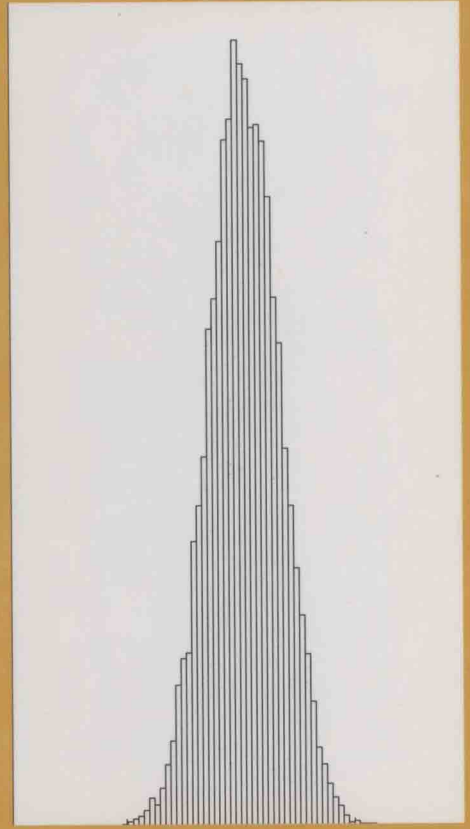
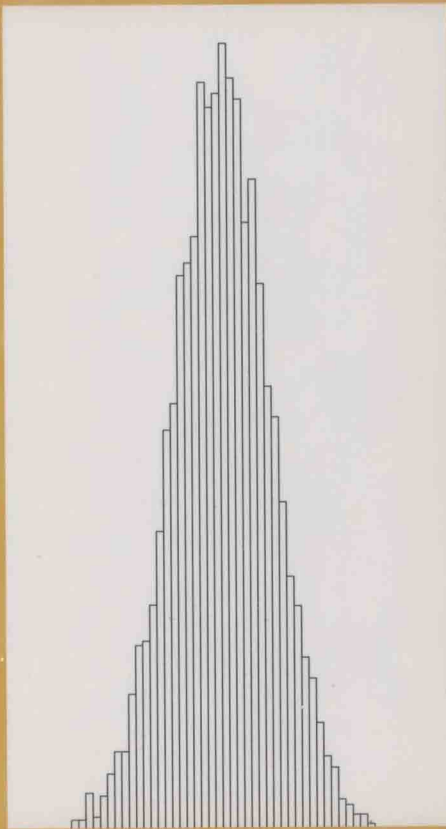


INTRODUCTION TO  
**Bayesian Econometrics**



EDWARD GREENBERG

CAMBRIDGE

# Introduction to Bayesian Econometrics

EDWARD GREENBERG

*Washington University, St. Louis*



CAMBRIDGE UNIVERSITY PRESS  
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press  
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org  
Information on this title: www.cambridge.org/9780521858717

© Edward Greenberg 2008

This publication is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without  
the written permission of Cambridge University Press.

First published 2008

Printed in the United States of America

*A catalog record for this publication is available from the British Library.*

*Library of Congress Cataloging in Publication Data*

Greenberg, Edward, 1936-  
Introduction to Bayesian econometrics / Edward Greenberg.  
p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-521-85871-7 (hardback)

ISBN-10: 0-521-85871-2 (hardback)

1. Econometrics. 2. Bayesian statistical decision theory. I. Title.

HB139.G732 2008

330.01'519542--dc22 2007024630

ISBN 978-0-521-85871-7 hardback

Cambridge University Press has no responsibility for  
the persistence or accuracy of URLs for external or  
third-party Internet Web sites referred to in this publication  
and does not guarantee that any content on such  
Web sites is, or will remain, accurate or appropriate.

