Evaluating Neighbourhood Policing using Bayesian Hierarchical Models: No Cold Calling in Peterborough, England

Guangquan Li *,[1], Robert Haining [2], Sylvia Richardson[1,3], Nicky Best[1,3]

[1] Department of Epidemiology and Biostatistics, Imperial College, London, UK

[2] Department of Geography, University of Cambridge, UK

[3] MRC-HPA Centre for Environment and Health, Imperial College, London, UK

[*] To whom correspondence should be addressed: guang.li@imperial.ac.uk

Word count: 5108

Abstract

As part of a wider Neighbourhood Policing strategy, Cambridgeshire Constabulary, in common with other Police Forces in the UK, instituted "No Cold Calling" (NCC) zones to reduce cold calling (unsolicited visits to sell products or services), which is often associated with rogue trading and distraction burglary. This paper evaluates the NCC targeted areas chosen in 2005-6 and reports whether they experienced a measurable impact on their burglary rates in the period up to 2008. Time series data for burglary at the Census Output Area level is analysed using a Bayesian hierarchical modelling approach, addressing issues often encountered in small area quantitative policy evaluation. Results reveal a positive NCC impact on stabilising burglary rates in the targeted areas.

Keywords: Policy evaluation; Neighbourhood Policing; Time-series data; Interrupted time series analysis; Bayesian hierarchical models

Introduction

Policy evaluation is often performed by comparing event occurrence before and after policy implementation. Differences between the two time periods are indicative of the influence of the policy. However, apart from genuine policy impacts, these changes can be a consequence of changes in other external factors. It is therefore recommended to employ a control group so that time series data from the exposed group(s) can be compared against what would be expected based on time series data from non-exposed but demographically similar groups. Controlling for systematic changes over time helps to differentiate how much of the change may be due to the policy impact and how much of the change may be due to other external factors. Often termed the interrupted time-series design (Campbell et al. 1963), this approach has an abundance of evaluation applications, ranging from health care interventions (Morgan et al. 2007; Studnicki et al. 1997), school reform (Thum and Bhattacharya 2001) to policing strategy assessments (MacDonald et al. 2010).

In relation to crime, several aspects of the interrupted time-series approach have been adopted by UK Government agencies as well as the research community. A recent Home Office study revealed an overall positive impact from the National Reassurance Policing program on crime rates, the public's perception of crime and anti-social behaviours by comparing results from six trial sites against six control sites (Tuffin et al. 2006). Skogan et al. (2008) evaluated the CeaseFire program in Chicago by contrasting changes in the selected areas after the introduction of the program with trends in matched comparison areas. To avoid overestimating any impact, MacDonald et al. (2010) used a regression approach to adjust for an overall downward trend when assessing the effect of forming business improvement districts on violent crimes. However, since the adjustment was estimated from the business improvement districts, the resulting evaluation of the program could have been conservative. In all the above studies, including the one presented in this paper, there is no random assignment of areas to either the control or the exposure group. Conclusions regarding policy impacts should be verified against different definitions of the control group.

Another issue often faced when evaluating neighbourhood scale programs is data sparsity. For example, the number of burglary cases observed during a year at some geographical level can be too small to provide a reliable trend estimate. Effect of the policy could potentially be masked by the excessive noise in the sparse data. Multilevel models (or equivalently termed hierarchical models or random effect models) provide a natural framework to combine information from multiple areas and periods and hence to strengthen the estimation of impact. Here, we adopt the Bayesian model formulation for parameter estimation and making inference. Parameter summaries such as posterior means, uncertainty intervals and posterior probabilities (e.g., what is the probability that the overall burglary rate in the exposed group is lower than that of the non-exposed group) can be readily obtained as all parameters are considered to be random variables with associated posterior distributions. Both Iversen (1984) and Gill (2007) provide some nice introductory material to Bayesian statistical inference in the social science context. For readers who are interested in Bayesian hierarchical/multilevel models (BHM), a good introduction can be found in Part II of Gelman et al. (2004) while Gelman and Hill (2007) gives an excellent overview of BHM in social research.

A further issue related to data sparsity often arises because targeted areas often tend to be small relative to the geographical scale at which much of the data necessary for the evaluation are reported. This data complication should be recognised in constructing a model, which can help to enrich the interpretations.

Models utilised in this paper are close in spirit to those of Thum and Bhattacharya (2001) and MacDonald et al. (2010). Typically, a BHM evaluates the area-specific policy impacts hierarchically through a random effects distribution, e.g., $\theta_i \sim N(\mu_\theta, \sigma^2)$, where θ_i quantifies the local impact for area i and μ_θ represents the overall impact. All local data are combined to estimate aspects (e.g., mean and variance) of the global distribution, which in turn provides estimates for the local impacts. Under this structure, these local impacts are modelled as correlated, as opposed to independent. Estimation of impact in one area can "borrow" information from data in other areas and this information borrowing can lead to a better estimate of the impact.

The objective of this paper is to present a generic Bayesian framework for evaluating policies targeted at the small area or neighbourhood scale, paying close attention to the issues described above. We tailor the approach in order to carry out a statistically rigorous assessment of the "No Cold Calling" (NCC) scheme in Peterborough which was initiated by the Cambridgeshire and Peterborough Distraction Burglary and Rogue Trader Task Force in 2005.

"Cold calling" is defined as a visit, or in the first place a telephone call that may be followed by a visit, by a trader (or someone linked to a trader whether or not they supply goods or services) and which takes place without the consumer expressly requesting the initial contact. In 2002 the Trading Standards Institute carried out a national survey of almost 9000 households[1]. Over 60% of consumers said they had received a cold call in the preceding 3 months and 25% reported a bad experience in the past two years. Over 95% of respondents said they did not want doorstep callers. Incidents linked to cold calling can be devastating with older people suffering in particular (Cambridgeshire-Constabulary 2008).

[1]The survey report can be downloaded from http://www.tradingstandards.gov.uk/policy /researchandreports.cfm (last accessed on May 6, 2011)

Potential areas to be targeted by the scheme were identified by Police and Trading Standards staff and Neighbourhood Watch co-ordinators. Areas consisting of between 20 and 40 properties in close proximity occupied by older or vulnerable people were identified in the first instance. Those areas with the highest reported incidents of distraction burglary and rogue trading were selected for inclusion in the scheme. The project's aim was to reduce the impacts of distraction burglary and rogue trading in terms of (i) the number of incidents and (ii) the public's fear of crime. The project in Peterborough was first implemented in selected areas during 2005 and extended to more areas in 2006. More NCC areas were defined in subsequent years.

In addition to setting up street signage and door stickers to discourage cold calling, every resident was visited in the targeted areas and given an information pack containing literature advising how to avoid becoming a victim and the steps to take when answering the door. The police also participated in informal follow-up meetings.

