"Turning Points" in the Iraq Conflict: Reversible Jump Markov Chain Monte Carlo in Political Science

'Turning Points' in the Iraq Conflict: Reversible Jump Markov Chain Monte Carlo in Political Science*

Arthur Spirling†

June 16, 2007

*For resubmission to *The American Statistician.*

†Department of Political Science, University of Rochester. Email: `spln@mail.rochester.edu` First version: October 13, 2006. This version: June 16, 2007. For comments and advice, I thank Jeff Gill, David Hastie, Michael Herron, Tasos Kalandrakis, Andrew Martin, Michael Peress, Matthew Platt, Bing Powell, David Primo, Larry Rothenberg, Curtis Signorino, an anonymous referee and the associate editor at *TAS*. The Comparative Politics Working Group provided helpful feedback. Funding and technical support from `the star lab` is gratefully acknowledged. The usual caveat applies.

Abstract

We consider and explore structural breaks in a day-by-day time series of civilian casualties for the current Iraq conflict: an undertaking of potential interest to scholars of international relations, Comparative and American politics. We review Bayesian change point techniques already used by political methodologists before advocating and briefly describing the use of reversible-jump Markov chain monte carlo techniques to solve the estimation problem at hand. We find evidence of four change points, all associated with increasing violence, approximately contemporaneous with some important state building events. We conclude with a discussion of avenues for future research.

Key words: Political methodology, sectarian violence, change points

1 INTRODUCTION

The study of *inter–* and *intra–* state conflict is a mainstay of political science. As an *international* conflict that increasingly resembles a *civil* war, the current situation in Iraq provides both a testing ground for theories on the duration and termination of different types of conflicts (e.g. Filson and Werner, 2004; Stam and Bennett, 2006), as well as a rich source of data for empirical work. This is quite separate from its obvious importance as a political, military and economic event in progress. In part due to its contemporaneous nature, political scientists have access to carefully, daily recorded, military and civilian casualty information: an unusual and excitingly fine level of detail. Of course, the utility of any data is only as good as the way it is explored and analyzed. Here, we suggest that a fruitful approach for political scientists lies in examining the time series for (potentially multiple) structural breaks and their effects. For scholars of American politics and public policy, the way that these change points correspond with administration statements on the progress of the war may be particularly intriguing. This notion extends to Comparative institutions scholars interested in the potentially pacifying effect of various post-war 'state-building' activities. In keeping with the increasing acceptance and popularity of Bayesian methods in political science, in undertaking our study we justify and adopt a novel (to political science) approach that uses a more general form of Markov chain monte carlo (MCMC) techniques, well-known to statisticians as 'reversible jump' MCMC (Green, 1995). We do so primarily for computational reasons.

Examining civilian casualty data from the official cessation of hostilities (May 2003) to May 2007, we find evidence of four change points. These breaks are approximately contemporaneous with (1) the capture of Saddam, and the emergence of the *Abu Graib* scandal (late 2003 to Spring 2004); (2) the installation of the Iraqi Interim Government, and the subsequent handover of power to the Iraqi Transitional Government (Summer 2004 to early 2005); (3) the legislative elections for, and negotiations to form, the first full-term Iraqi government (the early months of 2006); (4) the assumption of security and some military responsibilities by the Iraqi government (August/September 2006) . In every case, the frequency with which such incidents occur is *increasing* after the break.

