelling: accounting for RTM

2.1 *EB approach*

To date, most road safety scheme evaluation studies that have attempted to quantify the effects of RTM have made use of an EB procedure; see, for example, [12,17,22,24,25]. The model formulation is rather simple and, given a specific choice of prior distribution for the mean casualty rate at each treated site j, the resulting posterior mean for this rate is – conveniently – a weighted sum of the abnormally high *observed* casualty frequency at site j and what we might usually *expect* to see at this site. Assuming a Poisson distribution with mean m_j for the casualty frequency y_j at site j in any period *before* the implementation of a road safety scheme, and a gamma distribution for m_j itself, with mean μ_j and variance μ_j^2/γ, the posterior distribution of $m_j|y_j$ is also of gamma form:

$$m_j|y_j \sim \text{Gamma}\left(\gamma + y_j, \frac{\gamma}{\mu_j} + 1\right).$$

The mean of this posterior is then used as the EB estimate of casualty frequency, i.e.

$$\mathbb{E}(m_j|y_j) = \alpha_j\mu_j + (1 - \alpha_j)y_j, \quad \text{where} \tag{1}$$

$$\alpha_j = \frac{\gamma}{\gamma + \mu_j} \tag{2}$$

and $0 \leq \alpha_j \leq 1$. In the studies we refer to above, evaluation of the effectiveness of a particular road safety scheme at site j is based on a comparison of the observed number of casualties at that site *after* scheme implementation ($y_{j,\text{after}}$) with the EB estimate of casualty frequency for that site given by Equation (1). The percentage change from y_j to $\mathbb{E}(m_j|y_j)$ is taken to be the change that would have happened anyway, even without the implementation of any road safety measure, i.e. the RTM effect.

Generalised linear modelling is usually adopted to obtain the prior mean μ_j, i.e.

$$\hat{\mu}_j = \exp\left\{\hat{\beta}_0 + \sum_{p=1}^{n_p} \hat{\beta}_p x_{p_j}\right\}, \tag{3}$$

where x_{p_j} are variables associated with attributes at site j that could have an effect on the mean number of casualties at that site (e.g. traffic flow or average vehicular speed) and n_p is the number of such variables used. The estimated regression coefficients $\hat{\beta}_i$, $i = 0, \ldots, n_p$, are obtained from a set of reference sites that are representative of the sites at which the road safety scheme has been implemented – in terms of the explanatory variables x_p, but *not* in terms of their casualty frequency. Indeed, sites chosen for a road safety scheme are usually done so on the basis of their unusually high casualty frequency during some baseline period; for μ_j we require a model that will give us a more representative prediction of mean casualty frequency at each treated site j.

Estimation of the weight α_j also requires specification of the shape parameter γ. The unconditional distribution of y_j is negative binomial with mean μ_j and variance $\mu_j + \kappa\mu_j^2$, where $\kappa = 1/\gamma > 0$ is the negative binomial dispersion parameter. Thus, if we assume a negative binomial error structure in model (3), maximum-likelihood estimates of β_i, $i = 0, \ldots, n_p$ and γ can be obtained, leading to estimates of μ_j and α_j via Equations (3) and (2), respectively, and hence EB estimates of casualty frequency at each site j via Equation (1).

2.2 *Limitations of the standard EB approach*

The EB approach for estimating RTM, as outlined in Section 2.1, has become the standard tool for practitioners who wish to evaluate the true effectiveness of a road safety scheme. However, there are several limitations to this approach, as it stands.

2.2.1 *Assumption of exchangeability*

The standard EB approach is flawed if the set of reference sites (from which the estimated regression coefficients in Equation (3) are obtained) are not comparable with the sites at which the road safety scheme has been implemented (to which Equation (3) is applied). Rigorous testing of this assumption is not commonplace, and in this paper (Section 3.1) we consider some pre-analysis exploratory checks of the assumption.

2.2.2 *Choice of prior for m_j*

Although the use of the gamma distribution as a prior for the mean casualty rate m_j gives a convenient and appealing expression for $\mathbb{E}(m_j|y_j)$, this choice of prior distribution is borne out of mathematical convenience, the gamma distribution being the conjugate prior for the Poisson. Developments in computer-based simulation procedures since the first use of the EB method in the 1980s have revolutionised Bayesian modelling, with the result that there is no longer any need nor advantage of using conjugate or other rather artificial forms of prior distribution for m_j. Using Markov chain Monte Carlo (MCMC), we can now (approximately) simulate directly from the posterior distribution of interest no matter what the choice of prior distribution for m_j. In this paper (Section 4.1), we consider the sensitivity of the estimated RTM effect to the choice of prior for m_j in a fully Bayesian analysis.

2.2.3 *Over-optimistic quantification of variability*

By substituting μ_j in Equation (1) with point estimates obtained from Equation (3) it is implied that population-level estimates do not contribute to the uncertainty in the estimate of casualty frequency for a specific site. This is bound to lead to unrealistically low posterior standard deviations for EB estimates of casualty frequency. A fully Bayesian analysis would assign hyper-prior distributions to the regression coefficients β_i and the prior shape parameter γ; doing so explicitly recognises that population-level estimates of casualty frequencies are also uncertain and thus contribute to the variance of the site-level estimates of m_j. In this paper, we compare the standard EB analysis to fully Bayesian analyses, paying attention to the resulting estimates of posterior variability of certain quantities of interest. Unlike previous studies, we also consider the effects of using more informative priors for the regression coefficients in Equation (3).

3. Application to the NSCP safety camera data

At each treated site the number of casualties before, and after, the implementation of the safety cameras was observed, as well as various explanatory variables collected by the NSCP as part of their standard reporting procedure: average observed speed (x_1 miles per hour); percentage of drivers exceeding the speed limit (x_2); daily traffic flow (x_3); speed limit (x_4 miles per hour); eighty-fifth percentile speed (x_5 miles per hour); percentage of drivers at least 15 miles per hour over the limit (x_6); and road classification and road type ($x_7 =$ A/B/C/Unclassified roads, and $x_8 =$ single/dual/mixed carriageway, respectively).

For the purpose of the present paper, we work with a subset of 56 of the 67 mobile camera sites in the original study (due to missing values and some sites being decommissioned during the study period). In terms of the explanatory variables listed above, these sites are extremely varied; however, all sites have in common their unusually high casualty records from the observational period. The total number of casualties in the before period was 436; after the implementation of the safety cameras, this reduced to 298. To formulate a regression equation for μ_j to be used at each treated site j (Equation (3)), we also have data available for a set of 67 reference sites in the Northumbria police force area.

3.1 ***Exploratory analyses to check the RTM assumptions***

As discussed in Section 1.1, most published studies make use of the EB procedure without checking whether or not we can expect RTM to be present. Figure 2 shows total casualty frequencies at

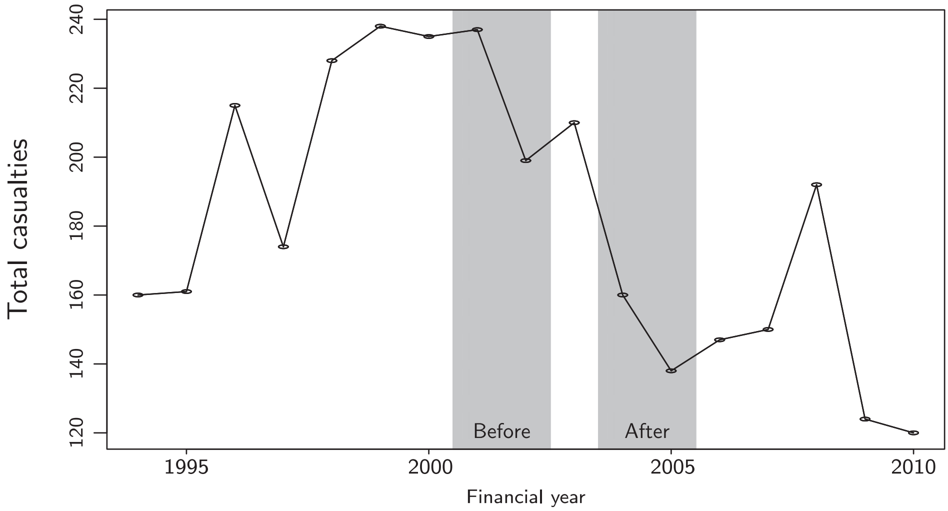

Figure 2. Time series plot of total casualty frequencies at the 56 mobile safety camera sites: 1995, for example, corresponds to the financial year April 1995 → March 1996.

the safety camera sites for the financial years 1994/1995–2010/2011. As can be seen, the safety cameras were commissioned after what may have been an unusually high run of casualties. Implementing Kendall's turning point test [15] gives a p-value of 0.180 for the hypothesis H_0 : Casualty frequencies are unsystematic, suggesting that what we see in Figure 2 could indeed be a 'blip'; the expectation of RTM thus seems feasible.