A "dip" sample interview-based evaluation of the NCC scheme by the Cambridgeshire Police reported high levels of satisfaction amongst residents in the NCC-targeted areas who generally expressed increased confidence in dealing with cold callers (Cambridgeshire-Constabulary 2008). This report recommended extending the scheme to other vulnerable locations dependent on securing further funding. However, to date, there has been no assessment of the scheme based on offence data so this is the question we address in this paper.

The paper is structured as follows. The following section describes data from the recorded crime database obtained from the Cambridgeshire Constabulary used in the evaluation. The section entitled "Evaluation framework" outlines various criteria used to construct the control group and discusses the two statistical models involved, one for estimating the control trend pattern and one for quantifying the policy impact. Two different ways to combine information across areas are also discussed in that section. Results are reported in the subsequent section and some concluding remarks are made in the last section.

The Data

Rogue trading is an under reported crime and the number of cases is small; distraction burglary events (Home Office classification Code 28, Sub-Code 10) are also very few in number. As a consequence, analysis here is based on all reported burglary in a dwelling events including aggravated burglary (Code 28 with Sub-Codes 0 and 10 and Code 29). We employ this as a surrogate measure for rogue trading and distraction burglary, two household-related acquisitive offences. The data were extracted from the 2001-2008 recorded crime database provided by the Cambridgeshire Constabulary.

During this period, there were in total 9388 burglary in a dwelling incidences recorded in Peterborough, which has a population of approximately 160,000. The annual burglary trend in Peterborough showed an initial increase in 2001/2002, followed by a drop in 2003, levelling off in 2005/2006, after which it increased again reaching close to 2002 levels by 2008. Some of the difference pre and post 2002 may be associated with the implementation of the National Crime Recording Standard (Berman 2008). We used the earliest time at which the incident might have taken place to define the burglary time. As the cases were then aggregated to the annual level, this uncertainty in the burglary time is unlikely to affect our policy assessment.

Table 1 summarises data on the NCC scheme implemented in 2005 and 2006 in Peterborough. A total of eleven areas were targeted, eight in 2005 and a further three in 2006. These targeted areas are much smaller than the corresponding Census Output Area (COA) that they are nested within. At the COA level, the median number of burglary cases per year was 2 over 2001-2008, with the 5^{th} and 95^{th} percentiles being 0 and 8, respectively. So, the burglary cases were aggregated to the COA level to avoid further problems associated with data sparsity and due to data availability for 2001-2003, where the offences records were only geo-referenced at the COA level. Note that there are two targeted areas, Boxgrove Close and Woad Court, that nest within the same COA, resulting in only 10 unique NCC-COAs. As evident in Figure 1, even at the COA level there is still substantial variability in annual burglary rates. Further aggregating the data for these 10 NCC-COAs into a single time series yields a clearer picture. The burglary rate for the NCC group remained lower during 2006-2008 than the overall burglary rates in Peterborough which were going up (Figure 1), a first indication of a positive impact from NCC. A simple Poisson test for comparing two rates showed that the overall burglary rates were similar between the NCC group and other non-NCC areas in Peterborough before the NCC scheme (2001-2004) with a rate ratio of 1.06 (p-value=0.56). After the introduction of the NCC scheme, the NCC group showed a non-significant reduction over the period 2005 to 2008 compared to the overall burglary rate from other non-NCC areas in Peterborough (rate ratio=0.85 and p-value=0.19). However, raw data may be too sparse to draw firm conclusions about the scheme's impact using simple significance tests such as this.

(Table 1 and Figure 1 here)

In what follows, our analysis is concerned with ways of strengthening inference on the differences between NCC area trends (aggregated or individually) and overall trends (for different sets of control, non-NCC, areas) and hence to reliably provide statistical evidence regarding the scheme's impact.

Evaluation framework

Constructing the control group

In this study, control groups are employed to set the reference trend for the before-after comparison. We used the Lower Super Output Area (LSOA) scale, each comprising several COAs, as the basic areal unit for the control group in order to obtain a reliable control trend pattern. To form the control group, LSOAs are selected on the basis of having similar local characteristics such as burglary rates or deprivation score to those areas in the NCC-targeted group. Six different control criteria have been considered, as tabulated in Table 2. While Criterion 1 includes all the 88 LSOAs in Peterborough, LSOAs under Criteria 2-4 are only selected if they are in the same burglary rate category as the NCC group prior to the NCC scheme. Criterion 5 uses spatial proximity to select controls whereas Criterion 6 selects LSOAs based on the LSOA-level multiple deprivation scores in 2004, obtained from the neighbourhood statistics website. Specifically, LSOAs in Criterion 6 are chosen to have multiple deprivation scores between 20 and 40, a range that 7 out of 10 NCC-exposed LSOAs fell within.

(Table 2 here)

The purpose of constructing these different control groups is to examine the robustness of our conclusions. All 88 LSOAs in Peterborough are eligible to be selected but in the case of those containing NCC-targeted COAs, we subtracted both burglary counts and dwelling numbers associated with the NCC-COAs.

The Models

The evaluation framework contains two models, one for constructing the reference trend pattern from the control group and one for comparing the before-after burglary rates in the policy areas with an adjustment for the reference trend. For the second model, we explore two different ways to synthesise information across areas, one by empirical grouping and the other using a fully model-based approach. One advantage of placing this analysis in the Bayesian framework is that uncertainty in the reference trend adjustments can be fully acknowledged when quantifying the NCC impact, which is discussed in more detail in both the Discussion section and Appendix C.

The following subsections summarise each model. Full specifications are given in Appendix A.

A model for the control group

To construct the reference temporal profile, a Poisson regression model is fitted to the count data in the control group (McCullagh and Nelder 1989). The number of burglary cases recorded in area i of year t, y_{it} ($i = 1, \ldots, N$ and $t = 1, \ldots, 8$), is assumed to be a Poisson random variable with mean $\lambda_{it} = n_i \cdot \theta_{it}$ where n_i denotes the number of dwellings in area i. On the log scale, the burglary rate θ_{it} is modelled as a linear combination of the place effect (u_i), the time effect (γ_t) and ϵ_{it}, which accounts for overdispersion. All these three terms are modelled as random effects. The place effects, $u_i \sim \mathrm{N}(\alpha, \sigma_u^2)$, are normally distributed around a grand mean α, and account for the differences in the total burglary rates from area to area. For the time effect, we acknowledge the temporal dependence structure in burglary counts by assigning a Gaussian random walk model of order 1 to $\gamma_{1:8}$, which is analogous to an autoregressive model. When the control group is small, i.e., the trend pattern is to be estimated from only a few areas, this temporal smoothing model can help to provide reliable trend estimates. Overdispersion is often encountered in count data analyses, whereby the observed variability exceeds that which can be explained by the Poisson model. This extra variability, captured by $\epsilon_{it} \sim \mathrm{N}(0, \sigma_\epsilon^2)$, might be associated with unobserved risk factors and/or the non-independence of burglaries up to the scale of the areal unit - associated with repeat victimization for example (Johnson and Bowers 2004; Tseloni and Pease 2004). The main aim of this model is to provide an estimate of $\gamma_{1:8}$, the control trend pattern.

