Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: https://www.tandfonline.com/loi/cjas20

Bayesian semi-parametric analysis of Poisson change-point regression models: application to policy-making in Cali, Colombia

Taeyoung Park , Robert T. Krafty & Alvaro I. Sánchez

Journal of Applied Statistics
Vol. 39, No. 10, October 2012, 2285–2298

Taylor & Francis
Taylor & Francis Group

Bayesian semi-parametric analysis of Poisson change-point regression models: application to policy-making in Cali, Colombia

Taeyoung Park[a]*, Robert T. Krafty[b] and Alvaro I. Sánchez[c,d]

[a]*Department of Applied Statistics, Yonsei University, Seoul 120-749, Korea;* [b]*Department of Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA;* [c]*Department of Surgery, University of Pittsburgh, Pittsburgh, PA 15260, USA;* [d]*CISALVA Institute, Universidad del Valle, Cali, Colombia*

(*Received 9 November 2011; final version received 2 July 2012*)

A Poisson regression model with an offset assumes a constant baseline rate after accounting for measured covariates, which may lead to biased estimates of coefficients in an inhomogeneous Poisson process. To correctly estimate the effect of time-dependent covariates, we propose a Poisson change-point regression model with an offset that allows a time-varying baseline rate. When the non-constant pattern of a log baseline rate is modeled with a non-parametric step function, the resulting semi-parametric model involves a model component of varying dimensions and thus requires a sophisticated varying-dimensional inference to obtain the correct estimates of model parameters of a fixed dimension. To fit the proposed varying-dimensional model, we devise a state-of-the-art Markov chain Monte Carlo-type algorithm based on partial collapse. The proposed model and methods are used to investigate the association between the daily homicide rates in Cali, Colombia, and the policies that restrict the hours during which the legal sale of alcoholic beverages is permitted. While simultaneously identifying the latent changes in the baseline homicide rate which correspond to the incidence of sociopolitical events, we explore the effect of policies governing the sale of alcohol on homicide rates and seek a policy that balances the economic and cultural dependencies on alcohol sales to the health of the public.

Keywords: Bayesian analysis; change-point model; inhomogeneous Poisson process; Markov chain Monte Carlo; partial collapse; Poisson regression

1. Introduction

According to the 2011 Global Study on Homicide by the United Nations Office on Drugs and Crime [26], the homicide rate in the Americas (15.6 per 100,000 population) is more than double the world average (6.9 per 100,000 population). Within the region of Americas, homicide rates are not homogeneous; the Central and South America sub-regions have considerably higher homicide rates than the North America sub-region. Field studies among countries where homicide rates are

*Corresponding author. Email: tpark@yonsei.ac.kr

ISSN 0266-4763 print/ISSN 1360-0532 online

http://dx.doi.org/10.1080/02664763.2012.709227
http://www.tandfonline.com

exorbitantly high have shown attribution to political, military, sociological, and public health factors [10,14]. In comparison to global rates, the distribution of homicide rates in Central and South Americas placed Colombia in the category of 'very high'. Despite a declining national trend in homicide rates in Colombia during the last decade [7], homicide remains a primary public health problem with the majority of interpersonal violent deaths being concentrated in the country's largest cities: Bogotá, Medellín, and Cali [10,16,28].

Cali is Colombia's third largest city. In 1993, a municipal fatal injury surveillance system was developed in Cali by the city's mayor [6,14]. This endeavor was maintained and further expanded under the name of 'crime observatories' by the CISALVA Institute at the Universidad del Valle [16]. Homicide in Cali became the leading cause of death in the general population over the 1990s and still remains the leading cause of death among people between 10 and 29 years of age [6,15]. Because it is known that alcohol consumption has been associated with an increase in interpersonal violent deaths [6,23], policies restricting the sale of alcoholic beverages after given hours have been implemented by Cali's mayor in an effort to reduce the likelihood of exposure to interpersonal violence [6,14,28]. However, evidence of a significant reduction in homicide rates in Cali has not been objectively assessed. Thus, the economic impact of the sales of alcoholic beverages, both directly on local merchants and indirectly on public works through alcohol taxes, in addition to cultural ties to the consumption of alcohol, has caused considerable debate about the benefits of polices that restrict the hours when alcohol may be sold.

Figure 1(a) shows the time-series plot of five distinct policies governing the legal sale of alcoholic beverages between January 1999 and August 2008. As shown in Figure 1(a), some policies on alcohol sales have remained in place for as few as 3 days, but many were continuously enforced for as long as 346 days. Hence, an accurate assessment of the association between policies on alcohol sales and homicide rates requires a long period of observation which is inherently subject to occurrences of unmeasurable sociopolitical events that affect homicide rates. The primary goal of this article is to use the data collected by the CISALVA Institute to correctly determine the effect that restrictions on alcohol sales had on the homicide rate in Cali, Colombia, from January 1999 until August 2008. In addition, we seek to further investigate the pattern of homicides by identifying and examining time points when a baseline homicide rate changed in the period of observation. Not only the identification of these change points avoids biases in the inference concerning the association of alcohol polices and homicide rates, but the investigation of sociopolitical events which correspond to these change points can also help illuminate the process driving homicide rates and can provide insight for the design of future studies.

The most common tool in the analysis of homicide rates over time is Poisson log-linear regression where, conditional on some covariates, a homicide rate is assumed to be constant [3,28]. Figure 1(b) shows the time-series plot of daily homicides collected by the CISALVA Institute from January 1999 to August 2008, where the dotted lines represent an estimate of the mean homicide rate via Gaussian kernel smoothing with a standard deviation of 14 days. A visual inspection of the daily homicides in Figure 1(b) raises questions about the existence of a constant baseline homicide rate across the observation period, where the overall sample mean and variance of the number of homicides are 5.06 and 9.82, respectively. To account for a potentially non-constant baseline rate, we characterize a baseline rate as a non-parametric step function whose width and height are both unknown. In this way, we assume that the underlying baseline rate is changed instantaneously rather than gradually. Although both instantaneous and gradual changes in the underlying baseline rate are plausible, the change-point model assumes that instantaneous changes can well approximate gradual changes in the baseline rate and has an advantage of identifying and examining change points for further research. Subsequently, we propose a Poisson change-point regression model with a piecewise-constant baseline homicide rate conditional on measurable covariates, where the number and location of change points for the baseline homicide rate are unknown. This Poisson change-point regression model does not only avoid bias in the estimation

