EERI

Economics and Econometrics Research Institute

# A Bayesian Reversible Jump Piecewise Hazard approach for modelling rate changes in mass shootings

Andrew G. Chapple

EERI Research Paper Series No 24/2016

ISSN: 2031-4892

EERI
**Economics and Econometrics Research Institute**
Avenue de Beaulieu
1160 Brussels
Belgium

Tel: +322 298 8491
Fax: +322 298 8490
www.eeri.eu

# A Bayesian Reversible Jump Piecewise Hazard approach for modeling rate changes in mass shootings

Andrew G. Chapple
Department of Statistics, Rice University, Houston, TX 77005, U.S.A.
agc6@rice.edu

**Abstract**

Time to event data for econometric tragedies, like mass shootings, have largely been ignored from a changepoint analysis standpoint. We outline a techniqnique for modeling economic changepoint problems using a piecewise constant hazard model to explain different economic phonomenon. Specifically, we investigate the rates of mass shootings in the United States since August 20th 1982 as a case study to examine changes in rates of these terrible events in an attempt to connect changes to the shooter's covariates or policy and societal changes.

## Keywords

Time-to-event Data, Bayesian Analyses, Piecewise Exponential, Reversible Jump, Mass Shooting

# A Bayesian Reversible Jump Piecewise Hazard approach for modeling rate changes in mass shootings

**Abstract**

Time to event data for econometric tragedies, like mass shootings, have largely been ignored from a changepoint analysis standpoint. We outline a techniqnique for modeling economic changepoint problems using a piecewise constant hazard model to explain different economic phonomenon. Specifically, we investigate the rates of mass shootings in the United States since August 20th 1982 as a case study to examine changes in rates of these terrible events in an attempt to connect changes to the shooter's covariates or policy and societal changes.

*Keywords:* Time-to-event Data, Survival Data, Mass Shooting

## 1. Introduction

In many economic social issues people wonder whether the incidences have increased or decreased over time. One particular focus is tragic events such as terrorist attacks or mass shootings. One of the issues with terrorist attacks is that they are not independent events due to the recent rise of ISIS. Lone wolves are more likely to commit these crimes since there exists a terrorist network which could encourage attacks and communicate other successful attacks. While this could also be argued for mass shootings, these are less dependent on each event as none of them have any general connection such as a network to incite them. One could argue there is a cult of personality and that seeing other attacks makes twisted individuals more brazen to commit mass attrocities against their fellow man, but this is a stronger assumption than the assumption of independence. Due to this, we consider the mass shootings to be statistically independent. Likewise, this methodology and interpretations could be applied to many different time to event data where independence is a reasonable assumption. We use a piecewise constant exponential hazard modeling scheme with reversible jump capabilities to model these events.

This allows us to draw inference on three quantities: the number of changepoints where these adverse events increase or decrease in quantity, the location of these changepoints in time, and the relative rates of these catastrophies within each of these intervals. We use the approach of Green [1] to these aims by placing a prior distribution on the number of changepoints and the hazard heights and location of the changepoints, allowing for a dimension varying MCMC. We adopt an Independent

Correllated Autoregresssive structure to model the spatial dependencies of adjacent interval hazards [2]. The coding for this approach is already available on CRAN in the BayesPiecewiseICAR package [3].

## 2. Methods

We present a Hierarchical Bayesian Analyses consisting of reversible jump Markov Chain Monte Carlo Methods, to adequately characterize the changing rates in the time to event data. This method assumes a piecewise constant hazard in each disjoint time interval with an Intrinsic Correllated Autoregression structure to model the dependencies in the hazard heights was used and described extensively by Lee et al but this approach also used covariates [2] [4]. While their model is computationally exceptional, this tends to estimate very small baseline hazards due to removing the covariate effects. Here we do not have data that seems to be driven by covariate differences, as the data compromises location, race of shooter, gender which is majority male and mental status which was largely unknown. None of these prove to suggest trends that are not already evident, such as most of the attacks being carried out by males. Furthermore since we are not observing every person in America to see if they committ a mass shooting (which would result in censoring), it is not feasible-nor statistically correct-to examine this type of data in a covariate based approach. Instead all we can do is model the changes over time, which we do here, regardless of the circumstances surrounding the shooting.

