The Evolution of International Political Risk 1956-2001

Ephraim Clark & Radu Tunaru[1]

Abstract

This paper deals with international political risk defined as political events with substantial, negative economic and financial repercussions that are felt world wide. We use Markov Chain Monte Carlo (MCMC) modelling techniques to measure the evolution between 1956 and 2001 of international political risk. For the first time to our knowledge international political events are investigated, a timely topic that fills a major hole in the literature. The Bayesian Hierarchical Markov Chain Monte Carlo modelling that we adopt also adds a new dimension to political risk assessment.

Key words: political risk, rare events, hierarchical Bayesian models, MCMC sampling

JEL: F37, C51, C53

[1] Corresponding author: Address: Cass Business School, City University, London EC1Y 8TZ, e-mail: r.tunaru@city.ac.uk

1. Introduction

International political events have been a fact of life since the end of the Korean War. The Suez crisis of 1956, the Cuban missile crisis of 1962, the oil embargo of 1973, and the debt crisis of 1982 are some of the most obvious examples. Combined with the ongoing "globalization" process, more recent events, such as September 11, 2001 and the war in Iraq, are powerful reminders of the importance of the international aspect of political risk. The political risk literature, however, either ignores this aspect or treats it indirectly. Root (1973), for example, focuses on country specific characteristics that he divides into transfer risks (potential restrictions on transfer of funds, products, technology and people), operational risks (uncertainty about policies, regulations, governmental administrative procedures which would hinder results and management of operations in the foreign country), and, finally, risks on control of capital (discrimination against foreign firms, expropriation, forced local shareholding, etc.). Robbock and Simmonds (1973) look at country specific political events that cause unanticipated discontinuities in the business environment. Brewer (1981) refers to political risk as miscellaneous risks from doing business abroad. The expropriation literature is country specific by definition (see, for example, Eaton and Gersovitz (1984), Andersson (1989), and Raff (1992)) and the contagion literature, which might be expected to consider international political risk, is also country focused.[2] Valdes (1997), for example, defines contagion as excess co-

[2] The thrust of the contagion literature is on how a crisis is transmitted from one country to another. Calvo (1998), for example, explains contagion as a result of liquidity and asymmetric information, whereby a leveraged investor facing margin calls must sell his assets to uninformed investors who cannot distinguish between good assets and bad (lemons problem). A variant of this scenario is leveraged investors facing

movement in asset returns across countries and Eichengreen, Rose and Wyplosz (1996) define it as a situation where knowledge of a crisis in one country increases the probability of a crisis in another country above that warranted by the fundamentals. Meldrum (2000) summarizes the definition of political risk as additional risks not present in domestic transactions that typically include risks arising from a variety of national differences in economic structures, policies, socio-political institutions, geography and currencies.[3]

In this paper we focus on international political risk, which we define as political events with substantial, negative economic and financial repercussions that are felt world wide as opposed to political events whose economic and financial consequences are limited to a specific country or region. More specifically, we use Markov Chain Monte Carlo (MCMC) modelling techniques to measure the evolution between 1956 and 2001 of international political risk. The novelty of our study is twofold. First, we focus on international political events, something that to our knowledge has never been done before. The second innovation is the use of MCMC modelling techniques to estimate the level of international political risk, where political risk is defined in Clark (1997) and Clark and Tunaru (2003) as the expected arrival rate of political events. This definition recognizes that political risk can arise from a wide range of sources, which are often

---

margin calls who sell assets whose price has not yet collapsed, thereby causing the collapse of these prices and spreading from market to market. Kaminsky and Reinhart (2000) emphasize the role of common lenders, such as commercial banks. In this explanation, the banks' need to rebalance their portfolios and recapitalize after initial losses causes an overall reduction in credit to most or all countries that rely on them for credit. The most plausible family of contagion models focuses on the role of trade in financial assets and information asymmetries. Calvo and Mendoza (2000), for example, show how the costs of gathering and processing country risk information can cause herding behavior even among rational investors.

[3] For a comprehensive, in-depth presentation of political risk and an extensive bibliography, see Bouchet et al. (2003).

mutually dependent. As such, it is very general. By looking specifically at international political risk, we deal with a timely topic and fill a major hole in the literature. The Bayesian Hierarchical Markov Chain Monte Carlo modelling technique that we adopt also adds a new dimension to political risk assessment.

Traditional methods for assessing political risk are generally country specific and range from the comparative techniques of rating and mapping systems to the analytical techniques of special reports, dynamic segmentation, expert systems, and probability determination to the econometric techniques of model building and discriminant and logit analysis.[4] The non-econometric techniques are generally very timely but have the shortcoming of also being very subjective. Rating systems, for example, reflect the latest information but the different factors and factor weights are generally subjective in the sense that they have no comprehensive statistical or theoretical underpinning. The econometric techniques are less subjective but have the shortcoming of not being very timely. For example, when conditions change, it can take a long time and many observations before the change is fully reflected in the estimated coefficients. MCMC simulation based on Bayesian hierarchical models has the advantage of being both timely and less subjective. It can also handle an important aspect of political risk that is widely acknowledged in the literature and modelled explicitly by Clark and Tunaru (2003), that is, that loss causing political events arise from a wide range of sources, which are often mutually dependent.

[4] See Bouchet et al. for a comprehensive presentation and analysis of assessment techniques.

Treating political risk as a loss causing event is clearly in the spirit of authors such as Root (1973), Simon (1982), Howell and Chaddick (1994), Roy and Roy (1994) and Meldrum (2000), who analyse risk as an explicit negative event that causes an actual loss or a reduction of the investment's expected return. This stands in contrast to other authors such as Robock (1971) and Haendel *et al.* (1975), Kobrin (1979) or more recently Feils and Sabac (2000), who focus on political risk as it affects the volatility of an investment's overall profitability both negatively and positively. Tests of political risk on investment outcomes reflect these two approaches. Kim and Mei (2001), Chan and Wei (1996), Cutler et al. (1989) and Bittlingmayer (1988) consider political risk with respect to stock market volatility. Other papers, such as Erb et al. (1995 and 1996), Cosset and Suret (1995), Bekaert (1995), and Bekaert and Harvey (1997) focus on losses and test political risk with respect to stock market performance. We choose the negative slant on political risk because we find it more intuitive and more in line with what investors generally understand by political risk.

The rest of the paper is organized as follows. Section 2 presents the data and section 3 contains an overview of Bayesian hierarchical models that can be applied for analysing rare events and MCMC techniques that are needed for extracting statistical inference from the economic time series. The main findings are discussed in Section 4 where empirical results are compared. The last section summarises and concludes.

2. Data

The data presented in table 1 is organized annually and comprises international political events between 1956 and 2001. We could find no general objective criteria on which to distinguish country specific or regional specific events from the truly international ones. Thus our choice of events is largely subjective, based on personal analysis with the help of professional historians. Although this subjective element is a potential weakness, we feel that the majority of events included in the study are non-controversial. These include events such as the Suez Crisis of 1956, the Cuban Missile Crisis of 1962, the stockmarket crash of 1987, the September 11 terrorist attack, the international debt crisis of 1982, the 1994 Mexican peso crisis, the 1997 Southeast Asian economic meltdown, the reorganization of the world monetary system in 1976 and the runs on gold and the international monetary crises of 1960, 1962, 1967, 1971, 1973, and 1985. We also feel that nuclearization is a world shaking event. Thus, we include France's first nuclear tests in 1960 and those of India and Pakistan in 1998. Other events that would normally be consigned to the regional event category are international events because of the oil factor. These include the 1967 Middle East War, the 1972 Arab terrorist attack at the Munich Olympic Games, the 1973 Middle East War that provoked the oil embargo, the 1979 Iranian Revolution, the 1980 Iran-Iraq war, and the 1990 Iraqi invasion of Kuwait followed by Operation Desert Storm in 1991. The Cold War is the direct source of other international events such as the U2 spy plane incident in 1960, the construction of the Berlin Wall in 1961, the 1968 invasion of Czechoslovakia by Warsaw Pact troops, and the US government’s Star Wars Initiative in 1983. It is also an indirect source for regional conflicts that would otherwise have been purely regional affairs. This is true for the Indo-Pakistan war of 1965, the Turkey-Cyprus wars of 1964 and 1974, the war between the Bengali rebels and Pakistan in 1971, and the Ethiopian-Eritrean war of 1975. Combined with the economic, military and political power of the United States in the world, it was also an important element in incidents such as Kennedy's assassination in 1963, and incidents in the Vietnam war, such as the US bombing of North Vietnam in 1966, and the Tet offensive of 1968. The Vietnam war itself is responsible for a series of US political events that rocked the country and, because of the US position of overwhelming international power, the world as well: the assassinations of Martin Luther King and Robert Kennedy and the upheaval at the Democratic convention in 1968, the massive anti-war demonstration in Washington DC in 1969 and the Kent State incident where US troops fired on and killed protesting students. Again, the importance of the US in the world makes any threat to its political stability a threat to the world. Thus we include Clinton's impeachment in 1998 and the contested presidential election in 2000. We include the interest equalization tax of 1963 and the mandatory controls on foreign investment in 1968 because of their effects on the international financial system and world capital flows.