2 BACKGROUND AND DATA

The United States and allied forces attacked Iraq with aerial bombardments, followed by a land invasion, on March 20th, 2003. By mid-April, Iraq's capital city, Baghdad and Saddam Hussein's home region of Tikrit was under allied control—bringing a *de facto* end to the war. A fortnight later, on May 3rd, 2003, President Bush declared that allied combat operations would now officially cease. As with all conflicts, the war has not been costless. What marks the Second Iraq War though, is the *continued* loss of life *after* the Iraqi army was formally defeated. At the time of writing, some 3,000 coalition force members had died in addition to at least 57,000 civilian fatalities since military operations began (sources are `http://icasualties.org/oif/` and `http://www.iraqbodycount.org/` respectively. Some studies have placed the number of civilian fatalities at a much higher number. For example, Burnham et al. (2006) claim up to 600,000 deaths). Violence has not yet abated despite the passing of some presumably important landmarks in what some characterize as the development of Iraq's polity and stability: for example, the capture of Saddam (December, 2003), the placing of the former dictator on trial from 'crimes against humanity' (July 2004) and his execution (December 2006); the killing of Saddam's sons, Uday and Qusay (July 2003); National Assembly elections (January 2005); the drafting (December 2003–March 2004) and subsequent referendum approval (October 2005) of a constitution; the election of a new president (April 2005) and the forming of a governing coalition (May 2006); the execution of an *Al-Qaeda* ringleader, Abu Musab al-Zarqawi, thought responsible for planning many terrorist attacks (June 2006); the assumption of security responsibility by the Iraqi government (September 2006). We are interested in violence for the post-(official) war period: although we certainly cannot make firm causal claims, our study will enable us, for example, to make statements about the plausibility of various events as "turning points" and allows us to pass some exploratory comments on how new democratic institutions and state apparatus developments are effecting Iraqis. Hence our study focuses on May 3rd 2003 through to the present time (May 2007).

Our data are drawn from `iraqbodycount.org` a (online) data base that records civilian deaths in Iraq "that have resulted from the 2003 military intervention by the USA and its allies. The count includes civilian deaths caused by coalition military action and by military or paramilitary responses to the coalition presence (e.g. insurgent and terrorist attacks)" (Dardagan and Sloboda, 2006). The data in raw form record deaths at the *day* level, from January 2003 through to the present and are compiled from (primarily Western) media reports and other sources. Since uncertainty often exists on precise numbers, especially when different agencies have conflicting figures for the same incident, the data base reports a range of possible death numbers from a 'minimum' to a 'maximum.' Potential 'over-counting' is a concern, so we use the 'minimum' and define a 'casualty incident' as involving five deaths or more (our findings below are similar when we define the incident threshold at ten or twenty deaths). For the purposes of this paper, we focus on the (changing) *frequency* of attacks, rather than their *size* (above our minimum). In part this is a behavioral assumption: we would contend that, at least initially, terrorists were able to control *how often* they planned to inflict casualties, rather than *how many.* There were 1682 such incidents in our time series, and we graph their occurrence in Figure 1; there, the solid line is the cumulative incident count, the solid dots are simply jittered incident occurrences (for which the y-axis is *not* the scale). We also report various dates that may of interest and to give readers a sense of timing perspective.

[Figure 1 about here.]

Although univariate time series work is not regularly encountered in political science, it is valuable in the current context as a 'first glance' exploration before covariate information becomes available. We think that such work helps to prompt both theorizing and data gathering for more nuanced and sophisticated analysis.

3 ESTIMATION PROBLEM

The *single* change point problem, estimated using Markov chain monte carlo techniques, has been discussed for and by political scientists elsewhere (see Western and Kleykamp, 2004). That treatment is similar to the (hierarchical) presentation given by Carlin, Gelfand and Smith (1992): suppose $\mathbf{y} = (y_1, \ldots, y_T)$ is a vector of observations of the random variable Y (casualty incidents) over time and let f and g be unknown densities in the same parametric family with $y_i \sim f(Y|\lambda_1), i = 1, \ldots, k$, $y_i \sim g(Y|\lambda_2), i = k+1, \ldots T$. We wish to estimate k the (single) change point which takes (discrete) values in $\{1, 2, \ldots, T\}$. A frequentist approach proceeds by maximizing

$$\mathcal{L}(\mathbf{y}) = \prod_{i=1}^{k} f(y_i|\lambda_1) \prod_{i=k+1}^{T} g(y_i|\lambda_2) \tag{1}$$