## Introduction to Bayesian Econometrics

This concise textbook is an introduction to econometrics from the Bayesian viewpoint. It begins with an explanation of the basic ideas of subjective probability and shows how subjective probabilities must obey the usual rules of probability to ensure coherency. It then turns to the definitions of the likelihood function, prior distributions, and posterior distributions. It explains how posterior distributions are the basis for inference and explores their basic properties. The Bernoulli distribution is used as a simple example. Various methods of specifying prior distributions are considered, with special emphasis on subject-matter considerations and exchangeability. The regression model is examined to show how analytical methods may fail in the derivation of marginal posterior distributions, which leads to an explanation of classical and Markov chain Monte Carlo (MCMC) methods of simulation. The latter is preceded by a brief introduction to Markov chains. The remainder of the book is concerned with applications of the theory to important models that are used in economics, political science, biostatistics, and other applied fields. These include the linear regression model and extensions to Tobit, probit, and logit models; time series models; and models involving endogenous variables.

Edward Greenberg is Professor Emeritus of Economics at Washington University, St. Louis, where he served as a Full Professor on the faculty from 1969 to 2005. Professor Greenberg also taught at the University of Wisconsin, Madison, and has been a Visiting Professor at the University of Warwick (UK), Technion University (Israel), and the University of Bergamo (Italy). A former holder of a Ford Foundation Faculty Fellowship, Professor Greenberg is the coauthor of four books: *Wages, Regime Switching, and Cycles* (1992), *The Labor Market and Business Cycle Theories* (1989), *Advanced Econometrics* (1983, revised 1991), and *Regulation, Market Prices, and Process Innovation* (1979). His published research has appeared in leading journals such as the *American Economic Review*, *Econometrica*, *Journal of Econometrics*, *Journal of the American Statistical Association*, *Biometrika*, and the *Journal of Economic Behavior and Organization*. Professor Greenberg's current research interests include dynamic macroeconomics as well as Bayesian econometrics.

# Preface

## To Instructors and Students

THIS BOOK IS a concise introduction to Bayesian statistics and econometrics. It can be used as a supplement to a frequentist course by instructors who wish to introduce the Bayesian viewpoint or as a text in a course in Bayesian econometrics supplemented by readings in the current literature.

While the student should have had some exposure to standard probability theory and statistics, the book does not make extensive use of statistical theory. Indeed, because of its reliance on simulation techniques, it requires less background in statistics and probability than do most books that take a frequentist approach. It is, however, strongly recommended that the students become familiar with the forms and properties of the standard probability distributions collected in Appendix A.

Since the advent of Markov chain Monte Carlo (MCMC) methods in the early 1990s, Bayesian methods have been extended to a large and growing number of applications. This book limits itself to explaining in detail a few important applications. Its main goal is to provide examples of MCMC algorithms to enable students and researchers to design algorithms for the models that arise in their own research. More attention is paid to the design of algorithms for the models than to the specification and interpretation of the models themselves because we assume that the student has been exposed to these models in other statistics and econometrics classes.

The decision to keep the book short has also meant that we have taken a stand on some controversial issues rather than discuss a large number of alternative methods. In some cases, alternative approaches are discussed in end of chapter notes.

Exercises have been included at the end of the chapters, but the best way to learn the material is for students to apply the ideas to empirical applications of their choice. Accordingly, even though it is not explicitly stated, the first exercise at the end of every chapter in Part III should direct students to formulate a model; collect

data; specify a prior distribution on the basis of previous research design and, if necessary, program an algorithm; and present the results.

A link to the Web site for the course may be found at my Web site: <http://edg.wustl.edu>. The site contains errata, links to data sources, some computer code, and other information.

### Acknowledgments

I would like to acknowledge and offer my sincere gratitude to some of the people who have helped me throughout my career. On the professional side, I start with my undergraduate years at the business school of New York University, where Abraham Gitlow awakened my interest in economics. My first statistics course was with F. J. Viser and my second with Ernest Kurnow, who encouraged me to continue my studies and guided me in the process.

At the University of Wisconsin–Madison, I was mentored by, among others, Peter Steiner and Guy Orcutt. Econometrics was taught by Jack Johnston, who was writing the first edition of his pathbreaking book, and I was fortunate to have Arthur Goldberger and Arnold Zellner as teachers and colleagues. My first mathematical statistics course was with Enders Robinson, and I later audited George Box's class, where I received my first exposure to Bayesian ideas. Soon afterward, Zellner began to apply the methods to econometrics in a workshop that I attended.