In Section 2.2.1 we also discussed that the standard procedure for estimating RTM is flawed if the sites treated with safety cameras cannot be considered exchangeable with sites in the reference set. We now suggest some simple methods practitioners could use to check this assumption. Consider the matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_8)$, where $\mathbf{x}_p$, $p = 1, \ldots, 8$, are column vectors of observations from the explanatory variables $x_1, \ldots, x_8$ (described above) of length 123, with each of the 56 treated and 67 reference sites having a row entry. Combining treated and reference sites into a single data matrix, we can then perform a multiple factor analysis [10] on $\mathbf{X}$; plotting scores on the first two dimensions against each other, using different plotting symbols for treated and reference sites, could help determine whether or not the assumption of exchangeability is plausible. Figure 3 shows this plot (which can be interpreted in the same way as a plot of scores from the first two principal components in a principal components analysis). As can be seen, there is no clear distinction between scores for the reference and treated sites.

To further investigate the plausibility of our assumption of exchangeability, we can perform permutation tests for various statistics that serve to compare the treated and reference sites. For example, suppose we wish to compare mean values of our explanatory variables at the treated sites to the corresponding averages observed in the reference set. The absolute differences are given by

$$\delta_p = |\bar{x}_p^{\mathrm{TRT}} - \bar{x}_p^{\mathrm{REF}}|, \quad p = 1, \ldots, 8, \tag{4}$$

where TRT and REF denote the treated and reference sets, respectively. If the treated and reference sites *are* exchangeable with respect to the explanatory variables, then the values calculated from Equation (4) would not be significantly different to those obtained after a random allocation of sites to the treated and reference sets. In a permutation test we find the statistic of interest for every possible allocation of sites to each of the reference and treatment sets; the exact p-value P for a test of the null hypothesis H_0 : sites are exchangeable, is then found as the proportion of allocations

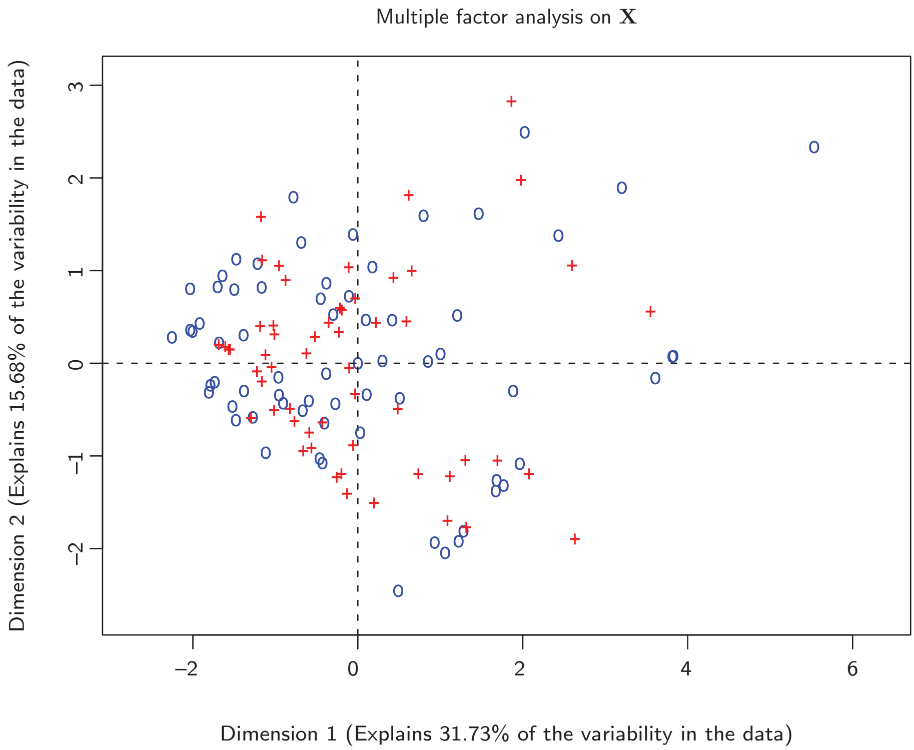

Figure 3. Multiple factor analysis on **X**: scores from the first dimension plotted against those from the second dimension for reference sites (**o**) and treated sites (**+**).

for whom the statistic of interest is at least as extreme as that found under the real allocation. In our example, there are $\binom{123}{56} \approx 4.68 \times 10^{35}$ permutations: thus, a Monte Carlo permutation test can be used, where N permutations are chosen randomly from all of those available. Let Π_k be one such random permutation, and let $\delta_p^{(\Pi_k)}$ be the absolute mean difference for variable x_p, as defined in Equation (4), obtained under permutation Π_k. Let I_k be an indicator variable such that

$$I_k = \begin{cases} 1 & \text{if } \delta_p^{(\Pi_k)} \geq \delta_p, \\ 0 & \text{otherwise;} \end{cases}$$

then an estimate of P is given by

$$\widehat{P} = \sum_{k=1}^{N} \frac{I_k}{N},$$

giving $\mathbb{E}[\widehat{P}] = P$ and $\mathsf{var}(\widehat{P}) = P(1-P)/N$. Performing such Monte Carlo permutation tests with $N = 10^6$ shows only a marginally significant difference between the reference and treated sets for x_5, the eighty-fifth percentile speed: here, the 95% confidence interval for P is (0.048, 0.051). For all other explanatory variables, $P \gg 0.05$.

Considering all explanatory variables together, we also perform a Monte Carlo permutation test on the mean Mahalanobis distance of sites in the treatment set to sites in the reference set with mean $\bar{\mathbf{M}}^{\text{REF}} = (\bar{x}_1^{\text{REF}}, \ldots, \bar{x}_8^{\text{REF}})$ and covariance matrix Σ, whose (s,t)th entry is given by

$\text{cov}(x_s^{\text{REF}}, x_t^{\text{REF}})$, $s, t = 1, \ldots, 8$; that is, we compare

$$\bar{D} = \frac{1}{56} \sum_{j=1}^{56} \sqrt{(\mathbf{X}_j^{\text{TRT}} - \bar{\mathbf{M}}^{\text{REF}})^T \Sigma^{-1} (\mathbf{X}_j^{\text{TRT}} - \bar{\mathbf{M}}^{\text{REF}})},$$

where $\mathbf{X}_j^{\text{TRT}}$ is the jth row of $\mathbf{X}$ for those sites treated with a safety camera, to a sample of size N from the permutation distribution of $\bar{D}$. Doing so, we get (0.165, 0.173) as the 95% confidence interval for P, further supporting the assumption of site exchangeability.

