Note on Transport for London Fixed Camera Data

Mike Maher, UCL

The aim is to investigate what effect speed cameras have had on accidents in the London area. The available data consists of annual numbers of KSIs and slight casualties at each of 497 fixed cameras over a period of 22 years (from 1990 to 2011). The cameras have been installed at varying times, ranging from 1992 to 2009. The annual numbers of KSIs and slights in the whole region are also known for the period 1990 – 2011. By subtracting the numbers of KSIs and slights at the 497 camera sites, we have the numbers of KSIs and slights at the non-camera sites across London. These are shown in Figure 1 below, and will be used to provide trends as a comparison for the numbers of slights and KSIs at the camera sites.

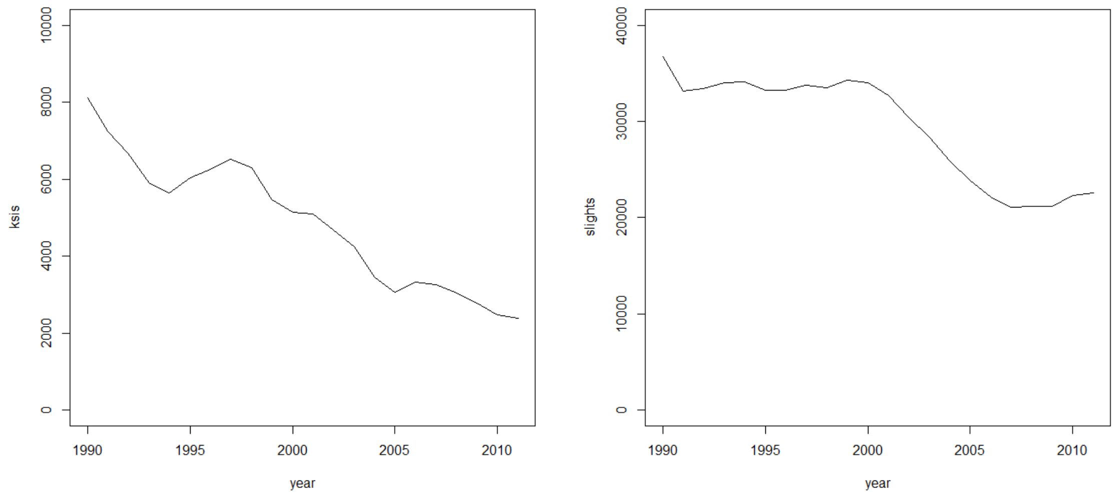

Figure 1: trends in KSIs (left) and slights (right) over the period 1990 – 2011 at non-camera sites

As with many forms of remedial treatment, sites are selected to have a camera installed because of a high number of collisions (especially KSIs) in the preceding period (the *site selection period*, or SSP, generally three years in length). Therefore there will generally be a fall, after treatment – *even if the treatment has no effect*. This is due to Regression to the Mean (RTM). The problem is therefore to be able to detect and estimate the camera effect separately from the RTM effect.

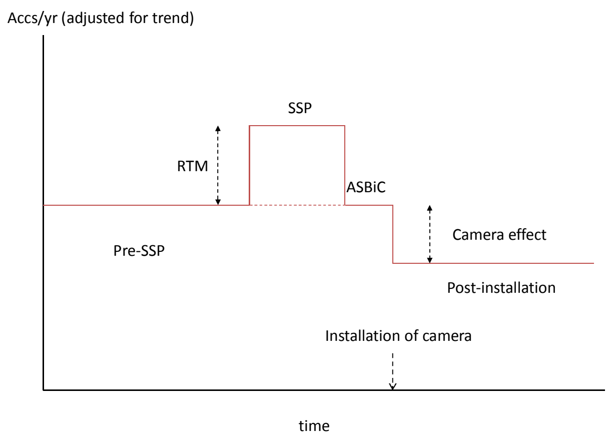

Figure 2: idealised form of plot of crashes per unit time, adjusted for trend

Figure 2 shows an idealised form of plot of annual crashes over time, with four periods identified: the pre-SSP (true before period), the SSP, the ASBIC (after selection but before installation of the camera), and the post-installation period. The ASBIC period is there because there is often a delay between the selection of sites and the installation of cameras. This delay can sometimes be a number of years in length. Unfortunately, although the date of installation is known, the SSP that has been used is not generally recorded. This presents a problem in the modelling process, as will be seen later. This model is sometimes known as the "four time periods method" (see, for example, Finney 2014). An earlier "three period" method was proposed and used previously by Mountain *et al* (1998), in which the period after selection and before installation was referred to as the "lag" period.