My interest in Bayesian methods was deepened at Washington University first by E. T. Jaynes and then by Siddhartha Chib. Sid Chib has been my teacher, collaborator, and friend for the last 15 years. His contributions to Bayesian statistics, econometrics, and MCMC methods have had enormous impact. I have been extremely fortunate to have had the opportunity to work with him. The students in my courses in Bayesian econometrics contributed to my understanding of the material by their blank stares and penetrating questions. I am most grateful to them.

My colleagues and the staff of the Economics Department at Washington University have always been extremely helpful to me. I am delighted to thank them for their support.

I am most grateful to my editor at Cambridge University Press, Scott Parris, for suggesting the book, and for his continuing encouragement and support, and to Kimberly Twist, Editorial Assistant at Cambridge, for her help in the publication process.

I am pleased to acknowledge the comments of Andrew Martin, James Morley, and two anonymous reviewers on various drafts of this book and, especially, those of Ivan Jeliaskov, who read it most carefully and thoughtfully and tested it on his students. All remaining errors are, of course, mine.

I am grateful to Professor Chang-Jin Kim for permission to utilize his software to compute some of the examples in Chapter 10.

On the personal side, I thank Arthur and Aida, Lisa and Howard, my grandchildren, and my colleagues and friends, particularly Sylvia Silver, Karen Rensing, Ingrid and Wilhelm Neufeind, Maureen Regan and Sid Chib, Jasmine and Steve Fazzari, and Camilla and Piero Ferri.

In December 2005, my wife of more than 46 years passed away. I dedicate this book to Joan's memory.

# Contents

<i>List of Figures</i>	<i>page</i> ix
<i>List of Tables</i>	xi
<i>Preface</i>	xiii

## **Part I Fundamentals of Bayesian Inference**

<b>1 Introduction</b>	<b>3</b>
1.1 Econometrics	3
1.2 Plan of the Book	4
1.3 Historical Note and Further Reading	5
<b>2 Basic Concepts of Probability and Inference</b>	<b>7</b>
2.1 Probability	7
2.1.1 Frequentist Probabilities	8
2.1.2 Subjective Probabilities	9
2.2 Prior, Likelihood, and Posterior	12
2.3 Summary	18
2.4 Further Reading and References	19
2.5 Exercises	19
<b>3 Posterior Distributions and Inference</b>	<b>20</b>
3.1 Properties of Posterior Distributions	20
3.1.1 The Likelihood Function	20
3.1.2 Vectors of Parameters	22
3.1.3 Bayesian Updating	24
3.1.4 Large Samples	25
3.1.5 Identification	28
3.2 Inference	29
3.2.1 Point Estimates	29



3.2.2	Interval Estimates	31
3.2.3	Prediction	32
3.2.4	Model Comparison	33
3.3	Summary	38
3.4	Further Reading and References	38
3.5	Exercises	39
<b>4</b>	<b>Prior Distributions</b>	<b>41</b>
4.1	Normal Linear Regression Model	41
4.2	Proper and Improper Priors	43
4.3	Conjugate Priors	44
4.4	Subject-Matter Considerations	47
4.5	Exchangeability	50
4.6	Hierarchical Models	52
4.7	Training Sample Priors	53
4.8	Sensitivity and Robustness	54
4.9	Conditionally Conjugate Priors	54
4.10	A Look Ahead	56
4.11	Further Reading and References	57
4.12	Exercises	58
 <b>Part II Simulation</b>		
<b>5</b>	<b>Classical Simulation</b>	<b>63</b>
5.1	Probability Integral Transformation Method	63
5.2	Method of Composition	65
5.3	Accept–Reject Algorithm	66
5.4	Importance Sampling	70
5.5	Multivariate Simulation	72
5.6	Using Simulated Output	72
5.7	Further Reading and References	74
5.8	Exercises	75
<b>6</b>	<b>Basics of Markov Chains</b>	<b>76</b>
6.1	Finite State Spaces	76
6.2	Countable State Spaces	81
6.3	Continuous State Spaces	85
6.4	Further Reading and References	87
6.5	Exercises	87
<b>7</b>	<b>Simulation by MCMC Methods</b>	<b>90</b>
7.1	Gibbs Algorithm	91