to obtain k and the parameters λ_1 and λ_2 (which for the count case are arrival rates for a Poisson) if they are of interest. A Bayesian approach proceeds by placing a prior $\tau(k)$ on the change point. There are computational advantages of a Bayesian MCMC approach here since (a) maximizing (1) requires optimization in a space that is not continuous (recall that k is discrete) which, say, Gibbs sampling does not; (b) the resultant non-nested models may be straightforwardly compared using Bayes factors (Chib, 1998); (c) missingness in $\mathbf{y}$ is handled systematically. This is quite apart from the philosophical appeal of Bayesian approaches of which political scientists are increasingly aware (for example, Gill see 2002, 1–6 and Jackman 2004, 486).

Here, we are interested in exploring *multiple* change point and such work (Bayesian or otherwise) is much less common in political science. In part this is because, with respect to the logic above, there are profound computational difficulties in generating proposals for situations where we suspect there are more than a couple of change points. One approach, suggested by Chib (1998) and applied to American politics by Park (2006), treats the change point model as a type of time series Markov mixture model, where the observations are (assumed) drawn from latent state variables. Notice that this approach requires *separate* Markov chain monte carlo runs for the different numbers of change points hypothesized (Leonte, Nott and Dunsmuir, 2003). An alternative solution is to use *reversible jump* Markov chain monte carlo which allows us to complete the computational operations in one 'go' as well as allowing us to be *a priori* agnostic over the number of parameters to be estimated.

Typically when MCMC is used in political science the parameter vector θ has a known number of components, denoted n. For the single change point problem $n = 3$ (these are k, λ_1 and λ_2). Now consider a very different scenario which arises for an unknown number of k change points: for every possible k, we need to estimate $2k+1$ parameters—the change points themselves and then parameters of the densities before, between and after them. That is, we have a set of $\mathbb{M}_k = \{1, \ldots, K\}$ candidate *models* of our data generating process, each with a *different* number of parameters. Otherwise put, the number of parameters is, of itself, a parameter. More formally, the k^{th} model in $\mathbb{M}_k$ has associated parameter vector θ_k which contains n_k parameters such that $\theta_k \in \mathbb{R}^{n_k}$.

Continuing to denote our data vector $\mathbf{y}$, the joint distribution becomes:

$$\begin{aligned} p(k, \theta_k, \mathbf{y}) &= p(\mathbf{y}|k, \theta_k)p(k, \theta_k) \\ &= p(\mathbf{y}|k, \theta_k)p(\theta_k|k)p(k). \end{aligned} \tag{2}$$

Since we have a constant of proportionality we can rearrange and reexpress (2) into the more familiar

$$\underbrace{p(k, \theta_k|\mathbf{y})}_{\text{posterior}} \propto \underbrace{p(k)p(\theta_k|k)}_{\text{prior(s)}} \underbrace{p(\mathbf{y}|k, \theta_k)}_{\text{likelihood}}. \tag{3}$$

Notice that $p(\mathbf{y}|k, \theta_k)$ is simply the likelihood, while $p(\theta_k|k)$ is the prior for the parameter vector, given a particular data generating process and $p(k)$ is the prior on the model itself. We wish to generate samples from (3). Setting up a Markov chain to do this may be difficult though, because it is required not simply to move around the parameter space for any particular θ_k, but to also 'jump' from space to space (from model to model) depending on the k in question.

This type of problem is given a general formulation by Green (1995), known as reversible jump MCMC (RJMCMC), of which standard MCMC algorithms are special cases. Green explicitly discusses a Poisson count change point problem and we followed his approach for our application (though we varied the priors somewhat to ensure that our results were robust to such alternative specifications). Although well known to statisticians, the details somewhat technical, and readers are guided to Brooks (1997) who gives an accessible overview for political scientists.