For the pieceiwse constant hazard approach, we assume that there exists a partition of the hazard of the form $s_0 = 0 < s_1 < ... < s_J < s_{J+1} = max(Y)$ similar to the approach of Lee et al [4]. If we have data with censoring, this is adjusted to $s_{J+1} = max(Y|I = 1)$, that is the largest observed non-censoring time. On each interval $(s_j, s_{j+1}]$ we assume that the hazard takes a piecewise constant value $exp(\lambda_j)$. Additionally, we follow the approach of Lee et al. in that we do not fix $J$ and rather allow it to vary in accordance to the data. We assume that $J \sim POI(\phi)$ and that the split points have a uniform prior, but so that we are not likely to observe no events within a given interval.

We assume a ICAR prior formulation for the heights $\lambda|\mathbf{s} \sim N_{J+1}(\mu, \sigma^2\Sigma_s)$ where $\sigma^2$ has an inverse gamma prior and $\mu$ has a flat prior. We define $\Sigma_s$ in the following manner. Denote $\Delta_j = s_j - s_{j-1}$ and let $W$ be an off-diagonal matrix with the entries $W_{j,j-1} = c_\lambda(\Delta_{j-1} + \Delta_j)/(\Delta_{j-1} + 2\Delta_j + \Delta_{j+1})$ and $W_{j,j+1} = c_\lambda(\Delta_{j+1} + \Delta_j)/(\Delta_{j-1} + 2\Delta_j + \Delta_{j+1})$ where $c_\lambda$ is a hyperparameter in the domain of $[0, 1]$ that characterizes the spatial dependence between adjacent interval heights. Then if $Q$ is a diagonal matrix with entries $2/(\Delta_{j-1} + 2\Delta_j + \Delta_{j+1})$ then we have that $\Sigma_s = (I - W)^{-1}Q$ The full list of the priors is as follows:

$$J \sim POI(\phi)$$

$$\mathbf{s}|J \sim \frac{(2J+1)! \prod_{j=1}^{J+1} (s_j - s_{j-1})}{s_{J+1}^{2J+1}}$$

$$\lambda|\mathbf{s} \sim N_{J+1}(\mu, \sigma^2 \Sigma_s)$$

$$\mu \propto 1 \qquad \sigma^{-2} \sim Gamma(a, b)$$

This formulation produces the likelihood (allowing the ability for censoring) of

$$L(\mathbf{Y}|\mathbf{s}, \lambda) = \prod_{j=1}^{J+1} \exp\left[\lambda_j d_j - \exp(\lambda_j) \sum_{m \in R_j} \Delta_{mj}\right]$$

where $d_j$ is the number of mass shootings in the interval $[s_{j-1}, s_j]$, $R_j$ is the risk set in the interval and $\Delta_{mj} = \max[\min(Y_i, s_j) - s_{j-1}, 0]$ where $Y_i$ is a mass shooting time. We allow birth and death moves similar to Lee et al., but we sample the random perturvationat each proposal rather than fixing it throughout the simulations [1] [4].

## 3. Markov Chain Monte Carlo

In our Markov Chain Monte Carlo sampling scheme, we perform seven different moves consisting of five Metropolis Hastings moves (two of which are Metropolis-Hastings-Green moves) and two Gibbs samplers. One generic iteration of the sampler proceeds as follows:

1. Sample $\mu|\lambda, \sigma^2, \mathbf{s}$ via a Gibbs step. Sample Directly from:

$$\mu|\lambda, \sigma^2, \mathbf{s} \sim N\left(\frac{\mathbf{1}^t \Sigma_s^{-1} \lambda}{\mathbf{1}^t \Sigma_s^{-1} \mathbf{1}}, \frac{\sigma^2}{\mathbf{1}^t \Sigma_s^{-1} \mathbf{1}}\right)$$

2. Sample $\sigma^{-2}|\mu, \lambda, \mathbf{s}$ via a Gibbs step. Sample Directly from:

$$\sigma^{-2}|\mu, \lambda, \mathbf{s} \sim Gamma\left(a + \frac{J+1}{2}, b + 0.5(\lambda - \mu\mathbf{1})^t \Sigma_s^{-1} (\lambda - \mu\mathbf{1})\right)$$

3. Sample $\lambda|\mathbf{s}, \mu, \sigma^2, \mathbf{Y}$ via a Metropolis-Hastings step. We sample each entry of $\lambda$ and accept or reject it by drawing $\lambda_k^* \sim U[\lambda_k - c, \lambda_k + c]$ where $c$ is the tuning parameter in our MCMC. $\lambda_k^*$ is accepted with probability $\alpha^* = min(1, \alpha)$ where

$$\alpha = \frac{L(\mathbf{Y}|\mathbf{s}, \lambda^*) N_{J+1}(\lambda^*|\mu\mathbf{1}, \sigma^2 \Sigma_s)}{L(\mathbf{Y}|\mathbf{s}, \lambda) N_{J+1}(\lambda|\mu\mathbf{1}, \sigma^2 \Sigma_s)}$$