7.1.1	Basic Algorithm	91
7.1.2	Calculation of Marginal Likelihood	95
7.2	Metropolis–Hastings Algorithm	96
7.2.1	Basic Algorithm	96
7.2.2	Calculation of Marginal Likelihood	101
7.3	Numerical Standard Errors and Convergence	102
7.4	Further Reading and References	103
7.5	Exercises	105

### Part III Applications

<b>8</b>	<b>Linear Regression and Extensions</b>	<b>111</b>
8.1	Continuous Dependent Variables	111
8.1.1	Normally Distributed Errors	111
8.1.2	Student- <i>t</i> Distributed Errors	114
8.2	Limited Dependent Variables	117
8.2.1	Tobit Model for Censored Data	117
8.2.2	Binary Probit Model	122
8.2.3	Binary Logit Model	126
8.3	Further Reading and References	129
8.4	Exercises	132
<b>9</b>	<b>Multivariate Responses</b>	<b>134</b>
9.1	SUR Model	134
9.2	Multivariate Probit Model	139
9.3	Panel Data	144
9.4	Further Reading and References	149
9.5	Exercises	151
<b>10</b>	<b>Time Series</b>	<b>153</b>
10.1	Autoregressive Models	153
10.2	Regime-Switching Models	158
10.3	Time-Varying Parameters	161
10.4	Time Series Properties of Models for Panel Data	165
10.5	Further Reading and References	166
10.6	Exercises	167
<b>11</b>	<b>Endogenous Covariates and Sample Selection</b>	<b>168</b>
11.1	Treatment Models	168
11.2	Endogenous Covariates	173
11.3	Incidental Truncation	175

11.4 Further Reading and References	179
11.5 Exercises	180
<b>A Probability Distributions and Matrix Theorems</b>	<b>183</b>
A.1 Probability Distributions	183
A.1.1 Bernoulli	183
A.1.2 Binomial	183
A.1.3 Negative Binomial	184
A.1.4 Multinomial	184
A.1.5 Poisson	184
A.1.6 Uniform	184
A.1.7 Gamma	185
A.1.8 Inverted or Inverse Gamma	185
A.1.9 Beta	186
A.1.10 Dirichlet	186
A.1.11 Normal or Gaussian	187
A.1.12 Multivariate and Matricvariate Normal or Gaussian	187
A.1.13 Truncated Normal	189
A.1.14 Univariate Student- $t$	189
A.1.15 Multivariate $t$	189
A.1.16 Wishart	190
A.1.17 Inverted or Inverse Wishart	191
A.1.18 Multiplication Rule of Probability	191
A.2 Matrix Theorems	192
<b>B Computer Programs for MCMC Calculations</b>	<b>193</b>
<i>Bibliography</i>	195
<i>Author Index</i>	201
<i>Subject Index</i>	203

## List of Figures

2.1	Beta distributions for various values of $\alpha$ and $\beta$	page 16
2.2	Prior, likelihood, and posterior for coin-tossing example	18
4.1	Prior and posterior distributions for coefficient of union membership	50
5.1	Target and proposal density to sample from Beta(3, 3)	68
5.2	Target and proposal density to sample from $N(0, 1)$	69
5.3	Simulation results for $Z = XY$ , where $X \sim \text{Beta}(3, 3)$ and $Y \sim \text{Beta}(5, 3)$	73
6.1	Random walk, $p = q = 0.5$	82
6.2	Random walk, $p = 0.55, q = 0.45$	83
7.1	Simulation results for MH sampling of Beta(3, 4) with $U(0, 1)$ proposal	100
7.2	Autocorrelations of $X^{(g)}$	104
8.1	Posterior distributions of $\beta_U$ and $\sigma^2$ , Gaussian errors	113
8.2	Autocorrelations of $\beta_U$ and $\sigma^2$ , Gaussian errors	114
8.3	Posterior distributions of $\beta_U$ and $\sigma^2$ , Student- $t$ errors	116
8.4	Autocorrelations of $\beta_U$ and $\sigma^2$ , Student- $t$ errors	117
8.5	Posterior distributions of $\beta$ : Tobit model, Mroz data	122
8.6	Predictive distributions of hours worked: Tobit model, Mroz data. Left: women with young children; right: women with older children	123
8.7	Posterior distributions of $\beta$ : computer ownership example, probit model	127
8.8	Posterior distributions of $\beta$ : computer ownership example, logit model	130
9.1	Summary of $\beta_F$	138
9.2	Summary of $\beta_C$	139

