Ecological Inference

Ecological Inference: New Methodological Strategies brings together a diverse group of scholars to survey the latest strategies for solving ecological inference problems in various fields. The last half decade has witnessed an explosion of research in ecological inference—the attempt to infer individual behavior from aggregate data. The uncertainties and the information lost in aggregation make ecological inference one of the most difficult areas of statistical inference, but such inferences are required in many academic fields, as well as by legislatures and the courts in redistricting, by businesses in marketing research, and by governments in policy analysis.

Gary King is the David Florence Professor of Government at Harvard University. He also serves as the director of the Harvard–MIT Data Center and as a member of the steering committee of Harvard’s Center for Basic Research in the Social Sciences. He was elected president of the Society for Political Methodology and Fellow of the American Academy of Arts and Sciences. Professor King received his Ph.D. from the University of Wisconsin. He has won numerous awards for his work, including the Gosnell Prize for his book A Solution to the Ecological Inference Problem (1997), on which the research in this book builds. His home page can be found at http://GKing.Harvard.edu.

Ori Rosen is Assistant Professor of Statistics at the University of Pittsburgh. His research includes work on semiparametric regression models, applications of mixtures-of-experts neural network models in regression, and applications of Markov chain Monte Carlo methods. Professor Rosen was educated at the Technion, and he later served as Mellon Postdoctoral Fellow at Northwestern University.

Martin A. Tanner is Professor of Statistics at Northwestern University. He has authored and coauthored nearly 100 research articles on wide-ranging topics in theoretical and applied statistics. His previous books include Investigations for a Course in Statistics (1990), Tools for Statistical Inference (1996), and Statistics for the 21st Century (2001). Professor Tanner is a Fellow of the American Statistical Association and the Royal Statistical Society, and he has been honored with the 1993 Mortimer Spiegelman Award as well as the American Statistical Association’s Continuing Education Excellence Award. He has served as editor of the Journal of the American Statistical Association (Theory and Methods). Professor Tanner received his Ph.D. from the University of Chicago.
Analytical Methods for Social Research

Series Editors

R. Michael Alvarez, California Institute of Technology
Nathaniel L. Beck, New York University
Lawrence L. Wu, New York University

*Analytical Methods for Social Research* presents texts on empirical and formal methods for the social sciences. Some series volumes are broad in scope, addressing multiple disciplines; others focus mainly on techniques applied within specific fields, such as political science, sociology, and demography.

*Previously published:*

*Event History Modeling: A Guide for Social Scientists*, Janet M. Box-Steffensmeier and Bradford S. Jones
Ecological Inference

New Methodological Strategies

Edited by

Gary King
Harvard University

Ori Rosen
University of Pittsburgh

Martin A. Tanner
Northwestern University
## Contents

*Contributors* page vii  
*Preface* ix  

**INTRODUCTION** 1  
*Information in Ecological Inference: An Introduction* 1  
Gary King, Ori Rosen, and Martin A. Tanner  

**PART ONE** 13  
1 *Prior and Likelihood Choices in the Analysis of Ecological Data* 13  
Jonathan Wakefield  
2 *The Information in Aggregate Data* 51  
David G. Steel, Eric J. Beh, and Ray L. Chambers  
3 *Using Ecological Inference for Contextual Research* 69  
D. Stephen Voss  

**PART TWO** 97  
4 *Extending King’s Ecological Inference Model to Multiple Elections Using Markov Chain Monte Carlo* 97  
Jeffrey B. Lewis  
5 *Ecological Regression and Ecological Inference* 123  
Bernard Grofman and Samuel Merrill  
6 *Using Prior Information to Aid Ecological Inference: A Bayesian Approach* 144  
J. Kevin Corder and Christina Wolbrecht  
7 *An Information Theoretic Approach to Ecological Estimation and Inference* 162  
George G. Judge, Douglas J. Miller, and Wendy K. Tam Cho  
8 *Ecological Panel Inference from Repeated Cross Sections* 188  
Ben Pelzer, Rob Eisinga, and Philip Hans Franses  

**PART THREE** 207  
9 *Ecological Inference in the Presence of Temporal Dependence* 207  
Kevin M. Quinn  
10 *A Spatial View of the Ecological Inference Problem* 233  
Carol A. Gotway Crawford and Linda J. Young
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Places and Relationships in Ecological Inference</td>
<td>245</td>
</tr>
<tr>
<td></td>
<td>Ernesto Calvo and Marcelo Escolar</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Ecological Inference Incorporating Spatial Dependence</td>
<td>266</td>
</tr>
<tr>
<td></td>
<td>Sebastien Haneuse and Jonathan Wakefield</td>
<td></td>
</tr>
<tr>
<td>PART FOUR</td>
<td></td>
<td>303</td>
</tr>
<tr>
<td>13</td>
<td>Common Framework for Ecological Inference in Epidemiology, Political Science, and Sociology</td>
<td>303</td>
</tr>
<tr>
<td></td>
<td>Ruth Salway and Jonathan Wakefield</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Multiparty Split-Ticket Voting Estimation as an Ecological Inference Problem</td>
<td>333</td>
</tr>
<tr>
<td></td>
<td>Kenneth Benoit, Michael Laver, and Daniela Giannetti</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>A Structured Comparison of the Goodman Regression, the Truncated Normal, and the Binomial–Beta Hierarchical Methods for Ecological Inference</td>
<td>351</td>
</tr>
<tr>
<td></td>
<td>Rogério Silva de Mattos and Álvaro Veiga</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>A Comparison of the Numerical Properties of EI Methods</td>
<td>383</td>
</tr>
<tr>
<td></td>
<td>Micah Altman, Jeff Gill, and Michael P. McDonald</td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td></td>
<td>409</td>
</tr>
</tbody>
</table>
Contributors

