



---

# *Journal of Statistical Software*

MMMMMM YYYY, Volume VV, Issue II.

<http://www.jstatsoft.org/>

---

## MCMCpack: Markov chain Monte Carlo in R

Andrew D. Martin  
Washington University in St. Louis

Kevin M. Quinn  
Harvard University

Jong Hee Park  
University of Chicago

---

### Abstract

We introduce **MCMCpack** (Martin and Quinn 2007), an R package that contains functions to perform Bayesian inference using posterior simulation for a number of statistical models. In addition to code that can be used to fit commonly used models, **MCMCpack** also contains some useful utility functions, including some additional density functions and pseudo-random number generators for statistical distributions, a general purpose Metropolis sampling algorithm, and tools for visualization.

*Keywords:* Bayesian inference, Markov chain Monte Carlo, R.

---

## 1. Introduction

The Bayesian paradigm for statistical inference is appealing to researchers on both theoretical (Jeffreys 1998 [1961]; Bernardo and Smith 1994) and practical (Gelman, Carlin, Stern, and Rubin 2003) grounds. The interest in Bayesian methods in the social sciences is growing, and a number of researchers are using these approaches in substantive applications dealing with: deliberative bodies (Clinton, Jackman, and Rivers 2004; Martin and Quinn 2002), economic performance (Western 1998), income dynamics (Geweke and Keane 2000), legislative redistricting (Gelman and King 1990), mass voting (King, Rosen, and Tanner 1999), party competition (Quinn, Martin, and Whitford 1999; Quinn and Martin 2002), and social networks (Hoff, Raftery, and Handcock 2002; Hoff and Ward 2004).

Despite this interest, the social scientific community has yet to take full advantage of Bayesian techniques when approaching research problems. This is due primarily to two distinct problems. The first problem, which has been largely solved, was the inability to compute the high dimensional integrals necessary to characterize the posterior distribution for most models. To a large extent, this has been solved by the advent of Markov chain Monte Carlo (MCMC) methods (Tanner and Wong 1987; Gelfand and Smith 1990; Besag, Green, Higdon, and Mengersen 1995) and the dramatic increases in computing power over the past twenty years. For a comprehensive treatment of MCMC methods, see Robert and Casella (2004).

MCMC methods are widely considered the most important development in statistical computing in recent history. MCMC has allowed statisticians to fit essentially *any* probability model—including those not even considered a few years ago. Unfortunately, statisticians have been about the *only* people who have been willing, and able, to write the computer code necessary to use MCMC methods to fit probability models.

Which brings us to the second problem; namely, the lack of flexible, yet easy-to-use software for social scientists unwilling or unable to devote substantial time and energy writing custom software to fit models via MCMC. Since reasonably efficient MCMC algorithms exist to sample from the posterior distribution for most classes of models, developing software to meet the needs of social scientists is feasible.

**MCMCpack** (Martin and Quinn 2007) is a computational environment that puts Bayesian tools (particularly MCMC methods) into the hands of social science researchers so that they (like statisticians) can fit innovative models of their choosing. Just as the advent of general purpose statistical software (like SPSS and SAS) on mainframe and then personal computers led to the widespread adoption of statistical approaches in the social sciences, providing easy-to-use general purpose software to perform Bayesian inference will bring Bayesian methods into mainstream social science.

**MCMCpack** currently contains the eighteen statistical models: linear regression models (linear regression with Gaussian errors, a singular value decomposition regression, and regression for a censored dependent variable), discrete choice models (logistic regression, multinomial logistic regression, ordinal probit regression, and probit regression), measurement models (a one-dimensional IRT model, a  $k$ -dimensional IRT model, a  $k$ -dimensional ordinal factor model, a  $k$ -dimensional linear factor model, a  $k$ -dimensional mixed factor model, and a  $k$ -dimensional robust IRT model), a model for count data (a Poisson regression model), models for ecological inference (a hierarchical ecological inference model and a dynamic ecological inference model), and a time-series model for changepoint problems (a Poisson changepoint model). Many of these models, especially the measurement models, are otherwise intractable unless one uses MCMC.

The package also contains the density functions and pseudo-random number generators for the Dirichlet, inverse Gamma, inverse Wishart, Noncentral Hypergeometric, and Wishart distributions. These functions are particularly useful for visualizing prior distributions. Finally, **MCMCpack** contains a number of utility functions for creating graphs, reading and writing data to external files, creating `mcmc` objects, and manipulating variance-covariance matrices. The `coda` package is currently used for posterior visualization and summarization (Plummer, Best, Cowles, and Vines 2006).

The remainder of this paper is organized as follows. In Section 2 we discuss the package environment and features of **MCMCpack**. The following sections will show how model fitting functions in **MCMCpack** are implemented in a Gaussian linear model, and a Poisson changepoint model. We conclude with a discussion of possible **MCMCpack** future developments.

## 2. The MCMCpack Environment

We have chosen to make the R system for statistical computation and graphics the home environment for our software. R has a number of features that make it ideal for our purposes. It is open-source, free software that is distributed under the GNU GPL. It is an extremely

powerful programming environment that is designed for statistical data analysis and data visualization. R is very similar to the S language (Becker, Chambers, and Wilks 1988), and provides an easy-to-use interface to compiled C, C++ or FORTRAN code, as well as excellent facilities for debugging, profiling, and documenting native R programs. R is already the general purpose tool of choice for most applied statisticians and is well documented and supported (Venables and Ripley 2000, 2002; Venables, Smith, and R Development Core Team 2003; R Development Core Team 2006). Evidence of the move toward R among social scientists can be found in the growing number of texts designed for social science graduate students that explicitly advocate R usage (Fox 2002; Gill 2002), and the decision of the Inter-University Consortium for Political and Social Research to require students to use R as a means of integrating its most advanced courses in their summer training program (Anderson, Fox, Frankin, and Gill 2003). Over the last five years R has become the *lingua franca* of applied statisticians working in the social sciences.

## 2.1. Design Philosophy

In building the **MCMCpack** package we have attempted to adhere to a design philosophy that emphasizes: a) widespread availability; b) model-specific, computationally efficient MCMC algorithms; c) compiled C++ code to maximize computational speed; d) an easy-to-use, standardized model interface that is very similar to the standard R model fitting functions; and e) compatibility with existing R packages (such as **coda**) for convergence assessment, posterior summarization, and data visualization.

From a purely practical perspective, the most important design goal has been the implementation of MCMC algorithms that are model-specific. The major advantage of such an approach is that the sampling algorithms, being hand-crafted to particular classes of models, can be made dramatically more efficient than black box approaches, such as those found in the WinBUGS software, while remaining robust to poorly conditioned or unusual data. In addition to using reasonably computationally efficient sampling algorithms, the **MCMCpack** model fitting functions are also designed to be fast implementations of particular algorithms. To this end, nearly all of the actual MCMC sampling takes place in C++ code that is called from within R.