9.3	Posterior distributions of $\beta_U$ and $\text{mean}(b_2)$	150
10.1	Probability of recession	161
10.2	Time-varying slope	165
11.1	Selected coefficients: incidental truncation model, Mroz data	180

## List of Tables

3.1	Jeffreys Guidelines	<i>page</i> 35
3.2	Bayes Factors for Selected Possible Outcomes	38
4.1	$\beta_U$ as a Function of Hyperparameters $\beta_{U0}$ and $B_{UU,0}$	54
8.1	Summary of Posterior Distribution: Tobit Model, Mroz Data	121
8.2	Summary of Posterior Distribution: Probit Model, Computer Example	126
8.3	Summary of Posterior Distribution: Logit Model, Computer Example	129
9.1	Summary of Posterior Distribution of $\beta_F$ : Grunfeld Data, SUR Model	138
9.2	Summary of Posterior Distribution of $\beta_C$ : Grunfeld Data, SUR Model	139
9.3	Means of Posterior Distribution of Contemporaneous Correlations: Grunfeld Data, SUR Model	140
9.4	Summary of Prior and Posterior Distributions of $\beta$ and $\sigma_{12}$ : Rubinfeld Data	144
9.5	Summary of Posterior Distribution: Panel Data Model, Vella–Verbeek Data	149
10.1	Summary of Posterior Distribution: AR(1) Errors	158
10.2	Parameter Estimates for GDP Data	161
10.3	Summary of Posterior Distribution: Time Varying Parameter Model	164
11.1	Summary of Posterior Distribution: Probit Selection Model, Mroz Data	179

# **Part I**

## **Fundamentals of Bayesian Inference**





# Chapter 1

## Introduction

THIS CHAPTER INTRODUCES several important concepts, provides a guide to the rest of the book, and offers some historical perspective and suggestions for further reading.

### 1.1 Econometrics

Econometrics is largely concerned with quantifying the relationship between one or more variables  $y$ , called the response variables or the dependent variables, and one or more variables  $x$ , called the regressors, independent variables, or covariates. The response variable or variables may be continuous or discrete; the latter case includes binary, multinomial, and count data. For example,  $y$  might represent the quantities demanded of a set of goods, and  $x$  could include income and the prices of the goods; or  $y$  might represent investment in capital equipment, and  $x$  could include measures of expected sales, cash flows, and borrowing costs; or  $y$  might represent a decision to travel by public transportation rather than private, and  $x$  could include income, fares, and travel time under various alternatives.

In addition to the covariates, it is assumed that unobservable random variables affect  $y$ , so that  $y$  itself is a random variable. It is characterized either by a probability density function (p.d.f.) for continuous  $y$  or a probability mass function (p.m.f.) for discrete  $y$ . The p.d.f. or p.m.f. depends on the values of unknown parameters, denoted by  $\theta$ . The notation  $y \sim f(y|\theta, x)$  means that  $y$  has the p.d.f. or p.m.f.  $f(y|\theta, x)$ , where the function depends on the parameters and covariates. It is customary to suppress dependence on the covariates when writing the p.d.f. of  $y$ , so we write  $y \sim f(y|\theta)$  unless it is necessary to mention the covariates explicitly.

The data may take the form of observations on a number of subjects at the same point in time – cross section data – or observations over a number of time periods – time series data. They may be a combination of cross-section and time-series observations: data over many subjects over a relatively short period of time