

# An R Companion for the Third Edition of The Fundamentals of Political Science Research

*An R Companion for the Third Edition of The Fundamentals of Political Science Research* offers students a chance to delve into the world of R using real political data sets and statistical analysis techniques directly from Paul M. Kellstedt and Guy D. Whitten's best-selling textbook. Built in parallel with the main text, this workbook teaches students to apply the techniques they learn in each chapter by reproducing the analyses and results from each lesson using R. Students will also learn to create all of the tables and figures found in the textbook, leading to an even greater mastery of the core material. This accessible, informative, and engaging companion walks through the use of R step-by-step, using command lines and screenshots to demonstrate proper use of the software. With the help of these guides, students will become comfortable creating, editing, and using data sets in R to produce original statistical analyses for evaluating causal claims. End-of-chapter exercises encourage this innovation by asking students to formulate and evaluate their own hypotheses.

Paul M. Kellstedt is Professor of Political Science at Texas A&M University. He is the author of *The Mass Media and the Dynamics of American Racial Attitudes* (Cambridge, 2003), winner of Harvard University's John F. Kennedy School of Government's 2004 Goldsmith Book Prize. In addition, he has published numerous articles in a variety of leading journals. He is the Editor-in-chief of *Political Science Research and Methods*, the flagship journal of the European Political Science Association.

Guy D. Whitten is Cullen-McFadden Professor of Political Science, as well as Director of the European Union Center, at Texas A&M University. He has published a variety of articles in leading peer-reviewed journals. He is on the editorial boards of *Political Analysis* and *Political Science Research and Methods*.



# **An R Companion for the Third Edition of The Fundamentals of Political Science Research**

Paul M. Kellstedt  
Texas A&M University

Guy D. Whitten  
Texas A&M University



**CAMBRIDGE**  
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

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](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781108446037](http://www.cambridge.org/9781108446037)

DOI: 10.1017/9781108601832

© Paul M. Kellstedt and Guy D. Whitten 2021

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 2021

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

ISBN 978-1-108-44603-7 Paperback

Additional resources for this publication at [www.cambridge.org/FPSR-R](http://www.cambridge.org/FPSR-R)

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

## BRIEF CONTENTS

	<i>Preface</i>	page xi
	<i>List of Figures</i>	xiii
<b>1</b>	<b>The Scientific Study of Politics</b>	1
<b>2</b>	<b>The Art of Theory Building</b>	9
<b>3</b>	<b>Evaluating Causal Relationships</b>	19
<b>4</b>	<b>Research Design</b>	22
<b>5</b>	<b>Measuring Concepts of Interest</b>	24
<b>6</b>	<b>Getting to Know Your Data</b>	27
<b>7</b>	<b>Probability and Statistical Inference</b>	35
<b>8</b>	<b>Bivariate Hypothesis Testing</b>	42
<b>9</b>	<b>Two-Variable Regression Models</b>	50
<b>10</b>	<b>Multiple Regression: The Basics</b>	53
<b>11</b>	<b>Multiple Regression Model Specification</b>	58
<b>12</b>	<b>Limited Dependent Variables and Time-Series Data</b>	66
	<i>Bibliography</i>	71
	<i>Index</i>	73



## CONTENTS

	<i>Preface</i>	page xi
	<i>List of Figures</i>	xiii
<b>1</b>	<b>The Scientific Study of Politics</b>	1
	1.1 Overview	1
	1.2 “A Workbook? Why Is There a Workbook?”	1
	1.2.1 Reading Commands in This Workbook	2
	1.3 Getting Started with R and RStudio	2
	1.3.1 Launching RStudio	3
	1.3.2 Getting R to Do Things	3
	1.3.3 Initially Examining Data in R	6
	1.3.4 Adding Notes to Script Files and Saving Them	7
	1.4 Exercises	8
<b>2</b>	<b>The Art of Theory Building</b>	9
	2.1 Overview	9
	2.2 R Packages	9
	2.3 Examining Variation across Time and across Space	10
	2.3.1 Producing a Bar Graph for Examining Cross-Section Variation	11
	2.3.2 Producing a Connected Plot for Examining Time-Series Variation	13
	2.4 Using Google Scholar to Search the Literature Effectively	14
	2.5 Wrapping Up	18
	2.6 Exercises	18
<b>3</b>	<b>Evaluating Causal Relationships</b>	19
	3.1 Overview	19
	3.2 Exercises	19
<b>4</b>	<b>Research Design</b>	22
	4.1 Overview	22
	4.2 Exercises	22

