**AN R COMPANION FOR THE FUNDAMENTALS OF SOCIAL RESEARCH**

*An R Companion for The Fundamentals of Social Research* offers students the opportunity to delve into the world of R using real data sets and statistical analysis techniques directly from Paul M. Kellstedt, Guy D. Whitten, and Steven A. Tuch's new textbook. Workbook sections parallel chapters in the main text, giving students a chance to apply the lessons and techniques learned in each chapter in a statistical software setting. Detailed chapters teach students to reproduce results presented in the textbook, allowing them to become comfortable performing statistical analyses for evaluating causal claims through repeated practice. Step-by-step instructions for using R are provided, along with command lines and screenshots to demonstrate proper use of the software. Instructions for producing the figures and tables in the main text are integrated throughout the workbook. End-of-chapter exercises encourage students to formulate and evaluate their own hypotheses.

Paul M. Kellstedt is a professor of political science at Texas A&M University.

Guy D. Whitten is a professor of political science and Director of the European Union Center at Texas A&M University.

Steven A. Tuch is a professor of sociology and public policy and public administration at George Washington University.
AN R COMPANION FOR

The Fundamentals of Social Research

Paul M. Kellstedt
Texas A&M University

Guy D. Whitten
Texas A&M University

Steven A. Tuch
George Washington University
Brief Contents

List of Figures
Preface

1 The Scientific Study of Society ........................................ 1
2 The Art of Theory Building ............................................ 11
3 Evaluating Causal Relationships ..................................... 22
4 Research Design .......................................................... 25
5 Survey Research ............................................................ 28
6 Measuring Concepts of Interest ....................................... 29
7 Getting to Know Your Data ............................................. 32
8 Probability and Statistical Inference ............................... 40
9 Bivariate Hypothesis Testing .......................................... 48
10 Two-Variable Regression Models ................................. 59
11 Multiple Regression ....................................................... 63
12 Putting It All Together to Produce Effective Research ........ 76

References 78
Index 79
## Contents

List of Figures ................................. x
Preface ........................................ xiii

### 1 The Scientific Study of Society
1.1 Overview .................................. 1
1.2 “A Workbook? Why Is There a Workbook?”
   1.2.1 Reading Commands in This Workbook 2
1.3 Getting Started with R and RStudio
   1.3.1 Launching RStudio .................. 4
   1.3.2 Getting R to Do Things .......... 4
   1.3.3 Initially Examining Data in R 7
   1.3.4 Adding Notes to Script Files and Saving Them 9
1.4 Exercises ................................. 10

### 2 The Art of Theory Building
2.1 Overview .................................. 11
2.2 R Packages ................................ 11
2.3 Examining Variation Across Time and Across Space
   2.3.1 Producing a Bar Graph for Examining Cross-Sectional Variation 13
   2.3.2 Producing a Connected Plot for Examining Time-Series Variation 16
2.4 Using Google Scholar to Search the Literature Effectively 17
2.5 Wrapping Up ............................... 20
2.6 Exercises ................................. 20

### 3 Evaluating Causal Relationships
3.1 Overview .................................. 22
3.2 Exercises ................................. 22
## Contents

### 4 Research Design ........................................ 25
  4.1 Overview ........................................ 25  
  4.2 Exercises ........................................ 25  

### 5 Survey Research ......................................... 28
  5.1 Overview ........................................ 28  
  5.2 Exercises ........................................ 28  

### 6 Measuring Concepts of Interest .......................... 29
  6.1 Overview ........................................ 29  
  6.2 Exercises ........................................ 29  

### 7 Getting to Know Your Data .............................. 32
  7.1 Overview ........................................ 32  
  7.2 Describing Categorical and Ordinal Variables .......... 32  
  7.3 Describing Continuous Variables ....................... 36  
  7.4 Putting Statistical Output into Tables, Documents, and Presentations ................................ 38  
  7.5 Exercises ........................................ 38  

### 8 Probability and Statistical Inference .................. 40
  8.1 Overview ........................................ 40  
  8.2 Dice Rolling in R ................................... 40  
  8.3 Exercises ........................................ 46  

### 9 Bivariate Hypothesis Testing ............................ 48
  9.1 Overview ........................................ 48  
  9.2 Tabular Analysis .................................... 48  
  9.2.1 Generating Test Statistics ....................... 50  
  9.2.2 Putting Tabular Results into Papers ................ 51  
  9.3 Difference of Means ................................ 51  
  9.3.1 Examining Differences Graphically ............... 51  
  9.3.2 Generating Test Statistics ....................... 52  
  9.4 Correlation Coefficients ............................ 53  
  9.4.1 Producing Scatter Plots ........................ 53  
  9.4.2 Generating Covariance Tables and Test Statistics .............................................. 54  
  9.5 Analysis of Variance ................................ 55  
  9.5.1 Examining Differences Graphically ............... 55  
  9.5.2 Generating ANOVA Test Statistics ................ 55  
  9.6 Exercises ........................................ 58  

