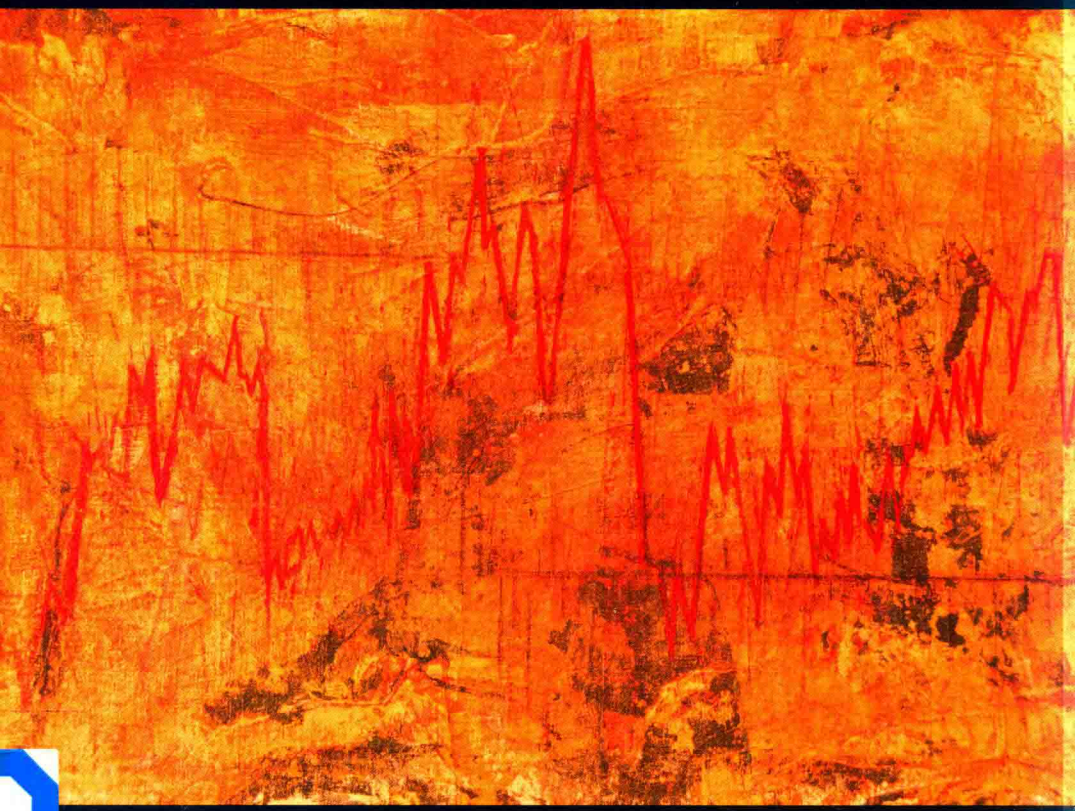
Analytical Methods for Social Research

Time Series Analysis for the Social Sciences



Janet M. Box-Steffensmeier
John R. Freeman
Matthew P. Hitt
Jon C. W. Pevehouse

Time Series Analysis for the Social Sciences

JANET M. BOX-STEFFENSMEIER
The Ohio State University

JOHN R. FREEMAN
University of Minnesota

MATTHEW P. HITT
Louisiana State University

JON C. W. PEVEHOUSE
University of Wisconsin, Madison

常州大学山书馆
藏书章

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

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

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521691550

© Janet M. Box-Steffensmeier, John R. Freeman, Matthew P. Hitt, and Jon C. W. Pevehouse 2014

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 2014

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
Box-Steffensmeier, Janet M., 1965–
Time series analysis for the social sciences / Janet M. Box-Steffensmeier, The Ohio State University, John R. Freeman, University of Minnesota, Matthew P. Hitt, Louisiana State University, Jon C. W. Pevehouse, University of Wisconsin Madison.
pages cm. – (Analytical methods for social research)
Includes index.
ISBN 978-0-521-87116-7 (hardback) – ISBN 978-0-521-69155-0 (paperback)
1. Time-series analysis. 2. Time-series analysis – Mathematical models. I. Title.
HA30.3.B69 2014
300.1′51955–dc23 2014010088

ISBN 978-0-521-87116-7 Hardback
ISBN 978-0-521-69155-0 Paperback

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.