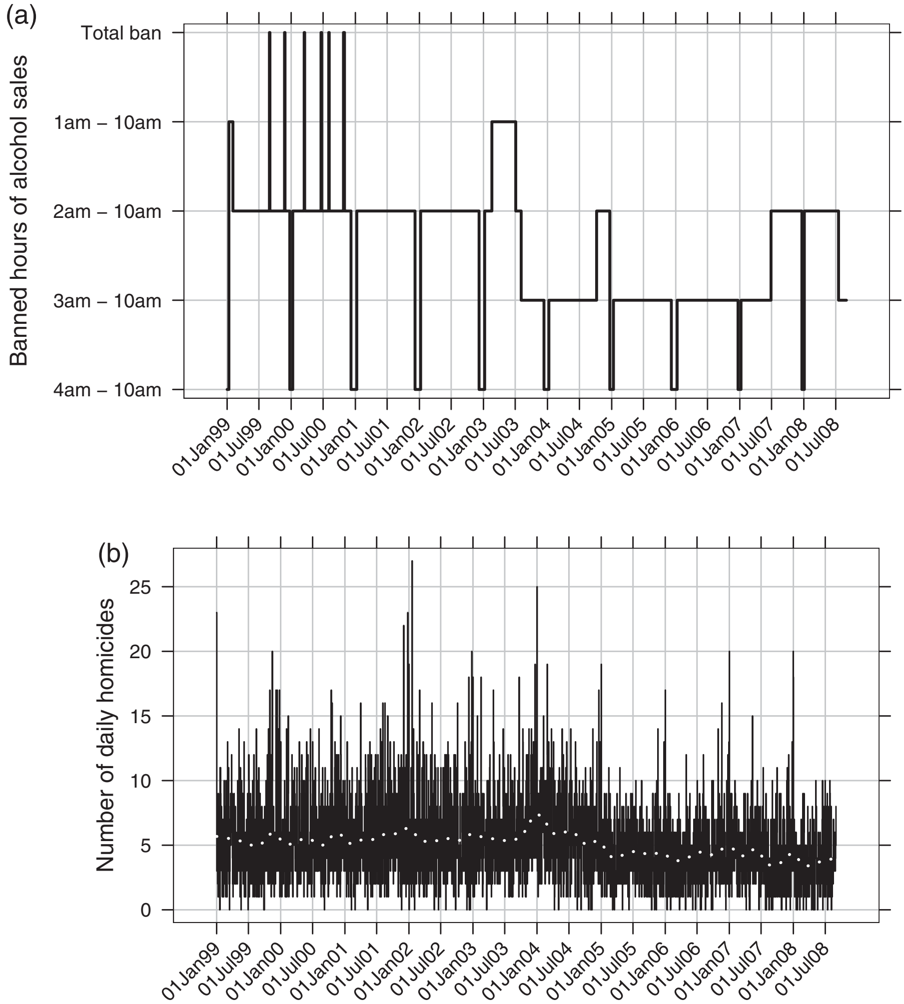

Figure 1. The time-series plots of (a) five distinct alcohol bans and (b) daily homicides in Cali, Colombia, from January 1999 to August 2008. An estimate of the mean homicide rate obtained by applying a Gaussian kernel smoother (dotted curve) is also shown.

of the association between alcohol restrictions and homicide rates incurred by misspecifying the model to have a constant baseline rate conditional on covariates, but also identify points in time when a baseline homicide rate changes.

In the absence of time-dependent covariates, much work has been devoted to the change-point analysis. Poisson change-point models where the number of change points is known have been studied by Smith [24], Raftery and Akman [22], and Carlin *et al.* [4]. Bayesian methods when the number of change points is unknown have been proposed through product partition models [1,2] and through a binary segmentation procedure [29]. These methods, however, assume that the change points partition the data into disjoint segments such that data from different segments are independent. This assumption makes the aforementioned methods difficult to accommodate time-dependent covariates which affect rates across different time segments. Recently, several researchers have proposed to estimate the unknown number of change points by fitting separate models with different numbers of change points and using an information criterion to select the optimal model [5,18,19]. These methods could be easily extended to include covariate effects, yet the inference for these covariate effects will be difficult to interpret as they depend on the number of change points.

Ideally, the inference for covariate effects should account for the fact that the dimension of a baseline rate is subject to the unknown number of change points. To avoid biases in the estimates of the covariate effects, we thus allow such a varying-dimensional component in a model and average the covariate effects over the unknown number and location of change points. In the absence of time-dependent covariates, the Bayesian Poisson change-point model that involves a varying-dimensional model component has been fitted through a reversible-jump Markov chain Monte Carlo (MCMC) algorithm [12], which can also be extended to include time-dependent covariates. Like our proposed method, it allows for the estimation of the posterior distribution of the number of change points, the posterior probability of a change point occurring at any point in time, and the posterior distribution of the intensity over time. The reversible-jump MCMC algorithm for the multiple change-point Poisson model is, however, computationally expensive and can have slow convergence, due to low acceptance rates, even in the absence of covariates [13]. By contrast, our proposed method exhibits faster convergence due to partial collapse; see [8,20] for the theoretical and empirical evidence of improved convergence.

This article proposes a Bayesian semi-parametric Poisson change-point regression model that assumes an unknown number of change points with unknown locations and allows for the assessment of time-dependent covariates. The number of homicides per day is modeled as an inhomogeneous Poisson process whose log rate is the sum of a log baseline rate that is constant for the block of time between consecutive change points and a linear function of a set of covariates. Possible over-dispersion is accounted for by assuming a gamma prior distribution on the baseline rate between change points. The number and location of the change points are flexibly modeled by assuming *a priori* that each day can represent a change point of the baseline rate with equal probability. The computational infeasibility of fitting such a model and possible intractability through conventional MCMC algorithms have hindered the use of similarly flexible and intuitive models. The partially collapsed Gibbs (PCG) sampler [8] is an iterative sampling algorithm that properly replaces some of the conditional distributions of its parent Gibbs sampler with conditional distributions that are conditioned on less, thereby expecting much better convergence properties than the Gibbs sampler. This strategy has been applied to fit a multilevel spectral model in high-energy astrophysics, a joint segmentation model for multivariate time-series data, and a joint imputation model for non-nested data [20]. Here, we extend the applicability of the PCG sampler to seemingly computationally intractable models by exploiting functional incompatibility among conditional distributions while improving the convergence characteristics of its parent Gibbs sampler. We thus develop a PCG sampler to efficiently fit the Poisson change-point regression model with an unknown number of change points. When coupled with a Metropolis algorithm for the estimation of coefficients, the resulting MCMC-type algorithm not only estimates the posterior distributions of the coefficients and the piecewise-constant baseline homicide rate, but also estimates the posterior probabilities of a change point occurring at each point in time. This procedure subsequently allows for the efficient evaluation of the association of homicide rates and measurable covariates, such as the restrictions on the sale of alcohol, while simultaneously identifying change points that can be examined to better understand changes in homicide rates in Cali.

We introduce the proposed Poisson change-point regression model in Section 2 and discuss its Bayesian analysis through PCG sampling in Section 3. Section 4 demonstrates the empirical properties of the proposed model through a simulation study. The daily homicide rates in Cali from January 1999 until August 2008 are analyzed in Section 5, while a discussion is presented in Section 6. The details of the MCMC-type algorithm are presented in Appendix 1.