The *implementation* of RJMCMC, in particular the efficiency of proposals, can be problematic in practice and Hastie (2005) devotes considerable attention to designing a technique to do this. We used his `Automix` sampler (with a maximum of ten possible change points) for our estimation. Though the full details are somewhat technical, drawing on Hastie (2005, 202–203), it is instructive to summarize the way that the model of the data generating process is selected. The first two stages of the sampler produce a Normal mixture distribution for every possible value of k. In the third stage, assuming the Markov chain is currently in state (k, θ_k), `Automix` allocates the parameter vector θ_k to a component l_k of the mixture and uses it to standardize θ_k. Then a new model k' is proposed, along with a commensurate (new) mixture which has component l'_k. To obtain the new state vector $\theta'_{k'}$, the standardized vector is transformed using the mean and the covariance matrix of the mixture component $l'_{k'}$. `Automix` then accepts the proposed state $(k', \theta'_{k'})$ with some specified acceptance probability. A particularly pleasing feature of this software is that issues such as burn in and the requisite number of post-burn iterations are handled automatically.

4 RESULTS

There are three sets of (posterior) distributions that interest us here:

1. the posterior of k: this enables us to answer the question "how many change points in the data?" This will have support $k = 1, \ldots, k_{\max}$ where $k_{\max} = 10$.

2. the posterior of change point *positions* conditional on some estimated k. More intuitively, this enables us to answer the question "given a particular number of change points, *when* did they occur in the data?"

3. the posteriors of the rates for each period, conditional on some estimated k: that is, given the *number* of change points, and *when* they occurred, we can answer "what were the *effects* of the change points?"

In Figure 2 we display the posterior for k, the number of change points. The strongest evidence (in the sense of Kass and Raftery (1995)) is for $k = 4$ and we will explore this possibility exclusively.

[Figure 2 about here.]

In Table 1 we summarize the results for $k = 4$ model in a way that answers questions 2 and 3 above. The first break, in late January 2003 occurs between incidents that may be of import. The first was the capture and arrest of Saddam at a farmhouse near Tikrit in December 2003. The subsequent months saw both an insurgency uprising lead by rebel Shia cleric Muqtada al-Sadr in Baghdad and the diffusion of abusive photographs taken at the *Abu Graib* prison where coalition forces were holding Iraqi detainees. The political fallout of the latter was profound, and criticism of the Bush administration by allied, Arab and other politicians was widespread. This event, arguably, rallied and spurred sectarian hatreds and violence. The break marks a sharp increase in the casualty rate, doubling from one incident every four days, to one every two days.

The second break occurs in August of 2004, a little while after the Iraqi Interim Government assumed power from the Coalition Provisional Authority (in June 2004). This new entity, under the Premiership of Iyad Allawi was subsequently recognized as the legitimate sovereign government of Iraq by both the United Nations and the Arab League (an important regional player). Allawi quickly announced new security measures to tackle insurgency forces and was criticized by some for their draconian nature. As part of this offensive, the Iraqi Interim Government began to censor the critical reports of media outlet *al Jazeera*. Included in the highest posterior density interval for this break is the January 2005 democratic elections for the Iraqi Transitional Government. This change point saw an increase in violence from one incident every two days, to four incidents every five days.

The third break itself, in February 2006, occurs not long after the elections for the first full term Iraqi government (December 2005) and at around the time of the protracted negotiations to form a new coalition government. These talks were deadlocked for some time, lasting from December through to April of the following year. Jawad al-Maliki, leader of the Islamic Dawa Party would become Prime Minister after the original candidate Ibrahim al-Jaafari proved unacceptable to the

Sunni and Kurdish representatives in parliament. Once again, violence surged after this point with, on average, five incidents occurring every three days.

The fourth and final break occurs in September of 2006, a time when the Iraqi government assumed control of national security for approximately 70 percent of the country. The first specific task of the Iraqi Security Forces was, and is, to tame insurgency (with coalition logistical and medical support). By now incident rates were approaching three per day.