Time Series Analysis for the Social Sciences

Time series or longitudinal data are ubiquitous in the social sciences. Unfortunately, analysts often treat the time series properties of their data as a nuisance rather than as a substantively meaningful dynamic process to be modeled and interpreted. *Time Series Analysis for the Social Sciences* provides accessible, up-to-date instruction and examples of the core methods in time series econometrics. Janet M. Box-Steffensmeier, John R. Freeman, Matthew P. Hitt, and Jon C. W. Pevehouse cover a wide range of topics, including ARIMA models, time series regression, unit root diagnosis, vector autoregressive models, error correction models, intervention models, fractional integration, ARCH models, structural breaks, and forecasting. This book is aimed at researchers and graduate students who have taken at least one course in multivariate regression. Examples are drawn from several areas of social science, including political behavior, elections, international conflict, criminology, and comparative political economy.

Janet M. Box-Steffensmeier is the Vernal Riffe Professor of Political Science and Professor of Sociology at The Ohio State University (courtesy), where she is a University Distinguished Scholar and directs the Program in Statistics and Methodology (PRISM). Box-Steffensmeier served as president of the Midwest Political Science Association and the Political Methodology Society and as treasurer of the American Political Science Association. She has twice received the Gosnell Prize for the best work in political methodology, and she received the Emerging Scholar Award from the Elections, Public Opinion, and Voting Behavior Subsection of the American Political Science Association and the Career Achievement Award from the Political Methodology Society. She was an inaugural Fellow of the Society for Political Methodology. The Box-Steffensmeier Graduate Student Award, given annually by the Interuniversity Consortium for Political and Social Research (ICPSR), is named after her in recognition of her contributions to political methodology and her support of women in this field.

John R. Freeman is the John Black Johnston Distinguished Professor in the College of Liberal Arts at the University of Minnesota and a Fellow of the American Academy of Arts and Sciences. Among his honors are the Morse-Alumni, All-University, and College of Liberal Arts Distinguished Teaching awards at the University of Minnesota. Freeman is the author of *Democracy and Markets: The Politics of Mixed Economies*, which won the International Studies Association's Quincy Wright Award, and the coauthor of *Three Way Street: Strategic Reciprocity in World Politics*. Freeman also edited three volumes of *Political Analysis*. He has (co)authored numerous research articles in academic journals. Freeman's research projects have been supported by the National Science Foundation, as well as by the Bank Austria Foundation and the Austrian Ministry of Science.

Matthew P. Hitt is an assistant professor of political science at Louisiana State University. His interests include judicial politics, legislative politics, interest groups, the presidency, and quantitative methodology. His research has been published in the *American Political Science Review* and *Presidential Studies Quarterly*.

Jon C. W. Pevehouse is a professor of political science at the University of Wisconsin. His work examines the relationship between domestic and international politics. Pevehouse is the author of *Democracy from Above* (Cambridge University Press, 2005) and *While Dangers Gather* (2007). He is the coauthor, with Joshua Goldstein, of *International Relations*, the leading textbook on international politics. He is the recipient of the Karl Deutch Award, given by the International Studies Association, and has received numerous teaching awards, including the Chancellor's Distinguished Teaching Award at the University of Wisconsin. Pevehouse is also the editor of the journal *International Organization*.

Analytical Methods for Social Research

Analytical Methods for Social Research presents texts on empirical and formal methods for the social sciences. Volumes in the series address both the theoretical underpinnings of analytical techniques as well as their application in social research. Some series volumes are broad in scope, cutting across a number of disciplines. Others focus mainly on methodological applications within specific fields such as political science, sociology, demography, and public health. The series serves a mix of students and researchers in the social sciences and statistics.