Of course, the actual plots for sites are quite different from the idealised form in Figure 2, because of the random nature of crash frequencies. For example, in Figure 3, we have a plot of the total number of KSIs over the period 1990 to 2011 at the 12 cameras that were installed in the year 2001 (the left plot being the raw totals KSIs and the right hand plot the trend-adjusted values).

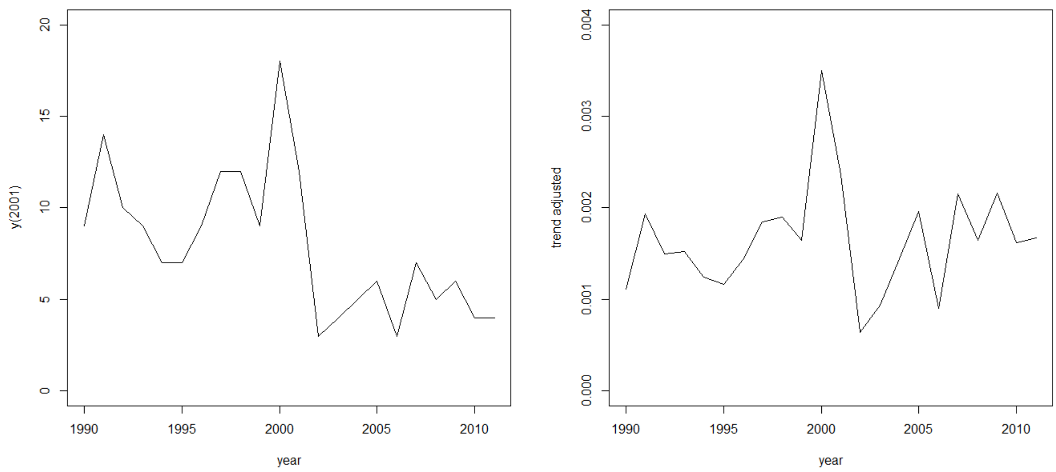

Figure 3: plots of the annual KSIs over the period 1990 – 2011 at cameras installed in 2001 (raw totals on the left and trend-adjusted on the right)

There are signs of a raised level of KSIs in the period immediately prior to 2001, and a fall from that raised level after installation, as would be expected but there is a fair amount of volatility. Although installation was in 2001, we cannot be sure that the SSP for all 12 of these sites was the immediately preceding three-year period 1998 – 2000 (even though the highest spike was in the year 2000).

In order to estimate the effect of the camera being installed it is therefore necessary, not only to allow for trend, but also to avoid the effect of RTM. This means we should compare the after installation collisions with the numbers in the combined pre-SSP and ASBIC periods, both adjusted for trend. However, as the SSP is not precisely known, some assumption must be made. The aim is to go back sufficiently far as to make it reasonably safe that the assumed pre-SSP period does not contain any part of the SSP. This is not entirely straightforward, as amongst other things the SSP for different cameras, or groups of cameras, may be different. So we need to set a value for the assumed start of the SSP relative to the installation date that is appropriate for the great majority of cameras in the data set. If we go too far back in time in setting this then we are losing valuable data in the assumed before period. If we do not go far back enough, we may include in the assumed before period some years that are part of the SSP for some camera sites.

In his original report Allsop (2013a) assumed the SSP to be the three calendar year period immediately preceding the year of installation. After comments and further discussion this was altered in the revised report (Allsop, 2013b) by pushing the start of the SSP back in time by one year, so as to be the first three years of the immediately preceding four year period. If we denote by d the year t relative to the installation year for a site $s(i)$ so that $d = t - s(i)$, then $d = 0$ denotes the installation year and $d < 0$ denotes the years before installation and $d > 0$ denotes the years after installation. So the SSP assumed by Allsop in the original report was (-3, -1) and this was modified in the revised report to (-4, -2). The problem is that this assumption may not necessarily be correct, and in any case not all sites may have the same SSP. So in any group of sites that are analysed together, there may be a mix of SSPs. It may therefore be safer, in order to avoid the effect of RTM, to assume a wider range of years than the standard three.

The model we assume is

$$E(y_{it}) = k_i \, R_t \, F_{t-s(i)}$$

where $E(y_{it})$ is the expected number of collisions at site i in year t, R_t is the regional total of collisions in year t (to allow for trend), and F is a factor to allow for the year t relative to the year of installation $s(i)$ of the camera at site i. If the first year of the SSP is d^* then the pre-SSP period consists of those years for which $d < d^*$. We fit a Negative Binomial model with log link, and $\log(R_t)$ as offset, and have a factor F that takes separate values for each value of d between some d_{min} and d_{max} (here we have used -10 and +6). (Values for $d < d_{min}$ are grouped in with those at d_{min} ; similarly those for $d > d_{max}$ are grouped in with those at d_{max}). We then expect to find values of F in excess of 1 for $d = d^*, d^*+1, \ldots 0$ (because of the RTM effect for these years) and, if the cameras are effective in reducing accidents, a value of less than 1 for $d > 0$, all relative to the base level (of 1) for the lowest value d_{min}.