2. The Poisson change-point regression model

In this section, we develop a Poisson change-point regression model with an unknown number of change points that is extended to include time-dependent covariates. After adjusting for some measurable covariates, we assume that an expected homicide rate is piecewise constant, allowing a non-constant baseline rate. In this way, T time points are partitioned into *unknown* $B(\leq T)$ time blocks within which an expected baseline homicide rate is constant. We consider index sets for the disjoint time blocks, denoted by $\{\mathcal{B}_b\}_{b=1}^{B}$, that are constructed by sequentially combining time points, where $\mathcal{B}_b$ is the set of indices for time points within the time block b. That is, $\mathcal{B}_b \subset \{1, \dots, T\}$ is disjoint and sequential with union $\{1, \dots, T\}$.

The number of homicides on day t is modeled as an inhomogeneous Poisson process:

$$y_t \overset{\text{ind}}{\sim} \text{Poisson}(P_t \xi_t) \quad \text{for } t = 1, \dots, T, \tag{1}$$

where P_t is the known population size on day t and ξ_t represents the expected homicide rate on day t. To account for the stochastic baseline homicide rate, we propose a Poisson change-point regression model that has the link function given by

$$\log(\xi_t) = \log \lambda(t) + \boldsymbol{X}_t^{\mathrm{T}} \boldsymbol{\beta},$$

where $\lambda(t)$ specifies a stochastic baseline rate on day t after accounting for covariate effects, $\boldsymbol{X}_t$ is a $(p \times 1)$ vector of covariates on day t, and $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^{\mathrm{T}}$ is a $(p \times 1)$ vector of the corresponding coefficients. In particular, the stochastic baseline rate on day t, $\lambda(t)$, is equal to λ_b for $t \in \mathcal{B}_b$ and $b = 1, \dots, B$, and $\boldsymbol{\lambda}_B = (\lambda_1, \dots, \lambda_B)$ denotes a collection of the block-specific baseline rates. When $B = 1$, the model in Equation (1) is reduced to a standard Poisson regression model that parameterizes the log rate as a linear function of covariates and treats the baseline rate as a constant for the entire time period. When $B = T$, model (1) corresponds to the Poisson generalized linear mixed model and the baseline rate continually changes over time t. In addition, $\boldsymbol{X}_t = \boldsymbol{0}$ results in the Bayesian Poisson change-point model [12] which ignores the effect of time-dependent covariates and focuses on structural changes solely in log rates.

Because both the location of block change points and the number of time blocks are unknown, we consider the method of data augmentation [25] to simplify the analysis of block change points. That is, we introduce latent block change-point indicator variables $\boldsymbol{Z} = \{Z_t\}_{t=1}^{T}$, where $Z_t = 1$ if a block change occurs on day t and 0 otherwise. The first and the last indicator variables are set to 0 and 1, respectively, so that the sum of the indicator variables corresponds to the number of time blocks, that is, $S(\boldsymbol{Z}) = \sum_{t=1}^{T} Z_t = B$; we sometimes substitute B with $S(\boldsymbol{Z})$ to emphasize that the number of time blocks is a function of latent variables $\boldsymbol{Z}$. We model the latent block change-point indicator variable with a Bernoulli distribution, that is,

$$Z_t \overset{\text{iid}}{\sim} \text{Bernoulli}(\rho) \quad \text{for } t = 1, \dots, T,$$

with $p(Z_t = 1) = \rho$. In this way, we *a priori* assume that each day can represent a change in the non-constant baseline log rate with equal probability.

3. Bayesian analysis via PCG sampling

The Bayesian hierarchical model is developed by specifying proper prior distributions on model parameters. The baseline rate in time block b is hierarchically modeled as

$$\lambda_b \overset{\text{ind}}{\sim} \text{Gamma}(1, \gamma) \quad \text{for } b = 1, \dots, S(\boldsymbol{Z}),$$

with rate parameter γ, so that $\mathrm{E}(\lambda_b) = 1/\gamma$. We assume that the hyper-parameter γ has a conjugate gamma prior distribution:

$$\gamma \sim \mathrm{Gamma}(g_1, g_2),$$

where g_1 and g_2 are *a priori* chosen to make γ over-dispersed. Proper prior distributions are also imposed on $(\boldsymbol{\beta}, \rho)$:

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0) \quad \text{and} \quad \rho \sim \mathrm{Beta}(a_1, a_2).$$

Here, the standard Gibbs sampler cannot be implemented because the dimensionality of the parameter vector $\boldsymbol{\lambda}_{S(\mathbf{Z})}$ is subject to another model component $\boldsymbol{Z}$ [12]. Specifically, $(S(\boldsymbol{Z}), \boldsymbol{\lambda}_{S(\mathbf{Z})})$ belongs to a product space $\cup_{K=1}^{T-1}(\{K\} \times \Re_+^K)$, which requires the Markov chain to jump between parameter spaces of different dimensionality. The complication with the parameter vector of a varying dimension can be, however, avoided by using what we call the PCG sampler [8,20]. After the block change probability ρ is completely collapsed out of the standard MCMC sampler, the parameter vector of the varying dimension $\boldsymbol{\lambda}_{S(\mathbf{Z})}$ can be partially collapsed out without complicating the updating of the other model components. To simulate $p(\boldsymbol{Z}, \boldsymbol{\beta}, \boldsymbol{\lambda}_{S(\mathbf{Z})}, \gamma|\boldsymbol{Y})$ with ρ collapsed, the resulting PCG sampler iteratively draws values from the following conditional distributions:

$$\begin{aligned}
&p(Z_t|\boldsymbol{Z}_{-t}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y}) \quad \text{for } t = 2, \ldots, T-1, \\
&p(\boldsymbol{\beta}|\boldsymbol{Z}, \gamma, \boldsymbol{Y}), \\
&p(\boldsymbol{\lambda}_{S(\mathbf{Z})}|\boldsymbol{Z}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y}), \\
&p(\gamma|\boldsymbol{Z}, \boldsymbol{\beta}, \boldsymbol{\lambda}_{S(\mathbf{Z})}, \boldsymbol{Y}),
\end{aligned} \tag{2}$$

where $\boldsymbol{Z}_{-t} = (Z_1, \ldots, Z_{t-1}, Z_{t+1}, \ldots, Z_T)$ and $\boldsymbol{Y} = \{y_t\}_{t=1}^T$. The details of the PCG sampler in Equation (2) are given in Appendix 1.

4. Simulation study

We conducted a simulation study to check the self-consistency of the proposed Poisson change-point regression model, and herein, we illustrate its flexibility to model an unknown baseline rate. Our test data set was generated with $T = 1000$ time bins and $B = 3$ time blocks, where two instantaneous changes occur at $t = 400$ and 700. Based on $\gamma = 0.1$, we simulated three block-specific baseline Poisson intensities, $\boldsymbol{\lambda}_3 = (\lambda_1, \lambda_2, \lambda_3) = (9.19, 5.13, 7.44)$. The covariate X_t is an indicator for 100 time bins between $t = 501$ and 600, and the corresponding coefficient was set to $\beta = 0.2$. For simplicity, the population sizes $\{P_t\}_{t=1}^{1000}$ were set to be 1 over the entire time period. Based on these true values, we generated a test data set under the model in Equation (1), that is,

$$y_t \overset{\text{ind}}{\sim} \mathrm{Poisson}\{\lambda(t) \exp(0.2 \times X_t)\}, \tag{3}$$

where $\lambda(t)$ equals λ_1 if $t \in [1, 400]$, λ_2 if $t \in [401, 700]$, and λ_3 if $t \in [701, 1000]$. Given the test data set, our goal was to estimate the coefficient β by correctly specifying an unknown baseline rate.