Series Editors:
R. Michael Alvarez, California Institute of Technology
Nathaniel L. Beck, New York University
Stephen L. Morgan, Johns Hopkins University
Lawrence L. Wu, New York University

Other Titles in the Series:

Counterfactuals and Causal Inference, Second Edition, by Stephen L. Morgan and Christopher Winship
Statistical Modeling and Inference for Social Science, by Sean Gailmard
Formal Models of Domestic Politics, by Scott Gehlbach
Counterfactuals and Causal Inference: Methods and Principles for Social Research, by Stephen L. Morgan and Christopher Winship
Data Analysis Using Regression and Multilevel/Hierarchical Models, by Andrew Gelman and Jennifer Hill
Political Game Theory: An Introduction, by Nolan McCarty and Adam Meirowitz
Essential Mathematics for Political and Social Research, by Jeff Gill
Spatial Models of Parliamentary Voting, by Keith T. Poole
Ecological Inference: New Methodological Strategies, edited by Gary King, Ori Rosen, and Martin A. Tanner
Event History Modeling: A Guide for Social Scientists, by Janet M. Box-Steffensmeier and Bradford S. Jones

此为试读，需要完整PDF请访问：www.ertongbook.com

All that really belongs to us is time; even he who has nothing else has that.

Balthasar Gracian

This book is dedicated to our families, who make time so valuable.

To Mike, Andrew, Zach, Nate, and Lizzy from Jan
To Tom from John
To Jen, Theodore, Shelley, and Larry from Matt
To Jessica, Claire, Ava, and Carl from Jon

Preface

Our work has several motivations. We think that longitudinal analysis provides infinitely more insight than does examining any one slice of time. As we show throughout the book, longitudinal analysis is essential for the study of normatively important problems such as democratic accountability and international conflict. Given the importance of dynamic analysis in answering new questions and providing new answers to old questions, we want to get more social scientists thinking in dynamic terms. Time series is one of the most useful tools for dynamic analysis, and our goal is to provide a more accessible treatment for this approach. We are also motivated by the burgeoning supply of new social science time series data. Sometimes this causes the opposite problem of too much data and figuring out how to analyze it, but that is a problem we gladly embrace. The proliferation of new social science data requires techniques that are designed to handle complexity, and time series analysis is one of the most applicable tools. The incorporation of time series analysis into standard statistical packages such as STATA and R, as well as the existence of specialized packages such as RATS and Eviews, provides an additional motivation because it enables more scholars to easily use time series in their work.

We have found over our years of teaching time series that, although many social science students have the brain power to learn time series methods, they often lack the training and motivation to use the most well-known books on the topic. We specifically wanted to write an accessible book for social scientists so that they too could leverage the power of time series analysis from the introductory material to current innovations. That said, we are not able to offer complete coverage. We do not address dynamic panel data analysis, Bayesian time series analysis, spectral analysis, or the event history approach to temporal data. We hope the foundation and discussion of recent advances we do provide result in a useful reference book for scholars.

Chapter 1 provides an intuitive motivation for the study of time series and social dynamics. Important issues such as measurement, fit and scale, and structural change are introduced here. Chapter 2 focuses on univariate models, which are important because we argue that understanding the nature of the data-generating process should be the first step in the data analysis process. Chapter 3 provides a discussion of conventional time series regression methods, in particular the workhorse model in which a single lag of the endogenous variable is included on the right side of a single regression equation. In contrast, Chapter 4 explores the specification, estimation, and interpretation of familiar multiequation regression models with strong restrictions, simultaneous structural equation models, and also weakly restricted multiequation dynamic models – vector autoregressions – that allow uncertainty about the specification. The chapter highlights the differences between the approaches. Chapter 5 introduces the concept of stationarity and discusses its methodological and substantive importance. Chapter 6 discusses cointegration, which is a cornerstone of current time series analysis. Cointegration is especially useful for studying equilibrium relationships. Chapter 7 concludes with discussions of four critical concepts in current time series analysis: fractional integration, heterogeneity, unknown structural break(s), and forecasting. The Appendix covers difference equations. This is a foundational concept needed for understanding the mathematical foundation of common time series approaches, such as vector autoregression. Although readers interested in directly estimating models without mathematical preliminaries may skip the Appendix, we strongly encourage readers interested in working in time series methodology to read it closely.