Models for the NCC data

Two approaches are considered in order to combine information for estimating the NCC impact. First, data from the ten NCC-targeted COAs were aggregated to form one single time series, from which the overall impact can be examined. Since there were only three NCC areas in 2006, we assume for this second NCC group that their schemes were also implemented in 2005. This assumption is likely to dilute the impact of NCC but is unavoidable if pooling all the areas together. In the second approach we will see how both global and local evaluations can be achieved simultaneously via Bayesian hierarchical models. Furthermore, we will incorporate the coverage rates, the proportion of properties in a COA visited by the Police, to explain heterogeneity in the local impacts. Including coverage rates is a way of recognising the fact that the NCC areas are much smaller than COAs and that if the policy effect increases with the size of the coverage rate this provides further evidence of an NCC effect. Importantly, the estimated trend pattern $\gamma_{1:8}$ from the control group model will be incorporated as an adjustment in both approaches.

In the description that follows and in the appendices, parameters associated with the NCC data are denoted with a superscript *.

A model for the aggregated NCC group

Similar to that for the control group, a Poisson model is fitted to the single time series data. The decomposition of the burglary rate, θ_t^*, comprises an intercept (α^*), the estimated control trend from the control group model (γ_t), the overdispersion parameter (ϵ_t^*) and, crucially, an impact function $f(t, \mathbf{\Omega})$.

Explicitly, we have $\log(\theta_t^*) = \alpha^* + \gamma_t + \epsilon_t^* + I_{t \geq 5} \cdot f(t, \boldsymbol{\Omega})$ where $I_{t \geq 5}$ denotes the indicator function so that the impact function only comes into play when $t \geq 5$, that is after NCC started. First introduced in Box and Tiao (1975), the impact function, $f(t, \boldsymbol{\Omega})$, assesses the nature (through examining various forms of f) and magnitude (through estimating the associated parameters $\boldsymbol{\Omega}$) of the departure of the NCC trend from the control trend. Specification of this impact function is discussed in the following section.

Impact functions

Various forms have been considered, summarised in Table 3. In the step change function, δ quantifies the level shift in (log) burglary rates after the NCC implementation. A negative estimate for δ is indicative of a positive policy impact. The linear function of time allows for a gradual change and $\eta < 0$ suggests a positive policy impact. Both the step and linear functions are special cases of the generalised function, which allows for non-linearity of change on the log scale. In the generalised function, ξ, constrained to be between 0 and 1, controls the degree of non-linearity in the resulting function while a negative ω suggests a positive impact in this application. Further discussion of these three functional forms can be found in Appendix B.

(Table 3 here)

It should be noted that since the before-after comparison is carried out after adjusting for the control trend pattern, these impact functions measure a change in burglary rates relative to the control trend. So, for example, an estimated reduction may not correspond to an actual reduction of burglary cases but rather a lower burglary rate than what would be expected from other non-targetted (non-policy) comparison areas.

Models for the COA-level NCC groups/areas

Through data aggregation, random variability is reduced and hence statistical inference is strengthened. However, information about local effects is lost. Here we extend the model for the aggregated NCC group by allowing parameters in the impact function to be area-specific. These area-specific parameters are treated as random effects and modelled hierarchically. Without aggregation, this model deals with the NCC data at the COA-level. So, for example, for the linear impact function model the slope parameter becomes η_i where i labels the i^{th} NCC-COA and $\eta_i \sim \mathrm{N}(\mu_\eta, \sigma^2_\eta)$. Here η_i quantifies the local impact while μ_η represents the overall effect of NCC.

Analysing the data at the COA level has two advantages in addition to the global-local evaluation. First, we can handle different NCC starting years by using area-specific indicators. Second, to explain the differential local impacts we can further model the area-specific impact parameters by, for example, the COA-specific coverage rates (see Appendix A.3 for details).

Results

Grouped analysis

The overall impact of the NCC scheme is measured by the parameters in the impact functions. Under various criteria for constructing the control group, Table 4 summarises the parameter estimates from the three impact functions. All three impact functions consistently reveal a positive effect associated with the NCC-targeted areas. The burglary rates of the NCC group during the post-NCC period were lower than expected from the control trends. While overall burglary rates were going up, the NCC policy had the effect of "stabilising" the rates in the targeted group, as previously observed in Figure 1. The negative slope estimate for the linear function suggests that the trend for the NCC group moved further away from the trend for the control group over time, implying a persistent and gradual shift. Although the conventionally-defined 95% credible intervals (CI) for both δ and η do not quite exclude zero, the revealed positive impact is robust against different comparison groups (Table 4).

(Table 4 here)

For the generalised function, the posterior means of ω are all negative, leading to a downward pattern. However, with only 4 data points (2005-2008), the rate parameter ξ was poorly estimated with the 95% CI covering virtually the whole constrained interval (0, 1). Additional post-scheme observations could help improve the estimate of this rate parameter.

(Figure 2 here)

Compared to the overall trend from all LSOAs in Peterborough (Control Criterion 1), Figure 2 illustrates the estimated relative reduction in the burglary rate after the implementation of NCC. The solid line represents the posterior mean and the grey area shows the 95% uncertainty region. The slight curvature which appears in the linear function is a result of the exponential back-transformation (the model is linear on the log rate scale). The large uncertainty of ξ led to a larger portion of the grey area lying above 0 for the generalised function, compared to that for the linear function. Comparisons with other definitions of control areas led to similar reduction patterns (results available from the authors on request).

The deviance information criterion (DIC, Spiegelhalter et al. 2002) is used to compare the three impact functions together with the one with no change. DIC is a Bayesian analogue of the Akaike information criterion (AIC) and a smaller DIC value is indicative of a more parsimonious model. The model with the linear impact function yields the smallest DIC value (DIC=11.54) when compared to the model with no change (DIC=15.58), the model with step change (DIC=15.50) and the model with the generalised function (DIC=13.34).

COA-level analysis

Instead of forming one aggregated NCC unit, a Bayesian hierarchical model is utilised to allow for assessment of local impacts. Based on results from the grouped analysis and for ease of interpretation, the linear function is used as the impact function.