The model fitting functions in **MCMCpack** have been written to be as similar as possible to the corresponding R functions for classical estimation of the models in question. This largely eliminates the need to learn a specialized model syntax for anyone who is even a novice user of R. For instance, to fit a linear regression model in R via ordinary least squares one could use the following syntax:

```
reg.out <- lm(y~x1+x2+x3, data=mydata)
```

Bayesian fitting of the same model with an improper uniform prior on the coefficient vector and the default settings for the parameters governing the MCMC algorithm can be accomplished with **MCMCpack** using the following syntax:

```
Bayes.reg.out <- MCMCregress(y~x1+x2+x3, data=mydata)
```

It is our experience that such similarities greatly decrease the amount of time it takes to become a competent user of the **MCMCpack** package.

In addition, the **MCMCpack** model fitting functions are designed to be as similar to each other as possible. As a result, once a user becomes familiar with the basics of using one model fitting function she can quickly move on to confidently use other model fitting functions. For instance, to fit a probit model in which the  $\beta$  vector is assumed to have a multivariate normal prior with mean  $m$  and precision  $P$  one would use the **MCMCpack** syntax:

```
Bayes.probit.out <- MCMCprobit(y~x1+x2+x3, b0=m, B0=P, data=mydata)
```

To fit a Poisson regression model in which the  $\beta$  vector is assumed to have a multivariate normal prior with mean  $m$  and precision  $P$  one would use the **MCMCpack** syntax:

```
Bayes.poisson.out <- MCMCpoisson(y~x1+x2+x3, b0=m, B0=P, data=mydata)
```

Using R as the home environment for **MCMCpack** allows us to make use of a wide range of software for MCMC convergence assessment, posterior summarization, and data visualization [such as the **coda** package of [Plummer \*et al.\* \(2006\)](#)] that is already part of the R system. Output objects from all of the **MCMCpack** model fitting functions are formatted as **coda** **mcmc** objects, sometimes with additional attributes to allow for other types of analyses. This provides for a very easy-to-use interface between the two packages. For instance, to generate trace plots and density plots of the output in `Bayes.poisson.out` using the `plot` method, one would enter

```
plot(Bayes.poisson.out)
```

at the R command prompt. Similarly, to compute posterior means, standard deviations, quantiles, etc. using the `summary` method one would enter

```
summary(Bayes.poisson.out)
```

at the R command prompt. All other **coda** functions can be used in conjunction with **MCMCpack** with similar ease.

## 2.2. Comparison with Existing Software

Perhaps the most widely used piece of software for applied Bayesian inference is the WinBUGS package, distributed by the MRC Biostatistics Unit at Cambridge ([Spiegelhalter, Thomas, Best, and Lunn 2004](#)). WinBUGS is an important contribution, and is ideally suited for small models (both in the number of parameters and in the amount of data). WinBUGS is based on a set of general computational algorithms that can be used to estimate models specified using a model definition syntax. WinBUGS oftentimes does not exploit model-specific information for estimation. All models are estimated using general purpose black-box algorithms. For the most part, sampling is done parameter-by-parameter. There are two other implementations of the BUGS language: JAGS ([Plummer 2005](#)) and OpenBUGS ([Thomas, O'Hara, Ligges, and Sturtz 2006](#)).

While the BUGS language is useful for fitting many types of models, it has its limitations. The generic estimation engine used in WinBUGS is quite flexible, but because it does not exploit model-specific information, it is often inefficient, and sometimes ineffective. It is well

known that WinBUGS has trouble with algorithms that require sampling from truncated distributions, such as the [Albert and Chib \(1993\)](#) algorithm for the binary probit model, and algorithms for fitting ordinal data models ([Cowles 1996](#); [Johnson and Albert 1999](#)). The parameter-by-parameter approach is computationally inefficient, both in terms of computation time and clock time ([Liu, Wong, and Kong 1994](#)). While we think **MCMCpack** has definite advantages over WinBUGS for many users, we emphasize that we view BUGS and **MCMCpack** as complementary tools for the applied researcher. In particular, the greater flexibility of the BUGS language is perfect for users who need to build and fit custom probability models.

Other general purpose software packages with designs similar to that of **MCMCpack** exist. One is the Bayesian Analysis, Computation, and Communication (**BACC**) package discussed in [Geweke \(1999\)](#). The development of BACC was supported by the National Science Foundation, and the newest release was made available on June 3, 2003. **BACC** is available as a computational engine for MATLAB, R, S-PLUS, GAUSS, and the UNIX or Windows console. **BACC** shares a design component with **MCMCpack**, as each algorithm is coded in compiled code for each individual model. Another general purpose software package is the recently released **bayesm** package by [Rossi and McCulloch \(2006\)](#). This package is geared to analyses in marketing and applied econometrics. Both **BACC** and **bayesm** use an interface where users provide appropriate scalars, vectors, and matrices for data and priors, and where posterior density samples are returned as matrices, requiring the user to perform additional computation to summarize results, assess model fit, and the like. **MCMCpack**, on the other hand, has a more R-like interface. We also view these packages as complementary to **MCMCpack**, as they provide estimation engines for useful models.

In addition to these packages that offer a comprehensive set of tools, there are a number of other available R packages that perform Bayesian inference for single classes of models. For a list of available packages, see the Bayesian Inference Task View on the Comprehensive R Archive Network ([Park, Martin, and Quinn 2006](#)).

### 2.3. An Overview of **MCMCpack** Features

In addition to the model fitting functions mentioned above, **MCMCpack** has numerous features aimed at both researchers and instructors. For the sake of space we do not comprehensively demonstrate this functionality here. Rather, we highlight three features that may be of interest. We encourage interested users to consult the documentation for further details.

The `MCMCmetrop1R` function allows users to sample from a user-defined continuous density using a random walk Metropolis algorithm with a multivariate normal proposal distribution. This can be used to explore a posterior (or log-posterior) distribution, as well as any other target distribution of interest. See [Gelman \*et al.\* \(2003\)](#) and [Robert and Casella \(2004\)](#) for details of the random walk Metropolis algorithm. The sampler itself is coded in C++, so the iterations run quite fast. Users only have to provide the target density as an R function.

**MCMCpack** is sometimes used on large problems where parallel computation might afford some advantages. While **MCMCpack** does not currently support parallelization within the Monte Carlo loop, for some problems it is useful to perform embarrassingly parallel simulations, e.g., sampling from the posterior density of the same model with twenty-five simultaneous processes. Doing so requires the availability of a random number generator that provides independent substreams of pseudo-random digits across processes. The default pseudo-random

number generator in **MCMCpack** is the Mersenne twister (Matsumoto and Nishimura 1998). **MCMCpack** provides a default seed, which can be changed in any model-fitting function. The package also provides the generator of L’Ecuyer, Simard, Chen, and Kelton (2002), which provides independent substreams suitable for parallel computation. Again, the seed can be changed by the user, who only needs to provide a unique substream number to ensure independence. This generator works just fine with the powerful **snow** package (Tierney, Rossini, Li, and Sevcikova 2006).