We have had the privilege of team-teaching time series together for about 15 years. John, an award-winning teacher (we think he has won every teaching award possible), graciously offered to team-teach with Jan; this was made possible by the innovation pioneered by Phil Shively and Pete Nardulli, the cross-campus interactive television (I.T.V.) program. We welcomed Jon to the team shortly after he finished graduate school. Indeed, he had been a student in the first iteration of the course, and later, while still in graduate school, he wrapped up a time series course for Jan when her third child was seriously ill. Team-teaching time series for the three of us is a biennial occurrence that does not come frequently enough for us, but is thoroughly enjoyed each time the occasion arises. Matt Hitt was a student in the course who later became a valuable collaborator who pushed the book to completion.

We have a long list of thank you's to our home departments, the I.T.V. program, and our many students. Brandon Bartels, Quintin Beazer, Patrick Brandt, Harold Clarke, Dave Darmofal, Suzie DeBoef, Charles Franklin, Jeff Gill, Tobin Grant, Agnar Helgason, Mel Hinich, Tana Johnson, Ben Jones, Luke Keele, Paul Kellstedt, Matt Lebo, Tse-Min Lin, Eleonora Mattiacci, Sara Mitchell, Jason Morgan, Dave Ohls, Heather Ondercin, Erica Owen, Dave

Peterson, Eve Ringsmuth, Darrell Steffensmeier, Anand Sokhey, Jim Stimson, Mark Toukan, Felicity Vabulas, Eve Van Sice, Herb Weisberg, Dan Wood, and especially Ben Jones and Corwin Smidt. In addition, we are grateful for the guidance and encouragement of the series editors, Mike Alvarez and Neal Beck, and the Cambridge editors, Ed Parsons and Lew Bateman.

Contents

Preface *page* xiii

1 Modeling Social Dynamics 1
1.1 Time Series Analysis 8
1.1.1 Contending Approaches and the Quasi-Experimental Tradition 9
1.1.2 The Architecture of the Book 10
1.2 Challenges in Time Series Modeling 11
1.2.1 Time Series Measurement 11
1.2.2 Fit: On the Use of Time Trends and Counters 18
1.2.3 Discerning Structural Change 19
1.3 The First Step 20
1.4 For Further Reading 21

2 Univariate Time Series Models 22
2.1 Understanding Univariate Processes 22
2.1.1 Why Stochastic Models of Time Series? 23
2.1.2 White Noise 26
2.1.3 A Note on Notation 27
2.2 Univariate Processes 29
2.2.1 The Autoregressive (AR) Process 29
2.2.2 The Moving Average (MA) Process 32
2.2.3 The Equivalence of AR and MA Processes 33
2.3 Diagnosing Stationary Univariate Processes 34
2.3.1 The Autocorrelation Function (ACF) 35
2.3.2 The Partial Autocorrelation Function (PACF) 38
2.3.3 Seasonality 43