Using all LSOAs in Peterborough as controls, the overall mean of the local slopes, namely μ_η, is estimated to be -0.17 (with 95% CI: -0.40, 0.03). This figure translates to a 15% (with 95% CI: -3%, 33%) reduction in burglary rate in NCC areas relative to control areas in the first year and 27% (with 95% CI: -5%, 55%) and 37% (with 95% CI: -8%, 70%) of reduction in the second and third years, respectively[1]. Figure 3 illustrates both these local and the overall (bottom of the plot) NCC impacts after the first year of implementation. In a Bayesian analysis, we can make probability statements about parameters easily using the corresponding posterior distributions. Here, it is useful to obtain the posterior probability that the overall mean of the local slopes is less than zero, $P(\mu_\eta < 0|\text{data})$, representing the probability of overall success. This posterior probability is consistently estimated to be 0.92 or higher across the six comparison groups, providing strong evidence of an overall success from the NCC scheme in reducing burglary rates compared to controls. In particular, constructing the control group using (non-NCC) areas with similar burglary rates

[1] For the sake of simplicity, we demonstrate the calculation using the model for the aggregated NCC group. In that model, the burglary rate is effectively modelled as a product of two terms, $\theta_t^* = \exp(\mathrm{A}) \cdot \exp(\mathrm{B})$ where A=$\alpha^* + \gamma_t + \epsilon_t^*$ and B=$I_{t \geq 5} \cdot \eta \cdot (t-4)$. Here we have replaced the generic function f by the linear function. The second term exp(B) measures the change to the annual burglary rates of the NCC group relative to the rates predicted from the control group, namely exp(A), and hence $\exp(\eta)$ represents the rate of such change (per year). Therefore, the percentage of reduction/increment t^* years after the policy was implemented can be obtained by $(\exp(\eta \cdot t^*) - 1) \times 100$. The associated credible intervals can be readily obtained from the posterior distributions of the transformed variables.

(Criteria 2 and 3) and deprivation scores (Criterion 6) leads to marginally bigger impacts. Also revealed in Figure 3, there exists different impacts between targeted-COAs. Four COAs, namely 00JANH0003, 00JANG0025, 00JANT0027 and 00JANQ0023, appear to have benefited most from the scheme, but others such as 00JANY0010 and 00JANC0016 show little impact with the slope estimates close to 0. These local estimates again are robust across different control criteria.

(Figures 3 and 4 here)

A simple correlation plot (Figure 4) suggests that some of the variability in impacts between COAs may be due to the coverage rates. To examine this, a model was fitted with area coverage as a linear predictor of the COA-specific slopes (see Appendix A.3). The associated regression coefficient was estimated to be negative, $\beta_1 = -1.05$ (with 95% CI: -2.47, 0.13), providing some evidence that the more properties that were visited, the greater the impact of the NCC scheme. As predicted by the model, an area with 10% households visited by the police would have a success probability, i.e., $p(\eta_i < 0|\text{data and coverage} = 10\%)$, of 0.45, which would increase to 0.98 and 0.99 if 30% or 60% of the households were visited. We return to this interpretation in the final section.

Including a covariate of LSOA-level multiple deprivation score (either with or without the coverage rate) shows no evidence of an association between the efficacy of the scheme and local deprivation.

Discussion

Reassurance policing draws to the attention of the police the importance of creating and responding to an "in-depth understanding of places and their problems." (Rix et al. 2009, p. 10). "No Cold Calling" schemes involve targeted policing activity to tackle crime and disorder problems that matter in neighbourhoods; they involve the community in identifying priorities; they produce visible expressions of the police's engagement (Tuffin et al. 2006). However their quantitative evaluation is at a geographical scale that presents a challenge both in terms of data availability and in terms of methodology for evaluating statistical evidence of impact.

Neighbourhood policing programs in general are best evaluated with a "local focus" (Mason 2009) and those evaluations based on quantitative data need to be able to examine such data rigorously to provide statistical evidence of changes over time and with reference to well defined control groups.

Crime reduction was not a stated aim of reassurance policing at the outset (Tuffin et al. 2006). Raising the public's confidence in the police was a priority as was raising people's perception of their safety at home or in the street. The effect of raising the public's confidence in the police could even have the effect of increasing recorded crime. But crime reduction ought to be one of the outcomes of a successful program. If neighbourhood programs do result in reducing crime rates then this may go someway to addressing the skepticism of some police officers with community engagement which has been identified as one of the pitfalls in the implementation of reassurance policing, though not specifically of "No Cold Calling" schemes (Rix et al. 2009).

We have presented a model-based approach to evaluate the "No Cold Calling" policy in Peterborough. This policy is shown to have had a positive impact on stabilising burglary rates in the targeted areas while the general burglary trend moved upwards. Setting a reference frame for comparison, namely the control group, led to a sounder conclusion, for otherwise this evaluation would have reported no change in the burglary rate. Furthermore, results are shown to be robust against different comparison groups, which strengthen the interpretation.

Placing the analysis in a Bayesian framework means that uncertainty associated with the reference trend estimates can be propagated into measuring the policy's impact. Typically, a frequentist method would employ a two-stage approach, where the first stage would involve obtaining point estimates for the trend pattern from the control group. These estimates then enter the NCC model as fixed adjustments. Uncertainty about the trend estimates is thus ignored. In the Bayesian framework, however, all quantities are treated as random variables, including the trend adjustments. As outlined in Appendix C, the trend adjustments are sampled from the corresponding posterior distributions (at each MCMC iteration). Thus, the uncertainty in the trend adjustments are fully accounted for in estimating parameters in the impact function, a distinctive advantage of the Bayesian paradigm. Another advantage of the Bayesian framework is that uncertainty intervals and probability statements can readily be obtained as all parameters are treated as random variables, associated with which are posterior distributions. This means that direct quantitative answers to policy-relevant questions can be obtained - for example, the probability that the NCC scheme was successful in a particular area.

Results from the BHM showed some degree of heterogeneity in the local impacts. These variable impacts could be attributable to factors such as that the NCC scheme might not be implemented uniformly across the targeted areas and/or different areas may have responded differently to the scheme because of differences in local characteristics. Although not associated with deprivation, the level of impact does appear to be strongly associated with the coverage rate. There are two possible explanations for this coverage effect. The scheme was not successful (in terms of reducing burglary rates) in certain areas because too few properties were visited, a genuine "dose level-response" effect. Alternatively, because the COA is the unit of analysis, the NCC impact, even if it is present and does not depend on the "dose level", could be masked when the original NCC area is small relative to the size of the COA within which it is located. Neither of these explanations for the coverage rate effect undermines our overall assessment of the policy's success. However, the possibility that the first explanation is valid means that any future implementation of this scheme should pay careful consideration to whether to continue to target many small areas or instead identify perhaps fewer but rather larger areas. This implies the need to also consider the trade-offs between scheme coverage on the one hand and manageability of the targeted areas on the other. There may be other dangers in adopting too wide coverage. The scheme may owe some of its effect precisely because relatively small areas are targeted and it is this feature which discourages would be offenders.

The results presented here cover quite a short period of time following the implementation of the scheme but nonetheless it is long enough to allow any changes to be assessed. Longer time series of data, when they become available, allow further assessments of existing schemes which may provide not only evidence for the robustness of the findings reported here but also on the longer term shape of any impact function. Analyses of other schemes in other parts of Cambridgeshire and Peterborough will also add to the evidence base and will be the subject of future work.