4. Sample the locations of $\mathbf{s}|\lambda, \mathbf{Y}, J$ via a Metropolis-Hastings move that shuffles the locations of $s_1, ..., s_J$. We sample $s_j^* \sim U[s_{j-1}, s_{j+1}]$ for $j = 1, ..., J$, which also adjusts $\lambda$ in the following manner. If $s_j^* > s_j$ then $\lambda_{j+1}^* = \lambda_{j+1}$ while

$$\lambda_j^* = \frac{(s_j^* - s_j)\lambda_{j+1} + (s_j - s_{j-1})\lambda_j}{s_j^* - s_{j-1}}.$$

Likewise if $s_j^* < s_j$, then $\lambda_j^* = \lambda_j$ and

$$\lambda_{j+1}^* = \frac{(s_{j+1} - s_j)\lambda_{j+1} + (s_j - s_j^*)\lambda_j}{s_{j+1} - s_j^*}.$$

We accept $(\mathbf{s}, \lambda)$ with probability $\alpha^* = \min(1, \alpha)$ where

$$\alpha = \frac{L(\mathbf{Y}|\mathbf{s}^*, \lambda^*)N_{J+1}(\lambda^*|\mu\mathbf{1}, \sigma^2\Sigma_s^*)(s_j^* - s_{j-1})(s_{j+1} - s_j^*)}{L(\mathbf{Y}|\mathbf{s}, \lambda)N_{J+1}(\lambda|\mu\mathbf{1}, \sigma^2\Sigma_s)(s_j - s_{j-1})(s_{j+1} - s_j)}$$

5. Sample $\mathbf{s}|\lambda, \mathbf{Y}$ via a Metropolis-Hastings-Green move by propsing adding a split point and deleting a split point.

   - **Birth Move:** Draw a random split point via a $Birth = U[0, s_{\max}]$ and set $\mathbf{s}^* = Sort(Birth, \mathbf{s})$. Changing the dimension of $\mathbf{s}$ also adjusts the entries of $\lambda$ in the following manner. Draw $U \sim U[0, 1]$ and assume that $Birth \in (s_{j-1}, s_j]$. Then we define the multiplicative perturbation as $\frac{\exp(\lambda_{j+1}^*)}{\exp(\lambda_j^*)} = \frac{1-U}{U}$ as in Green and Lee et al [1] [4]. Then the new heights are determined as

$$\lambda_j^* = \lambda_j - \frac{s_j - Birth}{s_j - s_{j-1}} \log\left(\frac{1-U}{U}\right)$$

$$\lambda_{j+1}^* = \lambda_j + \frac{Birth - s_{j-1}}{s_j - s_{j-1}} \log\left(\frac{1-U}{U}\right)$$

   The proposed vectors $(\mathbf{s}^*, \lambda^*)$ is accepted with probability $\alpha^* = \min(1, \alpha)$ where

$$\alpha = \frac{L(\mathbf{Y}|\mathbf{s}^*, \lambda^*)N_{J+1}(\lambda^*|\mu\mathbf{1}, \sigma^2\Sigma_s^*)Poi(J+1|\phi)(2J+3)(2J+2)(Birth - s_{j-1})(s_j - Birth)}{L(\mathbf{Y}|\mathbf{s}, \lambda)N_{J+1}(\lambda|\mu\mathbf{1}, \sigma^2\Sigma_s)Poi(J|\phi)s_{\max}^2 U(1-U)(s_j - s_{j-1})}$$

   - **Death Move:** Similar to a Birth move, a death move adjusts both $\mathbf{s}$ and $\lambda$. We propose deleting one entry $s_1, .., s_J$ equally likely. Assume we delete $s_j$ to obtain $\mathbf{s}^*$ then we delete $\lambda_{j+1}$ from $\lambda^*$ and set

$$\lambda_j^* = \frac{\lambda_j(s_j - s_{j-1}) + \lambda_{j+1}(s_{j+1} - s_j)}{s_{j+1} - s_{j-1}}$$