The test data set was fit with the PCG sampler devised in Section 3. To ensure the convergence of the PCG sampler, we ran two chains of 10,000 iterations with different sets of starting values. Our inference was based on the second halves of the two chains after assessing convergence by computing the $\hat{R}^{1/2}$ statistic [11] and confirming $\hat{R}^{1/2} < 1.1$ for all model parameters. Figure 2 shows the convergence characteristics and marginal posterior distributions of the selected model

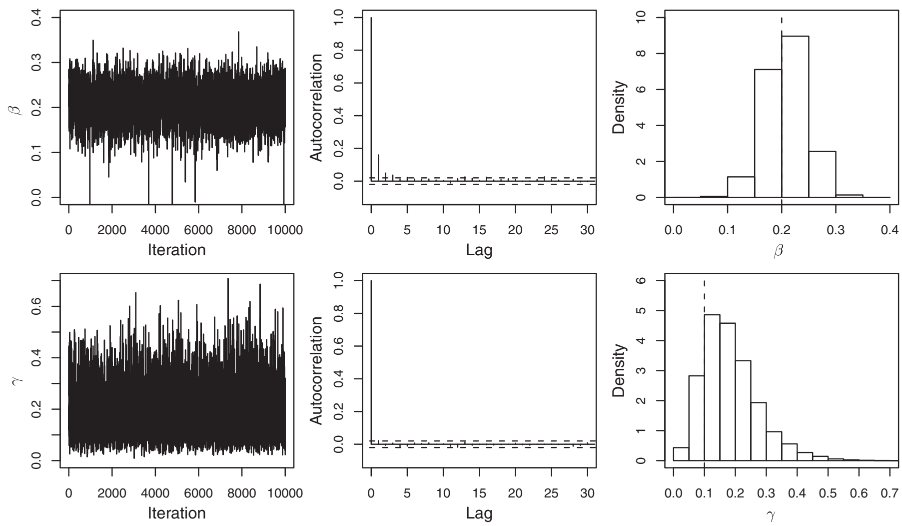

Figure 2. The mixing plots (left), autocorrelation plots (middle), and marginal posterior distributions (right) of model parameters, β and γ, simulated using the PCG sampler. The dashed vertical lines in the marginal posterior distributions represent the true values of the model parameters used to simulate a test data set, that is, $\beta = 0.2$ and $\gamma = 0.1$.

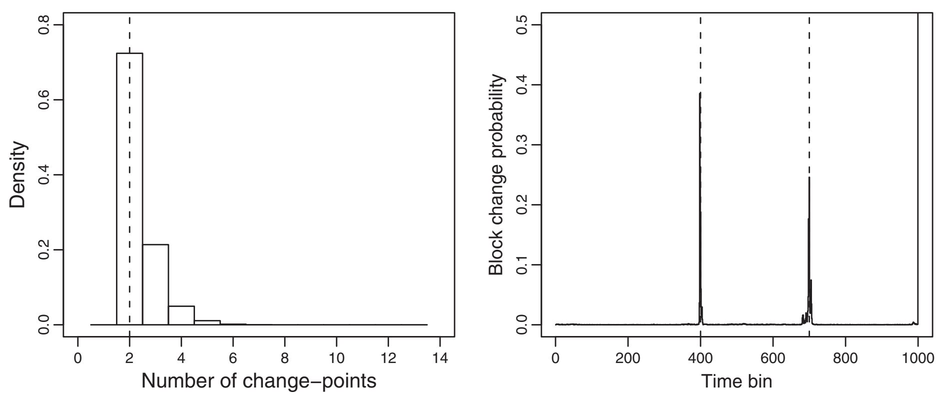

Figure 3. Posterior inference about the number and location of change points. The left panel shows the posterior distribution of the number of change points. The vertical dashed lines represent the true number of change points, that is, two. The pointwise posterior probability of a block change is shown in the right panel, conditional on the two change points.

parameters, β and γ. The left and middle columns of Figure 2 present the trace and autocorrelation plots of each parameter using a single chain of 10,000 iterations, respectively. The fast mixing and low autocorrelations imply the quick convergence of the PCG sampler. After collecting the posterior samples from the second halves of the two separate chains, the marginal posterior distribution of each parameter is shown in the right column of Figure 2. The dashed vertical lines in the panels represent the true values of the selected model parameters, which are well covered by the corresponding marginal posterior distributions.

Using the 10,000 posterior samples collected from the second halves of the two separate chains, various posterior inferences can be made about the number and location of change points. A block change-point indicator vector $\boldsymbol{Z}$ contains information about both the number and the location of

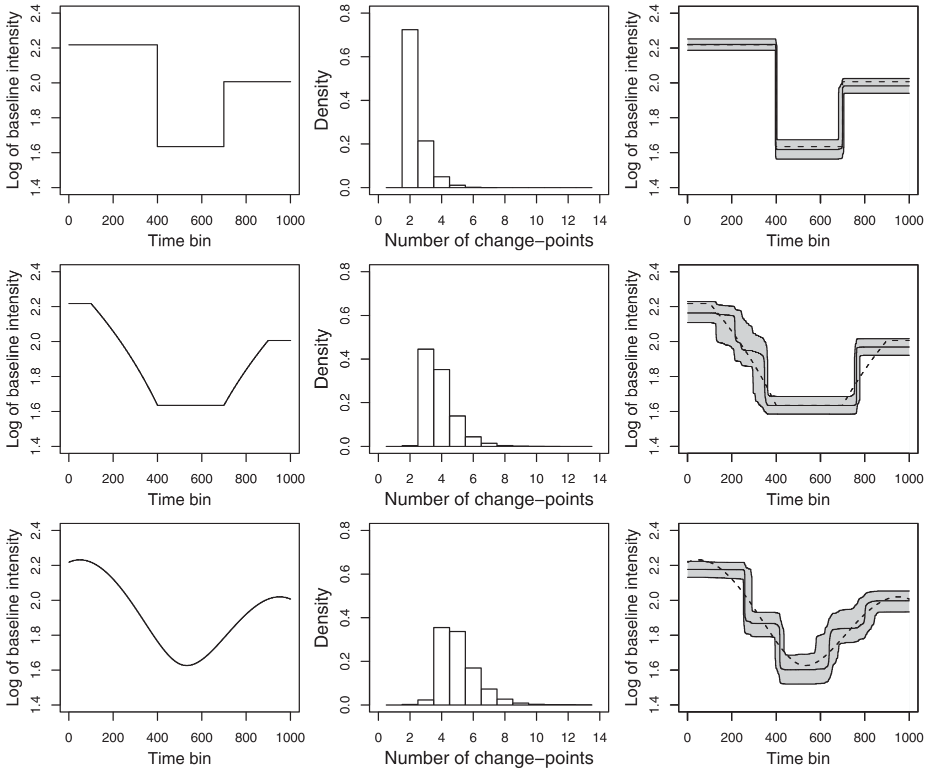

Figure 4. The pointwise posterior means (solid lines) and 95% posterior intervals (dotted lines) of the log baseline rate and its true value (dashed lines) when the number of change points equals two.