[Table 1 about here.]

In Figure 3 we summarize our findings in a different way: the open circles represent the median incident rate between the relevant breaks which are demarcated by the broken lines. For reference, we again draw the jittered incidents themselves on the plot.

[Figure 3 about here.]

5 DISCUSSION

Our study—to our knowledge the first that uses RJMCMC in a political science context—suggests that violence is increasing and that important state-building activities, like democratic elections, are contemporaneous with upticks in casualties. Apart from this rather grim substantive conclusion, we found that investigating time series on violence to be an interesting and fruitful exercise. If a Bayesian approach is pursued, then reversible jump techniques seem most helpful. We hope that our brief article will encourage others in political science to consider such methods in future.

As noted above, we do not establish causation in any sense: the events we noted were simply occurring at around the same time as the breaks in the time series and it is speculative that they may be of direct importance. This suggests some interesting avenues for future research: for example, one possibility is that increasing violence is a product of an increasingly organized insurgency. On this point, notice that the solid line in Figure 1 resembles a exponential curve of form $F(t) = a\mathrm{e}^{rt} + \epsilon$ (we are indebted to an anonymous *TAS* referee for this observation). Modeling—both theoretical and statistical—of this apparent pattern would allow us to think about the development of the conflict in Iraq in a more systematic way. We leave this for future research.

References

Brooks, Stephen P. 1997. "Markov chain Monte Carlo Method and its applications." *The Statistician* 47(1):67–100.

Burnham, Gilbert, Riyadh Lafta, Shannon Doocy and Les Roberts. 2006. "Mortality after the 2003 invasion of Iraq: a cross-sectional cluster sample survey." *The Lancet (online)* 10.1016/(S0140-6736(06)69491-9).

Carlin, Bradley, Alan Gelfand and Adrian Smith. 1992. "Hierarchical Bayesian Analysis of Changepoint Problems." *Applied Statistics* 41(2):389–405.

Chib, Siddhartha. 1998. "Estimation and comparison of multiple change point models." *Journal of Econometrics* 86:221–241.

Dardagan, Hamit and John Sloboda. 2006. "Iraq Body Count.".
URL: *http://www.iraqbodycount.net/*

Filson, Darren and Suzanne Werner. 2004. "Bargaining and Fighting: The Impact of Regime Type on War Onset, Duration, and Outcomes." *American Journal of Political Science* 48(2):296–313.

Gill, Jeff. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. New York: Chapman & Hall/CRC.

Green, Peter J. 1995. "Reversible Jump MCMC computation and Bayesian model determination." *Biometrika* 82(3):711–732.

Hastie, David. 2005. Towards Automatic Reversible Jump Markov Chain Monte Carlo PhD thesis University of Bristol.

Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7:483–505.

Kass, Robert E. and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(430):773–795.

Leonte, Daniela, David J. Nott and William T. M. Dunsmuir. 2003. "Smoothing and Change Point Detection for Gamma Ray Count Data." *Mathematical Geology* 35(2).

Park, Jong Hee. 2006. "Modeling Structural Changes: Bayesian Estimation of Multiple Change-point Models and State Space Models." *Prepared for American Political Science Association Meeting* August, 2006, Philadelphia.

Stam, Allan C. and D. Scott Bennett. 2006. "Predicting the Length of the 2003 U.S.-Iraq War." *Foreign Policy Analysis* 2(2):101–116.

Western, Bruce and Meredith Kleykamp. 2004. "A Bayesian Change Point Model for Historical Time Series Analysis." *Political Analysis* 12:354–374.