To get a better feel for what we have included as international political events, it is instructive to consider some events that were not included. For example, why did we exclude the Hungarian invasion by the Soviet Union in 1956 while including the invasion of Czechoslovakia in 1968 or why did we exclude the Bay of Pigs in 1961 but include the Cuban missile crisis of 1962? The invasion of Hungary was aimed at putting down a revolution, seen from the eyes of the recognized government, basically an internal affair, and treated as such by the US and the rest of the free world. The invasion of Hungary was aimed at toppling the recognized government, basically an act of war, and treated as such by the free world. The Bay of Pigs did not involve the superpowers directly while the missile crisis did. The same reasoning went for excluding the Soviet invasion of Afghanistan and the US invasions of Grenada and Panama.

Table 1. International Political Events 1956-2001
1956 Suez Crisis
1960 Run on gold causes creation of London gold pool
1960 Soviets shoot down American spy plane
1960 France tests its first atomic weapon
1961 Berlin Wall is constructed
1962 Cuban missile crisis
1962 France begins selling dollars for gold
1963 Kennedy assassinated
1963 interest equalization tax
1964 Turkey planes attack Cyprus (UN peace force takes over in Cyprus)
1965 Indo-Pakistan War
1966 B-52 bomb North Vietnam
1967 Devaluation of British pound followed by world monetary crisis
1967 Middle East War
1968 Tet offensive
1968 Czechoslovakia invaded by Warsaw Pact troops
1968 Martin Luther King assassinated
1968 Robert Kennedy assassinated
1968 Upheaval at Democratic convention
1968 mandatory controls on foreign investment by US residents
1969 Massive antiwar demonstration in DC

1970 Kent State incident troops fire on students

1971 Gold convertibility suspended
1971 War between Bengali rebels and Pakistan

1972 Arab terrorist attack at Munich Olympics

1973 Dollar devalued
1973 Oil embargo/Yom Kippur War

1974 Turks invade Cyprus

1975 Ethiopian-Eritrean War

1976 New international monetary system (gold demonetized, floating exchange rates agreed)

1979 Iranian Revolution

1980 Iran-Iraq War

1982 Mexico defaults

1983 Star Wars Initiative by US government

1985 Group of 5 announce policies to push down the dollar's value

1987 Stockmarket crash

1990 Iraq Invades Kuwait

1991 Operation Desert Storm

1994 Peso crisis

1997 Asian crisis

1998 India and Pakistan conduct nuclear tests

1998 Clinton impeached

2000 Contested election

2001 September 11

3. Bayesian Hierarchical Modelling with Markov Chain Monte Carlo Simulations

The data we analyse falls outside repeated random independent experiment framework. Hence, classical statistical inferential methods are not appropriate. Moreover, the data generation process undergoes major changes from time to time, which necessitates a more flexible tool than maximum likelihood inference or least-squares class of methods or the more empirical methods-of-moments approach. Ideally the analyst could use a tool that embeds all the above in an exploratory methodology that allows fitting any model that is correctly specified for the data investigated.

In addition, it is a well accepted fact that data cannot be the whole story *all the time*. Data is still a valuable source of information in finance but the imminent change in tax or increase or decrease of interest rates may shift the future data pattern to a direction that is not implied by the current state of affairs. Bayesian Hierarchical Modelling provides a solution for both problems identified above but the price to pay is the need for complex computational methods for practical implementation. The necessary computational methods do exist and are widely applied in other areas such as biostatistics, marketing, business, and epidemiology (see Gilks et al. (1996) for a wide range of applications).

In Bayesian hierarchical modelling the model is specified on several layers. For example, denoting generically the vector of all data by y, the vector of all parameters by

φ[5] and a probability density function by *p*, we first provide a likelihood distribution $p(y \mid \varphi)$ and an a priori distribution for the parameters $p(\varphi)$. Then, using the Bayes' law it is true that

$$p(\varphi \mid y) \propto p(y \mid \varphi)p(\varphi) \qquad (1)$$

where the $\propto$ signifies up to a proportionality constant. This process may continue hierarchically with further prior parameters associated with φ. The models in this paper are all hierarchical.

One of the simplest models for our data would be the following model (Model 1 from now on)

$$\begin{aligned} &Y_i|\theta \sim Pois(\theta) \\ &\theta|\alpha,\beta \sim Gamma(\alpha,\beta) \\ &\alpha \sim Gamma(a_1,a_2) \quad \beta \sim Gamma(b_1,b_2) \end{aligned} \qquad (2)$$

where a_1, a_2, b_1, b_2 are some constants that are chosen in order to specify the degree of information that the analyst has about the parameters α and β. Most of the time there is no precise information available so those values must be chosen such that the resulting Gamma distribution has a wide range of likely values. The model postulates that the number of events in each year are conditionally independent draws from the same Poisson distribution with arrival rate θ which is also a random draw from a Gamma distribution with parameters α and β.

For this model the joint posterior distribution of *all* parameters is

[5] A missing data observation can be considered as a parameter in the context of Bayesian modelling

$$
\begin{aligned}
p(\theta,\alpha,\beta \mid y) &\propto p(y \mid \theta)\,p(\theta \mid \alpha,\beta)\,p(\alpha)\,p(\beta) \\
&\propto \left[\prod_{i=1}^{N}\frac{\theta^{y_i}e^{-\theta}}{y_i!}\right]\left[\frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\theta\beta}\right]\left[\alpha^{a_1-1}e^{-\alpha a_2}\right]\left[\beta^{b_1}e^{-\beta b_2}\right]
\end{aligned}
\tag{3}
$$

The marginal posterior distribution for each parameter (or group of parameters) of interest can be identified by collecting all factors containing that parameter from the joint posterior distribution. Thus

$$
\begin{aligned}
&p(\theta \mid y,\alpha,\beta) \propto \theta^{\sum_{i=1}^{N} y_i+\alpha-1} e^{-\theta(\beta+1)} \propto Gamma(\sum_{i=1}^{N} y_i + \alpha, \beta+1) \\
&p(\alpha \mid y) \propto \frac{(\beta\theta)^{\alpha}}{\Gamma(\alpha)}\alpha^{a_1-1}e^{-\alpha a_2} \\
&p(\beta \mid y) \propto \beta^{\alpha+b_1-1}e^{-\beta(\theta+b_2)} \propto Gamma(\alpha+b_1, \theta+b_2)
\end{aligned}
\tag{4}
$$

In order to obtain inference, the analyst samples from values from the posterior distributions. While this may be an easy task when the obtained distributions are well-known distributions such as normal, Poisson or Gamma, this may be a daunting task when faced with distributions such as that obtained above for α. Moreover, it is intuitively obvious that with more sophisticated models hierarchically specified, the posterior combinations of probability density functions can be far from known.

The inference is easily obtained via simulation techniques such as Markov Chain Monte Carlo (MCMC). The first step is to ensure that the simulated chain or chains are stationary. Although it is theoretically impossible to be 100% sure that the chain has converged, a series of tests, measures and exploratory graphical investigations are conducted prior to any inferential calculations.

The results reported below for the models we use were obtained after a burn-in period of 30000 iterations. The first step for checking convergence is the trace plots of the simulated values, then the autocorrelation plots and then the Gelman-Rubin statistics. Two chains were used starting from overdispersed values and the inference sample is sometimes thinned (taking every 5th value from the sample) so that more independent values from the posterior densities are employed for calculations. The beauty of the MCMC methodology is that once a sample for the joint posterior distribution of all parameters and data is estimated[6].

4. Empirical Analysis

The model described above in Section 3 is just a starting point. One may wish to consider that each year the number of events is coming from a different Poisson distribution or that the arrival rate has a time trend. Two other more complex models using generalizations of the Poisson distribution are investigated later in this section.