To enhance the usability of **MCMCpack** in the classroom, the package also contains a number of “toy” instructional models. While the use of Monte Carlo simulation is not necessary to characterize the posterior density of these models, we find it useful to introduce Monte Carlo methods in contexts where analytical results are readily available. To wit, **MCMCpack** provides **MCmultinomdirichlet** (Monte Carlo simulation from a multinomial likelihood with a Dirichlet prior), **MCnormalnormal** (Monte Carlo simulation from a normal likelihood with known variance and a normal prior), and **MCpoissongamma** (Monte Carlo simulation from a Poisson likelihood with a gamma prior).

### 3. An Example of Bayesian Linear Regression

In this section, we look at how to fit a Bayesian linear model using **MCMCregress**. We use Wilkerson’s (1999) analysis of “killer” amendments in U.S. Congress as an example. A killer amendment causes a bill, that would pass absent the amendment, to fail. According to Wilkerson, a typical example of a killer amendment occurred in 1996,

...when the House was asked to approve an international shipbuilding trade agreement that had taken more than a decade to negotiate. During the debate, a majority party member, Herbert Bateman (R-VA), proposed adding 30 months of “transition assistance” for U.S. shipbuilders to the agreement... Jennifer Dunn (R-WA) argued “Some individuals argue that no agreement is better than this agreement. In reality, if the Bateman amendment is adopted, that is exactly what we would have: No agreement” (Congressional Record, June 13 1996, H6321)... Ignoring Dunn’s warning, the House passed the amendment (278-149) and voted for the bill (325-100). Recognizing that the shipbuilding agreement was dead on arrival... , the Senate never took it up. (pp. 544-545)

Thus, in order to achieve the preferred final result (an international shipbuilding trade agreement), legislative majorities are often forced to vote against an amendment which seems closer to their preferences (the Bateman amendment) than an original bill. Doing so is a classic example of strategic voting; i.e., voting for a less preferred choice at an early stage of a voting procedure to ultimately get a better outcome in the end.

Using a linear regression model estimated with ordinary least squares, Wilkerson shows that killer amendments are rare and hence legislative majorities seldom face situations that require strategic voting. To investigate the existence of strategic voting, Wilkerson relies upon Poole and Rosenthal’s (1997) measures that capture the fit of actual roll call voting positions to estimated unidimensional legislative preferences (NOMINATE). Among them, we focus on proportional reduction in error (PRE), which compares NOMINATE’s success at predicting roll call voting positions with the success rate of simply assuming that everyone votes with

the majority:

$$PRE = \frac{(\text{Minority vote} - \text{NOMINATE classification errors})}{(\text{Minority vote})}.$$

A high value of PRE for a particular roll call implies that the NOMINATE is accurately classifying the votes on that roll call compared to the prediction based on the majority voting.

Wilkerson identifies five types of killer amendments and includes them as covariates: major weakening, strengthening, political cover, minor weakening, and new issue. Wilkerson also considers the sponsor. Amendments proposed by a member of the majority party are much more likely to be viewed as sincere attempts to improve the content of the legislation, rather than deliberate efforts to sabotage it. Lastly, Wilkerson controls for the chamber of introduction, in case there are chamber effects on sincerity of voting. The model to be estimated is:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

With priors:

$$b \sim \mathcal{N}_K(\mathbf{b}_0, \mathbf{B}_0^{-1}) \quad \sigma^2 \sim \mathcal{IG}(c_0/2, d_0/2)$$

To simulate from the posterior distribution of the model using `MCMCregress` we load Wilkerson's data and call `MCMCregress`.

```
> library(MCMCpack)
> load("killamdt.Rda")
> wilkerson <- MCMCregress(APRE1 ~ STRENGTH + COVER + WEAKMIN +
+   NEWISSUE + SPONSOR + CHAMBER, data = killamdt, b0 = 0,
+   B0 = 0.1, c0 = 2, d0 = 0.11, marginal.likelihood = "Chib95")
```

Note that the output from the call to `MCMCregress` has been placed in an `mcmc` object called `wilkerson`. By default, `MCMCregress` uses a noninformative prior for the coefficient parameters. To facilitate model comparison, here we use a weakly informative prior. The prior for  $\boldsymbol{\beta}$  is specified using the `b0` and `B0` arguments, where `b0` is the prior mean and `B0` is the prior precision of  $\boldsymbol{\beta}$ . The prior for  $\sigma^2$  is governed by two parameters `c0` and `d0`. We use a weakly informative prior (`c0=2` and `d0=0.11`) with a mean equal to the marginal variance of the APRE1 (0.11).

Before examining the results of our model, it is necessary to examine some diagnostics to assess whether the Markov chain has converged to its stationary distribution. The `plot` method for an `mcmc` object produces a trace plot and a density plot for each parameter. To save space, we only report these plots for first two parameters in Figure 1. There are also a number of standard convergence diagnostics implemented in `coda` one should check before using the posterior density sample for inference.

The `coda summary` method can be used to summarize an the `mcmc` object produced by `MCMCregress`.

```
> summary(wilkerson)
```

```
Iterations = 1001:11000
Thinning interval = 1
```

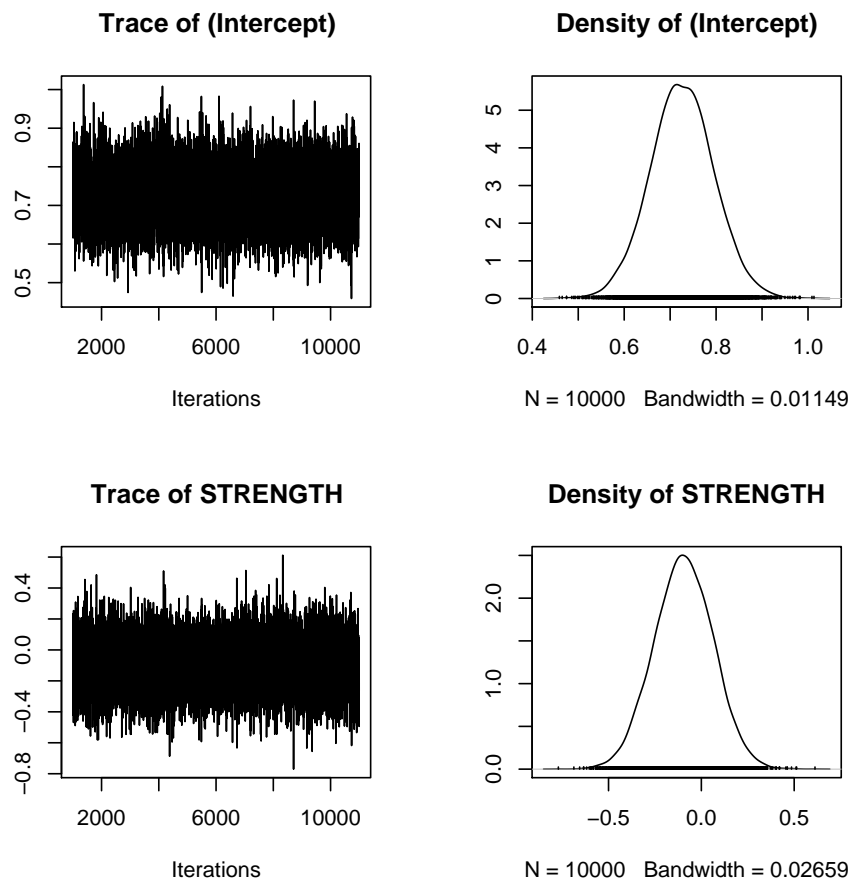


Figure 1: Trace and density plots for two coefficients in the `MCMCregress` example.