Acknowledgements

This work was supported by the Economic and Social Research Council (ESRC) through the National Centre for Research Methods, BIAS II node [grant number RES-576-25-0015]. Sylvia Richardson and Nicky Best are investigators in the MRC-HPA Centre for Environment and Health. The authors are grateful to Cambridgeshire Constabulary for making available to us its recorded offence data (at the COA level between 2001 and 2003 and at the postcode level between 2004 and 2008) for the period 2001-2008 and also wish to express their thanks to Shirley Simpson, Emilia Keen and Brin Hodgskiss for their advice and input into this evaluation. The views expressed in this paper are those of the authors and do not necessarily reflect those of Cambridgeshire Constabulary. *Conflict of Interest:* None declared.

References

Berman, G. (2008), Changes in crime recording practices., Technical report, Library of the House of Commons, Standard Note SN/SG/2607.

Box, G. and Tiao, G. (1975), 'Intervention analysis with applications to economic and environmental problems', *Journal of the American Statistical Association* **70**(349), 70–79.

Cambridgeshire-Constabulary (2008), Evaluation of "no cold calling" areas in Cambridgeshire, Technical report, Cambridgeshire Constabulary.

Campbell, D., Stanley, J. and Gage, N. (1963), *Experimental and quasi-experimental designs for research*, Rand McNally Chicago.

Gelman, A. (2006), 'Prior distributions for variance parameters in hierarchical models', *Bayesian analysis* **1**(3), 515–533.

Gelman, A., Carlin, J., Stern, H. S. and Rubin, D. B. (2004), *Bayesian data analysis*, CRC Pressress.

Gelman, A. and Hill, J. (2007), *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press New York, USA.

Gill, J. (2007), *Bayesian methods: A social and behavioral sciences approach*, second edn, Chapman and Hall/CRC press.

Gillings, D., Makuc, D. and Siegel, E. (1981), 'Analysis of interrupted time series mortality trends: an example to evaluate regionalized perinatal care.', *American Journal of Public Health* **71**(1), 38.

Iversen, G. (1984), *Bayesian statistical inference*, Sage Publications, Inc.

Johnson, S. and Bowers, K. (2004), 'The burglary as clue to the future: the beginnings of prospective hot spotting', *European Journal of Criminology* **1**(2), 237–255.

Lunn, D., Thomas, A., Best, N. and Spiegelhalter, D. (2000), 'Winbugs-a bayesian modelling framework: concepts, structure, and extensibility', *Statistics and Computing* **10**(4), 325–337.

MacDonald, J., Golinelli, D., Stokes, R. and Bluthenthal, R. (2010), 'The effect of business improvement districts on the incidence of violent crimes', *Injury Prevention* **16**(5), 327.

Mason, M. (2009), Findings from the second year of the national neighbourhood policing programme evaluation. home office, research report no. 14., Technical report, Research, Development and Statistics Directorate, London.

McCullagh, P. and Nelder, J. (1989), *Generalized Linear Models*, second edn, Chapman and Hall/CRC.

Morgan, O., Griffiths, C. and Majeed, A. (2007), 'Interrupted time-series analysis of regulations to reduce paracetamol (acetaminophen) poisoning', *PLoS medicine* **4**(4), 654.

Rix, A., Joshue, F., Maguire, M. and Morton, S. (2009), Improving public confidence in the police: a review of the evidence. home office research report no. 28. (2nd edition), Technical report, Research, Development and Statistics Directorate, London.

Skogan, W. G., Hartnett, S. M., Bump, N. and Dubois, J. (2008), Evaluation of ceasefire-chicago, Technical report, Institute for Policy Research, Northwestern University.

Spiegelhalter, D., Best, N., Carlin, B. and van der Linde, A. (2002), 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society Series B* **64**, 583–616.

Studnicki, J., Remmel, R., Campbell, R. and Werner, D. (1997), 'The impact of legislatively imposed practice guidelines on cesarean section rates: the florida experience', *American Journal of Medical Quality* **12**(1), 62.

Thum, Y. and Bhattacharya, S. (2001), 'Detecting a change in school performance: A bayesian analysis for a multilevel join point problem', *Journal of Educational and Behavioral Statistics* **26**(4), 443.

Tseloni, A. and Pease, K. (2004), 'Repeat personal victimization: random effects, event dependence and unexplained heterogeneity', *British Journal of Criminology* **44**(6), 931–945.

Tuffin, R., Morris, J. and Poole, A. (2006), An evaluation of the impact of the national reassurance policing programme, Technical report, Home Office Research Study 296.

Table 1: Summary of the NCC scheme in Peterborough started in 2005 and 2006 at the COA level

	COA_code	Postcode	Area name	Number of targeted dwellings	Started	Number of dwellings in COA	Coverage (%)
	NCC 2005						
1	00JANC0016	PE3 8JU	Hanover Court, Bretton	42	10/03/2005	122	34
2	00JANE0006	PE1 2NL	Kimbolton Court, Millfield	48	28/01/2005	150	32
3	00JANE0010	PE1 3RR	Eaglesthorpe, New England	28	28/01/2005	151	19
4	00JANG0013	PE1 5JD	Mellows Close, Eastfield	12	25/01/2005	131	9
5	00JANH0003	PE6 7TZ	Boxgrove Close, Eye	8	03/06/2005	126	6
6	00JANQ0023	PE2 6XN	Napier Place, Orton Wistow	54	17/03/2005	103	52
7	00JANT0027	PE4 7PS	Bevishall, Paston	45	20/07/2005	127	35
8	00JANY0010	PE4 6QT	Dudley Avenue, Walton	10	20/12/2005	122	8
	NCC 2006						
9	00JANG0025	PE1 4SL	Keys Park Mobile Home Park	100	22/08/2006	168	60
10	00JANH0003	PE6 7XF	Woad Court	28	19/10/2006	126	22
11	00JAPB0010	PE3 6LA	Thorpe Avenue	36	27/09/2006	128	28

Table 2: Construction of 6 different control groups based on similar local characteristics.

Control criterion	Description	No. of LSOAs
1	All LSOAs in Peterborough	88
2	Within 10% of the burglary rate of the NCC group in 2005	9
3	Within 20% of the burglary rate of the NCC group in 2005	20
4	Within 30% of the burglary rate of the NCC group in both 2004 and 2005	7
5	LSOAs containing the NCC-targeted COAs (but excluding the NCC-targeted COAs)	10
6	LSOAs that had "similar" multiple deprivation scores to those for the NCC LSOAs in 2004	46

Table 3: Various functional forms for the impact function. The parameters δ, η, ξ and ω describe the policy impact. t_0 indicates the start of policy.