The Model 1 is fitted for $a_1 = 10, a_2 = 1, b_1 = 0.001, b_2 = 0.001$ both sets of values ensuring a very wide spread Gamma distribution. Forecasting a future number of events can be done in two ways for Model 1. A new draw can be made from the fitted Poisson distribution with parameter θ. The table contains main summary statistics for all parameters describing the model. The forecast for the next future year is one event. Another idea that is quite easy to implement is to consider the next future event count as a missing data observation. The Bayesian Hierarchical modelling coupled with MCMC

[6] Note that this sample is made of values that are correlated. Nonetheless the sample is large enough to cover the whole density range and the lack of independence does not affect in any way the inference. If

techniques can handle this scenario quite easily. As shown in subtable b) the inference is very similar, confirming that next year is very likely to observe one event.

Table 2 Summary statistics for the Poisson model with a single rate of arrival The estimated quantities are obtained from two chains with over-dispersed initial values. after a burn-in period of 20000 from a thinned sample of 4000, that is every 5th value from the 10000 values part of the stationary Markov chains; a) model is fitted with 46 data points and prediction is made by drawing a new value from the fitted hierarchical Poisson model and b) model is fitted with 47 data points, the last one being declared a missing data and prediction is made by considering this missing observation as a parameter in the fitted hierarchical Poisson model

a)

node	mean	Sd	MC error	2.50%	median	97.50%
alpha	10.02	3.143	0.040	4.847	9.709	17.06
beta	10.68	5.075	0.066	3.132	9.886	22.73
theta	0.956	0.145	0.001	0.694	0.951	1.261
y.new	0.947	0.976	0.007	0.0	1.0	3.0
deviance	118.0	1.443	0.010	117.0	117.4	122.1

b)

alpha	10.02	3.156	0.024	4.827	9.691	17.08
beta	10.7	5.095	0.039	3.125	9.923	22.76
theta	0.957	0.144	0.0006	0.693	0.95	1.262
y[47]	0.946	0.979	0.004	0.0	1.0	3.0
deviance	118.0	1.446	0.006	117.0	117.4	122.2

Another advantage of MCMC modelling is that a whole posterior distribution can be estimated. In all this section for the majority of parameters the posterior density kernels are estimated. For example, in Figure 1 it can be seen that the mode of the posterior distribution for the future number of events is 0. In addition the distribution for the arrival rate is quite symmetric around 1.

---

some sort of independence in the sample is desired then the sample can be thinned by retaining from the sample every k-th value.

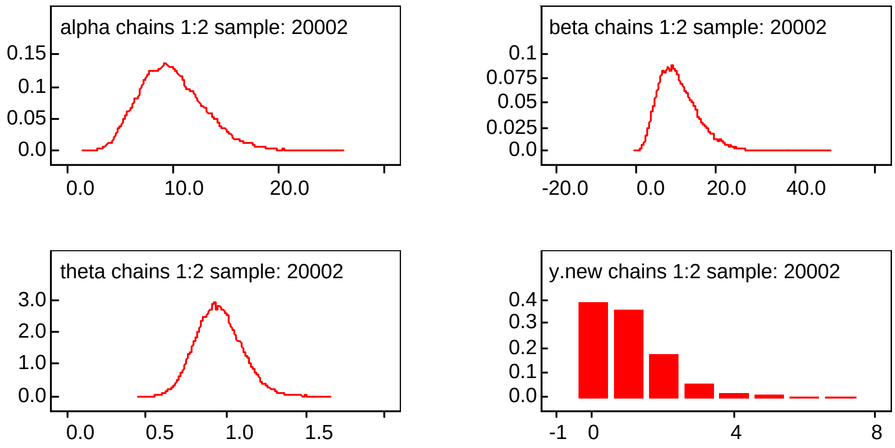

Figure. 1 Posterior density functions and histogram for the main parameters of interest of the Poisson model with a single rate of arrival gamma distributed.

This feature can be very useful to identify multimodal posterior distributions that cannot be included in the maximum likelihood framework, to explore possible correlations between parameters or to directly estimate any complicated function of subsets of parameters.

The next model investigated is a generalization of Model 1. Here we assume that every year the number of events comes from a Poisson distribution with individual arrival rate θ_i and all those rates are independent random draws from a Gamma distribution. Note that although here we have more parameters (θ's and other hyper-prior parameters) than data points, inference can be obtained because of the hierarchical structuring of the model on several layers. The Model 2 is given by

$$\begin{aligned} &Y_i|\theta_i \sim Pois(\theta_i) \\ &\theta_i|\alpha,\beta \sim Gamma(\alpha,\beta) \\ &\alpha \sim Gamma(3,1) \quad \beta \sim Gamma(3,1) \end{aligned} \tag{5}$$

The last line Gamma specification is not very restrictive, is till quite a wide spread distribution and it leads to conditional distributions that are easier to follow. The joint posterior distribution of all parameters is

$$
\begin{aligned}
p(\theta_1,...,\theta_N,\alpha,\beta \mid y) &\propto p(y \mid \theta_1,...,\theta_N)\,p(\theta_1,...,\theta_N \mid \alpha,\beta)\,p(\alpha)\,p(\beta) \\
&\propto \left[\prod_{i=1}^{N} \frac{\theta_i^{\,y_i} e^{-\theta_i}}{y_i!}\right]\left[\frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta_i^{\,\alpha-1} e^{-\theta_i\beta}\right]\left[\alpha^2 e^{-\alpha}\right]\left[\beta^2 e^{-\beta}\right]
\end{aligned}
\tag{6}
$$

The conditional distributions need for MCMC simulations are

$$
\begin{aligned}
p(\theta_i \mid y,\alpha,\beta) &\propto \theta_i^{\,y_i+\alpha-1} e^{-\theta_i(\beta+1)} \propto Gamma(y_i+\alpha, \beta+1) \\
p(\alpha \mid y) &\propto \frac{(\beta)^{N\alpha}}{\Gamma(\alpha)^N}\left[\prod_{i=1}^{N}\theta_i\right]^{\alpha} \alpha^2 e^{-\alpha} \\
p(\beta \mid y) &\propto \beta^{N\alpha} e^{-\beta(\sum_{i=1}^{N}\theta_i+1)} \propto Gamma(N\alpha+3, \sum_{i=1}^{N}\theta_i+1)
\end{aligned}
\tag{7}
$$

The two chains simulated with the help of WinBugs 1.4 become stationary quite rapidly. The autocorrelation plots, not provided here for lack of space, show no problems with convergence. The Gelman Rubin statistics are also very good, all being between 0.98 and 1.01. Various MCMC output produced in WinBUGS that is taken into consideration when extracting inference is presented in the Appendix.

The inference is summarised in Table 3. To forecast for the next year, first a new arrival rate θ_{new} is simulated and then a Poisson draw is made from this distribution. It is obvious that a good estimate for the future number of events is 1 with a credibility interval [0, 4].

Table 3 Summary statistics for the Poisson model with independent gamma rates of arrival. The estimated quantities are obtained from two chains with overdispersed initial values. after a burn-in period of 20000 from a thinned sample of 4000, that is every 5th value from the 10000 values part of the stationary Markov chains