Number of chains = 1  
 Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,  
 plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
(Intercept)	0.72631	0.06934	0.0006934	0.0007085
STRENGTH	-0.09573	0.15828	0.0015828	0.0015134
COVER	-0.47822	0.15122	0.0015122	0.0015871
WEAKMIN	-0.28246	0.07989	0.0007989	0.0008861
NEWISSUE	-0.49685	0.11002	0.0011002	0.0009653
SPONSOR	-0.14988	0.07425	0.0007425	0.0007096
CHAMBER	0.18924	0.07666	0.0007666	0.0007562
sigma2	0.07493	0.01387	0.0001387	0.0001681

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
(Intercept)	0.59017	0.68042	0.72596	0.77207	0.864040
STRENGTH	-0.40370	-0.20210	-0.09570	0.01191	0.211530
COVER	-0.77342	-0.57890	-0.48005	-0.37593	-0.184722
WEAKMIN	-0.43952	-0.33595	-0.28221	-0.22865	-0.127368
NEWISSUE	-0.71283	-0.57028	-0.49699	-0.42332	-0.281150
SPONSOR	-0.29571	-0.19903	-0.14914	-0.10004	-0.005032
CHAMBER	0.03841	0.13804	0.18912	0.24031	0.342175
sigma2	0.05232	0.06512	0.07343	0.08294	0.106865

One of main advantage of Bayesian methods is that posterior inference is quite straightforward using MCMC output since the methods provide direct samples of parameters of interest. In this example, Wilkerson's primary quantity of interest is in whether PRE is higher for minority-sponsored major weakening amendments than for majority-sponsored major weakening amendments. Wilkerson considers a higher PRE for minority-sponsored major weakening amendments than for majority-sponsored major weakening amendments as major evidence for killer amendments.

We can answer this question using our posterior samples. In our data set, SPONSOR is coded 1 if majority-sponsored amendments and WEAKMIN is coded 1 if minor weakening amendments, and 0 otherwise. Thus, the probability that PRE is higher for minority-sponsored major weakening amendments than for majority-sponsored major weakening amendments can be shown by comparing a posterior distribution of Constant with a posterior distribution of Constant + SPONSOR.

Figure 2 demonstrates that minority-sponsored major weakening amendments are better predicted by NOMINATE scores than majority-sponsored major weakening amendments. This result is consistent with Wilkerson's finding that members of the majority party are more likely to give serious consideration to the substance and strategic implications of amendments offered by fellow party members.

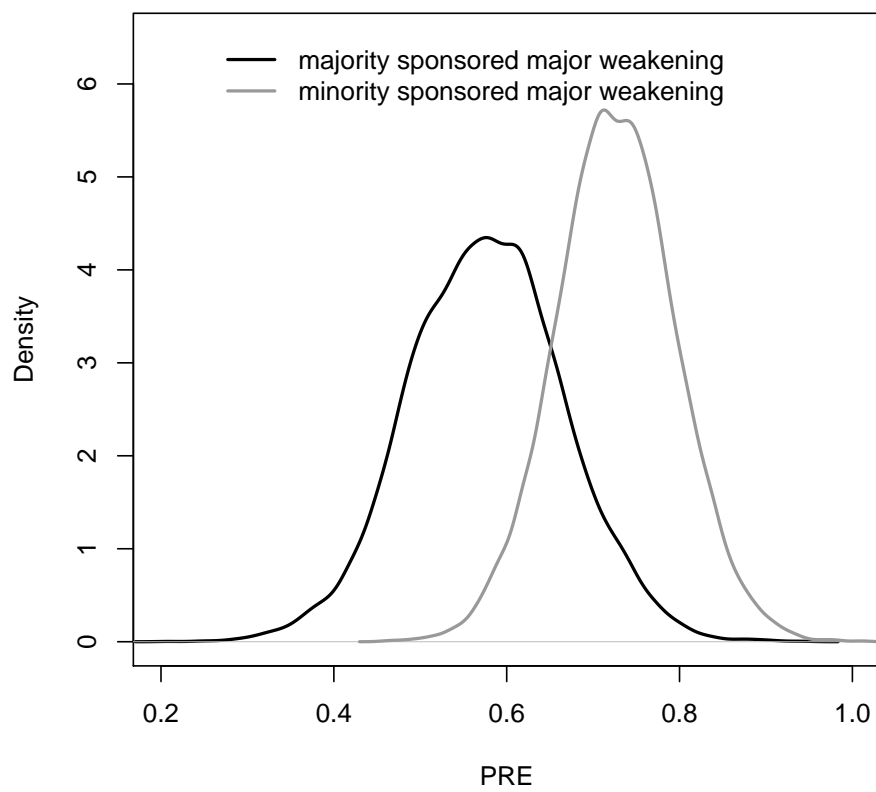


Figure 2: PRE for various types of amendments and sponsors.

	Mean	St. Dev	Lower	Upper
Constant	0.7263	0.0693	0.5902	0.8640
STRENGTH	-0.0957	0.1583	-0.4037	0.2115
COVER	-0.4782	0.1512	-0.7734	-0.1847
WEAKMIN	-0.2825	0.0799	-0.4395	-0.1274
NEWISSUE	-0.4968	0.1100	-0.7128	-0.2811
SPONSOR	-0.1499	0.0743	-0.2957	-0.0050
CHAMBER	0.1892	0.0767	0.0384	0.3422
$\sigma^2$	0.0749	0.0139	0.0523	0.1069

Table 1: Analysis of killer amendments (Wilkerson 1999). The dependent variable is proportional reduction in error (PRE) when the minority vote is used as benchmark. Lower and Upper indicate central 95 percent Bayesian credible intervals.

**MCMCpack** also provides a tool for Bayesian model comparison (for some of its models): Bayes factors. Bayes factors can be used to compare any model of the same data. Suppose that the observed data  $y$  could have been generated under one of two models  $\mathcal{A}_1$  and  $\mathcal{A}_2$ . A natural thing to ask from the Bayesian perspective is: “What is the posterior probability that  $\mathcal{A}_1$  is true assuming either  $\mathcal{A}_1$  or  $\mathcal{A}_2$  is true?” Using Bayes theorem we can write:

$$\Pr(\mathcal{A}_1|y) = \frac{p(y|\mathcal{A}_1) \Pr(\mathcal{A}_1)}{p(y|\mathcal{A}_1) \Pr(\mathcal{A}_1) + p(y|\mathcal{A}_2) \Pr(\mathcal{A}_2)}$$

It is instructive to look at the posterior odds in favor of one model (say  $\mathcal{A}_1$ ):

$$\frac{\Pr(\mathcal{A}_1|y)}{\Pr(\mathcal{A}_2|y)} = \frac{p(y|\mathcal{A}_1)}{p(y|\mathcal{A}_2)} \times \frac{\Pr(\mathcal{A}_1)}{\Pr(\mathcal{A}_2)}.$$

What this means is that if we want to move from the prior odds in favor of  $\mathcal{A}_1$  to the posterior odds in favor of  $\mathcal{A}_1$  we simply multiply the prior odds by:

$$B_{12} = \frac{p(y|\mathcal{A}_1)}{p(y|\mathcal{A}_2)}$$

which is called the *Bayes factor* for  $\mathcal{A}_1$  relative to  $\mathcal{A}_2$ .  $B_{12}$  provides a measure of the strength of evidence for  $\mathcal{A}_1$  over  $\mathcal{A}_2$ . For a thorough discussion of Bayes factors see Kass and Raftery (1995).