change points in the baseline rate. In particular, the sum of the block change-point indicator variables minus 1, that is, $S(\boldsymbol{Z}) - 1$, corresponds to the number of change points. The left panel of Figure 3 shows that the posterior distribution of $S(\boldsymbol{Z}) - 1$ correctly estimates the true number of change points (dashed lines) in the test data set, which is given as two. The corresponding location of the two change points can be estimated by computing the pointwise posterior average of a block change-point indicator variable given two change points, that is, three time blocks, denoted by $\mathrm{E}(Z_t|S(\boldsymbol{Z}) = 3, \boldsymbol{Y}) = p(Z_t = 1|S(\boldsymbol{Z}) = 3, \boldsymbol{Y})$. The right panel of Figure 3 presents the block change probability for each time point, conditional on the mostly likely number of change points, that is, two. The true location of each of the two change points is correctly specified by the high posterior probabilities of potential locations of a block change point.

To illustrate the robustness of the proposed semi-parametric Bayesian approach to modeling an unknown baseline rate, we consider three types of baseline rates with (1) instantaneous changes, (2) gradual changes, and (3) smoothly varying changes. The left panels of Figure 4 show the log of the true baseline rates of these kinds of changes, respectively. The middle panels of Figure 4 present the posterior distribution of the number of change points. As a true baseline rate becomes smoother, we naturally expect more change points in a fitted piecewise-constant baseline rate, which is confirmed in the posterior distribution of the number of change points. Given the most likely number of change points in each case, we can compute the pointwise posterior medians and 95% posterior intervals of a log baseline rate, as shown in the right panels of Figure 4; the dashed lines indicate the log of the true baseline rate. The pointwise posterior medians of a log baseline rate correctly specify the true baseline rate with instantaneous changes. In addition, the smoother true baseline rates are reasonably approximated by the fitted step function, thereby illustrating the flexibility of the change-point model to fit a general shape for a baseline rate.

5. Poisson change-point regression analysis of Cali homicide rates

5.1 *Model specification*

During the 3531 days from January 1999 until August 2008, five distinct policies that restrict the sale of alcoholic beverages after given hours were enforced in Cali. First, the sale of alcohol was allowed from the hours of 10 a.m. to 4 a.m. the following day. Such a most lenient policy was enforced during La Feria de Cali, the city's premier cultural event that is held annually during the end of December and beginning of January and includes bullfights, horse parades, and multiple Salsa concerts. In total, there were 237 non-consecutive days between January 1999 and August 2008 when the most lenient policy was enforced. Second, the sale of alcohol was allowed from the hours of 10 a.m. to 3 a.m. the following day for 1296 non-consecutive days in this period. Third, the sale of alcohol was allowed from the hours of 10 a.m. to 2 a.m. the following day for 1817 non-consecutive days in this period. Fourth, the sale of alcohol was allowed from the hours of 10 a.m. to some time before 2 a.m. (e.g. 12 a.m. or 1 a.m.) the following day for 159 non-consecutive days in this period. Last, the city's administration banned all alcohol sales on a number of occasions in response to extenuating events that escalated tension in the city to dangerously high levels, such as highly contested elections or labor strikes. In total, there were 22 non-consecutive days during the period of observation when all sales of alcoholic beverages were banned.

There are several covariates which are believed to be associated with homicide rates in Cali and should be adjusted for when assessing the relationship between the restricted hours of alcohol sale and homicide rates. The number of homicides was elevated on the weekends, that is, the average daily number of homicides on Fridays, Saturdays, and Sundays was 6.10 when compared with 4.28 for the remaining 4 days of the week. There was also an increase in the number of homicides on Christmas, New Year's Eve, and New Year's Day with an average of 13.64 homicides per day when compared with 5.00 on all the other days. In addition, there were 27 homicides on 19 January 2002. This day was the deadliest day in our observation period, and on no day other than Christmas, New Year's Eve, or New Year's Day were more than 19 people killed. The high number of homicides on 19 January 2002 is due in part to a clash between the Colombian Army and the militant guerrilla group Revolutionary Armed Forces of Colombia, which resulted in 16 deaths [17]. Despite their significant contribution to the national homicide rate, guerrilla activities are rarely experienced within the city limits of Cali [10], so we subsequently treated 19 January 2002 as an outlier and removed it from our analysis.

Let W_t be the indicator variable which is equal to 1 if day t is either a Friday, Saturday, or Sunday and zero otherwise, H_t be the indicator variable which has value 1 if day t is 25 December, 31 December, or 1 January and zero otherwise, and F_t be the indicator variable which has value 1 during La Feria de Cali when alcohol sales are banned after 4 a.m. the following day and zero otherwise. In our Poisson change-point regression, we set the reference level for the hours of restricted sale of alcohol to be after 2 a.m. the following day and define the indicator variables B_{1t}, B_{2t}, and B_{3t} to be 1 if alcohol sales on the morning of day t are banned after 3 a.m. the following day, if sales are banned some time before 2 a.m. the following day, and if the sale of alcohol is completely banned, respectively, and zero otherwise. If ξ_t is the expected homicide rate on day t, then we fit the Poisson change-point regression model described in Section 2 with the link function

$$\begin{aligned}\log(\xi_t) = \log\lambda(t) + \beta_1 W_t + \beta_2 H_t + \beta_3 F_t + \beta_4 B_{1t} + \beta_5 B_{2t} + \beta_6 B_{3t} \\ + \beta_7 F_t W_t + \beta_8 B_{1t} W_t + \beta_9 B_{2t} W_t + \beta_{10} B_{3t} W_t,\end{aligned}$$

which can be rewritten as

$$\log(\xi_t) = \begin{cases} \log \lambda(t) + \beta_2 H_t + \beta_3 F_t + \beta_4 B_{1t} + \beta_5 B_{2t} + \beta_6 B_{3t} & \text{during weekdays,} \\ \log \lambda(t) + \beta_1 + \beta_2 H_t + (\beta_3 + \beta_7) F_t + (\beta_4 + \beta_8) B_{1t} & \\ \quad + (\beta_5 + \beta_9) B_{2t} + (\beta_6 + \beta_{10}) B_{3t} & \text{during weekends.} \end{cases}$$

The proposed model was fit to the homicide rates in Cali by running the PCG sampler described in Section 3 and Appendix 1. The computer code with the C interface to R was used to run the PCG sampler.