node	Mean	sd	MC error	2.50%	Median	97.50%
alpha	2.821	1.272	0.030	1.13	2.581	5.963
beta	3.017	1.459	0.036	1.072	2.727	6.555
theta[1]	0.984	0.543	0.009	0.230	0.883	2.312
theta[2]	0.695	0.447	0.007	0.083	0.605	1.837
theta[3]	0.688	0.453	0.008	0.088	0.596	1.868
theta[4]	0.682	0.443	0.006	0.081	0.597	1.772
theta[5]	1.51	0.709	0.011	0.506	1.38	3.263
theta[6]	0.973	0.552	0.007	0.221	0.859	2.347
theta[7]	1.244	0.608	0.010	0.383	1.125	2.726
theta[8]	1.246	0.624	0.009	0.370	1.139	2.713
theta[9]	0.967	0.533	0.008	0.219	0.871	2.268
theta[10]	0.978	0.546	0.007	0.216	0.877	2.351
theta[11]	0.982	0.540	0.009	0.227	0.879	2.246
theta[12]	1.234	0.598	0.010	0.378	1.137	2.687
theta[13]	2.375	0.982	0.018	0.967	2.199	4.655
theta[14]	0.954	0.521	0.008	0.215	0.864	2.208
theta[15]	0.978	0.537	0.008	0.225	0.886	2.253
theta[16]	1.248	0.630	0.010	0.355	1.135	2.834
theta[17]	0.951	0.515	0.008	0.212	0.857	2.149
theta[18]	1.246	0.622	0.009	0.356	1.135	2.771
theta[19]	0.964	0.525	0.007	0.223	0.867	2.233
theta[20]	0.963	0.524	0.007	0.216	0.871	2.226
theta[21]	0.961	0.545	0.008	0.202	0.856	2.342
theta[22]	0.682	0.434	0.006	0.077	0.602	1.73
theta[23]	0.691	0.449	0.007	0.081	0.610	1.763
theta[24]	0.974	0.540	0.007	0.232	0.869	2.322
theta[25]	0.966	0.536	0.007	0.217	0.866	2.306
theta[26]	0.690	0.453	0.007	0.087	0.607	1.798
theta[27]	0.963	0.550	0.008	0.213	0.875	2.313
theta[28]	0.963	0.528	0.009	0.228	0.868	2.276
theta[29]	0.694	0.464	0.007	0.075	0.601	1.824
theta[30]	0.958	0.538	0.009	0.221	0.855	2.294
theta[31]	0.678	0.453	0.007	0.077	0.591	1.818
theta[32]	0.967	0.532	0.008	0.222	0.862	2.265
theta[33]	0.688	0.449	0.007	0.079	0.599	1.835
theta[34]	0.691	0.458	0.006	0.084	0.599	1.816
theta[35]	0.973	0.538	0.008	0.225	0.886	2.223
theta[36]	0.965	0.549	0.009	0.222	0.863	2.277
theta[37]	0.697	0.461	0.007	0.072	0.607	1.821
theta[38]	0.677	0.449	0.007	0.083	0.585	1.783
theta[39]	0.958	0.537	0.008	0.203	0.861	2.236
theta[40]	0.689	0.447	0.006	0.079	0.611	1.802
theta[41]	0.689	0.451	0.006	0.077	0.604	1.798
theta[42]	0.950	0.515	0.008	0.216	0.864	2.218
theta[43]	1.251	0.630	0.010	0.367	1.13	2.84
theta[44]	0.687	0.446	0.007	0.083	0.603	1.801
theta[45]	0.962	0.528	0.007	0.218	0.868	2.225
theta[46]	0.961	0.538	0.009	0.217	0.867	2.335
theta.new	0.961	0.657	0.012	0.133	0.813	2.651
y.new	0.971	1.175	0.017	0	1	4

Moreover, the most likely next future value is 0. This can be easily seen from the kernel mass density (histogram) of the future value in Figure 2. It would be difficult to distinguish between various estimates for this theoretical value. The posterior median is 1 with a credibility interval [0,4] that is not symmetric around the estimate. The estimate that is arguably most useful to the analyst is the posterior mode. This is 0 in this case and also it can be seen that the distribution where it is coming from is not multi-modal. This type of analysis is one of the strengths of the MCMC .

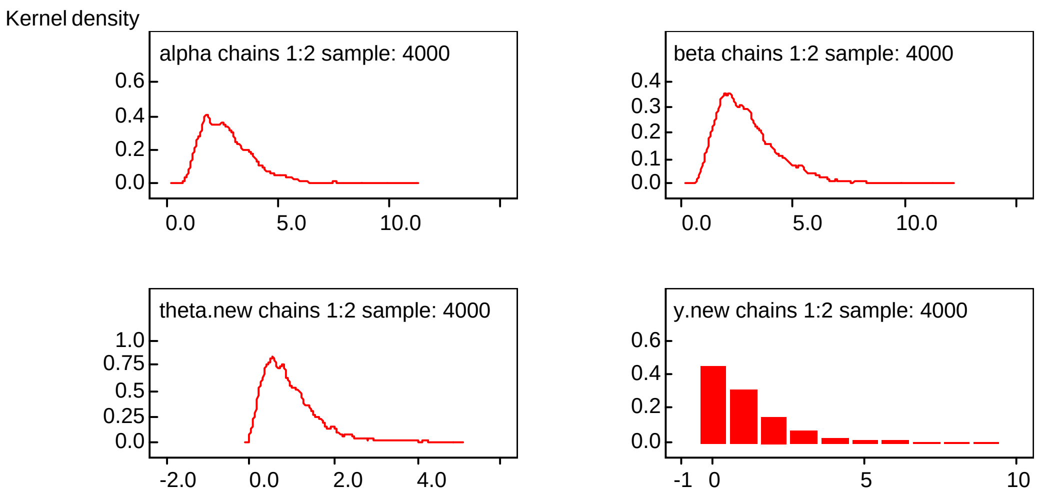

Figure 2. Posterior density functions and histogram for the main parameters of interest of the Poisson model with independent gamma rates of arrival

.

Dealing with a time series it seems natural to query whether the there is a time trend for the arrival rate of events. The next model (Model 3) deals exactly with this type of situation.

$$\begin{aligned}
&Y_i \left| \theta_i \right. \sim Pois(\theta_i) \\
&\theta_i = a + bi \\
&a \sim N(0, 0.0001) \quad \mathrm{b} \sim N(0, 0.0001)
\end{aligned} \tag{8}$$

The last line of the model specification acknowledges our lack of any prior information about the regression coefficients that are treated as random variables. The parameterisation of the normal distribution is in terms of precision, which is the inverse of variance. This is the way it is implemented in WinBugs and therefore a very small precision means a very large variance leading to a very flat normal distribution similar to an uniform distribution over a very large range. The joint posterior distribution of the parameters of interest, the regression coefficients *a* and *b* here, is

$$p(a, b \mid y) \propto p(y \mid a, b) p(a) p(b) \propto \left[\prod_{i=1}^{N} \frac{(a + bi)^{y_i} e^{-a+bi}}{y_i!} \right] \left[e^{-\frac{0.0001}{2} a^2} \right] \left[e^{-\frac{0.0001}{2} b^2} \right] \tag{9}$$

For MCMC sampling the conditional distributions are required. These are

$$\begin{aligned}
p(a \mid y, b) &\propto \left[\prod_{i=1}^{N} (a + bi)^{y_i} e^{-a+bi} \right] \left[e^{-\frac{0.0001}{2} a^2} \right] \\
p(b \mid y, a) &\propto \left[\prod_{i=1}^{N} (a + bi)^{y_i} e^{-a+bi} \right] \left[e^{-\frac{0.0001}{2} b^2} \right]
\end{aligned} \tag{10}$$

and it is obvious that powerful computational techniques are needed in order to simulate from these distributions. After checking that the simulated Markov chain looks stationary a large sample is used for inference. From Table 4 it can be seen that an *a posteriori* estimate is a = 0.451 and b = -0.024. Testing whether there is a time trend can be done by looking at the 95% credibility interval constructed from the 2.5% percentile and 97.5% percentiles. This is [-0.047, -0.0013] and it just misses out the zero value. Thus we can say that there is a time trend but only just. A 99% credibility interval may lead to the conclusion that there is no time trend.

Table 4. Summary statistics for the Poisson model with rates of arrival regressed on time

a)

Node	mean	sd	MC error	2.50%	median	97.50%
a	0.455	0.274	0.0046	-0.112	0.463	0.971
b	-0.024	0.012	1.94E-04	-0.047	-0.024	-0.001
mu_a	0.450	0.756	0.0073	-1.058	0.452	1.945
mu_b	-0.024	0.692	0.0046	-1.423	-0.021	1.352
V_a	2.993	1.714	0.0134	0.627	2.683	7.193
V_b	3.006	1.738	0.0125	0.627	2.669	7.307
Deviance	114.7	1.97	0.0244	112.8	114.1	120.1

b)

a	0.451	0.274	0.00439	-0.108	0.460	0.964
b	-0.024	0.011	1.88E-04	-0.047	-0.024	-0.0013
Deviance	114.7	2.022	0.02235	112.8	114.1	120.1

The entire posterior distributions of the parameters of interest are illustrated in Figure 3.