The key ingredients necessary for calculating Bayes factors (and posterior model probabilities) are the marginal likelihoods. There are several efficient methods (Tierney and Kadane 1986; Chib 1995; Chib and Jeliazkov 2001) for these calculations. **MCMCregress** computes marginal likelihoods using either the Laplace approximation (Tierney and Kadane 1986) or Chib’s (1995) method to compute the marginal likelihoods. Using Chib’s method, we compare three alternative models with Wilkerson’s model using comparable prior distributions.

```
> model1 <- MCMCregress(APRE1 ~ STRENGTH + COVER, data = killamdt,
+   mcmc = 10000, b0 = 0, BO = 0.1, c0 = 2, d0 = 0.11,
+   marginal.likelihood = "Chib95")
> model2 <- MCMCregress(APRE1 ~ STRENGTH + COVER + WEAKMIN +
+   NEWISSUE, data = killamdt, mcmc = 10000, b0 = 0,
```

```

+      B0 = 0.1, c0 = 2, d0 = 0.11, marginal.likelihood = "Chib95")
> model3 <- MCMCregress(APRE1 ~ SPONSOR + CHAMBER, data = killamdt,
+      mcmc = 10000, b0 = 0, B0 = 0.1, c0 = 2, d0 = 0.11,
+      marginal.likelihood = "Chib95")
> BF <- BayesFactor(model1, model2, model3, wilkerson)
> summary(BF)

```

The output from `BayesFactor` shows a matrix of Bayes Factors for all pairs of models specified in the argument. Table 2 shows that there is positive evidence or better to support model 2 over all other models considered. That is, information about the sponsorship and the chamber of introduction does not help explain the variation of PRE given the information on five types of killer amendments. We report a posterior summary of model 2 in Table 3 for interested readers.

	model 1	model 2	model 3	Wilkerson
model 1	0.00	-3.11	-1.39	-0.52
model 2	3.10	0.00	1.71	2.58
model 3	1.39	-1.72	0.00	0.86
Wilkerson	0.52	-2.58	-0.86	0.00

Table 2: A matrix of Bayes factors (on the natural log scale) for the [Wilkerson \(1999\)](#) replication. The dependent variable is PRE.

	Mean	St. Dev	Lower	Upper
(Intercept)	0.7888	0.0591	0.6720	0.9042
STRENGTH	-0.2681	0.1572	-0.5765	0.0385
COVER	-0.4741	0.1583	-0.7849	-0.1676
WEAKMIN	-0.3059	0.0839	-0.4687	-0.1411
NEWISSUE	-0.4518	0.1060	-0.6607	-0.2404
$\sigma^2$	0.0848	0.0157	0.0593	0.1210

Table 3: Killer amendments results from model 2. Lower and Upper indicate central 95 percent Bayesian credible intervals.

## 4. A Bayesian Poisson Changepoint Model Example

We show how to fit a Poisson changepoint model using `MCMCpoissonChangepoint` in this section. Generally speaking, a changepoint model is a time series model in which data are generated from multiple regimes or states. Main quantities of interests in the changpeoint analysis are the number, the timing, and the magnitude of regime changes.

The algorithm of `MCMCpoissonChangepoint` is based on [Chib \(1998\)](#). Chib transforms the changepoint problem to a special type of mixture model in which observations are assumed to be drawn from latent states. The movement of latent states which follows the first order

Markov process can be summarized by a transition matrix:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & 0 & \dots & 0 \\ 0 & p_{22} & p_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & p_{mm} & p_{m,m+1} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

where  $p_{ij} = Pr(s_t = j | s_{t-1} = i)$  is the probability of changing from state  $i$  at time  $t - 1$  to state  $j$  at time  $t$ .

If we let the prior distribution of the transition probabilities be Beta, and the prior mean of a Poisson model be Gamma, we can write a Bayesian Poisson changepoint model as follows.

$$\begin{aligned} y_t &\sim \text{Poisson}(\lambda_i), \quad i = 1, \dots, k \\ \lambda_i &\sim \mathcal{G}(c_0, d_0), \quad i = 1, \dots, k \\ p_{ii} &\sim \text{Beta}(\alpha, \beta), \quad i = 1, \dots, k. \end{aligned}$$

The sampling algorithm of `MCMCpoissonChangepoint` consists of the simulation of model specific parameters, latent state variables, and transition probabilities.

$$\begin{aligned} \mathbf{S}_T &| \mathbf{Y}_T, \boldsymbol{\lambda}, \mathbf{P} \\ \boldsymbol{\lambda} &| \mathbf{Y}_T, \mathbf{S}_T, \mathbf{P} \\ \mathbf{P} &| \mathbf{S}_T \end{aligned}$$

In this section, we employ `MCMCpoissonChangepoint` to revisit the empirical test of cyclical theories of great-power war by [Pollins \(1996\)](#). Pollins identifies four core frameworks in the study of cyclical patterns of international conflicts based on “(1) the relative importance of *global political versus economic subsystems*, (2) the *dating of the phases* they identify in the cycle, and (3) predictions regarding the *level of conflict* within particular phases of their respective cycles” ([Pollins 1996](#), 106). These four core frameworks are a theory of hegemony by [Gilpin \(1981\)](#), a theory of the world system by [Wallerstein \(1983\)](#), the long wave in the global economy by [Goldstein \(1988\)](#), and the world leadership cycle by [Modelski and Thompson \(1988\)](#).

These four theories give different weights on the role of global economic order and global political order in explaining the patterns of international conflicts. For example, while the theory of Long Wave and the theory of world system argue that the growth and stagnation of global economy ([Goldstein 1988](#)) or economic hegemon in the system ([Wallerstein 1983](#)) generate different levels of international conflicts, [Modelski and Thompson \(1988\)](#) and [Gilpin \(1981\)](#) consider the rise and decline of global leadership or a hegemonic power to play a central role in producing international conflicts. These four theories are further diverged in their predictions on the specific cycle phases of international conflicts as shown in Table ??.

Following [Pollins \(1996\)](#), we use the annual number of military conflicts from the Militarized Interstate Disputes (MID) data ([Jones, Bremer, and Singer 1996](#)). We can use this event count data on the level of international conflicts to detect the number and the timing of cycle phases in international conflicts.