<b>5</b>	<b>Measuring Concepts of Interest</b>	24
	5.1 Overview	24
	5.2 Exercises	24
<b>6</b>	<b>Getting to Know Your Data</b>	27
	6.1 Overview	27
	6.2 Describing Categorical and Ordinal Variables	27
	6.3 Describing Continuous Variables	30
	6.4 Putting Statistical Output into Tables, Documents, and Presentations	32
	6.5 Exercises	33
<b>7</b>	<b>Probability and Statistical Inference</b>	35
	7.1 Overview	35
	7.2 Dice Rolling in R	35
	7.3 Exercises	40
<b>8</b>	<b>Bivariate Hypothesis Testing</b>	42
	8.1 Overview	42
	8.2 Tabular Analysis	42
	8.2.1 Generating Test Statistics	43
	8.2.2 Putting Tabular Results into Papers	44
	8.3 Difference of Means	45
	8.3.1 Examining Differences Graphically	45
	8.3.2 Generating Test Statistics	45
	8.4 Correlation Coefficients	46
	8.4.1 Producing Scatter Plots	46
	8.4.2 Generating Covariance Tables and Test Statistics	47
	8.5 Exercises	48
<b>9</b>	<b>Two-Variable Regression Models</b>	50
	9.1 Overview	50
	9.2 Estimating a Two-Variable Regression	50
	9.3 Graphing a Two-Variable Regression	51
	9.4 Exercises	52
<b>10</b>	<b>Multiple Regression: The Basics</b>	53
	10.1 Overview	53
	10.2 Estimating a Multiple Regression	53
	10.3 From Regression Output to Table – Making Only One Type of Comparison	54
	10.3.1 Comparing Models with the Same Sample of Data, but Different Specifications	54
	10.3.2 Comparing Models with the Same Specification, but Different Samples of Data	55



## CONTENTS

ix

---

10.4	Standardized Coefficients	56
10.5	Exercises	56
<b>11</b>	<b>Multiple Regression Model Specification</b>	<b>58</b>
11.1	Overview	58
11.2	Dummy Variables	58
11.2.1	Creating New Variables	58
11.2.2	Estimating a Multiple Regression Model with a Single Dummy Independent Variable	60
11.2.3	Estimating a Multiple Regression Model with Multiple Dummy Independent Variables	60
11.3	Dummy Variables in Interactions	61
11.4	Post-estimation Diagnostics in R for OLS	61
11.4.1	Identifying Outliers and Influential Cases in OLS	61
11.5	Exercises	64
<b>12</b>	<b>Limited Dependent Variables and Time-Series Data</b>	<b>66</b>
12.1	Overview	66
12.2	Models with Dummy Dependent Variables	66
12.3	Exercises	70
	<i>Bibliography</i>	71
	<i>Index</i>	73



## Preface

We received a wealth of useful feedback from instructors and students about the first two editions of *The Fundamentals of Political Science Research*. In response to feedback on the first edition, we substantially increased the number of end-of-chapter exercises in the second edition. While the response to this increase was positive, we sensed a demand for even more exercises and, in particular, more hands-on material for how to put the techniques that we discuss in the book into action. This workbook is our attempt to meet these demands. It is one of three workbooks, each written to help students to work with the materials covered in *The Fundamentals of Political Science Research, Third Edition* using a particular piece of statistical software.

This workbook focuses on using the program R from another program, RStudio. Our expectation is that the modal user of this book will be using a relatively recent version of R and RStudio.

The chapter structure of this workbook mirrors the chapter structure of the third edition of *The Fundamentals of Political Science 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/fpsr](http://www.cambridge.org/fpsr)). 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 T<sub>E</sub>X/Beamer slides for each chapter, a test-bank, and answer keys for the exercises.



## FIGURES

1.1	RStudio initial launch	<i>page</i> 3
1.2	Data set successfully loaded into R	4
1.3	Script editor window open	5
1.4	Preparing to execute a command from the script-file editor	6
1.5	RStudio after having successfully run a command from the script-file editor	6
1.6	Initially examining data in RStudio	7
1.7	Script with comment line saved	8
2.1	Packages installed by script	10
2.2	Initial bar graph of military spending	12
2.3	Initial bar graph of military spending with angled nation names	13
2.4	The Google Scholar home page	15
2.5	Google Scholar results for “Ray Fair economic voting” search	16
2.6	Google Scholar search results of articles citing Ray Fair (1978) article	17
6.1	Raw output from CrossTable command	28
6.2	R pie graph	29
6.3	R bar chart	30
6.4	R summary statistics	31
7.1	Simulating a single roll of a six-sided die	36
7.2	Simulating 60 rolls of a six-sided die	36
7.3	Descriptive statistics from 60 rolls of a six-sided die	37
7.4	Simulating rolling two six-sided dice at once	39
7.5	Simulating rolling two six-sided dice at once 60 times	40
8.1	Raw output from CrossTable command	43
8.2	Raw output from CrossTable command with a chi-squared test	44
8.3	Raw output from a difference of means test	46
8.4	Basic output from scatter command	47
8.5	Output from cor.test command	48
9.1	RStudio after the estimation of a two-variable regression model	51
9.2	Model summary output for a two-variable regression	51
10.1	Basic output from a multiple regression	54
10.2	Basic output from a multiple regression with the “beta” coefficients	56

11.1 Output from checking whether we correctly coded the dummy variable female	59
11.2 Output from a regression model with a dummy variable	60
11.3 Output from a regression model with an interaction	61
11.4 Output from a regression model with an interaction	62
11.5 Examining DFBETA values	63
11.6 Output from a regression model and a vif command	64
12.1 Table of values for the variable “Bush”	67
12.2 Calculating and displaying the predicted values from the linear probability model	67
12.3 Summary statistics for predicted values from a binomial logit model and a binomial probit model	69
12.4 Raw output used to produce Table 12.3 of <i>FPSR</i>	70