Wed Appendix 1 - Details of statistical methods

A) Collisions

For each region, we estimated how much on average rates of collisions had changed in each road segment after lighting was changed adjusting for possible biases by using a conditional Poisson regression model (Armstrong, Gasparrini, and Tobias 2014). This approach is broadly similar to that used previously to study changes after introduction of 20mph speed limit zones (Grundy et al. 2009). We adjusted for seasonal variations and temporal trends using step functions (indicators) for calendar month and year, following exploratory analyses in which more parsimonious models fitted the data less well by Akaike's Information Criterion (AIC). All street lighting interventions (switch off, part-night lighting, dimming, and white light) were entered into the same model to avoid mutual confounding.

Because changes in street lighting would only be expected to have a direct impact at night, we restricted attention to night-time collisions. However, to guard against bias due to changes concurrent with lighting interventions that impact on the overall (i.e. day and night) collision rates, our model also estimated the change in day-time collisions rates associated with interventions and, as a refined measure of change in outcome rates following each intervention, the ratio between the night-time and day-time changes. We consider this 'day-time collision rate corrected' measure to be the most robust estimate of the change in collision rates following the changes to street lighting.

The final model was thus (expression (3) from Armstrong 2014):

$$E(Y_{i,s}) = \mu_{i,s} = \exp\{\alpha_s + \boldsymbol{\beta}^{\mathbf{T}}\mathbf{x_i}\}, Y \sim Poisson(\mu_i, conditional\ on\ total\ collisions\ in\ segment\ Y_{.,s})$$

Where

i = 1…I is the study month (months elapsed since the start of the study in 2000);

s = 1…S is the road segment;

$Y_{i,s}$ is the number of collisions in month i on road segment s; night-time and day-time collisions are included for each month, distinguished by an indicator variable.

The parameters α_s , which allow for variation in rates across road segments, are fitted implicitly by 'conditioning out', as discussed further in Armstrong 2014 and the references cited therein.

The vector $\mathbf{x_i}$ of explanatory variables has the following components (sub-vectors of x):

Potentially confounding variables:

- $\mathbf{x_{month}}$: Indicator variables for calendar month (1-12) to control for seasonal patterns and month duration;
- $\mathbf{x_{year}}$ Indicator variables for calendar year (2000 to 2013) to control for time trends.

To allow adjustment for patterns of day-time collisions:

- $\mathbf{x_{night}}$ An indicator variable for night-time collisions (1 for night-time; 0 for day-time);
- $\mathbf{x_{month*night}}$, $\mathbf{x_{year*night}}$ Interaction of the night-time indicator with each potentially confounding variable (above) to allow different seasonal patterns and time trends for night-time and day-time collisions.

Variables of interest:

- $\mathbf{x_{so}}$,$\mathbf{x_{pn}}$,$\mathbf{x_{dim}}$,$\mathbf{x_{white}}$ Indicator variables for each lighting intervention: switch off, part-night lighting, dimming, and white light (0 before intervention; 1 after intervention);
- $\mathbf{x_{so*night}}$, $\mathbf{x_{pn*night}}$, $\mathbf{x_{dim*night}}$, $\mathbf{x_{white*night}}$ Interaction of these with the night-time indicator; the coefficients of these variables estimate change in the night-time collision rate following the intervention adjusting for any changes observed in the day-time collision rate.

This model was fitted separately to each government region. The mean over regions of the coefficients of interest for each intervention was estimated using a standard meta-analysis model. Because we found no evidence for heterogeneity, fixed effect models were used.

B) Crime

The model for crime rates had the same basic form as that for road traffic collisions, but with the following differences, reflecting different geographical resolution (i.e. counts at MSOA level rather than road segment level), the data origins (i.e. data collected at police force level), the larger average counts, and complex background temporal patterns:

- Instead of indicator variables for pre- intervention and post-intervention months, intervention variables comprised, for each year-month, the proportion (0-1) of total kilometres of road length in an MSOA subject to the intervention; thus 0 would

represent no roads with the intervention and 1 would represent all roads with the intervention.

- Because time of day of the crime is not available from *Police.uk* data, total crime numbers were included, and no night-time interaction variables were needed; the coefficients of interest for each intervention were the 'main effects' of the intervention variables.
- We fitted models to each police force (of 31 police forces included in the data) separately, in view of their different data collection systems and evidence for different background time patterns across them.
- Given evidence for more complex temporal patterns than allowed for by month and year indicators, we fitted indicator variables for the number of months (1-36) elapsed from the start of the study data series (i.e. a step function for elapsed month from December 2010).
- Residual variance greater than that expected (i.e. 'over dispersion') in the Poisson statistical distribution was allowed for in standard errors using a scale factor estimated from the Pearson chi-squared statistic divided by the degrees of freedom of the residuals. [There was no evidence of over dispersion in the sparser collision analyses.]
- Because of the number of police forces and categories of crime considered, we do not present forest plots of the meta-analyses in the main text, only the means over England and Wales.

References

Armstrong, Ben, Antonio Gasparrini, and Aurelio Tobias. 2014. "Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis." *BMC Medical Research Methodology* no. 14 (1):122.

Grundy, C., R. Steinbach, P. Edwards, J. Green, B. Armstrong, and P. Wilkinson. 2009. "The effect of 20 mph traffic speed zones on road injuries in London, 1986-2006: a controlled interrupted time series analysis." *BMJ* no. 339:b4469.