To implement `MCMCpoissonChangepoint`, a user needs to specify data, the number of states ( $m$ ), and hyperparameters of  $\lambda$  ( $c_0$  and  $d_0$ ). [Chib \(1998\)](#)’s method identifies the number

Framework	Cycle Phases			
Goldstein	Stagnation	Rebirth	Expansion	War
	1816	1831	1849	1861
	1873	1885	1894	1911
Level of Conflict	1921	1934	1941	1969
	Low Conflict	High Conflict	High Conflict	Highest Conflict
Wallerstein	Victory	Maturity	Decline	Ascent
	1816	1850	1873	1897
	1913	1945	1967	
Level of Conflict	Scattered Conflict	Lowest Conflict	Competitiveness	Highest Conflict
Modelski and Thompson	World Power	Delegitimation	Deconcentration	Global War
	1816	1850	1874	1915
	1946	1974		
Level of Conflict	Lowest Conflict	High Conflict	Low Conflict	Highest Conflict
Gilpin	U.K. Hegemony	No Hegemony	U.S. Hegemony	
	1815	1873	1945	
Level of Conflict	Low Conflict	High Conflict	Low Conflict	

Table 4: Starting Dates of Cycle Phases and Predicted Level of Conflict. Source: [Pollins \(1996\)](#)

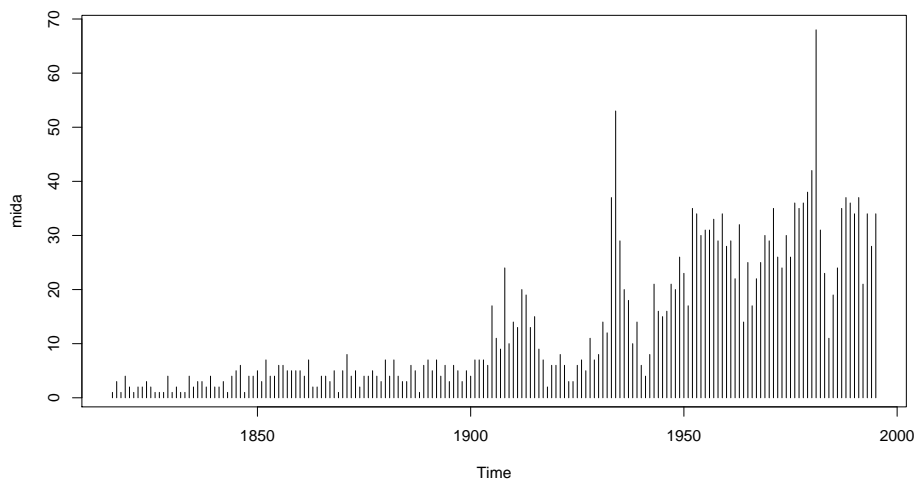


Figure 3: Militarized interstate disputes data.

of changepoints by Bayesian model comparison using Bayes factors, which are the ratio of marginal likelihoods. Thus, **MCMCpack** does not provide default values for hyperparameters of  $\lambda$  ( $c0$  and  $d0$ ). The user must specify them in order to facilitate marginal likelihood comparison. In this specific example, we set  $c0 = 13$  and  $d0 = 1$  since the mean of MID is about 13. The following R commands implement eight Poisson changepoint models for the Militarized Interstate Disputes data.

```
> model1 <- MCMCpoissonChangepoint(mida, m = 1, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model2 <- MCMCpoissonChangepoint(mida, m = 2, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model3 <- MCMCpoissonChangepoint(mida, m = 3, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model4 <- MCMCpoissonChangepoint(mida, m = 4, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model5 <- MCMCpoissonChangepoint(mida, m = 5, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model6 <- MCMCpoissonChangepoint(mida, m = 6, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model7 <- MCMCpoissonChangepoint(mida, m = 7, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
> model8 <- MCMCpoissonChangepoint(mida, m = 8, c0 = 13,
+   d0 = 1, marginal.likelihood = c("Chib95"))
```

The results of the model comparison are summarized in Table 5. The ratio of marginal likelihoods shows that the sixth changepoint model is favored over the alternatives. Using Jeffrey's rule, which provides guidelines as to the strength of a Bayes factor, we can conclude that there is positive evidence for the sixth changepoint model over the alternatives considered.

	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8
model 1	0.0	-60.5	-80.0	-93.3	-96.0	-111.61	-109.02	-106.46
model 2	60.5	0.0	-19.5	-32.9	-35.6	-51.12	-48.53	-45.96
model 3	80.0	19.5	0.0	-13.3	-16.0	-31.59	-29.00	-26.44
model 4	93.3	32.9	13.3	0.0	-2.7	-18.26	-15.67	-13.11
model 5	96.0	35.6	16.0	2.7	0.0	-15.56	-12.97	-10.41
model 6	111.6	51.1	31.6	18.3	15.6	0.00	2.59	5.15
model 7	109.0	48.5	29.0	15.7	13.0	-2.59	0.00	2.56
model 8	106.5	46.0	26.4	13.1	10.4	-5.15	-2.56	0.00

Table 5: A matrix of Bayes factors (on the natural log scale) for the Poisson changepoint analysis of militarized interstate disputes militarized interstate disputes between 1816 and 1995.

`plotState()` generates a plot of posterior state probabilities and `plotChangepoint()` plots of posterior probabilities of six changepoints as shown in Figures 4 and 5, respectively. Both graphs indicate that critical shifts in the level of international conflicts are concentrated in the first half of the twentieth century and cyclical patterns predicted by the four theoretical models are not supported by the evidence. However, the picture provided by the Poisson

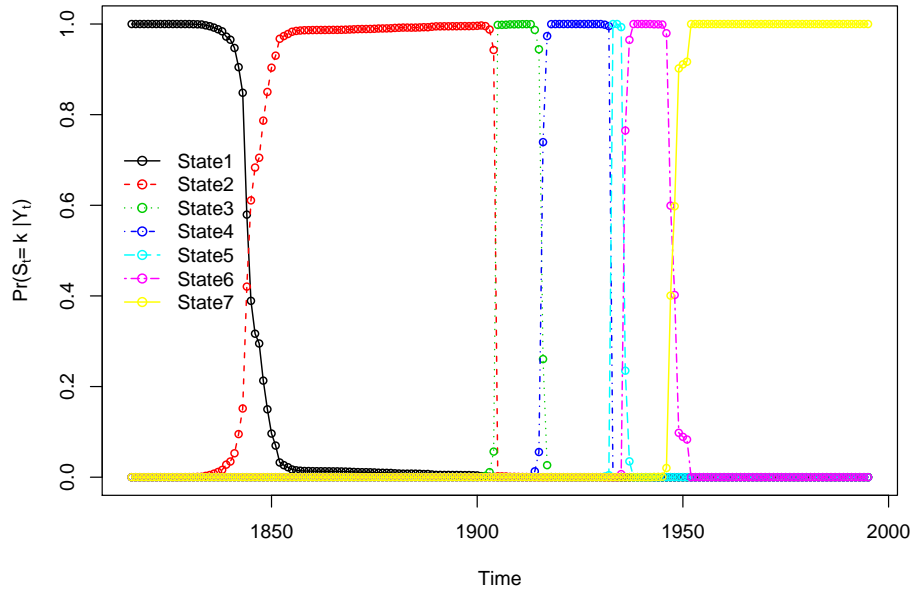


Figure 4: Posterior probabilities of states from the Poisson changepoint analysis of militarized interstate disputes militarized interstate disputes between 1816 and 1995.

changepoint analysis is closest to Gilpin’s theory of hegemony which predicts high levels of conflicts in the absence of a world hegemon between 1873 and 1945.

