Preface

Bayesian analysis has become increasingly popular in recent years and proved useful in many areas, including economics, medicine, science, and the social sciences. Much of this growth has been fueled not by philosophical concerns about the nature of statistical analysis or by a genuine desire to use prior information but by very practical considerations; the fact is that many complex models are much easier to handle within a Bayesian framework.

The chief limiting factor in adopting Bayesian methods is the availability of software. Thanks to programs such as Stata, most frequentist statistical methods are easy to implement, but Bayesian analyses require their own specialist software. Following theoretical developments in the 1980s, a major step forward in Bayesian computing came about with the creation of the BUGS project, which provided flexible, free software for fitting Bayesian models. The original BUGS program was eventually replaced by WinBUGS, which in turn has given way to an open-source version called OpenBUGS. Unfortunately, while these programs will fit a model, they are not particularly user friendly, and they do not offer the many facilities for data handling, exploratory statistics, and graphics that Stata users take for granted. Consequently, even investigators familiar with WinBUGS tend to do most of their work in traditional statistics packages and only switch to WinBUGS to fit the model.

Thompson, Palmer, and Moreno (2006) describe a suite of Stata ado-files, all beginning with the prefix wb, that help Stata users integrate WinBUGS into their analyses. The idea is that users should be able to store their data in Stata and analyze them in the usual way. Then when they want to fit a Bayesian model, they should be able to prepare it, send it to WinBUGS, and read back the results without ever leaving Stata. The results provided by WinBUGS or OpenBUGS consist of a very large file of simulations, so other programs are provided to help the user read those results into Stata, check the analysis, and summarize the findings. The commands that operate on the results of the WinBUGS run work equally well with Markov chain Monte Carlo (MCMC) simulations produced by any other software, so they are not restricted to use with output from WinBUGS.

Since 2006, OpenBUGS has become available, the original wb ado-files have been extended by adding extra options, and further programs have been added to the collection. New and old versions of the commands can be distinguished between those required to run WinBUGS or OpenBUGS and those used to investigate the resulting MCMC simulations; the former have been collected together as the new commands beginning with the letters wbs, and the latter now form the new commands beginning with the letters mcmc. Thus the updated version of wbrun that runs WinBUGS from within Stata is called wbsrun, and the updated version of wbtrace, the program for drawing a trace or history plot, is called mcmctrace.

In this book, I describe the updated commands and introduce other programs for running Bayesian analyses that do not need WinBUGS or OpenBUGS. Chapter 3 presents a set of new programs collected under the names beginning with the letters mhs for running Metropolis–Hastings samplers, and chapter 4 introduces programs beginning with the letters gbs for running Gibbs samplers. These programs can create MCMC simulations and fit Bayesian models without using WinBUGS. For multiparameter models, it is convenient to have a housekeeping program that cycles through the parameters and stores the MCMC updates. This job is performed by the program mcmcrun, which oversees the fitting process by calling any user-specified combination of the mhs and gbs samplers or even the user's own samplers.

Creating MCMC simulations independently of WinBUGS serves two purposes: first, it can be used as a teaching aid to demonstrate what happens when a Bayesian model is fit; second, it offers an alternative practical method by which users can tackle real problems. The only limitation of using Stata on its own is the time taken by simulation-based Bayesian model-fitting algorithms. While the samplers with command names beginning mhs and gbs work well on small- or moderately sized problems, efficient programming is essential when there are more than about 10 parameters in the model or when the dataset is very large. This means either using Stata's matrix language, Mata, or exporting the problem to WinBUGS; these two approaches are broadly equivalent in terms of speed and are introduced in chapters 7 and 8. Users who are only interested in using WinBUGS or OpenBUGS can skip chapters 2, 3, 4, 7, and 12, though in doing so, they will miss out on the explanation of how MCMC samplers work.

The author's blog can be found at http://staffblogs.le.ac.uk/bayeswithstata/. It is dedicated to the discussion of the use of Stata for Bayesian analysis and to describing future developments of the ado-files introduced in this book.

Downloading the user-written commands

The user-written commands discussed in this book can be downloaded from within Stata using the following commands:

```
. net from http://www.stata-press.com/data/bas/
. net install bas
```

To download the do-file to install the user-written Mata code, type

```
. net get bas
```

and then run the mcmclibrarymata.do file.