Change	Functional form	Prob. of success
No impact	$f(t) = 0$	–
Step change	$f(t, \delta) = \delta$	$p(\delta < 0 \mid \text{data})$
Linear function of time	$f(t, \eta) = \eta \cdot (t - t_0 + 1)$	$p(\eta < 0 \mid \text{data})$
A generalised function	$f(t, \xi, \omega) = \xi \cdot f(t-1) + \omega$ if $t \geq t_0$ $f(t) = 0$ otherwise.	$p(\omega < 0 \mid \text{data})$

Table 4: Posterior means and 95% credible intervals of parameters in the three impact functions fitted to the aggregated NCC data.

Control	Step change	Linear function	Generalised function	
criterion	δ (after-before)	η (slope)	ξ (rate)	ω
1	-0.15 (-0.60, 0.33)	-0.11 (-0.27, 0.04)	0.48 (0.04, 0.96)	-0.15 (-0.48, 0.17)
2	-0.22 (-0.75, 0.31)	-0.13 (-0.30, 0.04)	0.48 (0.04, 0.96)	-0.19 (-0.57, 0.15)
3	-0.18 (-0.70, 0.30)	-0.11 (-0.29, 0.05)	0.47 (0.03, 0.96)	-0.15 (-0.51, 0.17)
4	-0.20 (-0.66, 0.27)	-0.09 (-0.24, 0.07)	0.43 (0.03, 0.95)	-0.14 (-0.48, 0.17)
5	-0.14 (-0.73, 0.44)	-0.11 (-0.31, 0.10)	0.47 (0.03, 0.96)	-0.14 (-0.57, 0.28)
6	-0.21 (-0.64, 0.19)	-0.13 (-0.28, 0.01)	0.50 (0.03, 0.97)	-0.19 (-0.52, 0.10)

Figure 1: Temporal burglary profiles of the individual NCC-targeted COAs. Superimposed are the overall burglary trend from all LSOAs in Peterborough excluding the NCC areas and the trend from the NCC group, including both 2005 and 2006 NCC COAs.

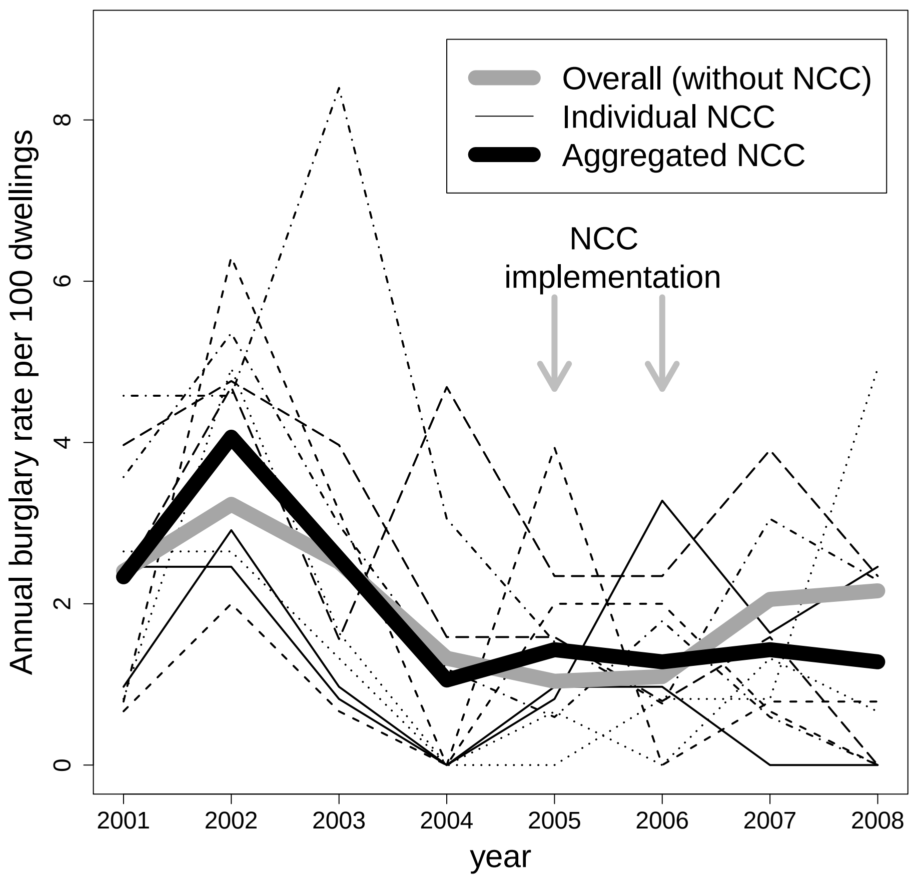


Figure 2: Percentage change in burglary rate in the NCC group relative to control group 1, which consists of all 88 LSOAs in Peterborough. These changes are estimated from the aggregated NCC data. The solid lines are the posterior means of the exponentiated impact function with the associated 95% credible intervals in grey. The horizontal dotted line represents the case where the NCC scheme had no impact.

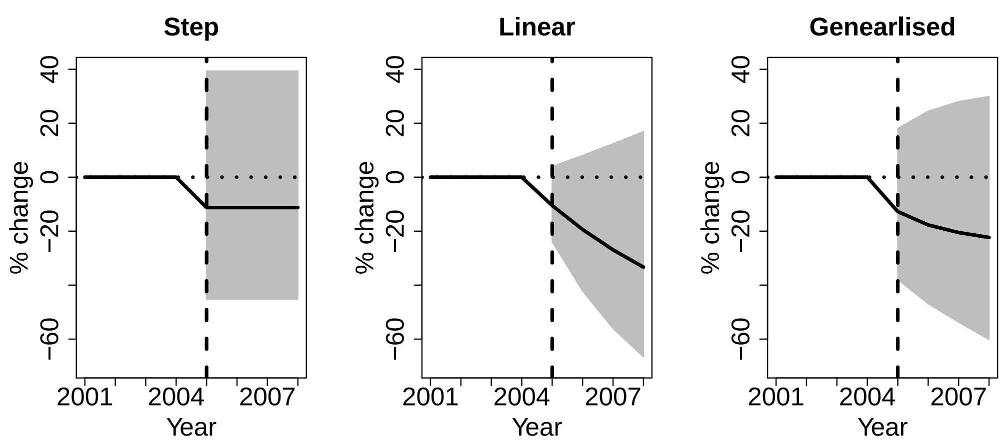


Figure 3: Percentage change in burglary rates in the NCC areas after the first year of NCC implementation compared to various control groups (a negative value indicates a relative reduction). These changes are estimated using the NCC data at the COA-level. The points, where different symbols correspond to different control groups, represent the posterior means and the horizontal bars are the 95% CIs. The number next to each area label is the posterior probability of local success, $P(\eta_i < 0|\text{data})$ and $P(\mu_\eta < 0|\text{data})$ (at the bottom) is the overall success relative to control group 1.