List of Figures

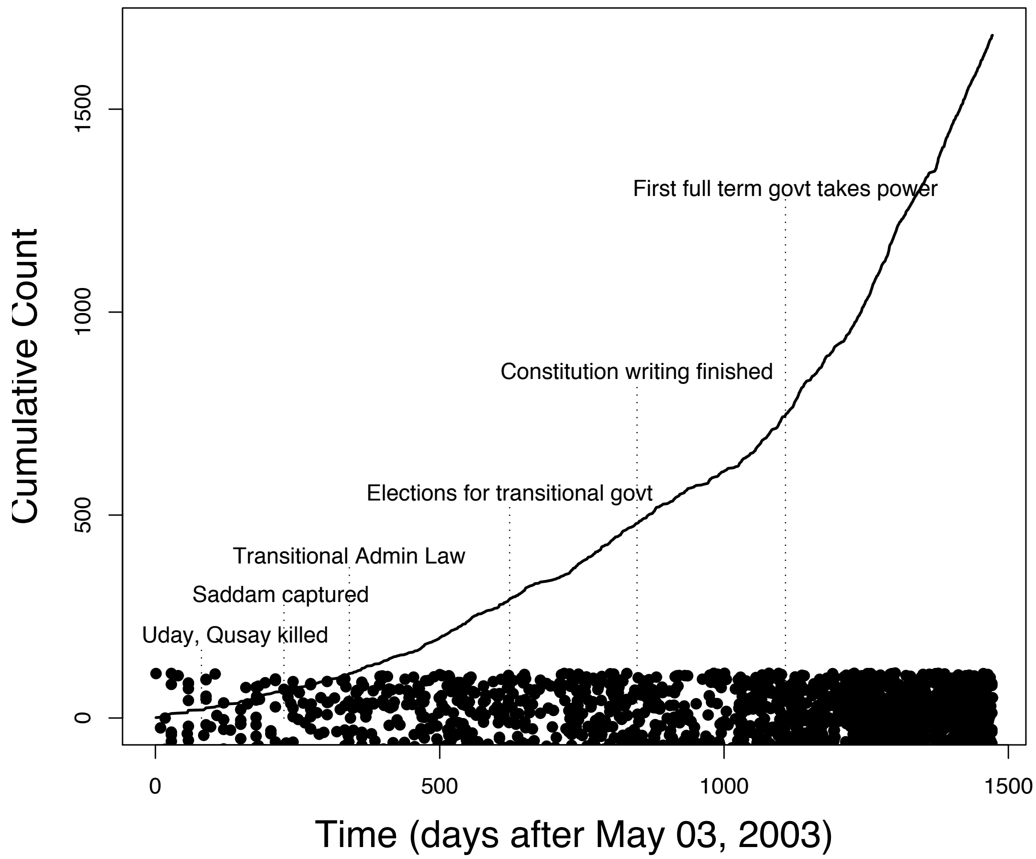

Figure 1: Iraq Casualty Incidents, May 2003–May 2007. Solid line represents cumulative counts; solid dots are jittered incident occurrences; various dates of interest demarcated.

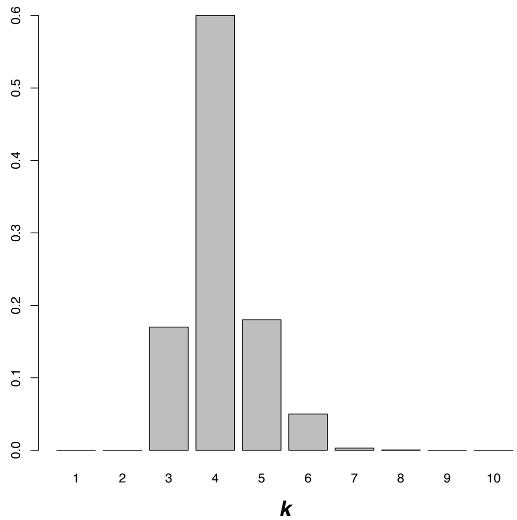

Figure 2: Posterior of k: number of change points, Iraq casualty data.