2.4 Model Estimation, Interpretation, and Diagnosis 44
2.4.1 Residual Analysis 46
2.4.2 Metadiagnosis 46
2.5 Application: U.S. Homicide Rates 48
2.6 Application: U.S. Relations with the Middle East 52
2.7 Intervention Models 58
2.7.1 Impact Assessment: Continuing the Middle East Example 64
2.7.2 Structural Breaks 65
2.8 Conclusion 66
2.9 For Further Reading 67
3 Dynamic Regression Models 68
3.1 Variations of the General Dynamic Regression Model 72
3.1.1 Building Blocks 73
3.1.2 Autoregressive Models/Lagged Endogenous Variables 74
3.1.3 The Workhorse Model 75
3.1.4 Serial Correlation: Tests and Solutions 76
3.2 Obtaining Consistent Estimates 78
3.2.1 Approach Number One: Pseudo-GLS Estimation 78
3.2.2 Approach Number Two: Hatanaka's Method 80
3.3 Illustration: Explaining Voting Intentions in the United Kingdom 81
3.3.1 Voting Intentions and Exchange Rates: Results 83
3.4 Dynamic Regression Models with Multiple Causal Variable Lags 87
3.4.1 Distributed Lag Models 87
3.4.2 Autoregressive Distributed Lag Models 88
3.5 Conclusion 90
3.6 For Further Reading 91
4 Modeling the Dynamics of Social Systems 92
4.1 Two Approaches to Multivariate Time Series Analysis 95
4.2 The Structural Equation Approach 97
4.2.1 The Workhorse Model of Time Series Regression Revisited 100
4.2.2 The Structural Equation Approach in Practice 101
4.2.3 Application: The Political Economy of Monetary Policy in the United Kingdom 102
4.3 Vector Autoregression 106
4.3.1 Terminology and Stationarity 107
4.3.2 Specification and Estimation 110
4.3.3 Tools for Inference in VAR Modeling 112
4.3.4 VAR Modeling in Practice 118
4.4 The Political Economy of Monetary Policy in the United Kingdom Revisited 119

此为试读,需要完整PDF请访问: www.ertongbook.com

4.5 *Conclusion: SEQ and VAR Modeling Compared* 122
4.6 *For Further Reading* 124
5 Univariate, Nonstationary Processes: Tests and Modeling 125
5.1 *Stationary Data* 125
5.2 *Stationarity and Unit Root Tests* 130
5.2.1 *Dickey-Fuller and Augmented Dickey-Fuller Tests* 132
5.2.2 *Variance Ratio Test* 135
5.2.3 *Modified Rescaled Range Test* 136
5.2.4 *KPSS Test* 137
5.2.5 *Double Unit Roots* 138
5.3 *Application: Macropartisanship* 138
5.3.1 *Macropartisanship: Diagnostic Plots* 139
5.3.2 *Macropartisanship: Formal Statistical Tests* 142
5.3.3 *Macropartisanship: An ARIMA Model* 147
5.4 *Conclusion* 148
5.5 *For Further Reading* 149
6 Cointegration and Error Correction Models 150
6.1 *Introduction* 150
6.2 *The Intuition behind Cointegration* 152
6.3 *Estimating an ECM* 160
6.3.1 *Engle-Granger Regression-Based Approach* 161
6.3.2 *Johansen's VAR Approach* 164
6.3.3 *Application: The Engle-Granger Methodology and the Indo-Pakistani Arms Race* 166
6.4 *ECMs for Non-Integrated Data?* 169
6.5 *Conclusion* 171
6.6 *For Further Reading* 171
7 Selections on Time Series Analysis 173
7.1 *Fractional Integration* 175
7.1.1 *Fractional Cointegration* 178
7.1.2 *Near and Fractional Integration* 179
7.2 *Incorporating Heterogeneity* 181
7.3 *Forecasting* 187
7.3.1 *Purpose and Definitions* 188
7.3.2 *Approaches to Forecasting* 193
7.3.3 *Criteria for Determining Optimum Forecast* 199
7.3.4 *Illustration: Forecasts of the U.S. Uniform Crime Reports Homicide Data, 1975–1993* 200
7.3.5 *Forecasting Discussion* 202
7.4 *Estimating and Modeling with Unknown Structural Breaks* 205
7.4.1 *Identifying a Single Unknown Structural Break* 206
7.4.2 *Identifying Multiple Unknown Structural Breaks* 207