3.2 *EB analysis*

Having checked the assumption of exchangeability between the reference and treated sites, we now apply the EB method, as outlined in Section 2.1, to the data collected by the NSCP. Fitting the full model – to include information on *all* explanatory variables – reveals potential problems of multicollinearity by including variable x_5. After removing x_5, and using a backwards elimination procedure for the selection of suitable explanatory variables, we obtain the following model for data at the 67 reference sites:

$$\hat{\mu} = \exp\left\{ \underset{(0.534)}{1.93} - \underset{(0.015)}{0.04}x_1 - \underset{(0.004)}{0.01}x_2 + \underset{(0.193)}{0.44}x_3 + \underset{(0.382)}{0.67}x_{4I} + \underset{(0.422)}{0.85}x_{5I} + \underset{(0.380)}{1.06}x_{6I} \right\}, \quad (5)$$

where x_1, x_2 and x_3 correspond to the average observed speed (miles per hour), the percentage of drivers exceeding the speed limit and traffic flow (respectively, as defined earlier), and x_{4I}, x_{5I} and x_{6I} are indicator variables associated with road classification (variable x_7), where: $x_{4I} = 1$ for road classification ‘A’, $x_{5I} = 1$ for road classification ‘B’ and $x_{6I} = 1$ for road classification ‘C’, each taking the value 0 otherwise; standard errors for each of the estimated regression coefficients are given in parentheses underneath. The maximum-likelihood estimate for the negative binomial dispersion parameter is $\hat{\kappa} = 0.401$ (0.015), giving $\hat{\gamma} = 1/0.401 = 2.494$ (0.774); again, standard errors are given in parentheses. There were no significant pairwise dependencies between any of the remaining explanatory variables, suggesting no real problems of multicollinearity at this stage of the analysis. The usual diagnostic tools for assessing the fit of such regression models can be used to confirm the adequacy of this model for data collected at the reference group of sites. The standard EB procedure uses Equation (5) on data x_{1_j}, x_{2_j}, x_{3_j}, x_{4I_j}, x_{5I_j} and x_{6I_j} collected at each safety camera site j, $j = 1, \ldots, 56$, and treats each resulting $\hat{\mu}_j$ as the ‘true value’, substituting this into Equation (1) along with $\hat{\gamma}$ and the observed casualty frequency in the before period (y_j) to obtain the EB estimate of casualty frequency at each site j.

Table 1 shows numerical results for some of the individual safety camera sites (chosen because their observed casualty frequencies and estimated RTM effects give a good indication of the

Table 1. Results of the EB analysis to account for RTM for four sites treated with safety cameras, as well as totals for all 56 safety camera sites.

	EB method						Difference	
	y_j	μ_j	α_j	$\mathbb{E}(m_j \mid y_j)$	$\text{SD}(m_j \mid y_j)$	$y_{j,\text{after}}$	Observed	After RTM
Site $j = 2$	4	1.43	0.63	2.38	0.936	0	−4	−2.38
Site $j = 13$	12	1.71	0.59	5.95	1.564	2	−10	−3.95
Site $j = 39$	7	1.31	0.65	3.29	1.069	2	−5	−1.29
Site $j = 47$	16	7.84	0.24	14.06	3.273	5	−11	−9.06
Total	436			321		298	−138	−23

range of values observed across all treated sites), as well as overall totals for all sites in the NSCP study. For example, at site 13 there was an observed reduction in casualties from 12 in the before period to 2 in the after period; however, the EB estimate suggests that this would have reduced to about 6 anyway, giving a more realistic reduction, after RTM, of just 4 casualties. Across all 56 safety camera sites, the total observed reduction of 138 casualties between the before and after periods is reduced to just 23 after taking RTM into account. This suggests an RTM effect of $100(\sum_{\forall j}\mathbb{E}(m_j|y_j)-y_j)/\sum_{\forall j}y_j=-26.4\%$, towards the middle of the range reported in [13], discussed in Section 1.1.

3.3 *Fully Bayesian analysis*

We now formulate a fully Bayesian (FB) modelling framework to assess the effectiveness of the safety cameras. In this section, we work with exactly the same Poisson–Gamma hierarchy as outlined in Section 2.1. However, we now unify the entire modelling procedure by assigning prior distributions to the regression coefficients β_i, $i=0,\ldots,6$, and the negative binomial dispersion parameter κ. We use diffuse, independent priors:

$$\beta_i \sim N(0,100), \quad i=0,\ldots,6 \text{ and}$$
$$\rho=\log(\kappa)\sim N(0,100),$$

working with the natural logarithm of the negative binomial dispersion parameter to retain the positivity of κ. Inference proceeds by initialising each of the regression coefficients β_i and $\rho=\log(\kappa)$ at their prior means, and then using a random walk Metropolis–Hastings scheme (with data from the reference set) to update the chains. At each iteration R in the MCMC, the current values of the regression coefficients $\beta_i^{(R)}$ are used to obtain the posterior draw $\mu_j^{(R)}$ for each safety camera site j via Equation (3); the current values $\mu_j^{(R)}$ and $\gamma^{(R)}=\exp\{-\rho^{(R)}\}$ are then used as the mean and shape (respectively) of the gamma prior distribution for m_j. Since the gamma distribution is the conjugate prior for the Poisson distribution, Gibbs sampling can then be used to sample from the full conditional for m_j. After initial pilot runs to tune the efficiency of the sampler, the MCMC was run for 500,000 iterations, and various starting values for β_i and ρ were used to help assess convergence.

Table 2 shows some posterior summaries after the removal of the burn-in period (the first 5000 iterations). Also shown are the corresponding posterior summaries for μ_j and m_j for the four safety camera sites we reported in the EB analysis (Table 1). At each iteration R we have also computed the total expected casualties $T^{(R)}$ across all 56 sites, given by

$$T^{(R)}=\sum_{j=1}^{56} m_j^{(R)}|y_j.$$

The posterior means for the regression coefficients (β_i) match up quite closely to the MLEs for these parameters as given in Equation (5); none of the 95% credible intervals for these parameters include zero, which also agrees with the earlier frequentist analysis in which the regression coefficients were all significant. The posterior means for β_i are all quite close to the medians, indicating fairly symmetric marginal posteriors. The posterior means for m_j for the four sites reported here also compare quite closely to the posterior means for m_j obtained in the EB analysis and reported in Table 1; however, as can be seen when comparing the standard deviations for m_j in Table 2 to those from the EB analysis in Table 1, posterior variability is substantially greater in the FB analysis. This is, of course, because we have now acknowledged the variability of the regression parameters β_i and hence the mean casualty rate μ_j, $j=1,\ldots,56$, through the prior

Table 2. Posterior summaries for a fully Bayesian analysis to account for RTM.

		Posterior			
		Mean	St. dev.	Median	95% Credible interval
	β_0	1.981	0.544	1.981	(0.900, 3.044)
	β_1	−0.042	0.016	−0.042	(−0.074, −0.012)
	β_2	−0.013	0.004	−0.013	(−0.021, −0.005)
	β_3	0.476	0.218	0.474	(0.063, 0.912)
	β_4	0.648	0.437	0.651	(0.012, 1.506)
	β_5	0.840	0.451	0.839	(0.043, 1.733)
	β_6	1.059	0.400	1.058	(0.284, 1.845)
	$\gamma = \exp\{-\rho\}$	2.281	0.756	2.146	(1.201, 4.111)
μ_j	Site $j = 2$	1.534	0.639	1.416	(0.649, 3.105)
	Site $j = 13$	1.903	0.908	1.716	(0.725, 4.209)
	Site $j = 39$	1.411	0.601	1.301	(0.586, 2.907)
	Site $j = 47$	8.195	1.767	7.986	(5.355, 12.245)
m_j	Site $j = 2$	2.465	1.188	2.255	(0.780, 5.369)
	Site $j = 13$	6.283	2.447	5.946	(2.502, 12.039)
	Site $j = 39$	3.475	1.527	3.239	(1.232, 7.133)
	Site $j = 47$	14.225	3.466	11.032	(8.320, 21.842)
	Total T	322	23.833	308	(289.92, 369.97)

Notes: The posterior means and medians of the regression parameters β_i are comparable to the MLEs from the empirical Bayes analysis; posterior summaries for the casualty rate at site j (m_j) and its corresponding mean (μ_j) for the four sites reported here can also be compared to those from the earlier empirical Bayes analysis.

distributions for β_i. Although our prior distributions for the regression coefficients are rather vague, and so probably act to *over*-state these sources of variability, this is still potentially more attractive than assuming that the MLEs for β_i (and hence μ_j) are the true values, as is the case in the EB approach.