5.2 *Results*

Figure 5 shows various posterior inferences from the Poisson change-point regression analysis of Cali homicide data. The top left panel of Figure 5 displays the posterior distribution of the number of change points, which has a mode at four. The plausible location of these four change points is illustrated in the top right panel of Figure 5 that shows the posterior probability of a block change conditional on four change points. Based on the block change probability, we estimated that the four change points occurred in November 2003, March 2004, December 2004, and July 2007. The bottom panel of Figure 5 shows the pointwise posterior means and 95% posterior intervals of a baseline rate conditional on four change points and suggests a non-constant baseline homicide rate across 1999–2008. To illustrate the presence of a non-constant baseline, we fit two negative binomial log-linear regression models. Both models include the coefficients for weekend,

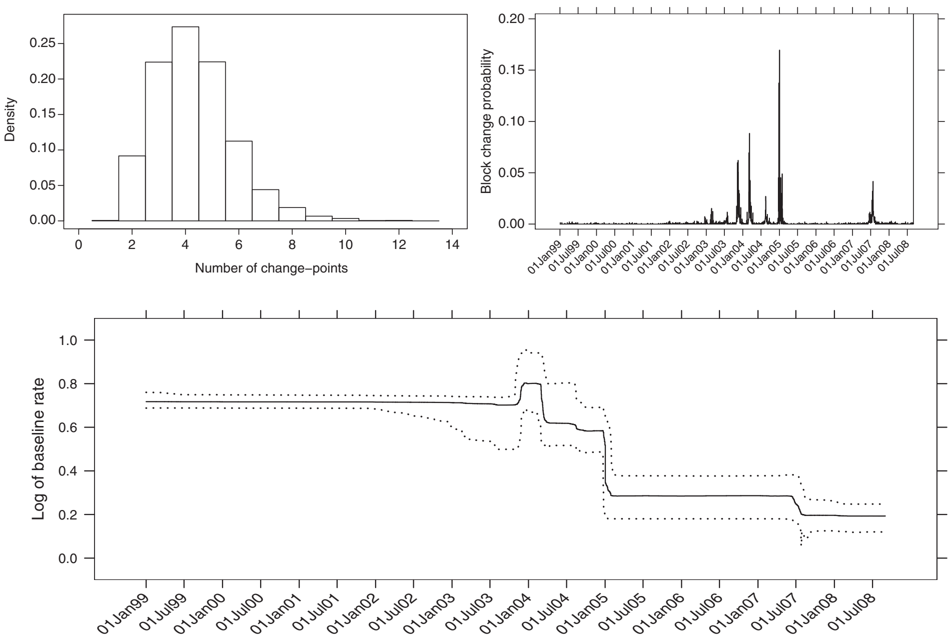

Figure 5. Various posterior inferences from the change-point analysis of Cali homicide data. The top left panel shows the posterior distribution of the number of change points, which suggests four change points during the time period of observation. The top right panel shows the posterior probability of a block change, conditional on the four change points. The bottom panels present the pointwise posterior means (solid lines) and 95% posterior intervals (dotted lines) of the log baseline homicide rate per million people when the number of change points is four.

holiday, alcohol policies, and interactions. The first model assumes a constant baseline, while the second uses a piecewise-constant baseline with change points on 15 November 2003, 6 March 2004, 29 December 2004, and 22 July 2007. Using the parameterization for the negative binomial distribution with the expected value μ and variance $\mu + \mu^2/\nu$, the maximum likelihood estimate for ν is 8.89 in the model with a constant baseline and it is 11.35 for the model with a piecewise-constant baseline. The likelihood ratio test comparing these two models has a p-value less than 0.001, indicating that a significant proportion of the over-dispersion is accounted for by using a piecewise-constant baseline. Failing to account for the non-constant baseline rate can potentially bias the estimates of the effects of policies restricting the sale of alcohol on homicide rates.

The summary statistics for the coefficients are reported in Table 1. As anticipated, our analysis found that the homicide rate increased on the weekends and on 25 December, 31 December, and 1 January. The estimated risk ratio of homicide when the sale of alcohol was banned after 2 a.m. the following day on the weekends when compared with that on the weekdays was $\exp(\hat{\beta}_1) = 3.08$ with a 95% credible interval of (2.79, 3.38). After adjusting for other covariate effects, the estimated risk ratio of homicide for 25 December, 31 December, and 1 January when compared with that for other days in La Feria de Cali was $\exp(\hat{\beta}_2) = 3.95$ with a 95% credible interval of (2.99, 5.19). The discrepancy in the homicide rate during the days in La Feria de Cali, excluding 25 December, 31 December, and 1 January, when alcohol sales were banned after 4 a.m. the following day relative to the homicide rate when alcohol sales were banned after 2 a.m. the following day differed between the weekends and weekdays. We found an increase in the expected homicide rate during the weekdays with a risk ratio of $\exp(\hat{\beta}_3) = 1.10$ and 95% credible interval of (1.01, 1.18), while we found no significant difference in the homicide rate on the weekends. During the weekdays, the risk ratio of homicide when the sale of alcohol was banned after 3 a.m. the following day versus when it was banned after 2 a.m. the following day was estimated as $\exp(\hat{\beta}_4) = 1.14$ with a 95% credible interval of (1.04, 1.25), whereas there was no significant difference on the weekends. At the 0.05 level, we found no significant difference, either on the weekends or during the weekdays, in the expected homicide rate when legal alcohol sales ended some time before 2 a.m. the following day versus when legal alcohol sales ended at 2 a.m. the following day. When the sale of alcohol was completely banned, we found no significant difference in the homicide rate compared with when the sale of alcohol was banned after 2 a.m. the following day during the weekdays, but a decrease in the expected homicide rate on the weekends with a risk ratio of $\exp(\hat{\beta}_6 + \hat{\beta}_{10}) = 0.73$ and a 95% credible interval of (0.57, 0.96).

Table 1. Posterior summary statistics for coefficients.

Parameter	Mean	Std. dev.	Median	2.5%	97.5%
β_1	1.1237	0.0497	1.1230	1.0251	1.2178
β_2	1.3743	0.1411	1.3781	1.0954	1.6467
β_3	0.0919	0.0391	0.0918	0.0132	0.1689
β_4	0.1286	0.0487	0.1233	0.0433	0.2206
β_5	0.0070	0.0502	0.0025	−0.0797	0.1144
β_6	−0.1216	0.0943	−0.1197	−0.3085	0.0648
β_7	−0.1073	0.0471	−0.1061	−0.1961	−0.0089
β_8	−0.0918	0.0248	−0.0914	−0.1394	−0.0429
β_9	−0.0057	0.0522	−0.0035	−0.1108	0.0989
β_{10}	−0.1913	0.1640	−0.1876	−0.5184	0.1375
$\beta_3 + \beta_7$	−0.0154	0.0371	−0.0149	−0.0869	0.0585
$\beta_4 + \beta_8$	0.0367	0.0486	0.0293	−0.0498	0.1268
$\beta_5 + \beta_9$	0.0014	0.0501	−0.0027	−0.0808	0.1195
$\beta_6 + \beta_{10}$	−0.3129	0.1325	−0.3057	−0.5692	−0.0397

5.3 *Implications*

From a policy perspective, the most important findings of this analysis are the differences in the estimated homicide rates when the sale of alcoholic beverages was banned after 3 a.m. the following day versus when it was banned after 2 a.m. the following day. A majority of the debate about policies dictating the sale of alcohol in Cali has centered on this additional hour. The estimated risk ratio of homicide when the sale of alcohol was banned after 3 a.m. the following day versus when it was banned after 2 a.m. the following day was $\exp(\hat{\beta}_4) = 1.14$ during the weekdays, but the risk ratio was not significant on the weekends. With a population of approximately 2.5 million, allowing the sale of alcohol during the weekdays until 3 a.m. the following day rather than until 2 a.m. the following day would result in estimated 88 additional deaths in 2009.