Results

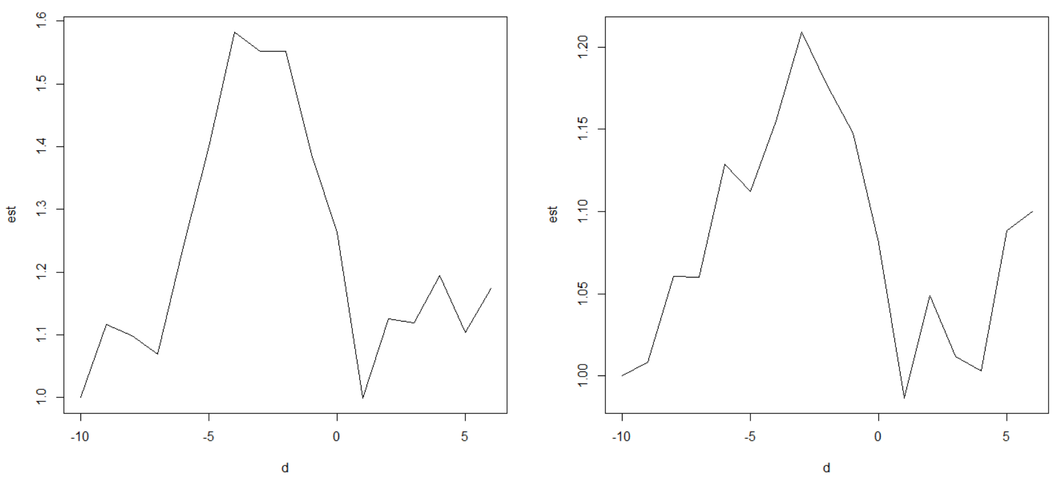

Figure 3: plots of estimated values of the factor *F* for KSIs (left) and slights (right)

The results are as shown below in Figure 3, for d_{min} = -10 and d_{max} = 6 (the plot of KSIs on the left and for slights on the right). There are clear raised values of F between between -6 and 0 in both graphs (with higher values as expected in the KSI plot), indicating that the most appropriate value for d^* = -6 (that is, the start of the SSP is six years before the installation date). So, to avoid the RTM effect we should model with d^* = -6, and have common values of F for $d < d^*$ (the before period: the base level, so that F = 1) and $d > 0$ (the after period). The results are then (estimates of log(F), relative to the base of 0, and standard errors):

Table 1: estimates of log(*F*) and st. errors for *d* = *d, ... 0 and the camera effect (*d* > 0) for KSIs**

d	-6	-5	-4	-3	-2	-1	0	camera
Estimate	0.179	0.268	0.391	0.353	0.370	0.283	0.195	0.067
St error	0.056	0.054	0.053	0.053	0.053	0.054	0.056	0.033

Table2: estimates of log(*F*) and st. errors for *d* = *d, ... 0 and the camera effect (*d* > 0) for slights**

d	-6	-5	-4	-3	-2	-1	0	camera
Estimate	0.116	0.100	0.121	0.169	0.120	0.113	0.049	0.026
St error	0.030	0.029	0.029	0.029	0.028	0.029	0.029	0.017

The camera effect for KSIs is then exp(0.067) = 1.07, an increase of 7%, with a 95% CI of (0%, 14%), and for slights is exp(0.026) = 1.026, an increase of 2.6%, with a 95% CI of (-1%, 6%). So, there is no clear evidence of the cameras having had any significant effect in reducing KSIs or slights.

Alternative approach

We can consider the approach taken by Allsop in the RACF report, with an assumed SSP of (-4,-2) – that is, the SSP consisting of the three years with d = -4, -3, -2. Then we can count up, for each site, the number of crashes (i) in the SSP, (ii) in the before period: d = -1 and d < -4, and (iii) the camera period, d > 0. We then count up the regional totals R_t over the corresponding periods, to get R_{bef}, R_{ssp} and R_{cam} for each site.