7.4.3 Hypothesis Tests for Multiple Structural Breaks 209
7.5 Conclusion 211
7.6 For Further Reading 212
8 Concluding Thoughts for the Time Series Analyst 214
8.1 A Review 215
8.2 Looking Ahead 217
Appendix: Time Series Models as Difference Equations 219
A.1 Introduction 219
A.2 The Nature of Difference Equation Models 226
A.2.1 Expressing Difference Equations 226
A.2.2 The Solutions of Difference Equations 229
A.3 Finding the Solutions of Difference Equation Models 240
A.3.1 Solving the Most Simple First-Order Difference Equations 240
A.3.2 Solving More Complicated Difference Equations 246
A.3.3 Solving Systems of Difference Equations 259
A.4 Bring on the Data 261

Bibliography 263
Index 277

1

Modeling Social Dynamics

Many social processes are best understood in terms of changes over time. Current social conditions are outgrowths of people's memories of past conditions and, sometimes, also of people's expectations about future conditions. Social processes are rarely, if ever, at rest. Rather they are moving in time, trending in a regular way or exhibiting patterns that recur in time. Understanding the causal relationships between social variables across time and the forces that drive these variables through time is essential to explaining and forecasting social processes. We call these relationships and forces social dynamics. Time series methods are a powerful means by which to analyze these dynamics. In his essay on the history of time series, Tsay (2002, 121) emphasizes that studies of dynamic structures and dynamic relationships between variables have played a vital role in the development of the field, especially through applications in business and economics.

Many important social problems are conceived in terms of dynamics. Indeed, a broad array of questions across the social sciences cannot be addressed without utilizing data comprised of measurements of variables over time. The substantive bite of many empirical findings from the social world comes not from knowing the static level of an important variable at one frozen moment, but rather from understanding the dynamics of that variable. For instance: Is not is the phenomenon of interest rising? Falling? Repeating a pattern seasonally? Drifting randomly? Reverting to a stable value that we can forecast? And how do the dynamics in the phenomenon relate and react to the dynamics in other phenomena? Substantively meaningful answers to questions of this sort require that we conceive our data as a dynamic process and utilize techniques such as those we present in this book.

Moving beyond the abstract, what sorts of phenomena can be conceptualized and studied under the framework we promote here? Consider economic

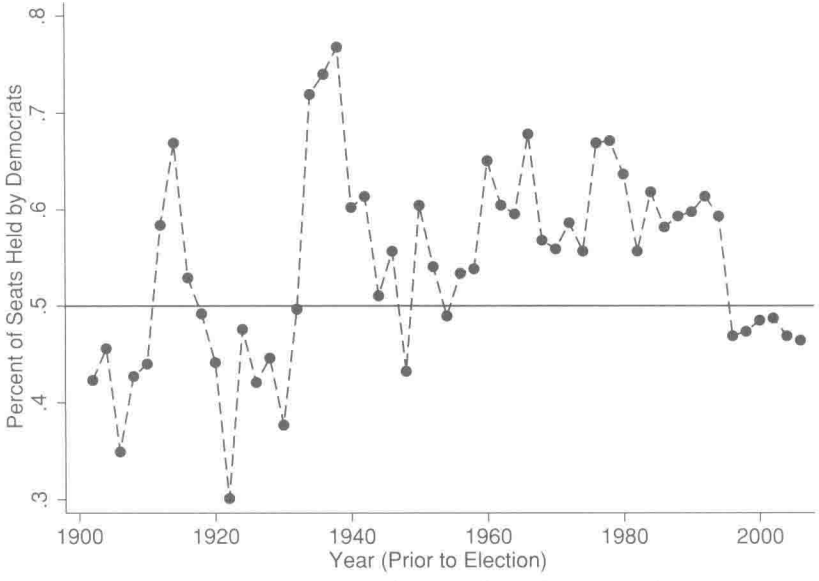


FIGURE 1.1. Democratic Party Seat Share in the U.S. House of Representatives, 1990–2004. Compiled by the authors from data supplied by the Clerk of the House of Representatives.