The positive skew of the marginal posteriors for m_j is noticeable when we inspect the posterior draws from the FB analysis; for four sites in particular, this gives a substantial difference between the posterior mean and median (site 47, as reported in Table 2, is one such site). Studies that we are aware of in the road safety literature do not consider any posterior summary other than the mean given by Equation (1). Clearly, for locations such as site 47, this could be misleading. Of course, this is not to say that other posterior summaries from an EB analysis cannot be obtained; rather, in a *typical* EB analysis the posterior mean is taken as *the* estimate of mean casualty frequency, without any regard to the shape of the posterior distribution itself. Working within a fully Bayesian framework gives access to *any* posterior summary via our MCMC sample. Figure 4 shows posterior summaries for m_j for *all* 56 safety camera sites. The effect of using the mean at the sites with a relatively large mean/median discrepancy (as is done in a typical EB approach) could be to understate the effect of RTM at those sites: as the bottom row of Table 2 shows, when we sum casualty frequencies for all sites at each iteration in the MCMC, the posterior median total expected casualty frequency is just 308, compared to a mean of 322 (and an analytical mean, as found in the EB analysis, of 321, as given in Table 1).

3.4 *Implications for the demand on secondary healthcare*

One of the aims of the original NSCP study was to estimate the financial consequences of the implementation of the safety cameras for local NHS secondary healthcare providers. A large, multi-stage data linking exercise took place in the NSCP study to link police collision data

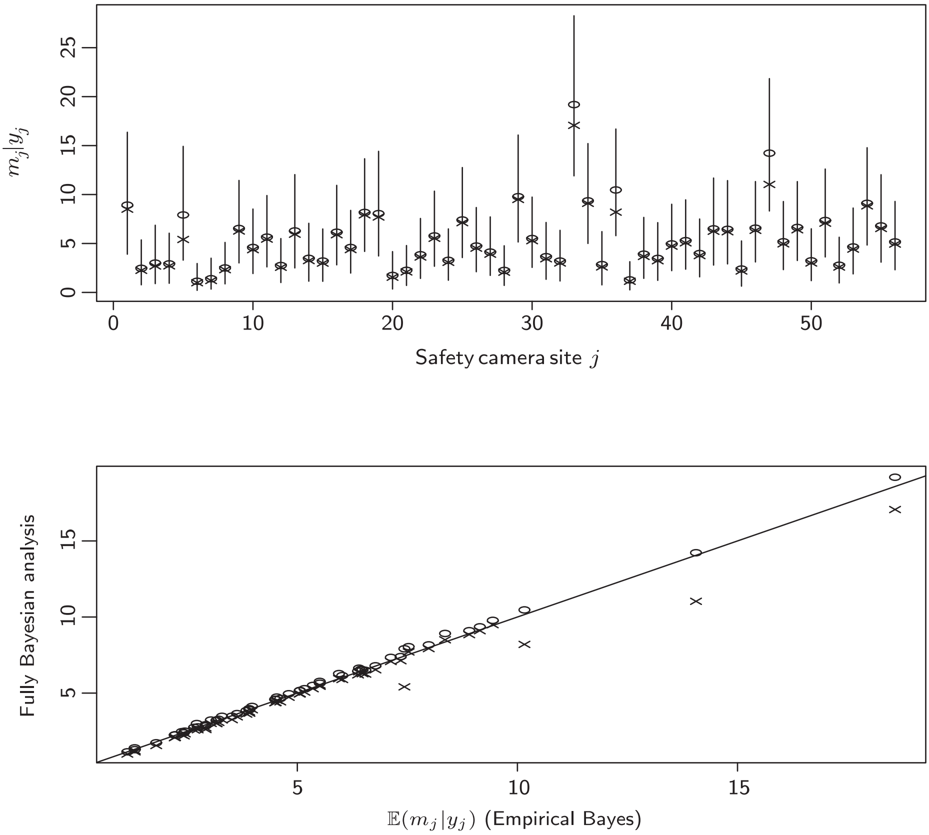

Figure 4. Posterior means (circles) and medians (crosses) from the fully Bayesian analysis: along with 95% credible intervals, for each safety camera site (top); plotted against the corresponding EB estimate of casualty frequency (bottom).

from the before period to NHS casualty data. This was attempted using unique identifiers which are collected by both the police, at the scene of the accident, and the hospital involved. Rather disappointingly, it was only possible to match about 44% of casualties in the before period to corresponding hospital records. However, as pointed out in [4], this is in line with matching rates obtained in other studies that have attempted to link police records to hospital databases; for example, in a similar exercise, Simpson [26] achieved a 46% success rate. Reasons for differences between hospital and police data, which can lead to low matching rates, are suggested in [5].

NHS Accident and Emergency (A&E) patients fall into one of eight Health Resource Group (HRG) categories depending on the severity of their injuries and the extent of treatment required. These A&E HRGs range from high cost categories which include computerised tomography scans and magnetic resonance imaging scans, to relatively low cost categories, involving more routine urine/bacteriological investigations. Some A&E patients would then require admission to hospital for inpatient treatment; inpatients are allocated to one of some 700 inpatient HRGs (see [4] for full details). Each A&E HRG has an associated financial tariff, as does each inpatient HRG. Overall individual inpatient tariffs are calculated as a function of time, and so instead of considering the 700 inpatient HRGs themselves we discretise the inpatient tariffs into groups of £500 (where £0 is used for A&E admissions not requiring inpatient treatment).

Each hospital admission will fall into one of our A&E HRG/inpatient tariff category combinations τ; thus, we consider each τ, with associated combined A&E/inpatient financial tariff £C_τ, as a multinomial outcome whose probabilities p_τ are just the observed proportions falling into each τ in the before period. The estimated saving to the NHS by implementing the safety cameras can then be obtained, in the EB analysis, by multiplying the total change in casualty frequency after RTM by each p_τ; these expected frequencies are then converted into expected financial savings by multiplying by the financial tariff associated with each corresponding category τ, and so the total expected financial saving £S is found as

$$\mathbb{E}(S) = \left(\sum_{j=1}^{56} \left\{ \mathbb{E}(m_j|y_j) - y_{j,\text{after}} \right\} \right) \sum_{\forall \tau} p_\tau C_\tau. \tag{6}$$

The A&E contributions to each C_τ take a fixed value; however, since we have partitioned the inpatient tariffs into groups of £500, we use the minimum, midpoint and maximum values of each inpatient tariff to obtain a combined tariff £C_τ for each τ. Here, we assume that casualties involving the police, but not requiring any hospital treatment at all, cost £$C_\tau = 0$; below, we discuss an alternative approach which attempts to incorporate the cost of hospital treatment with all other costs that we might expect to be incurred, per casualty.

In the FB analysis we can obtain posterior draws for the expected number of casualties falling into each category τ, by multiplying each posterior draw for the total change in casualty frequency by the corresponding p_τ. Then, posterior draws for the expected financial saving to the NHS can be obtained by multiplying each expected frequency by the associated financial tariff for category τ, giving, at each iteration R,

$$S^{(R)} = \left(T^{(R)} - \sum_{j=1}^{56} y_{j,\text{after}} \right) \sum_{\forall \tau} p_\tau C_\tau. \tag{7}$$

Table 3 compares results from the standard EB analysis to posterior summaries from the FB approach. Of course, all values represent cost savings to the NHS over the two year operational period of the safety cameras in this study. Although it could be the case that there is an association between the severity of a casualty, and whether or not the police/hospital records of this casualty are successfully matched, it is difficult to understand the effect this might have on estimates of S: there is no direct relationship between the police classification of casualty severity and the associated HRG tariff category τ.