We can then apply a multinomial model in which we model the split of the total number of crashes y_{tot} at each site between the three periods: before, SSP and camera. This is a multinomial model with probabilities

$$p_{bef} = \frac{R_{bef}}{R_{bef} + \alpha R_{ssp} + \beta R_{cam}}$$

$$p_{rtm} = \frac{\alpha R_{ssp}}{R_{bef} + \alpha R_{ssp} + \beta R_{cam}}$$

$$p_{cam} = \frac{\beta R_{cam}}{R_{bef} + \alpha R_{ssp} + \beta R_{cam}}$$

where the R are the regional totals over the relevant periods, and α and β are respectively the factors by which the before accident rate (adjusted for trend) is multiplied in the SSP and camera periods. Hence the probability that the y_{tot} accidents at any site will be split y_{bef}, y_{ssp}, y_{cam} is given by the multinomial probability function:

$$P(y_{bef}, y_{ssp}, y_{cam}) = \frac{y_{tot}!}{y_{bef}!\, y_{ssp}!\, y_{cam}!} p_{bef}^{y_{bef}} p_{ssp}^{y_{ssp}} p_{cam}^{y_{cam}}$$

Hence to find the maximum likelihood estimate of α and β, we need to maximise

$$z(\alpha, \beta) = \sum_i \sum_t y_{it} \log(p_{it})$$

where i is over the camera sites, and $t = bef, ssp, cam$ for the three periods at each site, with respect to the two parameters α and β. For the cameras to be beneficial in reducing crashes, the value of β needs to be significantly less than 1.

Using a reduced data set, with the 412 sites for which there is at least six years of data before installation, the estimates (with standard errors) for KSIs are: $\hat{\alpha}$ = 1.428 (0.035), and $\hat{\beta}$ = 0.999 (0.025). For slights, the results are: $\hat{\alpha}$ = 1.123 (0.012), and $\hat{\beta}$ = 1.001 (0.010). So, the camera effect is non-significant in both cases.

Given the evidence that the SSPs stretch out over the period between -6 and -1, we can try out a "most likely" SSP, by considering four possible SSPs: (-6, -4), (-5, -3), (-4, -2) and (-3, -1) and choosing for each site the SSP that contains the highest number of KSIs. The before period then consists of the remaining years before installation. Applying this we find that of the 412 sites, 95 have (-6, -4) as the most likely SSP, 84 have (-5, -3), 93 have (-4, -2), and the remaining 140 have (-3, -1) (but note that ties are broken by choosing the later period). Then the estimates are: $\hat{\alpha}$ = 1.989 (0.045), and $\hat{\beta}$ = 1.123 (0.028), so appreciably higher as would be expected as this choice of SSP leads to a minimisation of the crashes in the before period, so that the after-before ratio is always going to be increased. This "most likely" SSP method is therefore very unfavourable to the camera estimate, and probably only has a role to play when other methods show a substantial camera *benefit* and this most likely SSP method gives a conservative camera benefit to provide a lower bound on camera benefit.

Alternatively, we could avoid the choice of SSP completely by only using as before data those years for which $d < -6$. In that case, we get the following estimates for KSIs: $\hat{\alpha}$ = 1.405 (0.030), and $\hat{\beta}$ = 1.084 (0.029). For slights, we get: $\hat{\alpha}$ = 1.127 (0.011), and $\hat{\beta}$ = 1.033 (0.011). The results here for the camera effect are then very close to those we obtained from the analysis in Part 1 (the only real difference being due to the fact that here we have a slightly reduced data set of 412 sites whereas there we had the full set of 497 sites). So the average effect of cameras is to increase KSIs by around 8%, and to increase slights by 3%.

Conclusions

We have tried three variants of the basic model. The first is the one used by Allsop in his revised report, with the assumed SSP being (-4, -2). The second is the "most likely" SSP being identified for each site. The third is where the whole of the (-6, -1) period is avoided, so that the camera estimate is achieved by comparison of the post-installation data ($d > 0$) with the data from before the earliest SSP ($d < -6$), adjusted for trend. The conclusion is that, although the estimates of the average camera effect from these variants of the basic model are different, they all agree that there is no discernible beneficial effect from the fixed cameras on either slights or KSIs. All calculations carried out were done in *R*.

References

Allsop R.E (2013a). Guidance on use of speed camera transparency data. RAC Foundation, original version June 2013.
Allsop R.E (2013b). Guidance on use of speed camera transparency data. RAC Foundation, revised version November 2013.
Finney D. (2014) The effect of fixed speed cameras on road safety.
http://speedcamerareport.co.uk/09_fixed_report.htm
Mountain L.J., Maher M.J. and Fawaz B. (1998). Improved estimates of the safety effects of accident remedial schemes. *Traffic Engineering and Control* **39**(10), 554-548, October 1998.
London speed camera collision data: http://data.london.gov.uk/dataset/safety-camera-data-london