Micah Altman  Harvard-MIT Data Center (HMDC), Harvard University, Cambridge, Massachusetts

Eric J. Beh  School of Quantitative Methods and Mathematical Sciences, University of Western Sydney, Sydney, Australia

Kenneth Benoit  Department of Political Science, Trinity College, Dublin, Ireland

Ernesto Calvo  Department of Political Science, University of Houston, Houston, Texas

Ray L. Chambers  Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, United Kingdom

Wendy K. Tam Cho  Departments of Political Science and Statistics, University of Illinois at Urbana-Champaign, Urbana, Illinois

J. Kevin Corder  Department of Political Science, Western Michigan University, Kalamazoo, Michigan

Carol A. Gotway Crawford  National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, Georgia

Rob Eisinga  Department of Social Science Research Methods, University of Nijmegen, Nijmegen, The Netherlands

Marcelo Escolar  Department of Geography, Universidad de Buenos Aires, Buenos Aires, Argentina

Philip Hans Franses  Econometric Institute, Erasmus University, Rotterdam, The Netherlands

Daniela Giannetti  Department of Political Science, University of Bologna, Bologna, Italy

Jeff Gill  Department of Political Science, University of California, Davis, California

Bernard Grofman  Department of Political Science, University of California, Irvine, California

Sebastien Haneuse  Department of Biostatistics, University of Washington, Seattle, Washington

George G. Judge  Department of Agricultural and Resource Economics, University of California, Berkeley, California
 Contributors

Gary King  Center for Basic Research in the Social Sciences, Harvard University, Cambridge, Massachusetts

Michael Laver  Department of Political Science, Trinity College, Dublin, Ireland

Jeffrey B. Lewis  Department of Political Science, University of California, Los Angeles, California

Rogério Silva de Mattos  Department of Economic Analysis, Federal University of Juiz de Fora, Minas Gerais, Brazil

Michael P. McDonald  Department of Public & International Affairs, George Mason University, Fairfax, Virginia

Samuel Merrill  Department of Mathematics and Computer Science, Wilkes University, Wilkes-Barre, Pennsylvania

Douglas J. Miller  Department of Agricultural Economics, Purdue University, West Lafayette, Indiana

Ben Pelzer  Department of Social Sciences Research Methods, University of Nijmegen, Nijmegen, The Netherlands

Kevin M. Quinn  Center for Basic Research in the Social Sciences, Harvard University, Cambridge, Massachusetts

Ori Rosen  Department of Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania

Ruth Salway  Department of Mathematical Sciences, University of Bath, Bath, United Kingdom

David G. Steel  School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, Australia

Martin A. Tanner  Department of Statistics, Northwestern University, Evanston, Illinois

Álvaro Veiga  Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil

D. Stephen Voss  Department of Political Science, University of Kentucky, Lexington, Kentucky

Jonathan Wakefield  Departments of Statistics and Biostatistics, University of Washington, Seattle, Washington

Christina Wolbrecht  Department of Political Science, University of Notre Dame, Notre Dame, Indiana

Linda J. Young  Department of Statistics, College of Medicine, University of Florida, Gainesville, Florida
Preface

The authors of the chapters in this volume hail from academic disciplines with markedly different substantive concerns. Indeed, we suspect that not many books have been written with contributions from political scientists, electrical engineers, economists, agricultural economists, geographers, statisticians, applied statisticians, mathematicians, public health researchers, biostatisticians, and computer scientists. Yet, while the substantive problems pursued by the diverse disciplinary origins of these researchers vary enormously, they all have a deep, if not widely recognized, methodological common ground. Although the style and terminology often obscure this fact, they all use roughly the same theories of inference and many of the same statistical methods. The subject of this book is ecological inference, the problem of reconstructing individual behavior from group-level data, which indeed turns out to be a key problem in all these fields, as well as a variety of others, which we were not able to include. Not only is ecological inference required in a growing number of applications, it has a large number of scholars working on the methods of ecological inference – now larger than at any time in history.

Because our work seems to have had a particularly visible role in the renewed interest in ecological inference, we found ourselves in a unique position of getting to know many otherwise unconnected scholars from this vast array of scholarly fields. So that our new scholarly acquaintances could get to know each other, begin to build on each other’s work, and start to create a more densely connected scholarly network that spans everyone’s traditional discipline, substantive concerns, and methodological commonalities, we held an intensive conference at Harvard University in June of 2002. The early version of the chapters herein were first presented during lively discussions at that conference. We hope the publication of this book enables methodologists whom we were not able to invite to the conference, or who find themselves working on similar issues, to now be able to join in the discussion. Others interested in different statistical problems may also find this work of some interest.

The limited information available means that ecological inference is an especially difficult area of statistical inference, and so we have found that studying it illuminates fundamental problems that do not surface as clearly when learning statistics in the context of other applications.

We thank Kim Schader and Jaronica Fuller for their help in organizing a superb conference, and we appreciate the outstanding help of Kim, Colleen McMahon, Cindy Munroe, and Alison Ney in preparing this volume. Special thanks go to the Center for Basic Research in the Social Sciences for providing financial and logistical support for the conference. Our appreciation goes to the anonymous reviewers from Cambridge University Press for their insightful comments on the project. We especially appreciate the patience and efforts of our authors on whom we inflicted proposals, travel, presentations, discussants, and revisions.
and then called on them for subsequent rounds of review to provide written comments on each other’s chapters. Finally, we are indebted to the crew at Cambridge University Press, including Ed Parsons and Eleanor Umali, for everything they have done to facilitate the entire publication process.

Gary King
Ori Rosen
Martin Tanner