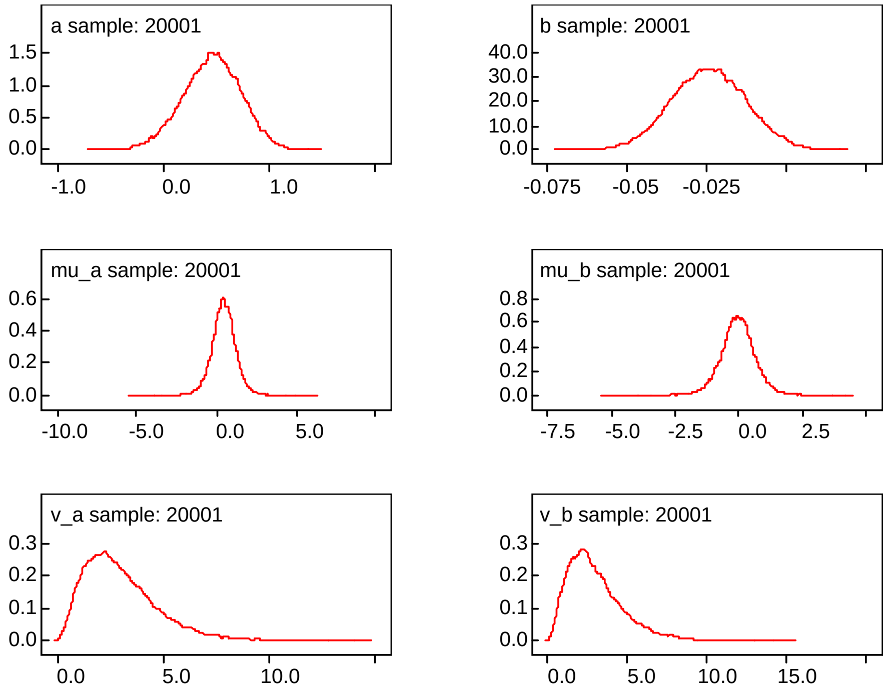

Figure 3. Posterior densities for the parameters of the Poisson model with rates of arrival regressed on time

The residual analysis points to a possible outlier, see Figures 4 and 5, the 1968 year when 6 events were observed. Maybe this was just a signal for the troubles to come in the financial markets of the 1970's. Apart from this data point the fit looks good and it seems that there are no problems with the (conditional) independence assumption or the distributional assumptions made.

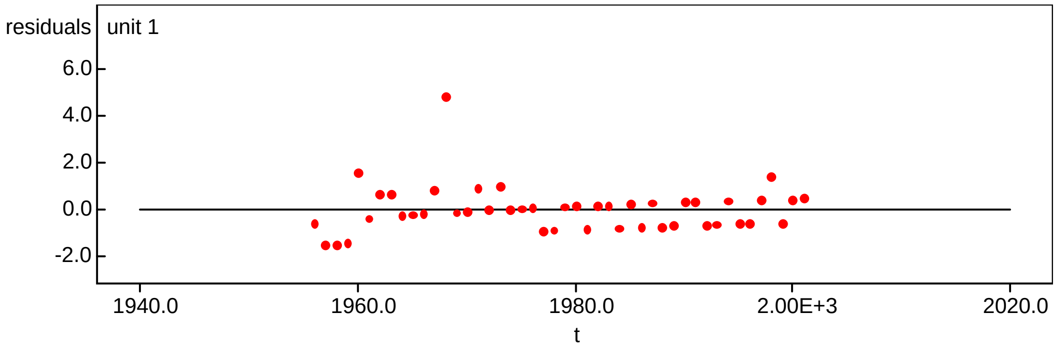

Figure 4. Residuals versus time for the hierarchical Poisson model with rates of arrival regressed on time and noninformative priors for the regression coefficients

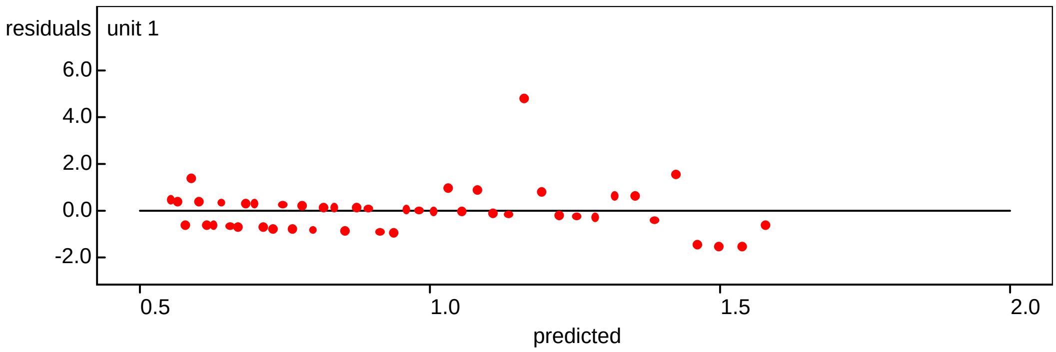

Figure 5. Residuals versus the expected values of observed data for the hierarchical Poisson model with rates of arrival regressed on time and noninformative priors for the regression coefficients

The usefulness of MCMC can also be seen when dealing with non-standard models. Here we explore also two such models applied also by Scollnik (2001) in an actuarial context. The first one employs the zero-inflated Poisson (ZIP) distribution given by

$$\Pr(X = x \mid \lambda, \mu) = \begin{cases} \lambda + (1-\lambda)e^{-\mu}, & x = 0 \\ (1-\lambda)\dfrac{\mu^x e^{-\mu}}{x!}, & x = 1,2,3,\ldots \end{cases} \tag{11}$$

with $0 < \lambda < 1$ and $\mu > 0$.

Table 5. MCMC inferential summary statistics for the ZIP model.

node	mean	sd	MC error	2.5%	median	97.5%
NLL	59.74	1.0710	0.01258	58.57	59.45	62.6
lambda	0.0778	0.0637	7.589E-4	0.0026	0.0627	0.2363
mu	1.0200	0.1644	0.00136	0.7319	1.0080	1.3770
p[1]	0.4159	0.0576	4.472E-4	0.3059	0.4146	0.5321
p[2]	0.3350	0.0246	3.008E-4	0.2734	0.3407	0.3654
p[3]	0.1701	0.0255	1.796E-4	0.1198	0.1705	0.2192
p[4]	0.0590	0.0176	1.307E-4	0.0296	0.0573	0.0978
p[5]	0.0157	0.0073	5.816E-5	0.0054	0.0145	0.0335
p[6]	0.0034	0.0022	1.877E-5	8.001E-4	0.0029	0.0092
p[7]	6.44E-4	5.541E-4	4.845E-6	9.721E-5	4.891E-4	0.0021
p[8]	1.237E-4	1.445E-4	1.298E-6	1.119E-5	8.036E-5	4.967E-4

The second is based on the generalised Poisson distribution (GP) as proposed by Consul (1989) and is described by

$$\Pr(X = x \mid \lambda, \mu) = \frac{\mu(\mu + \lambda x)^{x-1}}{x!} e^{-(\mu+\lambda x)}, \quad x = 0,1,2,... \qquad (12)$$

again with $0 < \lambda < 1$ and $\mu > 0$.

Table 6. MCMC inferential summary statistics for the GP model.

node	mean	sd	MC error	2.5%	median	97.5%
NLL	59.32	1.0130	0.0118	58.3900	59.0	62.020
lambda	0.0957	0.0701	7.382E-4	0.0043	0.0823	0.2636
mu	0.8871	0.1479	0.0011	0.6178	0.8800	1.2010
p[1]	0.4163	0.0606	4.487E-4	0.3010	0.4148	0.5392
p[2]	0.3284	0.0255	2.722E-4	0.2693	0.3325	0.3643
p[3]	0.1604	0.0256	1.952E-4	0.1110	0.1602	0.2117
p[4]	0.0622	0.0166	1.162E-4	0.0327	0.0611	0.0974
p[5]	0.0215	0.0093	7.694E-5	0.0074	0.0202	0.0432
p[6]	0.0072	0.0049	4.609E-5	0.0014	0.0059	0.0198
p[7]	0.0024	0.0025	2.52E-5	2.136E-4	0.0016	0.0094
p[8]	0.0015	0.0028	3.05E-5	3.395E-5	5.549E-4	0.0087

These models can be used to estimate the probability to see in the future 0 events, 1 event, 2 events and so on. From a practical perspective the probability to see a large number of events, although non-zero mathematically, is practically speaking zero. In other words the probability mass distribution function produced by these models has smaller and smaller probabilities in the right tail. Therefore, one may decide to consider only a sufficient number of probabilities[7]. Here only the first eight probabilities are reported. As can be easily seen from Tables 5 and 6 the probabilities to see 0, 1, 2,…, 7 events are decreasing as a sequence and the last probability reported is much less than 1%. Looking at the posterior distributions of these probabilities illustrated in Figures 6 and 8 it can be seen that the distribution gets very peaked around a value very close to zero. However, from a methodology point of view there is no problem in estimating more probabilities if needed.

The posterior distributions of the two parameters describing the ZIP model are illustrated in Figure 7. Visually an analyst may consider inferring a value of 1 for μ. For λ it is a different story as its distribution is heavily skewed. Standard statistics are reported in Table 5.