Since 1993, the valuation of casualties has been based on a consistent willingness to pay approach, which encompasses *all* aspects of the valuation of casualties, including police costs, human costs (representing pain, grief and suffering to the casualty and their relatives/friends) and loss of output due to injury. When combined with the direct cost of medical treatment, the UK DfT [7] puts the average total cost of a road casualty at £52,850. Thus, replacing $\sum_{\forall\tau} p_\tau C_\tau$ in

Table 3. Estimates of total expected financial savings to the NHS (£S thousand), and total value of prevention including human costs and lost output (£S^* thousand), owing to the implementation of the safety camera scheme in Northumbria: April 2004–March 2006.

Thousand £		EB	Full Bayes: Posterior Mean	St. dev.	Median	95% credible interval
	Midpoint	25.6	24.9	13.2	24.4	(0.3, 57.5)
S	Minimum	23.5	22.8	12.1	22.3	(0.1, 52.5)
	Maximum	27.7	27.1	14.4	26.5	(0.6, 62.5)
S^*		1215.6	1529.8	786.3	1479.8	(45.6, 4122.3)

Equations (6) and (7) with 52850 will give an EB estimate/posterior draws (respectively) of S^*, the *total* (average) value of prevention owing to the implementation of the safety cameras in the Northumbria region. The value of £52,850 per prevented casualty is, of course, an average; this will vary depending on the severity of the casualty's injuries, and the type of road user (e.g. car occupant, goods vehicle occupant, motorcycle user etc.). Results for S^* are given in the bottom row of Table 3.

4. Further modelling considerations

In the previous section, the standard EB procedure was used to estimate the contribution of RTM to the reduction in casualty frequency at 56 sites treated with safety cameras. Relative to an FB analysis, the standard EB approach was shown to be over-optimistic in its estimation of the variability of casualty frequency at each site. We provided a like-for-like comparison between the typical EB approach for estimating the RTM effect and an FB analysis. We now investigate the sensitivity of our results from this FB treatment to the choice of prior distributions for m_j. We then investigate the use of more informative priors for the regression coefficients in Equation (3), giving some thought to the issue of variable selection in the Bayesian framework. We also quantify the contribution of trend to the reduction in casualty figures after the implementation of the safety cameras.

4.1 *Sensitivity to other forms of prior distribution for m_j*

We now assess the sensitivity of the results in Section 3 to the choice of prior distribution for m_j. Recall that the gamma prior mean and variance for m_j were μ_j and μ_j^2/γ, respectively. We now use the lognormal (mean $= \lambda_j$, variance $= \sigma^2$) and Weibull (shape $= \omega$, scale $= \nu_j$) distributions as priors, keeping the mean and variances the same as in the original gamma prior to allow relative comparisons of the effects of using these different priors. This gives

$$\lambda_j = \log(\mu_j) - \tfrac{1}{2}\log(1+\gamma^{-1}) \quad \text{and}$$
$$\sigma^2 = \log(1+\gamma^{-1})$$

for the lognormal prior. For the Weibull prior, we solve

$$\omega\frac{\Gamma(2\omega^{-1})}{\Gamma^2(\omega^{-1})} = \frac{1}{2}(1+\gamma^{-1})$$

for ω and then use

$$\nu_j = \frac{\mu_j}{\Gamma(1+\omega^{-1})}.$$

A summary of results based on the lognormal and Weibull priors for m_j is shown in Table 4, along with the results for the gamma prior from Section 3. There is clearly some agreement between results obtained using the original gamma prior and the Weibull prior. However, using the lognormal prior gives a considerably higher total number of expected casualties. For example, the posterior mean total casualties is 355; comparing this with the number of observed casualties in the after period (298) would suggest the safety cameras had been more effective than if we had used the gamma or Weibull priors, with less contribution to any change in casualty frequency attributed to RTM. In fact, the 95% credible interval for the effect of RTM when using the lognormal prior does not include the value given in the EB analysis, or indeed the posterior summaries of average from the analyses using the other two forms of prior for m_j. This is reflected in the estimation of S and S^* with, for example, the median total value of prevention due to the safety cameras, when

Table 4. Posterior summaries for: the total expected number of casualties (T); the RTM; the expected financial savings to the NHS secondary healthcare providers as a result of implementing the safety cameras (S); and the total average value of prevention including human costs and lost output (S^*).

		Gamma		Lognormal		Weibull	
	EB	Mean	Median	Mean	Median	Mean	Median
		(95% credible interval)		(95% credible interval)		(95% credible interval)	
T	321	322	308	355	338	317	303
		(290, 370)		(309, 394)		(296, 371)	
RTM (%)	−26.4	−26.5	−29.7	−18.9	−22.8	−27.6	−30.8
		(−35.6 , −14.2)		(−26.3, −9.0)		(−39.3, −15.3)	
S (Thousand £)	25.6	24.9	24.4	29.3	29.3	25.3	24.9
		(0.3, 57.5)		(6.1, 73.5)		(0.7, 70.9)	
S^* (Thousand £)	1215.6	1529.8	1479.8	2803.0	2801.0	986.3	951.3
		(45.6, 4122.3)		(581.4, 5126.5)		(69.1, 4910.9)	

using the lognormal prior (about £2.8 million), being almost twice that when using the gamma prior (£1.5 million) and even greater still than when using the Weibull prior (less than £1 million).

The deviance information criterion (DIC), as discussed in [27], can be used to compare the three model formulations tried here. The DIC is akin to the Akaike information criterion and is based on an estimate of the log-likelihood, but includes a penalty for the number of parameters; hence, it can be used to compare alternative models. For the Poisson–gamma, Poisson–lognormal and Poisson–Weibull, we have a DIC of 693.3, 787.2 and 645.6 (respectively), suggesting the Weibull prior for m_j might be the most appropriate distribution (of the three tried) to use here. In practical terms, the lognormal prior might be the least suitable because of its heavy upper tail relative to the other two priors tried. For example, taking $\hat{\gamma} = 2.494$ (as estimated in Section 3.2) and $\hat{\mu}_{47} = 7.84$ (as estimated for site 47 in Table 1), gives a value for $\Pr(m_j > 40)$ under the lognormal prior that is 7 times larger than that under the gamma prior, and 77 times larger than that under the Weibull prior! Even for a relatively high casualty frequency site like site 47, it would be extremely unusual to observe a mean casualty rate of more than 40, and so we might trust most the prior with smaller tail probabilities (the Weibull).

4.2 *Choice of prior distribution for the regression coefficients*

The FB analysis in Section 3.3 used independent Normal priors with large variances for the regression coefficients β_i. We now consider prior specifications for the regression coefficients that more suitably capture the variability of these parameters, as well as any dependencies between them. We consider two forms of prior for the regression coefficients: a genuinely informative prior developed by eliciting priors for the mean number of casualties at observed levels of covariates, known as a *conditional mean prior* (CMP); and a *data augmentation prior* (DAP), making use of the covariate values themselves.

4.2.1 *Using a DAP*

An attractive method for augmenting a Bayesian analysis in the absence of any external or expert prior information is to adopt a DAP. For Poisson regression, Ntzoufras [23] suggests the following n_p-variate Normal DAP for $\boldsymbol{\beta}_{\setminus 0,n_p}$, the n_p-dimensional parameter vector of regression coefficients β_i excluding β_0:

$$\boldsymbol{\beta}_{\setminus 0,n_p} \sim N_{n_p}(\mathbf{0}, n(\mathbf{X}_{n_p}^{\mathrm{T}}\mathbf{X}_{n_p})^{-1}). \tag{8}$$

Here, $\mathbf{X}_{n_p}$ is the design matrix without the first column that corresponds to the constant β_0, $\mathbf{0}$ is the zero vector (of length n_p) and n is the sample size. Ntzoufras [23] suggests using a vague prior for β_0 such as that previously used, e.g. a zero mean Normal with large variance. Using the prior for $\boldsymbol{\beta}_{\backslash 0,n_p}$ given in Equation (8), and using gamma, lognormal and Weibull priors for m_j, $j = 1, \ldots, 56$, gives DIC values of 733.4, 813.7 and 685.4 (respectively), indicating, as before, that the Poisson–Weibull structure provides the best fit. The columns in Table 6 labelled 'DAP' summarise the MCMC runs for the Poisson–Weibull model; as before, the chains were allowed to run for 500,000 iterations, and the first 5000 were discarded as burn-in. Comparing inferences for the regression coefficients β_i to those in Table 2, where independent zero-mean Normal priors with large variances were used, we see a reduction in posterior variability. This follows through to inferences for the total number of expected casualties (T), as well as the RTM effect, both of these having narrower 95% credible intervals than when the non-informative priors were used (see the right-hand side of Table 4); similarly, the 95% credible intervals for S and S^* are narrower when using the DAP.