performance indicators such as unemployment, inflation, or gross domestic product. As both independent and dependent variables, they are integral to empirical analysis in many disciplines, including economics and political science. Measurements of aggregate fertility, obesity, smoking prevalence, and mortality over time are of keen interest in population studies and public health. In the field of education, fundamental data such as rates of literacy, graduation, and dropouts can all be well conceptualized as dynamic processes. A sophisticated understanding of the complex dynamic processes undergirding the rates of various crimes, incarceration, and recidivism is nothing short of foundational to the field of criminology. Adolescent fertility and drug use rates, along with rates of suicide and interpersonal violence across time, are all dynamic processes of great interest in sociology. These are, of course, but a few examples. The point is that, across the social sciences, some of our most interesting and fundamental empirical questions can be best addressed by properly understanding that our data are composed of dynamic processes and then modeling them as such.

What do data composed of dynamic processes look like in practice? We now turn to several illustrations of interesting time series data from a variety of fields. To begin, an important idea in political science is that of representation in democracies. Among other things, this concept implies a recurring alternation of power between contending groups. Consider, for example, Figure 1.1, which shows the balance of power in the U.S. House of Representatives over the

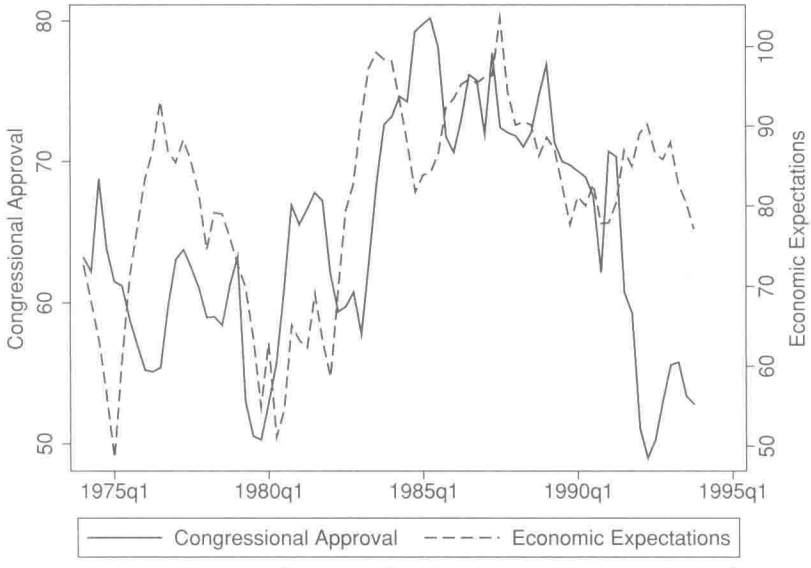


FIGURE 1.2. Congressional Approval and Economic Expectations in the U.S., 1974q1–1993q4. Compiled by the authors from replication data on website of Christina Wolbrecht (Durr, Gilmour, and Wolbrecht, 1997).

previous century.[1] The data are proportions of the seats held by the Democratic Party in the House ordered over time. Do the workings of our institutions produce regular shifts in the balance of power between the Republicans and Democrats? Is there a pattern of alternating control over time? The figure suggests that there might be, but the timing between changes in partisan control is not clear. There was a prolonged period of Democratic control in the middle of the century, but then power shifted to the Republicans. So conceivably there are recurring alterations of power between competing groups, but only over long periods of time. Time series methods help us characterize this process of representation and predict its behavior. We discuss methods for modeling and forecasting from a univariate time series like that in Figure 1.1 in Chapters 2, 5, and 7.