## 5. Conclusion

**MCMCpack** remains a work in progress. While it is difficult to commit to specific future development efforts, we hope to implement additional models others will find useful in their research and teaching. The ability to estimate marginal likelihoods and compute Bayes factors is a new feature in **MCMCpack**, and is only currently available for a handful of models. We hope to implement this for a wider range of models in the future. Of course, since **MCMCpack** is open source, others can build on this code-base to implement models of their choice.

In the next few years Jong Hee Park will develop a number of models to be used for Bayesian time series analysis, including changepoint models and state space models. He intends to expand some of these models to include covariates, and to provide various methods to enhance usability for this class of models. Finally, while the **coda** package provides a great deal of power for convergence checking and posterior density summarization, it does not provide adequate methods for formal model assessment, including prior and predictive checks, Bayes factors, and Bayesian model averaging (Hoeting, Madigan, Raftery, and Volinsky 1999). Our hope is develop software to make these tools as easy to use as possible.

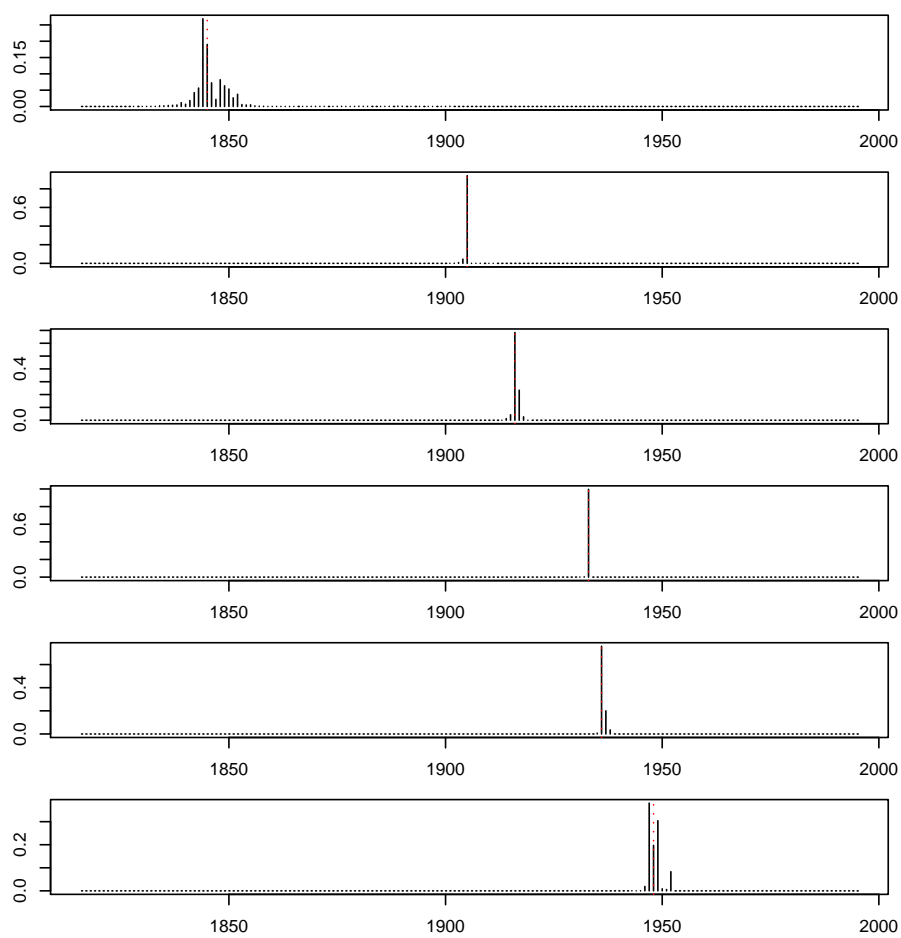


Figure 5: Posterior probabilities of changepoints: Dotted lines indicate expected changepoints: Local means for each regime are 3.8, 15.0, 7.06, 34.8, 10.0, 19.4, and 30.2, sequentially.

## 6. Acknowledgements

We have received a great deal of support for the **MCMCpack** project. First, we thank all of those in the R community who have used our software, found bugs, and provided suggestions, feedback, and patches. Second, we would like to thank all of the research assistants who have worked with us at Harvard University and Washington University over the years on this project: Lucy Barnes, Matthew Fasman, Ben Goodrich, Steve Haptonstahl, Kate Jensen, Laura Keys, Dan Pemstein, and Ellie Powell. Finally, we gratefully acknowledge financial support from: the U.S. National Science Foundation, Program in Methodology, Measurement, and Statistics, Grants SES-0350646 and SES-0350613, the Institute for Quantitative Social Sciences (<http://iq.harvard.edu>) at Harvard University, the Weidenbaum Center on the Economy, Government, and Public Policy (<http://wc.wustl.edu>) at Washington University, and the Center for Empirical Research in the Law (<http://cerl.wustl.edu>) at Washington University. Neither the National Science Foundation, Washington University, or Harvard University bear any responsibility for the content of this package. In addition, Quinn thanks the Center for Advanced Study in the Behavioral Sciences for its hospitality and support.

This work is licensed under a Creative Commons Attribution-No Derivative Works 3.0 License.

## References

- Albert JH, Chib S (1993). “Bayesian Analysis of Binary and Polychotomous Response Data.” *Journal of the American Statistical Association*, **88**, 669–679.
- Anderson B, Fox J, Frankin C, Gill J (2003). “The Times They R Achanging: The State of Advanced Methodology Curriculum at ICPSR.” *The Political Methodologist*, **11**, 22–24.
- Becker RA, Chambers JM, Wilks AR (1988). *The New S Language*. Chapman & Hall, London.
- Bernardo J, Smith AFM (1994). *Bayesian Theory*. Wiley, New York.
- Besag J, Green P, Higdon D, Mengersen K (1995). “Bayesian Computation and Stochastic Systems (with discussion).” *Statistical Science*, **10**, 3–66.
- Chib S (1995). “Marginal Likelihood From the Gibbs Output.” *Journal of the American Statistical Association*, **90**, 1313–1321.
- Chib S (1998). “Estimation and Comparison of Multiple Change-Point Models.” *Journal of Econometrics*, **86**(2), 221–241.
- Chib S, Jeliazkov I (2001). “Marginal Likelihood From the Metropolis Hastings Output.” *Journal of the American Statistical Association*, **96**, 270–281.
- Clinton J, Jackman S, Rivers D (2004). “The Statistical Analysis of Roll Call Data.” *American Political Science Review*, **98**, 355–370.
- Cowles MK (1996). “Accelerating Monte Carlo Markov Chain Convergence for Cumulative-link Generalized Linear Models.” *Statistics and Computing*, **6**, 101–110.
- Fox J (2002). *An R and S-Plus Companion to Applied Regression*. Sage Publications, Thousand Oaks, CA.