4.2.2 *Using a CMP*

We now attempt to construct a truly informative prior for the vector of regression coefficients $\boldsymbol{\beta}_{\backslash 0,n_p}$. Following [2], we elicit a prior on $\tilde{\mathbf{M}} = (\tilde{M}_1, \ldots, \tilde{M}_{n_p})$ where the $\tilde{M}_p$'s are mean responses at covariates $\mathbf{x}_p$, $p = 1, \ldots, n_p$. We denote by $\tilde{\mathbf{X}}$ the $n_p \times n_p$ non-singular matrix with $\mathbf{x}_p^T$ in the pth row (see Section 3.1). Following the notation of Bedrick *et al*. [2], $\mathbf{G}$ and $\mathbf{G}^{-1}$ are vector transformations that apply g and g^{-1} to each element; for example, $g(\cdot) = \log(\cdot)$ and $g(\cdot) = \text{logit}(\cdot)$ for Poisson and logistic regression, respectively. Assessing the $\tilde{M}_p$'s independently, the CMP is

$$\pi_0(\tilde{\mathbf{M}}) = \prod_{p=1}^{n_p} \pi_{0_p}(\tilde{M}_p).$$

Writing

$$\tilde{\mathbf{M}} = \mathbf{G}^{-1}(\tilde{\mathbf{X}}\boldsymbol{\beta}_{\backslash 0,n_p}) \quad \text{and} \quad \boldsymbol{\beta}_{\backslash 0,n_p} = \tilde{\mathbf{X}}^{-1}\mathbf{G}(\tilde{\mathbf{M}})$$

induces a prior on $\boldsymbol{\beta}_{\backslash 0,n_p}$ of the form

$$\pi(\boldsymbol{\beta}_{\backslash 0,n_p}) = \frac{\prod_{p=1}^{n_p} \pi_{0_p} g^{-1}(\tilde{\mathbf{x}}_p^{\mathrm{T}}\boldsymbol{\beta}_{\backslash 0,n_p})}{|\tilde{\mathbf{X}}^{-1}| \prod_{p=1}^{n_p} \dot{g}(\tilde{M}_p)},$$

where, generically, $\dot{f}(x) = \partial f(x)/\partial x$.

Using the same four covariates as in the original analyses in Section 3 requires us to elicit priors on $\tilde{\mathbf{M}} = (\tilde{M}_1, \ldots, \tilde{M}_6)$, since road classification is a factorial variable requiring three indicators. A regression analysis from a previous study of casualty frequencies at another group of sites in the Northumbria region gives a regression equation of the form in Equation (5); covariates $\mathbf{x}_p$, $p = 1, \ldots, 6$, at six of these sites can be used to suggest means a_p and variances b_p for each $\tilde{M}_p$. Suitable priors for $\tilde{M}_p$ can then be proposed – for example, gamma distributions with means and variances a_p and b_p (respectively), or indeed lognormal or Weibull distributions as used in Section 4.1. For six sites used in this previous study, we obtain means and variances as reported in the first two columns of Table 5. Using these means and variances we can obtain hyper-parameters for the priors on $\tilde{M}_p$: these are also given in Table 5 for the three priors we consider.

Using the DIC to assess fit when using gamma, lognormal and Weibull distributions for π_{0_p} and m_j, $p = 1, \ldots, 6$ and $j = 1, \ldots, 56$, once again suggests the Poisson–Weibull structure is best. Thus, Table 6 also reports posterior summaries from a Poisson–Weibull analysis using the CMP for

Table 5. Elicited prior parameters for $\tilde{M}_p$, as used in the CMP prior for $\boldsymbol{\beta}_{\backslash 0,n_p}$.

	a_p	b_p	Gamma		Lognormal		Weibull	
			Shape	Scale	Mean	Variance	Shape	Scale
$\tilde{M}_1$	9.93	11.63	8.48	0.85	2.24	0.11	3.20	11.09
$\tilde{M}_2$	2.77	1.64	4.68	1.69	0.92	0.19	2.29	3.13
$\tilde{M}_3$	3.59	3.10	4.16	1.16	1.17	0.22	2.15	4.05
$\tilde{M}_4$	3.12	1.87	5.21	1.67	1.05	0.18	2.43	3.52
$\tilde{M}_5$	8.19	6.25	10.73	1.31	2.06	0.09	3.64	9.08
$\tilde{M}_6$	5.42	3.79	7.75	1.43	1.63	0.12	3.04	6.07

Table 6. Posterior summaries for: the regression coefficients β_p; the total expected number of casualties across all safety camera sites T; the RTM; the estimated financial savings to NHS secondary healthcare providers as a result of implementing the safety cameras (£S thousand); and the total average value of prevention including human costs and lost output (£S^* thousand).

	Posterior							
	Mean		St. dev.		Median		95% Credible interval	
	DAP	CMP	DAP	CMP	DAP	CMP	DAP	CMP
β_0	1.785	1.543	0.463	0.521	1.790	1.549	(0.86, 2.69)	(0.51, 2.58)
β_1	−0.038	−0.025	0.014	0.013	−0.038	−0.025	(−0.07, −0.01)	(−0.05, 0.00)
β_2	−0.012	−0.013	0.004	0.004	−0.012	−0.013	(−0.02, −0.01)	(−0.02, −0.01)
β_3	0.464	0.318	0.217	0.188	0.458	0.313	(0.06, 0.90)	(0.04, 0.70)
β_4	0.701	0.799	0.426	0.412	0.701	0.802	(0.01, 1.55)	(0.03, 1.61)
β_5	0.886	0.913	0.442	0.432	0.882	0.913	(0.03, 1.77)	(0.06, 1.76)
β_6	1.101	1.142	0.385	0.395	1.097	1.140	(0.35, 1.87)	(0.37, 1.93)
γ	2.194	2.083	0.725	0.671	2.072	1.969	(1.16, 3.95)	(1.12, 3.71)
T	327	327	23.316	25.528	314	313	(286, 352)	(285, 354)
RTM %	−25.3	−25.3	6.265	5.855	−28.3	−28.5	(−35.3, −26.0)	(−36.6, −23.6)
S	31.4	30.9	15.128	14.092	31.0	30.9	(0.9, 61.9)	(0.9, 58.5)
S^*	1584.9	1541.5	448.4	349.1	1547.7	1541.7	(73.4, 4507.3)	(72.6, 4183.2)

Notes: Results are shown for the data augmentation prior (DAP) and the conditional mean prior (CMP) for $\boldsymbol{\beta}_{\backslash 0,n_p}$, using Weibull priors on m_j.

$\boldsymbol{\beta}_{\backslash 0,n_p}$ (columns labelled 'CMP'). As when using the DAP, we see that we have smaller posterior standard deviations for the regression coefficients than in the analysis using non-informative priors; again, this follows through to give greater posterior precision to our estimates of RTM and T, S and S^*.