[7] Note that the first probability corresponds to the event that there is no event, the second probability to see one event and so on. Thus, the censoring is done at 7 events. The notation is somehow shifted with p[1] for the probability of no events and p[8] the probability of 7 events. This is a result of using WinBUGS.

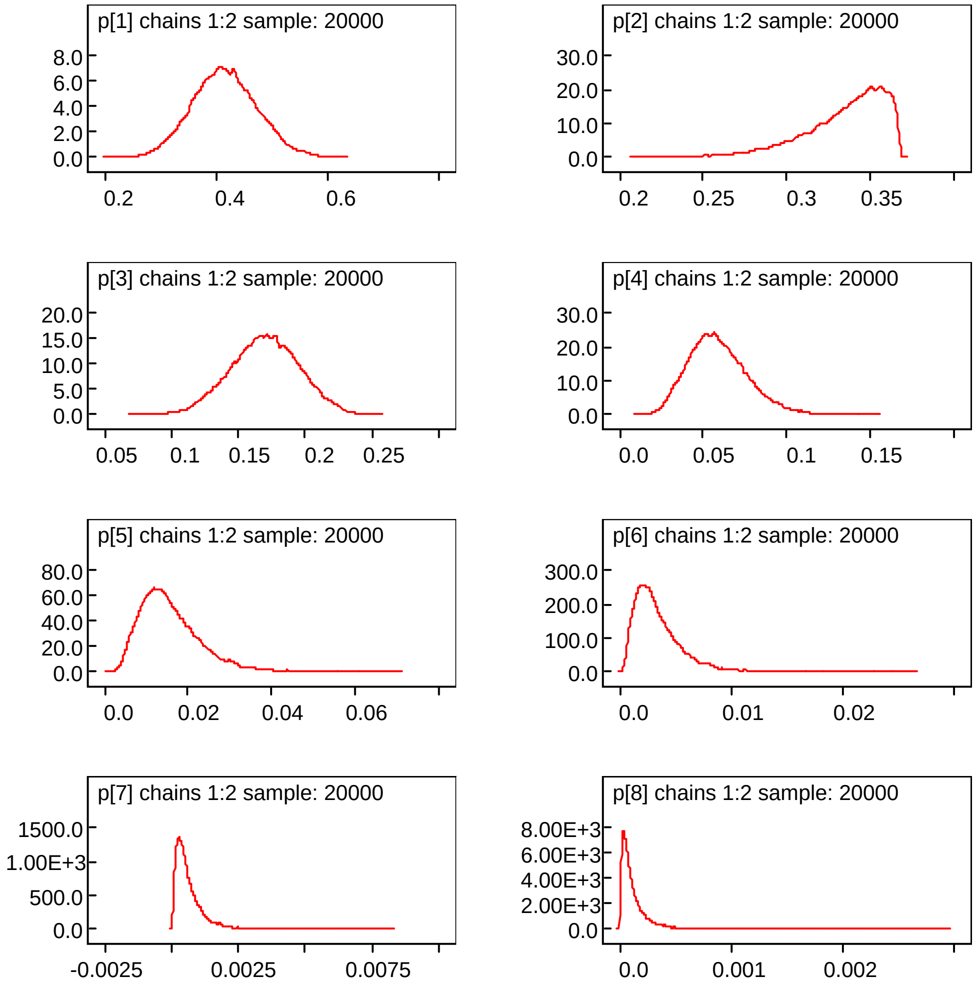

Figure 6. Posterior density plots for the probabilities of the ZIP model censored on the right at 7 events or more.

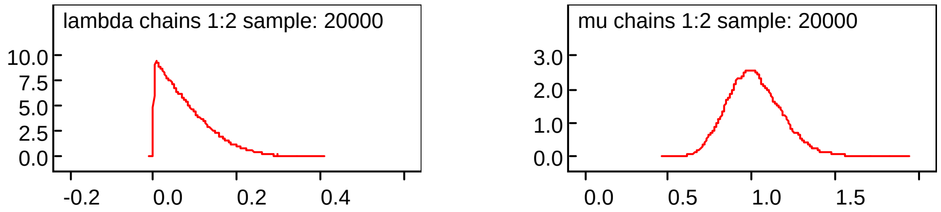

Figure 7. Posterior distributions of λ and μ the two parameters of the ZIP model.

Similarly, the posterior distributions of the two parameters describing the GP model are illustrated in Figure 9. Specific statistics such as the mean, median and credibility intervals are reported in Table 6.

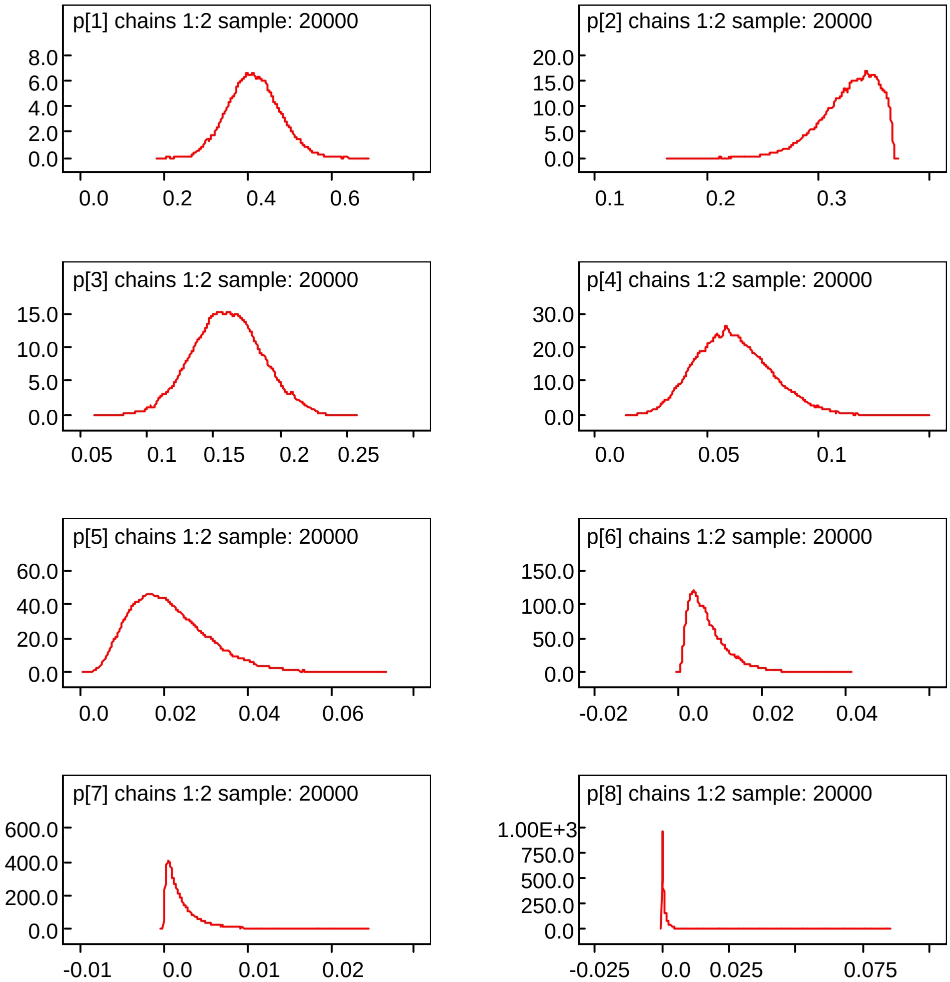

Figure 8. Posterior density plots for the probabilities of the GP model censored on the right at 7 events or more.

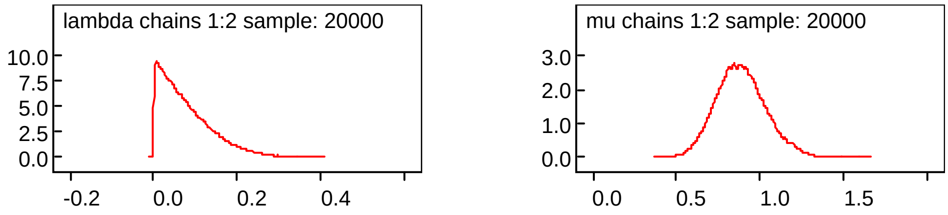

Figure 9. Posterior distributions of λ and μ the two parameters of the GP model.

It can be observed that while the distribution of μ is quite symmetric the distribution of λ is much skewed.