Patterns of political accountability, a concept closely related to representation, provide shorter term insights into this subject. Figure 1.2 shows time series for the proportion of citizens who approve of Congress in each quarter from the beginning of 1974 to the end of 1993. It also depicts the proportion of respondents who have positive expectations about the economy in the same time period. The two series appear to move together, suggesting that the public dispenses approval on the basis of how it evaluates economic policy. If so, the

[1] We denote annual time series by the first and last year of the series, for instance, 1985–1993. We denote quarterly data by the year and quarter, e.g., 1995q1 (see Figure 1.2 above). Monthly time series are denoted by the year and month of the observation, for example, 1997:5–2006:9. Data and STATA replication code for all figures and tables in this book are available at www.politicalscience.osu.edu/faculty/jbox/tsass.

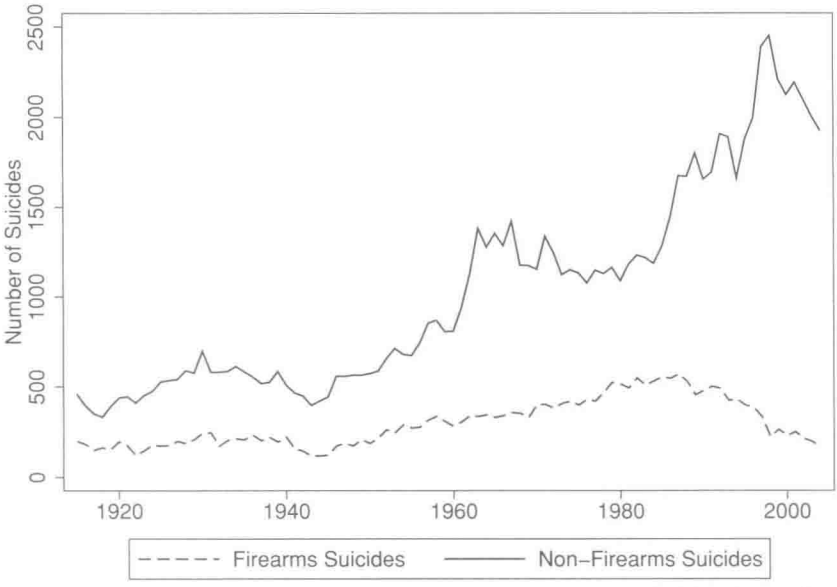


FIGURE 1.3. Suicide Deaths in Australia, 1914–2004. Compiled by the authors from replication data on website of Andrew Leigh (Neill and Leigh, 2008).

data indicates that Congress is held accountable for its policies. But what exactly is the nature of this relationship? Are the evaluations of the economy temporally prior to changes in representation? Is the effect immediate or lagged? Time series methods, such as those we discuss in Chapters 3, 4, and 6, enable us to describe exactly how these variables trend together.

Let us turn to a different field: since the pioneering work of one of the discipline's founding thinkers, Emile Durkheim, sociologists have studied the question of suicide. In very broad strokes, some sociologists study questions such as the following: What are the different types of suicidal behavior? What, if anything, can (or should) governments do to discourage people from committing suicide? Figure 1.3 shows deaths from firearms and non-firearms suicides in Australia from 1915–2004, first collected and analyzed by Neill and Leigh (2010). Non-firearms suicides seem to spike in the 1960s and again in the 1990s, whereas firearms suicides seem to decline in the mid-1990s. After several mass killings, the Australian government implemented a massive firearms buyback program in 1997 to reduce the number of guns in circulation in Australia. Did this program reduce the amount of suicides by firearms in Australia? Using time series intervention models (discussed in Chapter 2), and endogenous structural break models (discussed in Chapter 7), analysts can explore this and other similar questions.[2]

[2] For more on the sociology of suicide, Wray, Colen, and Pescosolido (2011) provide an excellent overview covering research predating Durkheim to modern times. Interested readers are

此为试读,需要完整PDF请访问: www.ertongbook.com