### 10 Two-Variable Regression Models ....................... 59
  10.1 Overview ........................................ 59  
  10.2 Estimating a Two-Variable Regression ................ 59
## Contents

10.3 Graphing a Two-Variable Regression  
10.4 Exercises  

### 11 Multiple Regression  
11.1 Overview  
11.2 Estimating a Multiple Regression  
11.3 From Regression Output to Table – Making Only One Type of Comparison  
11.3.1 Comparing Models with the Same Sample of Data, but Different Specifications  
11.3.2 Comparing Models with the Same Specification, but Different Samples of Data  
11.4 Standardized Coefficients  
11.5 Dummy Variables  
11.5.1 Creating a Dummy Variable in R  
11.5.2 Estimating a Multiple Regression Model with a Single Dummy Independent Variable  
11.5.3 Estimating a Multiple Regression Model with Multiple Dummy Independent Variables  
11.6 Dummy Variables in Interactions  
11.7 Models with Dummy Dependent Variables  
11.8 Exercises  

### 12 Putting It All Together to Produce Effective Research  
12.1 Overview  
12.2 Exercises  

References  
Index
Figures

1.1 RStudio initial launch  
1.2 Data set successfully loaded into R  
1.3 Script editor window open  
1.4 Preparing to execute a command from the script-file editor  
1.5 RStudio after having successfully run a command from the script-file editor  
1.6 Initially examining data in RStudio  
1.7 Script with comment line saved  
2.1 Packages installed by script  
2.2 Initial bar graph of Gini coefficients  
2.3 Initial bar graph of Gini coefficients with angled country names  
2.4 The Google Scholar home page  
2.5 Google Scholar results for “economic threat and attitudes toward immigrants” search  
7.1 Raw output from CrossTable command  
7.2 R pie graph  
7.3 R bar chart  
7.4 R summary statistics  
8.1 Simulating a single roll of a six-sided die  
8.2 Simulating sixty rolls of a six-sided die  
8.3 Descriptive statistics from sixty rolls of a six-sided die  
8.4 Simulating rolling two six-sided dice at once  
8.5 Simulating rolling two six-sided dice at once sixty times  
9.1 Raw output from CrossTable command  
9.2 Raw output from CrossTable command with a chi-squared test  
9.3 Raw output from a difference of means test  
9.4 Basic output from scatter command  
9.5 Output from cor.test command  
9.6 Box–whisker plot of racial/ethnic differences in education
List of Figures

9.7 ANOVA descriptive statistics 56
9.8 ANOVA output 57
9.9 Output from a Scheffe ANOVA multiple comparison test procedure 58
10.1 RStudio after the estimation of a two-variable regression model 60
10.2 Model summary output for a two-variable regression 61
11.1 Basic output from a multiple regression 64
11.2 Basic output from a multiple regression with the “beta” coefficients 67
11.3 Output from checking whether we correctly coded the dummy variable “female” 69
11.4 Output from a regression model with a dummy variable 70
11.5 Table of values for the variable “Bush” 71
11.6 Calculating and displaying the predicted values from the linear probability model 72
11.7 Summary statistics for predicted values from a binomial logit model and a binomial probit model 74
11.8 Raw output used to produce Table 11.12 of FSR 74
Preface

This software companion book represents an effort to provide both extra exercises, as well as hands-on material for how to put the techniques that we discuss in *The Fundamentals of Social Research* into action. It is one of three workbooks, each written to help students to work with the materials covered in *The Fundamentals of Social Research* using a particular piece of statistical software.

This workbook focuses on using the program R from another program, RStudio. (The other companion books are designed to work with SPSS and Stata.) Our expectation is that the typical user of this book will be using a relatively recent version of R and RStudio. We have written this workbook on a computer running the Windows operating system, but, with a very small number of exceptions, it should work identically for users of different operating systems. An online appendix available at www.cambridge.org/fsr provides users with instructions for installing R and RStudio and other technical issues relating to the use of these programs for the purposes covered in this book.

The chapter structure of this workbook mirrors the chapter structure of *The Fundamentals of Social Research*. We have written with the expectation that students will read the chapters of this companion after they have read the chapters of the book.

We owe a special thank you to Andrea Junqueira for her help with the writing of R code for this workbook.

We continue to update both the general and instructor-only sections of the webpage for our book (www.cambridge.org/fsr). As before, the general section contains data sets available in formats compatible with SPSS, Stata, and R. The instructor-only section contains several additional resources, including PowerPoint and \TeX/Beamer slides for each chapter, a test-bank, and answer keys for the exercises.