The Poisson change-point regression model not only avoids potential biases in the inference of the coefficients due to model misspecification, but also illustrates the changes in the homicide rate over time while adjusting for measured covariates. The estimated change points temporally coincide with sociopolitical events which can be examined to illuminate factors which have strong relationships with homicide rates. The estimated posterior mean of $\log \lambda(t)$ remained constant from January 1999 until November 2003 where an increase that persisted for approximately 6 months occurred. An analysis conducted by the Colombian National Police indicates that the elevated number of homicides during this period of time could have been influenced by a struggle between two new drug-trafficking organizations to gain control of the locations in the city which are strategically important to the illegal drug and firearms trade [27]. The decrease in the baseline homicide rate in March 2004 coincides with a 20,000 million Colombian pesos increase in the budget of the Cali security institutions issued from the national government to provide new equipment and increase law enforcement personnel. Colombian national news organizations have reported on an increase in the number of captures and prosecutions of violent criminals between the end of 2004 and beginning of 2005, which increased the public's perception of its security [9]. These events are consistent with the estimated change point in the late December of 2004, which marks a large decrease in the baseline homicide rate. We observed an additional decrease in the baseline homicide rate in July of 2007. In May of 2007, the mayor who was elected to serve from 2004 to 2008 was deposed amid a corruption scandal and the new administration applied a series of reforms over the ensuing months [21]. It could be highly beneficial to explore the policies implemented by the new administration and conduct a study to determine the association between these reforms and the decrease in homicide rates in July of 2007.

6. Discussion

In this article, we have proposed a Poisson change-point regression model in the presence of covariates and an unknown number of change points with unknown locations and fitted this model to the daily homicide rates in Cali, Colombia. Our proposed model allows for a non-constant baseline rate while accounting for covariates. The standard MCMC sampler constructed for fitting the model is, however, infeasible to implement because the dimensionality of a certain parameter vector is not fixed and depends on other model components. To circumvent the difficulty, we considered partially collapsing a parameter vector of varying dimensions in the MCMC sampler. This partial collapse results in the so-called PCG sampler that makes it feasible to fit the proposed model and exhibits better convergence characteristics.

When fit to the Cali homicide data collected by the CISALVA Institute, the proposed model provides powerful insight into the pattern of homicide rates in Cali, Colombia. The change-point analysis indicates points in time when, conditional on measured covariates, the baseline homicide rate changed. These changes correspond to sociopolitical events and can provide insight for the design of future studies. The inference on the relative effectiveness of policies controlling the hours when the sale of alcoholic beverages is restricted accounts for these structural changes and subsequently avoids many potential biases. The analysis points toward a realistic policy that is less restrictive on the weekends than during the weekdays, which balances public health with the economic and cultural dependencies on alcohol sales in Cali, Colombia.

Acknowledgements

The authors thank Maria Isabel Gutierrez Martinez, director of the CISALVA Institute at the Universidad del Valle, for her support. This work was supported by the Korea Science and Engineering Foundation grant (KOSEF-2011-8-0882), the National Science Foundation grant (DMS-08-05050), and the National Institute of Health grant (D43-TW007560-01).

References

[1] D. Barry and J.A. Hartigan, *Product partition models for change point problems*, Ann. Statist. 20 (1992), pp. 260–279.
[2] D. Barry and J.A. Hartigan, *A Bayesian analysis for change point problems*, J. Amer. Statist. Assoc. 88 (1993), pp. 309–319.
[3] C.C. Branas, M.L. Nance, M.R. Elliott, T.S. Richmond, and W. Schwab, *Urban-rural shifts in intentional firearm death: Different causes, same results*, Am. J. Public Health 94 (2004), pp. 1750–1755.
[4] B.P. Carlin, A.E. Gelfand, and A.F.M. Smith, *Hierarchical Bayesian analysis of changepoint problems*, Appl. Stat. 41 (1992), pp. 389–405.
[5] S. Chib, *Estimation and comparison of multiple change-point models*, J. Econometrics 86 (1998), pp. 221–241.
[6] A. Concha-Eastman, V. Espitia, R. Espinosa, and R. Guerrero, *La epidemiologia de los homicidios en Cali, 1993–1998: seis aios de un modelo poblacional*, Pan Am. J. Public Health 12 (2002), pp. 230–239.
[7] G. De la hoz-Bohórquez, *Homicidio. Colombia 2007*, in *Forensis 2007 Datos para la Vida*, Instituto de Medicina Legal y Ciencias Forenses, eds. Bogota, D.C., 2008, pp. 19–56.
[8] D. van Dyk and T. Park, *Partially collapsed Gibbs samplers: Theory and methods*, J. Amer. Statist. Assoc. 103 (2008), pp. 790–796.
[9] N. El Tiempo, *Cali y Medellín tuvieron 1036 muertes menos en el (2005)*, ELTIEMPO.COM. Available at http://www.eltiempo.com/archivo/documento/MAM–1875701.
[10] S. Franco, *A social-medical approach to violence in Colombia*, Am. J. Public Health 93 (2003), pp. 2032–2036.
[11] A. Gelman and D.B. Rubin, *Inference from iterative simulations using multiple sequences (with discussion)*, Statist. Sci. 7 (1992), pp. 457–472.
[12] P.J. Green, *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika 82 (1995), pp. 711–732.
[13] P.J. Green, *Transdimensional Markov chain Monte Carlo*, in *Highly Structured Stochastic Systems*, P.J. Green, N.L. Hjort and S. Richardson, eds., Oxford University Press, Oxford, 2003, pp. 179–198.
[14] R. Guerrero and A. Concha-Eastman, *An epidemiological approach for the prevention of urban violence: The case of Cali, Colombia*, J. Health and Population in Developing Countries 4 (2003), pp. 1–11.
[15] M. Gutierrez-Martinez, A. Sanchez, E. Muñoz, A. Fandiño, and E. Muñoz, *Sistema de vigilancia de muertes por lesiones de causa externa*, Homicidios ocurridos en el año 2004 (Online Article), 2005. Available at http://vigilesiones.univalle.edu.co/informes/anual/homicidios/homic_2004.pdf.
[16] M. Gutierrez-Martinez, R. Villin, A. Fandiño, and R. Oliver, *The evaluation of a surveillance system for violent and non-intentional injury mortality in Colombian cities*, Int. J. Inj. Control Saf. Promot. 14 (2007), pp. 77–84.
[17] A. Isacson, *Was failure avoidable? Learning from Colombia's 1998–2002 peace process*, Dante B. Fascell North-South Center Working Paper Series, Paper No. 14, University of Miami North South Center, Miami, 2003.
[18] G. Koop and S. Potter, *Estimation and forecasting in models with multiple breaks*, Rev. Econom. Stud. 74 (2007), pp. 763–789.
[19] J. Pan and J. Chen, *Application of modified information criterion to multiple change point problems*, J. Multivariate Anal. 97 (2006), pp. 2221–2241.
[20] T. Park and D. van Dyk, *Partially collapsed Gibbs samplers: Illustrations and applications*, J. Comput. Graph. Statist. 18 (2009), pp. 283–305.
[21] R.C. Procuraduría General de la Nación, *Confirmada destitución de Alcalde de Cali*, República de Colombia, 2007. Available at http://www.procuraduria.gov.co/html/noticias_2007/noticias_156.html.
[22] A.E. Raftery and V.E. Akman, *Bayesian analysis of a Poisson process with a change-point*, Biometrika 73 (1986), pp. 85–89.