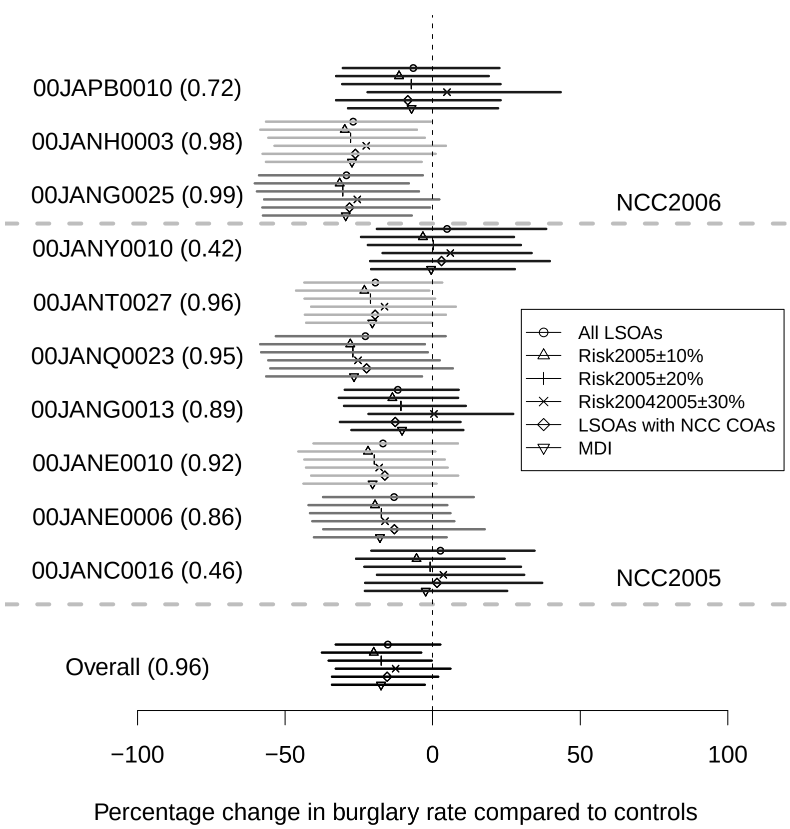
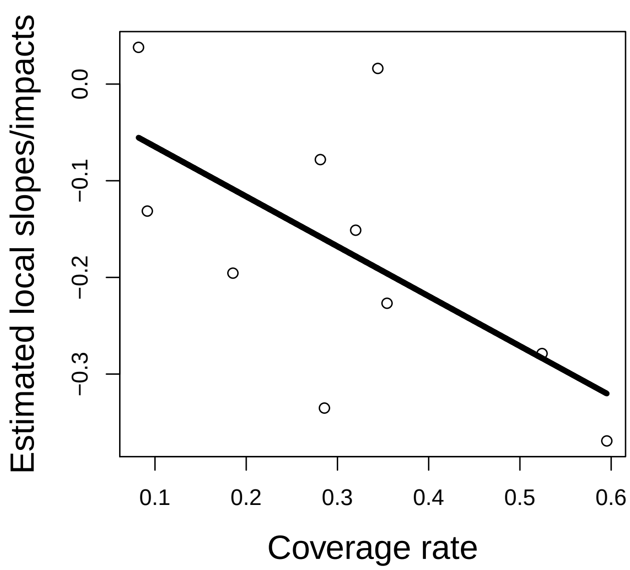


Figure 4: Initial assessment of the relationship between local impacts (the posterior means of the linear slopes η_i) and coverage rates. The fitted least squares line is shown.



A Full specifications of various models

A.1 A model for the control group

Let y_{it} be the number of recorded burglary cases in area i of year t. We then model these count data using a Poisson model,

$$\begin{aligned} y_{it} &\sim \text{Poisson}(n_i \cdot \theta_{it}) \\ log(\theta_{it}) &= u_i + \gamma_t + \epsilon_{it} \\ u_i &\sim \text{N}(\alpha, \sigma_u^2) \\ \gamma_{1:8} &\sim \text{RW}_1(\mathbf{W}, \sigma_\gamma^2) \\ \epsilon_{it} &\sim \text{N}(0, \sigma_\epsilon^2) \end{aligned}$$

where n_i is the number of dwellings in area i, assumed to be unchanged over time and θ_{it} is the burglary rate. On the log scale, the burglary rate is a linear combination of three terms. u_i and γ_t represent the place and time effects, respectively, both modelled as random effects. A Gaussian random walk model of order 1 (RW_1) is assigned to $\gamma_{1:8}$. The temporal dependence is imposed through the weight matrix $\mathbf{W}$ of dimension 8×8. The diagonal entries $w_{ii} = 0$ and the off-diagonal entries $w_{ht} = 1$ if $|h - t| = 1$ and $w_{ht} = 0$ otherwise. This setting specifies a "neighbouring" structure, where the previous and next time points are neighbours of the current time point (apart from the first and last points in the time series where the second and the penultimate time points are their only neighbours, respectively). So information can be "borrowed" in estimating these random effects. ϵ_{it} captures the overdispersion.

To complete the model, we assign a vague prior for the overall intercept $\alpha \sim \text{N}(0, 10000)$ and a weakly informative half Normal prior $\text{N}(0, 100)$ bounded strictly below by 0, as suggested in Gelman (2006), for the random effect standard deviations σ_u, σ_γ and σ_ϵ.

A.2 A model for the aggregated NCC data

Denoting y_t^* the number of recorded burglary cases in the NCC group in year t and n^* the total number of dwellings, we have

$$\begin{aligned} y_t^* &\sim \text{Poisson}(n^* \cdot \theta_t^*) \\ log(\theta_t^*) &= \alpha^* + \gamma_t + \epsilon_t^* + I_{t\geq 5} \cdot f(t, \mathbf{\Omega}) \\ \alpha^* &\sim \text{N}(\alpha, \sigma_u^2) \\ \epsilon_t^* &\sim \text{N}(0, \sigma_\epsilon^2) \end{aligned}$$

where the time trend γ_t and the overdispersion variance σ_ϵ^2 are both estimated from the control group model. The impact function $f(t, \mathbf{\Omega})$ is described in Appendix B. Multiplied with the impact function is an indicator function which takes 1 if $t \geq 5$ (corresponding to 2005) and 0 otherwise. Both the intercept α^* and the overdispersion random effects ϵ_t^* are assumed to follow the corresponding distributions, namely, for u_i and for ϵ_{it} respectively, in the control group model. The prior for ϵ_t^* reflects the assumption that apart from the exposure to NCC, both the control areas and the NCC areas are similar in local characteristics, including factors associated with overdispersion. The assumption on the intercept minimises the differences in the estimated α^* between different impact functions.