Both models predict that the most likely number of events to occur is zero. However the probability for this is less than half, about 41%, so there is enough mass probability attached to more than one event. For example, the probability to have one, two or three events is estimated using the posterior means with the ZIP model as

$$p[2] + p[3] + p[4] = 0.3350 + 0.1701 + 0.0590 = 0.5641 \tag{13}$$

and with the GP model as

$$p[2] + p[3] + p[4] = 0.3284 + 0.1604 + 0.0622 = 0.5510 \tag{14}$$

This type of inference would have been almost impossible to obtain by standard econometrics techniques. The methodology outlined here can provide various answers like that using the same MCMC output. This opens new possibilities of extracting valuable information from sparse data that can be used as further inputs in decision taking under uncertainty, policy making, pricing projects and products.

5. Conclusion

International political risk has received a renewed interest given the latest major world events. There is little literature focusing on this special issue and empirical investigations are difficult given the sparse nature of the data. We showed here how Markov Chain Monte Carlo (MCMC) modelling techniques can be usefully applied to quantify the evolution between 1956 and 2001 of international political risk.

The major tool is the arrival rate of political events. We have taken into consideration through our modelling the fact that political risk can arise from a wide range of sources, which are often mutually dependent.

Bayesian hierarchical models are fitted via Markov Chain Monte Carlo and a wide and interesting set of statistical inference is extracted. The approach presented here can be easily adapted to fit complex models in the same area. Here a zero-inflated model and a generalized Poisson model are also fitted on the same dataset.

References

Andersson, T., 1989, "Foreign direct investment in competing host countries", Doctoral Dissertation, Stockholm School of Economics.

Bouchet, M. H., E. Clark and B. Groslambert, (2003), *Country Risk Assessment: A Guide to Global Investment Strategy*, (London: John Wiley & Sons, Ltd).

Bekaert, G. (1995) Market Integration and Investment Barriers in Emerging Equity Markets, *World Bank Economic Review*, 9, no. 1, 75-107.

Bekaert, G. and Harvey, C. (1997), 'Emerging equity market volatility', *Journal of Financial Economics*, **43**, no. 1, January, 29-77.

Bittlingmayer, G., (1998), 'Output, Stock Volatility and Political Uncertainty in a Natural Experiment: Germany, 1880-1940', *Journal of Finance*, no. 6, December, 2243-2257.

Brewer T. L., (1981), Political Risk Assessment for Foreign Direct Investment Decisions: Better Methods for Better Results. *Columbia Journal of World Business*, Spring, 16 (1), 5-12.

Brooks, S P (1998) Markov chain Monte Carlo method and its application. *The Statistician*. **47**, 69-100.

Brooks, S P and Gelman, A., (1998), Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. **7** , 434-455.

Calvo, G. A., (1998), Varieties of Capital Market Crises, in G. A. Calvo and M. King, eds., *The Debt Burden and its Consequences for Monetary Policy*, (New York: Macmillan Press).

Calvo, G. A. and E. Mendoza, (2000), Rational Contagion and the Globalization of Securities Markets, Journal of International Economics, 51, pp. 79-113.

Carlin, B.P. and Louis, T.A. (1996) *Bayes and Empirical Bayes Methods for Data Analysis*, Monographs on Statistics and Applied Probability 69, Chapman & Hall, London.

Chan, Y. and Wei, J., (1996), 'Political Risk and Stock Price Volatility: the Case of Hong Kong', *Pacific Basin Finance Journal*, 259-275.

Clark, E., (1997), "Valuing political risk as an insurance policy", *Journal of International Money and Finance*, 16, pp. 477-490.

Clark, E. and R. Tunaru, (2003), “Quantification of political risk with multiple dependent sources”, *Journal of Economics and Finance*, vol. 27, no. 2, pp 125-135.

Consul, P.C. (1989) *Generalized Poisson Distributions: Properties and Applications*. New York: Marcel Dekker.

Cosset, J.C. and Suret, J.M. (1995), ‘Political Risk and the Benefits of International Portfolio Diversification’, *Journal of International Business Studies*, 2nd quarter, 301-318.

Cutler, D., Poterba, J. and Summers, L., (1989), ‘What Moves Stock Prices?’, *Journal of Portfolio Management*, **15**, 4-11.

Eaton, J. and M. Gersovitz, (1984), "A theory of expropriation and deviations from perfect capital mobility", *Economic Journal*, 94, pp. 16-40.

Eichengreen, B., A. Rose and C. Wyplosz, (1996), Contagious Currency Crises, NBER Working Paper no. 5681, (Cambridge MA, National Bureau of Economic Research).

Erb, C.B., Harvey, C.R. and Viskanta, T.E. (1996), ‘Political Risk, Economic Risk and Financial Risk’, *Financial Analysts Journal*, November-December.

Erb, C.B., Harvey, C.R. and Viskanta, T.E. (1995), ‘Country Risk and Global Equity Selection’, *The Journal of Portfolio Management*, Winter.

Feils, D.J. and Sabac, F.M. (2000), The Impact of Political Risk on the Foreign Direct Investment Decision: a Capital Budgeting Analysis. *Engineering Economist*, 45 (2), 129-43.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (1995) *Bayesian Data Analysis*, Chapman & Hall, London.

Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (eds) (1996) *Markov Chain Monte Carlo In Practice*, Chapman & Hall, London.

Haendel D, West G.T. and Meadow, R.G. (1975), Overseas Investment and Political Risk, Philadelphia: Foreign Policy Research Institute Monograph Series, No. 21, 1975.

Howell, L. D. and Chaddick B., (1994), Models of Political Risk for Foreign Investment and Trade: An Assessment of Three Approaches. *Columbia Journal of World Business*, Fall, 29 (3), 70-91.

Kaminsky, G. and C. M. Reinhart, (2000), On Crises, Contagion and Confusion, *Journal of International Economics*, 51, pp. 145-168.

Kim, H.Y. and Mei, J.P. (2001), 'What Makes the Stock Market Jump? An Analysis of Political Risk on Hong Kong Stock Returns', *Journal of International Money and Finance*, 20, 1003-1016.

Kobrin S. J., (1979), Political Risk: a Review and Reconsideration. *Journal of International Business Studies*, Spring/Summer, 10 (1), 67-80.

Meldrum DH, (2000), 'Country Risk and Foreign Direct Investment', *Business Economics*, Jan, 35 (1), 33-40.

Raff, Horst, "A model of expropriation with asymmetric information", *Journal of International Economics*, 33, (1992) pp. 245-265.

Robock, S. H., (1971), Political Risk: Identification and Assessment. *Columbia Journal of World Business*, Jul/Aug, 6 (4), 6-20.

Robock, Stefan H. and Simmonds, K. (1973). *International business and multinational enterprise*. R. Irwin, Homewood.

Root, F. (1973), Analysing political risks in International business, *in* A. Kapoor and P. Grub, eds., *Multinational enterprise in Transition*, Darwin Press, Princeton.

Roy, A. and Roy, P.G., (1994), Despite Past Debacles, Predicting Sovereign Risk Still Presents Problems, *Commercial Lending Review*, Summer, 9 (3), 92-95.

Scollnik, D.P.M. (2001) Actuarial Modelling with MCMC and BUGS, *North American Actuarial Journal*, 5 (2), 96-125.

Simon J.D., (1982), Political Risk Assessment: Past Trends And Future Prospects, *Columbia Journal of World Business*, Fall, 17 (3), 62-70.

Valdes, R. O., (1997), Emerging Markets Contagion: Evidence and Theory, Working Paper no. 7, (Santiago: Banco Central de Chile).