We draw $U \sim U[0,1]$ as the random perturbation to maintain balance between the two parameter spaces and we accept the proposed vectors $(\mathbf{s}^*, \lambda^*)$ with probability $\alpha^* = \min(1, \alpha)$ where

$$\alpha = \frac{L(\mathbf{Y}|\mathbf{s}^*, \lambda^*) N_{J+1}(\lambda^*|\mu\mathbf{1}, \sigma^2\Sigma_s^*) Poi(J-1|\phi)(s_{j+1} - s_{j-1}) s_{\max}^2 U(1-U)}{L(\mathbf{Y}|\mathbf{s}, \lambda) N_{J+1}(\lambda|\mu\mathbf{1}, \sigma^2\Sigma_s) Poi(J|\phi)(2J+1)2J(s_{j+1} - s_j)(s_j - s_{j-1})}$$

## 4. Application

The application of primary interest for this paper comes from motherjones.com consisting of the mass shooting dates and related information for each shooter since 1982 in the United States [5]. In this analysis, we do not consider any covariates related to the shooting and instead consider the date of each shooting since 1982 in an attempt to analyze the changes in mass shooting rates over time. To be clear, this data set considers a mass shooting to be a shooting that results in the death of at least four people. This eliminates cases of double murder-suicide in cases of infidelity and other crimes of passion.

The analyis was perfomed via 100,000 iterations on the package BayesPiecewiseICAR with hyperparameters $(.7, .7)$ for $\sigma^{-2}$, a prior mean of 5 different split points, a maximum allowed split point allocation of 50 and a spatial dependency of 0.5. Additionally, we used a tuning parameter of 0.25 (the default) and a starting value of 5 different split points in the hazard. The poserior results without burnin as given by the program ICARBHSampler from the package BayesPiecewiseICAR are shown in figure 1.

As we can see from the posterior distribution, there appears to be significant indication of the existence of only once split point. Economically this indicates that there is one time where the hazard of mass shootings has changed since 1982. After burning in the half of the samples, there are 47209 samples with one split point, 2771 with two split points and 21 with 3 split points. We examined the posterior location of this single split point by looking at samples with $J = 1$. The density of this posterior split point location is shown in figure 2. We see that the density is mostly concentrated slightly past the middle of 2011 indicating that the rate of mass shootings changed around August 2011. The posterior mean of this location is 2011.592, which falls in early August. Next we can examine the posterior rates on the time intervals [1982-August 2011), [August 2011-Today) to see how the rate of mass shootings changed. Figure 3 displays the posterior densities of the two log hazard rates, here a smaller loghazard indicates a decreased risk of a mass shooting occurance. We see that the posterior densities of the two log hazard rates have no overlap, indicating that the hazard of a mass shooting occuring increased significantly after August 2011. The posterior means of these two log hazards are $-9.42$ and $-6.89$, respecively, indicating that the hazard of a mass shooting increased by over 12 times after August 2011.

This allows us to conjecture about what is causing this large increase in the hazard of a mass shooting. Gun sales increased since 2008 as reflected by the number of FBI background checks, but increased drastically in 2012-2013. Since the rates of shootings changed drastically in August 2011, one might question whether this is the driving factor. Since August 2011, there have been 30 mass shootings, with all but 5 weapons being obtained legally. However, this is an increase in the proportion of shootings involving a legal weapon prior to August 2011 (77% compared to 83%). 15 of these 30 shooters had known mental illnesses while 32 of the shooters prior to August 2011 had mknown mental illnesses (60%). There is also not a discrepancy between the age of the shooters, with the mean (median) ages being 35.07 (33.5) and 34.92 (36), respectively. The data does not draw a clear line to what's causing this change.

However, we do know that the coverage of these instances has increased along with the public's ability to gather information about them through the internet, which could be a factor in the increased shootings. People may have become so desensitized to them through comprehensive coverage that these rates are increasing. Without additional data on the coverage of these events over time, we're unable to confirm this conjecture.

## 5. Conclusion

We outlined a method to assess the changes in rates for different economic events. This method allows us to model the number of changes, the location where the rate changes occur, and how the rates change probabilistically. Code for this method is provided on CRAN in the function BayesPiecewise-ICAR [3]. We applied this method to mass shooting data in attempts to connect rate changes to the shooter's covariates or policy decisions, showing that the hazard of mass shootings increased over 12 times around August 2011. However, covariate information of shooters prior to and after August 2011 showed inconclusive differences. Additionally, there were no federal changes in gun legislation around this time so it appears the change is not due to policy implications. One extension to this method include those that separate the time scale by both rate of occurance and severity, such as the number of fatalities. An additional interpretable extension could relax the piecewise constant assumption and allow for some interpretable function on each interval.

## 6. Acknowledgement

Andrew Chapple was partially supported by the National Institute of Health (Grant 5T32-CA096520-07).

## References


[1] Peter Green. *Reversible jump markov chain monte carlo computation and Bayesian model determination*. Biometrika, 82(4):711-731, 1995.