A.3 Models for the COA-level NCC groups/areas

Instead of data aggregation, data at the COA-level can be combined using the following hierarchical model. For the sake of illustration, the linear function of time is used as the impact function but the BHM can be specified with other functional forms. For each NCC-targeted COA, we have

$$\begin{aligned}
y_{it}^* &\sim \text{Poisson}(n_i^* \cdot \theta_{it}^*) \\
log(\theta_{it}^*) &= \alpha_i^* + \gamma_t + \epsilon_{it}^* + I_{t \geq t_{0,i}} \cdot \eta_i \cdot (t - t_{0,i} + 1) \\
\alpha_i^* &\sim \text{N}(\alpha, \sigma_u^2) \\
\eta_i &\sim \text{N}(\mu_\eta, \sigma_\eta^2) \\
\epsilon_{it}^* &\sim \text{N}(0, \sigma_\epsilon^2)
\end{aligned}$$

Here, the NCC starting year $t_{0,i}$ (=5 or 6) as well as the slope in the linear function become area-specific. A vague hyperprior is assigned to μ_η, the global impact, and a moderately informative half Normal prior N(0,1) bounded strictly below by 0 is assigned to the slope standard deviation σ_η. This half Normal prior corresponds to our expectation that the impact of NCC, if present, after the first year would result in no more than a five-fold reduction/increment of the burglary rate in the NCC area compared to that of the control areas.

Similar to the model for the single time series data, the area-specific intercepts, α_i^*, are assumed to follow the same hyper-distribution as that for the control areas, likewise, for the overdispersion parameters ϵ_{it}^*.

The coverage rate can be readily incorporated as a linear predictor of the local impacts:

$$\begin{aligned}
\eta_i &\sim \text{N}(\mu_{\eta,i}, \sigma_\eta^2) \\
\mu_{\eta,i} &= \beta_0 + \beta_1 \cdot \text{coverage}_i
\end{aligned}$$

Again, a vague prior, N(0,10000), is assigned to the regression parameters β_0 and β_1.

B Impact functions

The linear function considered in this paper is a simplified version of the so called segmented regression (or piecewise linear regression) models considered in Gillings et al. (1981) but with only one evaluation period. A slightly more complex Bayesian model by Thum and Bhattacharya (2001) attempted to estimate when the change in slope occurred. The complexity of the second model, however, may not be required here since we know the starting years of NCC. The generalised function considered is an autoregressive model of order 1 with a deterministic innovation, namely ω. For continuity, we set $\omega = \alpha^* - b$ and assign a vague prior, N(0,10000), on b. The parameter ξ is constrained, i.e., $0 < \xi < 1$, so that the resulting function is continuous and possesses an asymptote that $f(t) \to \frac{\omega}{1-\xi}$ as $t \to \infty$. In the limiting cases, when $\xi = 0$, the generalised function reduces to the step change function while when $\xi = 1$, we have the linear function. So ξ is interpreted as a rate parameter that controls how fast ($\xi \to 0$) or slow ($\xi \to 1$) the generalised function reaches its asymptote. Although this generalised function provides a better description of change, it requires more post-scheme observations in order to estimate the rate parameter ξ well. Nevertheless, being able to allow for non-linearity has its practical attractiveness.

C Implementation of the two models

Both the control and NCC models are implemented in WinBUGS (Lunn et al. 2000), a software that is specially designed for performing Bayesian analyses. Markov chain Monte Carlo (MCMC) methods are used to sample from the posterior distributions in an iterative manner. Posterior summaries, e.g., means, uncertainty intervals and posterior probabilities, are then obtained from these posterior samples. The WinBUGS code is provided in Appendix D. It is worth mentioning that while both models are simultaneously fitted, estimation of all model parameters for the control group is independent to the fitting of the NCC data. Through the use of the WinBUGS *cut* function, the trend $\gamma_{1:8}$, the overall intercept, α, the intercept variance, σ_u^2, and the overdispersion variance, σ_ϵ^2, in the control model are "fed" into the NCC model at each iteration but no information is feeding back to the control model. This ensures the estimation of the control trend (and other parameters) is independent to the policy quantification but, importantly, uncertainty of these estimates are properly accounted for in the impact estimation.

D WinBUGS code for fitting the BHM with the linear impact function

```
#
#    Area-specific NCC impacts are allowed in this model
#    The linear impact function is used
#
model {
#       constructing the control trend pattern
  for (i in 1:N) {
    for (t in 1:T) {
      y[i,t] ~ dpois(mu[i,t])
      mu[i,t] <- n[i]*p[i,t]
      log(p[i,t]) <- u[i] + gamma[t] + eps[i,t]
#       overdispersion
      eps[i,t] ~ dnorm(0,prec.eps)
    }
  }
#       prior for the overall intercept
  alpha ~ dnorm(0,0.0001)
#       RW prior for the control trend (time effects)
  gamma[1:T] ~ car.normal(adj.tm[],weights.tm[],num.tm[],prec.gamma)
#       place effects
  for (i in 1:N) {u[i] ~ dnorm(alpha,prec.u)}
```

```
#       priors for the standard deviations
  prec.gamma <- pow(sigma.gamma,-2)
  sigma.gamma ~ dnorm(0,0.01)I(0,)
  prec.u <- pow(sigma.u,-2)
  sigma.u ~ dnorm(0,0.01)I(0,)
  prec.eps <- pow(sigma.eps,-2)
  sigma.eps ~ dnorm(0,0.01)I(0,)

#       'feeding' the overdispersion variance to the NCC model
  prec.eps.ncc <- cut(prec.eps)

#       model for the NCC data at the COA level
  for (t in 1:T) {
    for (i in 1:all.NCC) {
      yN[i,t] ~ dpois(muN[i,t])
      muN[i,t] <- nN[i]*pN[i,t]

#       linear change
      log(pN[i,t]) <- alphaN[i] + gammaN[t]
                   + step(t-ncc.year[i])*eta[i]*(t-ncc.year[i]+1) + eps.ncc[i,t]

#       overdispersion
      eps.ncc[i,t] ~ dnorm(0,prec.eps.ncc)
    }
#       assigning trend adjustments from the control group model
    gammaN[t] <- cut(gamma[t])
  }

#    prior on the place effects for the NCC COAs
  for (i in 1:all.NCC) {alphaN[i] ~ dnorm(mu.alphaN,prec.alphaN)}

#    imposing an informative prior (estimated from the control group)
#    on the NCC intercepts
  mu.alphaN <- cut(alpha)
  prec.alphaN <- cut(prec.u)

#      prior for the COA-specific impacts/slopes
  for (i in 1:all.NCC){eta[i] ~ dnorm(mu.eta,prec.eta)}

  mu.eta ~ dnorm(0,0.0001)
  prec.eta <- pow(sigma.eta,-2)
  sigma.eta ~ dnorm(0,1)I(0,)

#    for plotting
  for (t in 1:T) {time[t] <- t}
}
```