Appendix

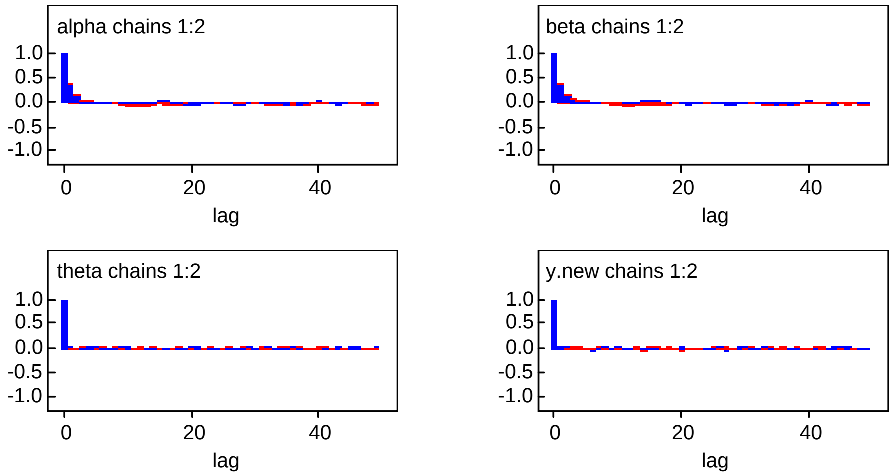

Figure A1. Autocorrelation plots for the main parameters of interest of the Poisson model with a single rate of arrival gamma distributed

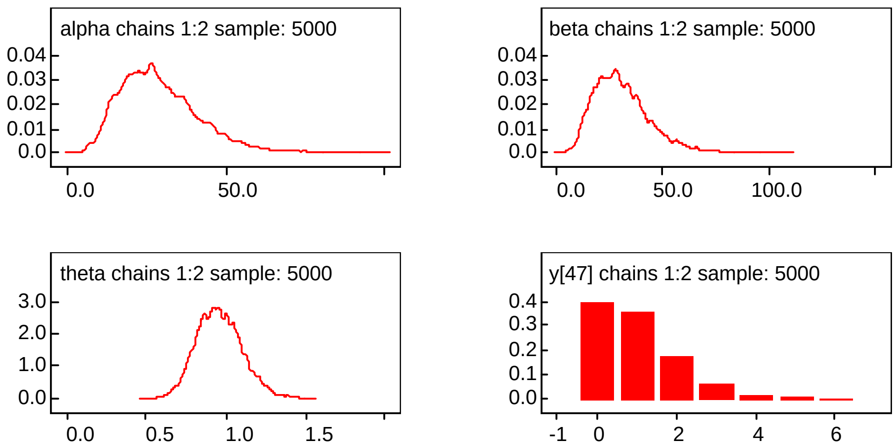

Figure A2. Posterior density functions and histogram for the main parameters of interest of the Poisson model with a single rate of arrival gamma distributed

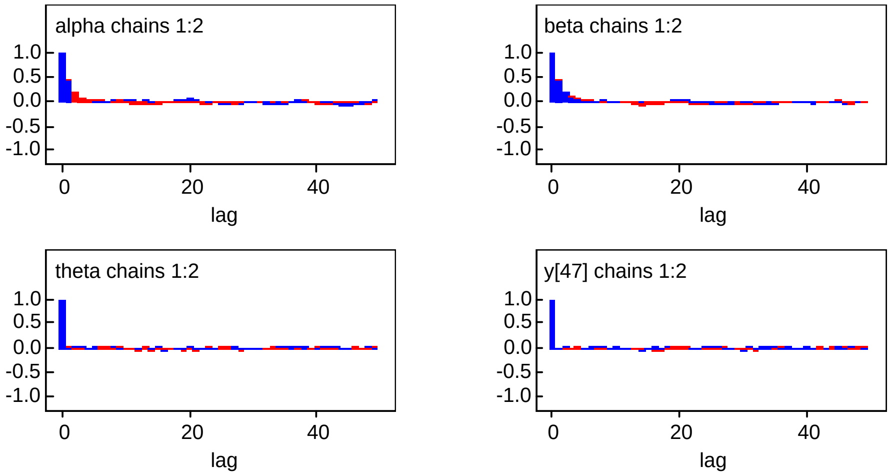

Figure A3. Autocorrelation plots for all parameters of interest of the Poisson model with a single rate of arrival gamma distributed

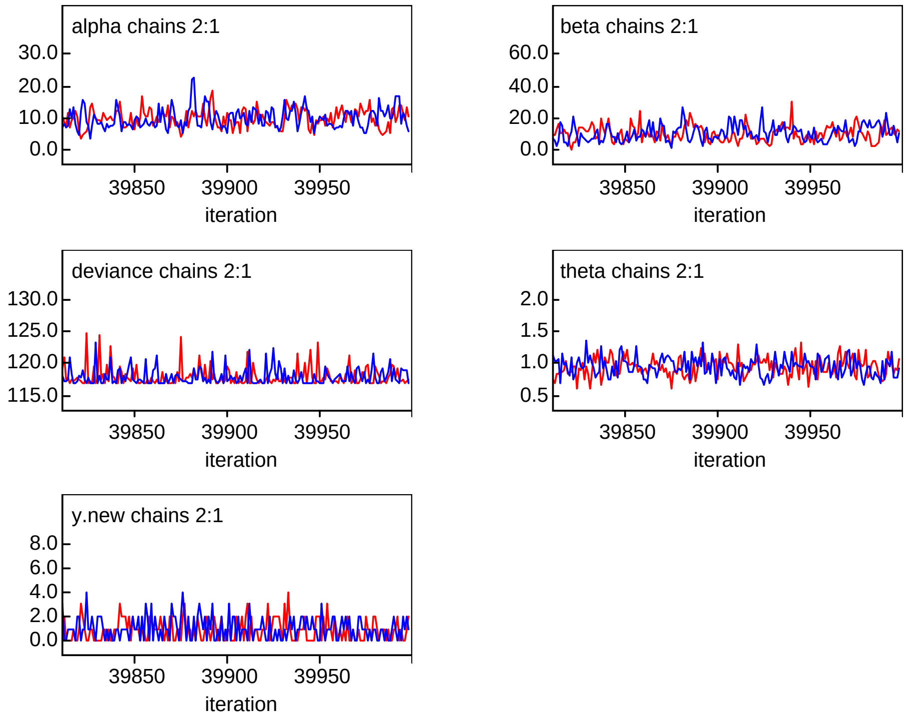

Figure A4. Trace plots of the simulated Markov chain for all parameters of interest and the deviance of the Poisson model with a single rate of arrival gamma distributed

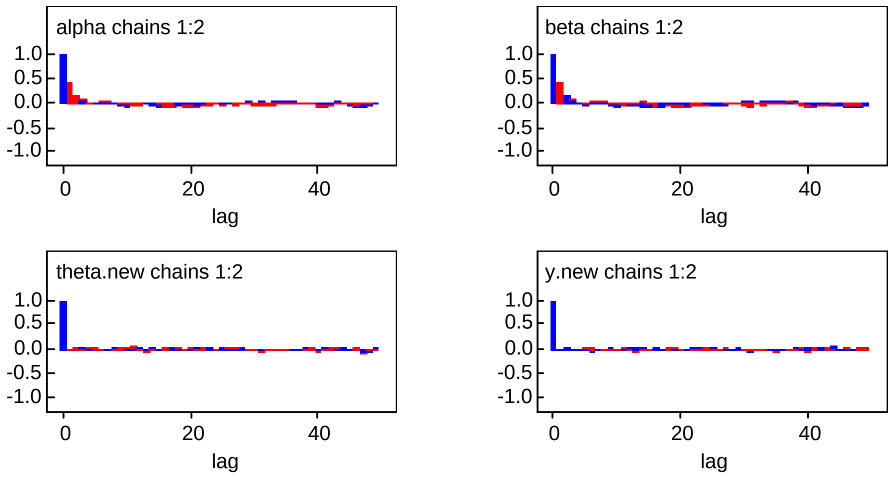

Figure A5. Autocorrelation plots for the parameters of interest of the Poisson model with independent gamma rates of arrival showing that the simulated Markov chain is stationary

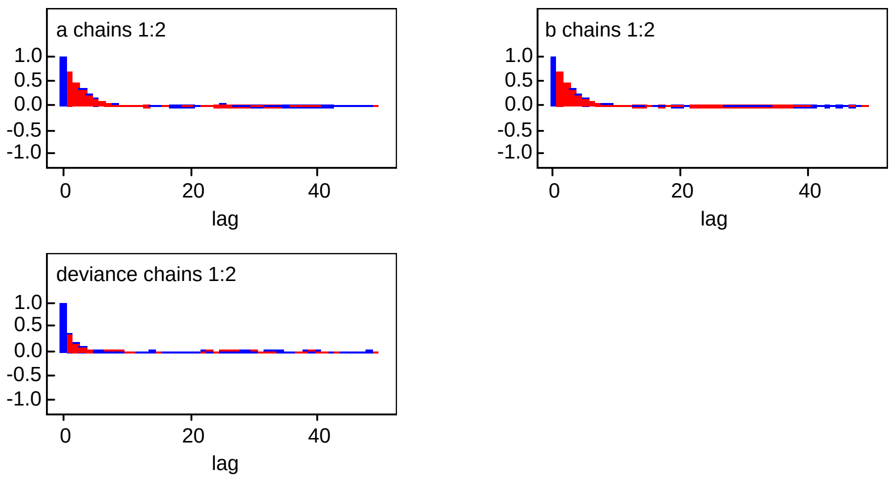

Figure A6. Autocorrelation plots for the parameters of the hierarchical Poisson model with rates of arrival regressed on time and noninformative priors for the regression coefficients