- Gelfand AE, Smith AFM (1990). “Sampling-Based Approaches to Calculating Marginal Densities.” *Journal of the American Statistical Association*, **85**, 398–409.
- Gelman A, Carlin JB, Stern HS, Rubin DB (2003). *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, second edition.
- Gelman A, King G (1990). “Estimating the Electoral Consequences of Legislative Redistricting.” *Journal of the American Statistical Association*, **85**, 274–282.
- Geweke J (1999). “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication.” *Econometric Reviews*, **18**, 1–126.
- Geweke J, Keane M (2000). “An Empirical Analysis of Income Dynamics Among Men in the PSID: 1968-1989.” *Journal of Econometrics*, **96**, 293–356.
- Gill J (2002). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Chapman & Hall, Boca Raton.
- Gilpin R (1981). *War and Change in World Politics*. Cambridge University Press, Cambridge.
- Goldstein J (1988). *Long Cycles: Prosperity and War in the Modern Age*. Yale University Press, New Haven.
- Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999). “Bayesian Model Averaging: A Tutorial.” *Statistical Science*, **14**, 382–417.
- Hoff PD, Raftery AE, Handcock MS (2002). “Latent Space Approaches to Social Network Analysis.” *Journal of the American Statistical Association*, **97**, 1090–1098.
- Hoff PD, Ward MD (2004). “Modeling Dependencies in International Relations Networks.” *Political Analysis*, **12**, 160–175.
- Jeffreys SH (1998 [1961]). *Theory of Probability*. Clarendon Press, Oxford, third edition.
- Johnson VE, Albert JH (1999). *Ordinal Data Modeling*. Springer, New York.
- Jones DM, Bremer SA, Singer JD (1996). “Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns.” *Conflict Management and Peace Science*, **15**, 163–213.
- Kass RE, Raftery AE (1995). “Bayes Factors.” *Journal of the American Statistical Association*, **90**, 773–795.
- King G, Rosen O, Tanner MA (1999). “Binomial-beta Hierarchical Models for Ecological Inference.” *Sociological Research and Methods*, **28**, 61–90.
- L’Ecuyer P, Simard R, Chen E, Kelton W (2002). “An Object-Oriented Random-Number Package With Many Long Streams and Substreams.” *Operations Research*, **50**, 1073–1075.
- Liu JS, Wong WH, Kong A (1994). “Covariance Structure of the Gibbs Sampler with Applications to the Comparisons of Estimators and Augmentation Schemes.” *Journal of the American Statistical Association*, **89**, 958–966.

- Martin AD, Quinn KM (2002). “Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999.” *Political Analysis*, **10**, 134–153.
- Martin AD, Quinn KM (2007). *MCMCpack: Markov chain Monte Carlo (MCMC) Package*. R package version 0.8-1, URL <http://mcmcpack.wustl.edu>.
- Matsumoto M, Nishimura T (1998). “Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator.” *ACM Transactions on Modeling and Computer Simulation*, **8**, 3–30.
- Modelski G, Thompson WR (1988). *Seapower in Global Politics*. University of Washington Press, Seattle.
- Park JH, Martin AD, Quinn KM (2006). “CRAN Task View: Bayesian Inference.” The Comprehensive R Network (<http://cran.r-project.org>).
- Plummer M (2005). *JAGS: Just Another Gibbs Sampler*. Version 0.8, URL <http://www-fis.iarc.fr/~martyn/software/jags/>.
- Plummer M, Best N, Cowles K, Vines K (2006). “CODA: Convergence Diagnosis and Output Analysis for MCMC.” *R News*, **6**, 7–11.
- Pollins BM (1996). “Global Political Order, Economic Change, and Armed Conflict: Coevolving Systems and the Use of Force.” *American Political Science Review*, **90**, 103–117.
- Poole KT, Rosenthal H (1997). *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press, Oxford.
- Quinn KM, Martin AD (2002). “An Integrated Computational Model of Electoral Competition.” *Statistical Science*, **17**, 405–419.
- Quinn KM, Martin AD, Whitford AB (1999). “Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models.” *American Journal of Political Science*, **43**, 1231–1247.
- R Development Core Team (2006). *R Language Definition, Version 2.4.1*. Comprehensive R Archive Network (<http://cran.r-project.org>).
- Robert C, Casella G (2004). *Monte Carlo Statistical Methods*. Springer, New York, second edition.
- Rossi P, McCulloch R (2006). *bayesm: Bayesian Inference for Marketing/Micro-econometrics*. R package version 2.0-8, URL <http://gsbwww.uchicago.edu/fac/peter.rossi/research/bsm.html>.
- Spiegelhalter D, Thomas A, Best N, Lunn D (2004). *WinBUGS*. Version 1.4.1, URL <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/>.
- Tanner MA, Wong W (1987). “The Calculation of Posterior Distributions by Data Augmentation.” *Journal of the American Statistical Association*, **82**, 528–550.
- Thomas A, O’Hara B, Ligges U, Sturtz S (2006). “Making BUGS Open.” *R News*, **6**, 12–17. URL <http://mathstat.helsinki.fi/openbugs/>.

- Tierney L, Kadane JB (1986). “Accurate Approximation for Posterior Moments and Marginal Densities.” *Journal of the American Statistical Association*, **81**, 82–86.
- Tierney L, Rossini AJ, Li N, Sevcikova H (2006). *snow: Simple Network of Workstations*. R package version 0.2-2.
- Venables WN, Ripley BD (2000). *S Programming*. Springer, New York.
- Venables WN, Ripley BD (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition.
- Venables WN, Smith DM, R Development Core Team (2003). *An Introduction to R: Notes on R: A Programming Environment for Data Analysis and Graphics, Version 1.7.1*. Comprehensive R Archive Network (<http://cran.r-project.org>).
- Wallerstein I (1983). “Three Instances of Hegemony in the History of the World Economy.” *International Journal of Comparative Sociology*, **24**, 100–108.
- Western B (1998). “Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach.” *American Journal of Political Science*, **42**, 1233–1259.
- Wilkerson JD (1999). “‘Killer’ Amendments in Congress.” *American Political Science Review*, **93**, 535–552.

**Affiliation:**

Andrew D. Martin  
Washington University School of Law  
Campus Box 1120  
One Brookings Drive  
St. Louis, MO 63130  
E-mail: [admartin@wustl.edu](mailto:admartin@wustl.edu)  
URL: <http://adm.wustl.edu>

Kevin M. Quinn  
The Institute for Quantitative Social Science  
Harvard University  
1737 Cambridge Street  
Cambridge, MA 02138  
E-mail: [kevin\\_quinn@harvard.edu](mailto:kevin_quinn@harvard.edu)  
URL: <http://www.people.fas.harvard.edu/~kquinn/>

Jong Hee Park  
Department of Political Science  
University of Chicago  
5828 S. University Avenue  
Chicago, IL 60637  
E-mail: [jhp@uchicago.edu](mailto:jhp@uchicago.edu)  
URL: <http://home.uchicago.edu/~jhp/>