4.2.3 *Issues of model selection*

In all the analyses performed so far, only four of the original eight predictor variables have been used to estimate μ_j. These four were selected in the original EB analysis using frequentist regression techniques. In a Bayesian setting, the DAP, given by Equation (8), can be used to aid variable selection. If the full model, defined by design matrix $\mathbf{X}_{n_p}$, has the prior for $\boldsymbol{\beta}_{\backslash 0,n_p}$ given in Equation (8), then for any submodel defined by a reduced design matrix $\mathbf{X}_{n'_p}$,

$$\boldsymbol{\beta}_{\backslash 0,n'_p} \sim N_{n'_p}(\mathbf{0}, n(\mathbf{X}^{\mathrm{T}}_{n'_p}\mathbf{X}_{n'_p})^{-1}),$$

where $n'_p < n_p$. Then models including/excluding each covariate can be compared by the computation of associated predictive probabilities. Our full model has 11 possible covariates (recall that two of the original eight covariates – road classification and road type – require three and two indicators respectively) giving $2^{11} = 2048$ possible models. The marginal log-likelihoods are calculated for each model; the model with the largest marginal log-likelihood is then identified as the 'best'. The R package LearnBayes [1] can be used to implement this method of model selection. Doing so reveals that the model with the largest associated marginal log-likelihood is that which includes exactly the same covariates as identified in the EB analysis.

Some might argue that in a truly Bayesian analysis, there is no fundamental reason why *any* of the predictors should be removed. In fact, it might be argued that it is not coherent for a Bayesian to believe that their predictions will be improved by ignoring some information. Thus, the default Bayesian position might be to use all covariates in the log-linear model for μ_j. Repeating the analyses in Sections 4.2.1 and 4.2.2 but using information on *all* covariates barely changes inferences on the effects of RTM, S and S^*.

4.3 *Accounting for trend*

In the original NSCP report [4] nothing was done to allow for other non-treatment effects such as general trends in casualty frequencies over time. To account for this, we now specify the following modified form for μ_j:

$$\mu_j = \xi \exp\left\{\beta_0 + \sum_{p=1}^{n_p} \beta_p x_{p_j}\right\}, \tag{9}$$

where ξ is a trend effect constant across all sites j. Reports detailing the number of casualties as a result of road traffic accidents, for sites in the Northumbria region *not* treated with safety cameras, suggest that casualty frequencies fell by (on average) 4.7% per year during our after period. Since the after period is two years long (April 2004–March 2006), we adopt the following uniform prior for ξ:

$$\xi \sim U(0.906, 1.000).$$

Although rather simple, we believe this prior represents fairly our uncertainty about whether to relate changes in casualty frequencies at the treated sites to overall regional figures, or to figures from our before period (no reduction per year), or to something in between.

Table 7. Comparison of posterior distributions for: the total number of expected casualties at sites treated with safety cameras (T); expected financial saving to the NHS as a result of implementing the safety cameras (£S thousand); and the total average value of prevention (£S^* thousand), when accounting for RTM both with and without trend.

		Posterior			
		Mean	St. dev.	Median	95% Credible interval
T	*Without trend*	327	25.528	313	(285, 354)
	With trend	319	25.716	306	(271, 340)
S	*Without trend*	30.9	14.092	30.9	(0.9, 58.5)
	With trend	26.5	14.195	26.7	(0.9, 56.2)
S^*	*Without trend*	1541.5	349.136	1541.7	(72.6, 4183.2)
	With trend	1096.5	358.728	1201.3	(67.2, 3978.1)

Notes: Results are shown for an analysis using the conditional mean prior for $\boldsymbol{\beta}$ and Weibull priors for m_j.

Table 7 reports posterior summaries for the total number of expected casualties across the safety camera sites (T), the expected financial saving to the NHS as a result of implementing the safety camera scheme in the Northumbria region (S) and the total average value of prevention (S^*), now accounting for trend through the modified form for μ_j given in Equation (9). A Poisson–Weibull model structure was used, with the CMP for $\boldsymbol{\beta}$ as outlined in Section 4.2.2. Also shown in Table 7, for comparison, are the corresponding posterior summaries when trend has *not* been accounted for, as given in Table 6. The effect of including the trend parameter ξ is obvious: the median number of expected total casualties decreases from 313 to 306 (a mean decrease from 327 to 319), with corresponding decreases in expected values of prevention (both S and S^*).

4.4 *Further modelling of count data*

4.4.1 *Generalised linear mixed models*

Throughout Section 4 we have investigated the sensitivity of the RTM effect, and the corresponding monetary value of prevention of the safety cameras, to the choice of prior for the Poisson mean and the associated regression coefficients. More sophisticated modelling of the casualty counts themselves could also be considered. For example, writing the linear predictor in matrix form, we might replace the terms inside the braces in Equation (9) with

$$\boldsymbol{\eta} = \mathbf{X}_{\mathrm{D}}\boldsymbol{\beta} + \mathbf{Z}\mathbf{U}.$$

Here, $\mathbf{X}_{\mathrm{D}}$ is the full design matrix $(\mathbf{1}, \mathbf{x}_1, \ldots, \mathbf{x}_{n_p})$ and $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_{n_p})^{\mathrm{T}}$; the matrix $\mathbf{Z}$ includes some – or all – of the covariates in $\mathbf{X}_{\mathrm{D}}$, whilst $\mathbf{U}$ is a vector of unobservable random effects associated with the covariates in $\mathbf{Z}$. Often, $\mathbf{U}$ is taken to be multivariate Normal with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$, the elements of which are estimated.

Such generalised linear mixed models (GLMMs) can be used to account for correlation and/or heterogeneity in the data. For example, in a study of accident prediction models for road corridors in Vancouver [9], sites are partitioned into homogeneous groups based on districts; they then account for corridor variation by allowing random intercept and parameter terms for each group, using MCMC within a fully Bayesian analysis to estimate the variances of these random effects. Of course, GLMMs could be considered within the fully Bayesian analyses of Sections 4.1–4.3, provided we adopt suitable priors for the elements of $\boldsymbol{\Sigma}$. However, since we only have aggregate casualty counts for each camera site, we would need to group our sites as in [9] to incorporate random group effects, or disaggregate in an attempt to incorporate random effects for the cameras themselves.

4.4.2 *Zero-inflation*

There are now several examples in the road safety literature of studies modelling vehicle accident frequencies using zero-inflated models; e.g. [19]. These models provide a way of accounting for zeros in excess of what we might expect using a Poisson model, for example. They operate under the assumption of two states existing for the data – the 'zero' state and the 'normal' count state, occurring with probabilities φ and $1 - \varphi$ (respectively); thus, if $\varphi = 0$ the zero-inflated model reduces to the standard Poisson set-up. Under the normal count state, modelling could proceed as throughout this paper. For our data in the reference set, about 16% of sites have zero casualties; under a Poisson assumption with an observed mean of 4.28 casualties per site, we would *expect* to see far fewer zeros (about 1.4% of sites) and so, naively, we may expect zero-inflation to be a problem. In fact, a fully Bayesian analysis as performed in Section 3.3, but now allowing for zero-inflation, gives a 95% credible interval for φ of (0.393, 0.557).

However, comparing the posterior predictive distribution of zeros from the standard model (as used in Section 3.3) to the observed number of zeros, suggests a zero-inflated model may not be required: the posterior predictive distribution covers the observed number comfortably, and this coverage only improves when we move to a Weibull prior for m_j (see Section 4.1). Practitioners can perform such checks for zero-inflation easily using the standard MCMC output. If the posterior predictive distribution of zeros is *not* consistent with the observed number of zeros, then models allowing for zero-inflation can be investigated. However, care should be taken to check that the source of the problem is indeed zero-inflation, and not an inappropriate distribution to describe the overdispersion. Of course, zero-inflated GLMMs can be used to accommodate both overdispersion and zero-inflation; see, for example, [11].

4.4.3 *Further considerations*

For the 56 camera sites reported in this study, mobile safety units were deployed for two hour slots across the intervention period. It is possible that the time of day could contribute to variability in our data; it is also possible that casualty frequencies could change during periods when the cameras are in operation. Although difficult to investigate given the data we were provided with, as far as we are aware little is known of such effects and this could provide an interesting avenue for future research.