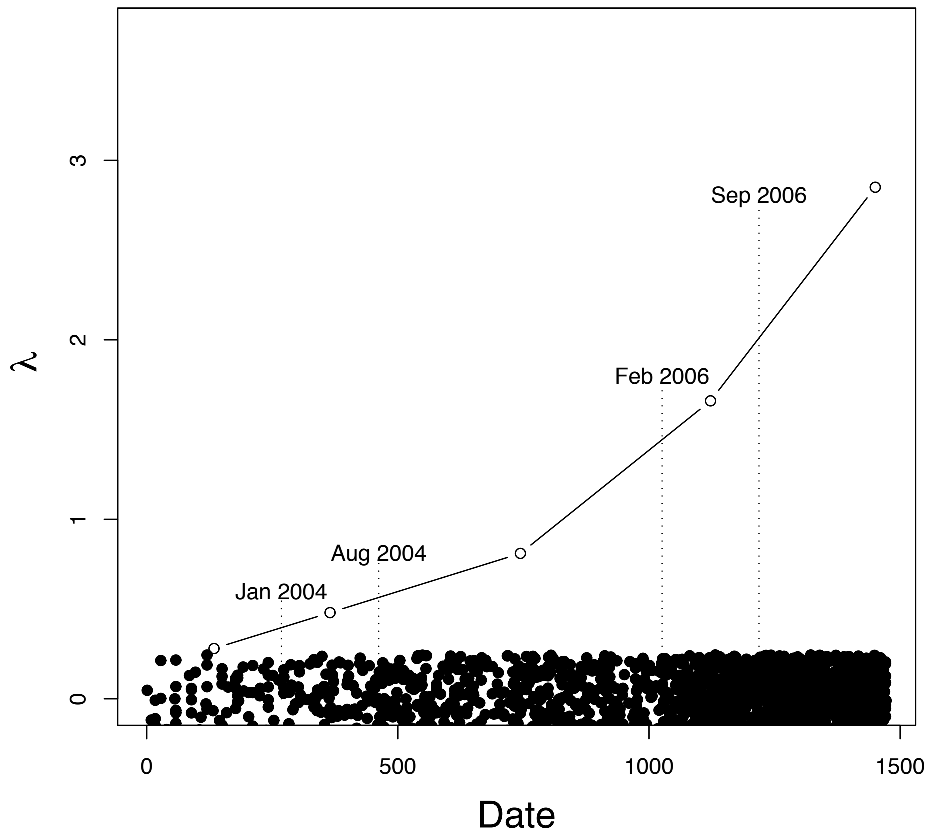

Figure 3: Change in rate of casualty incidents over time: rate on y-axis, actual incidents as jittered points. Dates correspond to breaks found.

List of Tables

Period	Time	Rate (λ)	Comments
before break 1		0.283	1 incident every 4 days
break 1	Jan 26, 2004 [Jul 18, 2003–Apr 30, 2004]		Capture of Saddam (Dec 2003) *Abu Graib* scandal (Apr 2004)
btwn break 1 and 2		0.476	1 incident every 2 days
break 2	Aug 7, 2004 [Jul 10, 2004–May 4, 2005]		Iraqi Interim Government assumes power (Jun 2004) elections for Iraqi Transitional Govt (Jan 2005)
btwn break 2 and 3		0.814	4 incidents every 5 days
break 3	Feb 22, 2006 [Feb 16, 2006–Mar 24, 2006]		elections for first full term Iraqi government (Dec 2005)
btwn break 3 and 4		1.660	5 incidents every 3 days
break 4	Sept 4, 2006 [Aug 23, 2006–Sept 20, 2006]		Iraqi govt assumes control of counter-insurgency operations for much of country (Sept 2006)
after break 4		2.850	8 incidents every 3 days

Table 1: Change point median dates [90% HPD] in Iraq casualty data, with rates between breaks. Comments note political and other activities approximately contemporaneous to break.