[23] F.P. Rivara, B.A. Mueller, G. Somes, C.T. Mendoza, N.B. Rushforth, and L.A. Kellermann, *Alcohol and illicit drug abuse and the risk of violent death in the home*, JAMA 278 (1997), pp. 569–575.
[24] A.F.M. Smith, *A Bayesian approach to inference about a change-point in a sequence of random variables*, Biometrika 62 (1975), pp. 407–416.
[25] M.A. Tanner and W.H. Wong, *The calculation of posterior distributions by data augmentation (with discussion)*, J. Amer. Statist. Assoc. 82 (1987), pp. 528–550.
[26] UNODC, *The 2011 Global Study on Homicide*, United Nations Office of Drugs and Crime (Online Article), 2011. Available at http://www.unodc.org/documents/data-and-analysis/statistics/Homicide/Globa_study_on_homicide_2011_web.pdf.
[27] Programa Presidencial de Derechos Humanos y DIH, Vicepresidencia de la Republica de Colombia, Dinámica reciente de la violencia en Cali (Online article), Bogotá, D.C., 2006. Available at http://www.derechoshumanos.gov.co/Observatorio/Publicaciones/documents/2010/Estu_Regionales/cali.pdf.
[28] A. Villaveces, P. Cummings, V.E. Espitia, T.D. Koepsell, B. McKnight, and A.L. Kellermann, *Effect of a ban on carrying firearms on homicide rates in 2 Colombian cities*, JAMA 283 (2000), pp. 1205–1209.
[29] T.Y. Yang and L. Kuo, *Bayesian binary segmentation procedure for a Poisson process with multiple changepoints*, J. Comput. Graph. Statist. 10 (2001), pp. 772–785.

Appendix 1. Implementing the PCG sampler

In this appendix, we describe the details of the PCG sampler used to fit the Poisson change-point regression model. After completely collapsing ρ out of the joint posterior distribution, the PCG sampler simulates $p(\boldsymbol{Z}, \boldsymbol{\beta}, \boldsymbol{\lambda}_{S(\boldsymbol{Z})}, \gamma | \boldsymbol{Y})$ by iteratively drawing values in the following order:

(1) Given $(\boldsymbol{Z}_{-t}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y})$, we draw Z_t from a Bernoulli distribution:

$$Z_t | (\boldsymbol{Z}_{-t}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y}) \sim \text{Bernoulli}\left(\frac{f(Z_t = 1, \boldsymbol{Z}_{-t} | \boldsymbol{\beta}, \gamma, \boldsymbol{Y})}{f(Z_t = 1, \boldsymbol{Z}_{-t} | \boldsymbol{\beta}, \gamma, \boldsymbol{Y}) + f(Z_t = 0, \boldsymbol{Z}_{-t} | \boldsymbol{\beta}, \gamma, \boldsymbol{Y})}\right),$$

where

$$f(\boldsymbol{Z} | \boldsymbol{\beta}, \gamma, \boldsymbol{Y}) = \left[\prod_{b=1}^{S(\boldsymbol{Z})} \frac{\Gamma(\sum_{t \in \mathcal{B}_b} y_t + 1)}{\{\sum_{t \in \mathcal{B}_b} P_t \exp(\boldsymbol{X}_t^{\mathrm{T}} \boldsymbol{\beta}) + \gamma\}^{\sum_{t \in \mathcal{B}_b} y_t + 1}}\right] \Gamma(S(\boldsymbol{Z}) + a_1) \Gamma(T - S(\boldsymbol{Z}) + a_2) \gamma^{S(\boldsymbol{Z})},$$

for $t = 2, \ldots, T-1$.

(2) Given $(\boldsymbol{Z}, \gamma, \boldsymbol{Y})$, we draw $\boldsymbol{\beta}$ from the conditional distribution:

$$p(\boldsymbol{\beta} | \boldsymbol{Z}, \gamma, \boldsymbol{Y}) \propto \left[\prod_{b=1}^{S(\boldsymbol{Z})} \frac{\Gamma(\sum_{t \in \mathcal{B}_b} y_t + 1)}{\{\sum_{t \in \mathcal{B}_b} P_t \exp(\boldsymbol{X}_t^{\mathrm{T}} \boldsymbol{\beta}) + \gamma\}^{\sum_{t \in \mathcal{B}_b} y_t + 1}}\right] \times \exp\left\{\sum_{t=1}^{T} y_t \boldsymbol{X}_t^{\mathrm{T}} \boldsymbol{\beta} - \frac{1}{2} (\boldsymbol{\beta} - \boldsymbol{\beta}_0)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}_0)\right\},$$

which is non-standard but log-concave. Thus, we simulate $\boldsymbol{\beta}$ using an efficient Metropolis algorithm with a multivariate normal distribution with matching mode and curvature at the mode as a proposal distribution.

(3) Given $(\boldsymbol{Z}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y})$, $\boldsymbol{\lambda}_{S(\boldsymbol{Z})}$ is sampled from independent gamma distributions, where

$$\lambda_b | (\boldsymbol{Z}, \boldsymbol{\beta}, \gamma, \boldsymbol{Y}) \overset{\text{ind}}{\sim} \text{Gamma}\left(\sum_{t \in \mathcal{B}_b} y_t + 1, \sum_{t \in \mathcal{B}_b} P_t \exp(\boldsymbol{X}_t^{\mathrm{T}} \boldsymbol{\beta}) + \gamma\right),$$

for $b = 1, \ldots, S(\boldsymbol{Z})$.

(4) Given $(\boldsymbol{Z}, \boldsymbol{\lambda}_{S(\boldsymbol{Z})}, \boldsymbol{\beta}, \boldsymbol{Y})$, γ is sampled from a gamma distribution,

$$\gamma | (\boldsymbol{Z}, \boldsymbol{\lambda}_{S(\boldsymbol{Z})}, \boldsymbol{\beta}, \boldsymbol{Y}) \sim \text{Gamma}\left(S(\boldsymbol{Z}) + g_1, \sum_{b=1}^{S(\boldsymbol{Z})} \lambda_b + g_2\right).$$