In Figure 1, and the associated discussion in Section 1.1, we acknowledge that RTM might not necessarily be a feature of any observational study. In our modelling approach we assume it is, and through the exploratory analyses of Section 3.1 we provide some justification for this assumption. However, any apparent RTM effect could instead be the result of *errors-in-variables*, whereby measurement error in the explanatory variables results in estimation bias of the regression coefficients in the linear predictor of μ_j. Instrumental variables can be constructed to overcome this problem and help identify the treatment effect. Such approaches are now commonplace in the econometrics literature; see, for example, [28]. These instrumental variable techniques could be used within the DAP of Section 4.2.1.

5. Conclusions

In this paper, we have illustrated the shortcomings of the standard tool for assessing the effectiveness of road safety schemes, using a case-study of mobile safety cameras in the UK. We have shown the EB method – the ‘gold standard’ in road safety scheme evaluation [16] – to be over-optimistic in its assessment of variability of estimates of casualty frequency at individual sites treated with mobile safety cameras. A fully Bayesian treatment can be used to appropriately account for all sources of variability by specifying prior distributions for the regression coefficients β_i, giving a more realistic assessment of the variability of casualty frequency estimates, as well as the value of prevention, in monetary terms, owing to the implementation of the safety cameras.

Using MCMC techniques within a fully Bayesian framework also gives the practitioner much more flexibility – both in terms of the prior distributions used and the way in which the resulting posteriors are summarised. For too long, assessment of the effectiveness of road safety schemes has relied on a Poisson–gamma structure, using only the resulting closed-form expression for the associated posterior mean as the counterfactual estimate. In fact, as we show, estimates of casualty frequency and RTM can be sensitive to the choice of prior used for the Poisson mean; other prior distributions can provide a better fit to the data, and the posterior mean might not always be the most appropriate summary of casualty frequency to use. We have also shown that the extra variability induced by a fully Bayesian treatment can be reduced by implementing more informative priors in our analysis.

After accounting for trend, we have shown that we might expect the total number of casualties to have reduced from 436 to about 306 anyway (posterior median in Table 7), even if the safety cameras had not been used. Comparing the posterior distribution for the total number of expected casualties T with the number of casualties in the period *after* implementation gives a median total value of prevention of just over £1.2 million for the 56 sites in this study, over the two year treatment period.

Applying fully Bayesian techniques opens up a field of opportunities for those required to assess the effectiveness of road safety schemes although, as we discuss, any assumptions underpinning the approach used should be carefully checked. Although the dispute between empirical and full Bayes is not new, relatively few authors in the transport field have given fully Bayesian techniques the attention they deserve – though [18,20,21] all make for accessible reading in this area. More realistic model structures can be used in the fully Bayesian framework that could otherwise prove difficult to draw inferences from; obtaining more accurate estimates of the impacts of road safety measures is crucial for guiding increasingly limited investment opportunities.

Acknowledgements

We would like to thank Ian Abernethy (Gateshead Council) and the NSCP for providing the data used in this work. We would also like to thank Professors Rob Henderson and Mike Maher for providing feedback on earlier versions of this paper, and Professor Richard Allsop for helpful discussions.

References

[1] J. Albert, *Bayesian Computation with R*, Springer, London, 2009.

[2] E.J. Bedrick, R. Christensen, and W. Johnson, *A new perspective on priors for generalized linear models*, J. Amer. Statist. Assoc. 51 (1996), pp. 211–218.

[3] J. Broughton, *The benefits of improved car secondary safety*, Acci. Anal. Prev. 35 (2003), pp. 527–535.

[4] J. Colligan, L.E. Goulbourne, N. Thorpe, L. Fawcett, and A. McNay, *Mobile Speed Cameras and the Demand for Secondary Health Care in Northumbria*, Northumbria Safety Camera Partnership, Newcastle upon Tyne, 2008.

[5] P.C. Cryer, S. Westrup, A.C. Cook, V. Ashwell, P. Bridger, and C. Clarke, *Investigation of bias after data linkage of hospital admissions data to police road traffic crash reports*, Inj. Prev. 7 (2001), pp. 234–241.

[6] Department for Transport, *Tomorrow's Roads – Safer for Everyone*, The Stationery Office, London, 2000.

[7] Department for Transport, *Benefits to Society Arising from Prevention of Road Accidents and Casualties*, The Stationery Office, London, 2009.

[8] Department for Transport, *Reported Road Casualties in Great Britain 2011*, The Stationery Office, London, 2012.

[9] K. El-Basyouny and T. Sayed, *Accident prediction models with random corridor parameters*, Acci. Anal. Prev. 41 (2009), pp. 1118–1123.

[10] B. Escofier and J. Pagés, *Multiple factor analysis*, Comput. Statist. Data Anal. 18 (1990), pp. 121–140.

[11] J.D. Hadfield, *MCMC methods for multi-response, generalised linear mixed models: The MCMCglmm R package*, J. Statist. Softw. 33 (2010), pp. 1–22.

[12] E. Hauer, J. Ng, and J. Lovell, *Estimation of safety at signalised intersections*, Transp. Res. Rec. 1467 (1988), pp. 30–37.

[13] W.M. Hirst, L.J. Mountain, and M.J. Maher, *Sources of error in road safety scheme evaluation: A quantified comparison of current methods*, Acci. Anal. Prev. 36 (2004), pp. 705–715.

[14] A. Hooke, J. Knox, and D. Portas, *Cost Benefit Analysis of Traffic Light and Speed Cameras*, Police Research Series, Paper 20, The Stationery Office, London, 1996.

[15] M.G. Kendall, *Time Series*, 2nd ed., Hodder–Arnold, London, 1976.

[16] J. Kennedy and B. Sexton, *Literature review of road safety at traffic signals and signalised crossings*, TRL Report No. 436, Transport Research Laboratory, Crowthorne, 2009.

[17] R. Kulmala, *Measuring the safety effect of road accident measures at junctions*, Acci. Anal. Prev. 6 (1994), pp. 781–794.

[18] W. Li, A. Carriquiry, M. Pawlovich, and T. Welch, *The choice of statistical models in road safety ountermeasure effectiveness studies in Iowa*, Acci. Anal. Prev. 40 (2008), pp. 1531–1542.

[19] D. Lord, S.P. Washington, and J.N. Ivan, *Poisson, Poisson–gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory*, Acci. Anal. Prev. 37 (2005), pp. 35–46.

[20] M.J. Maher and L.J. Mountain, *The sensitivity of estimates of regression to the mean*, Acci. Anal. Prev. 41 (2009), pp. 861–868.

[21] S.P. Miaou and D. Lord, *Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form, and Bayes versus empirical Bayes methods*, Transp. Res. Rec. 1840 (2003), pp. 1310–1340.
[22] L. Mountain, B. Fawaz, and L. Sineng, *The assessment of changes in accident frequencies on link segments: A comparison of four methods*, Traffic Eng. Control 33 (1992), pp. 429–431.
[23] I. Ntzoufras, *Bayesian Modelling Using WinBUGS*, Wiley, Hoboken, NJ, 2009.
[24] O.J. Pendleton, *Application of new accident analysis methodologies, Vol.1: New methodologies*, US Department of Transportation, Report FHWA-RD-90-091, Washington, DC, 1991.
[25] B. Persaud and L. Dzbik, *Accident prediction models for freeways*, Transp. Res. Rec. 1401 (1993), pp. 55–60.
[26] H.F. Simpson, *Comparison of hospital and police casualty data: A national study*, TRL Report No. 173, Transport Research Laboratory, Crowthorne, 1996.
[27] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. Van der Linde, *Bayesian measures of model complexity and fit (with discussion)*, J. R. Statist. Soc. B 64 (2002), pp. 583–616.
[28] T. Wansbeek and E. Meijer, *Measurement Error and Latent Variables in Econometrics*, Elsevier, Amsterdam, 2000.