[2] Julian Besag and Charles Kooperberg. *On conditional and intrinsic autoregression*. Biometrika, 82(4):733-746, 1995.

[3] Andrew Chapple. *Package: BayesPiecewiseICAR*. CRAN. 2016.

[4] Kyu Lee, Sebastien Haneuse, Deborah Schrag and Francesca Dominici. *Bayesian semiparametric analysis of semicompeting risks data: investigating hospital readmission after a pancreatic cancer diagnosis*. Journal of the Royal Statistical Society Series C Applied Statistics, 64(2):253-273, 2015.

[5] Mother Jones Mass Shooting Database,
`http://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data`


## Figures

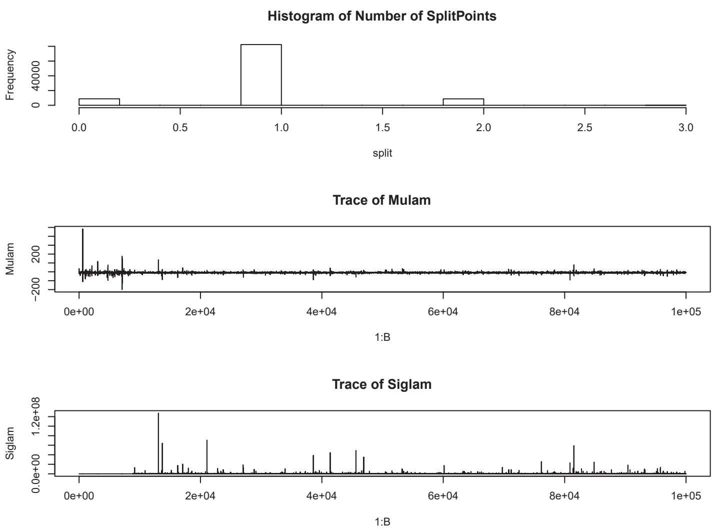


Figure 1: Posterior Summary of Piecewise Constant Model Quantities.

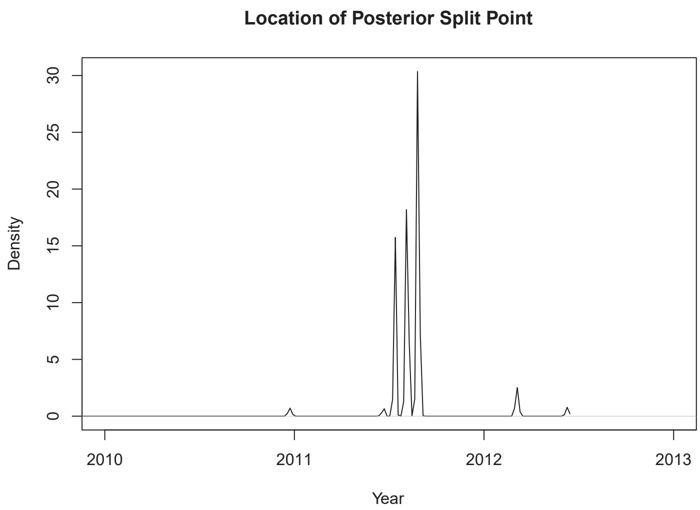


Figure 2: Posterior Density of the Location of the single split point.

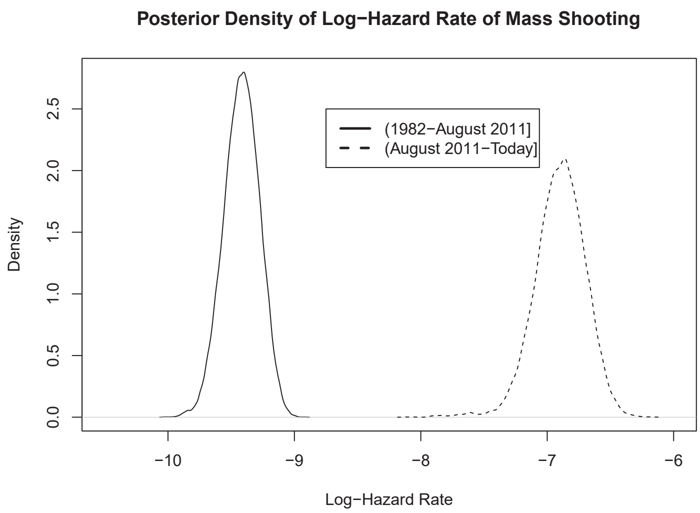


Figure 3: Posterior Density Log Hazard Rates with one split point .