Contents

List of tables xv

List of figures xvii

Preface xxvii

Acknowledgments xxix

1 Introduction 1

1.1 Overview of the book . . . 1

1.2 Getting the most out of this book . . . 3

1.3 Downloading the example datasets and programs . . . 4

1.4 The GSS dataset . . . 4

1.4.1 Income . . . 5

1.4.2 Age . . . 6

1.4.3 Education . . . 10

1.4.4 Gender . . . 12

1.5 The pain datasets . . . 12

1.6 The optimism datasets . . . 12

1.7 The school datasets . . . 13

1.8 The sleep datasets . . . 13

I Continuous predictors 15

2 Continuous predictors: Linear 17

2.1 Chapter overview . . . 17

2.2 Simple linear regression . . . 17

2.2.1 Computing predicted means using the margins command . . 20

2.2.2 Graphing predicted means using the marginsplot command 22

2.3 Multiple regression . . . 25

2.3.1 Computing adjusted means using the margins command . . 26
2.3.2 Some technical details about adjusted means 28
2.3.3 Graphing adjusted means using the marginsplot command . 29
2.4 Checking for nonlinearity graphically 30
2.4.1 Using scatterplots to check for nonlinearity 31
2.4.2 Checking for nonlinearity using residuals 31
2.4.3 Checking for nonlinearity using locally weighted smoother . 33
2.4.4 Graphing outcome mean at each level of predictor 34
2.4.5 Summary 37
2.5 Checking for nonlinearity analytically 37
2.5.1 Adding power terms 38
2.5.2 Using factor variables 39
2.6 Summary 43
3 Continuous predictors: Polynomials 45
3.1 Chapter overview 45
3.2 Quadratic (squared) terms 45
3.2.1 Overview 45
3.2.2 Examples 49
3.3 Cubic (third power) terms 55
3.3.1 Overview 55
3.3.2 Examples 56
3.4 Fractional polynomial regression 62
3.4.1 Overview 62
3.4.2 Example using fractional polynomial regression 66
3.5 Main effects with polynomial terms 75
3.6 Summary 77
4 Continuous predictors: Piecewise models 79
4.1 Chapter overview 79
4.2 Introduction to piecewise regression models 80
4.3 Piecewise with one known knot 82

4.3.1 Overview . . . 82
4.3.2 Examples using the GSS . . . 83
4.4 Piecewise with two known knots . . . 91
4.4.1 Overview . . . 91
4.4.2 Examples using the GSS . . . 91
4.5 Piecewise with one knot and one jump . . . 96
4.5.1 Overview . . . 96
4.5.2 Examples using the GSS . . . 97
4.6 Piecewise with two knots and two jumps . . . 102
4.6.1 Overview . . . 102
4.6.2 Examples using the GSS . . . 102
4.7 Piecewise with an unknown knot . . . 109
4.8 Piecewise model with multiple unknown knots . . . 113
4.9 Piecewise models and the marginsplot command . . . 120
4.10 Automating graphs of piecewise models . . . 123
4.11 Summary . . . 126
5 Continuous by continuous interactions 127
5.1 Chapter overview . . . 127
5.2 Linear by linear interactions . . . 127
5.2.1 Overview . . . 127
5.2.2 Example using GSS data . . . 132
5.2.3 Interpreting the interaction in terms of age . . . 133
5.2.4 Interpreting the interaction in terms of education . . . 135
5.2.5 Interpreting the interaction in terms of age slope . . . 137
5.2.6 Interpreting the interaction in terms of the educ slope . . . 138
5.3 Linear by quadratic interactions . . . 140
5.3.1 Overview . . . 140
5.3.2 Example using GSS data . . . 143
5.4 Summary . . . 148

6 Continuous by continuous by continuous interactions 149

6.1 Chapter overview . . . 149

6.2 Overview . . . 149

6.3 Examples using the GSS data . . . 154

6.3.1 A model without a three-way interaction . . . 154

6.3.2 A three-way interaction model . . . 158

6.4 Summary . . . 164

II Categorical predictors 165

7 Categorical predictors 167

7.1 Chapter overview . . . 167

7.2 Comparing two groups using a t test . . . 168

7.3 More groups and more predictors . . . 169

7.4 Overview of contrast operators . . . 175

7.5 Compare each group against a reference group . . . 176

7.5.1 Selecting a specific contrast . . . 177

7.5.2 Selecting a different reference group . . . 178

7.5.3 Selecting a contrast and reference group . . . 179

7.6 Compare each group against the grand mean . . . 179

7.6.1 Selecting a specific contrast . . . 181

7.7 Compare adjacent means . . . 182

7.7.1 Reverse adjacent contrasts . . . 185

7.7.2 Selecting a specific contrast . . . 186

7.8 Comparing the mean of subsequent or previous levels . . . 187

7.8.1 Comparing the mean of previous levels . . . 191

7.8.2 Selecting a specific contrast . . . 192

7.9 Polynomial contrasts . . . 193

7.10 Custom contrasts . . . 195

7.11 Weighted contrasts . . . 198

7.12 Pairwise comparisons . . . 200

7.13 Interpreting confidence intervals 202
7.14 Testing categorical variables using regression 205
7.15 Summary 208
8 Categorical by categorical interactions 209
8.1 Chapter overview 209
8.2 Two by two models: Example 1 211
8.2.1 Simple effects 215
8.2.2 Estimating the size of the interaction 216
8.2.3 More about interaction 217
8.2.4 Summary 218
8.3 Two by three models 218
8.3.1 Example 2 218
8.3.2 Example 3 223
8.3.3 Summary 228
8.4 Three by three models: Example 4 228
8.4.1 Simple effects 230
8.4.2 Simple contrasts 231
8.4.3 Partial interaction 233
8.4.4 Interaction contrasts 234
8.4.5 Summary 236
8.5 Unbalanced designs 236
8.6 Main effects with interactions: anova versus regress 241
8.7 Interpreting confidence intervals 244
8.8 Summary 246
9 Categorical by categorical by categorical interactions 249
9.1 Chapter overview 249
9.2 Two by two by two models 250
9.2.1 Simple interactions by season 252
9.2.2 Simple interactions by depression status 253
9.2.3 Simple effects 255

9.3 Two by two by three models 255
9.3.1 Simple interactions by depression status 258
9.3.2 Simple partial interaction by depression status 258
9.3.3 Simple contrasts 260
9.3.4 Partial interactions 260
9.4 Three by three by three models and beyond 262
9.4.1 Partial interactions and interaction contrasts 264
9.4.2 Simple interactions 268
9.4.3 Simple effects and simple comparisons 271
9.5 Summary 272

III Continuous and categorical predictors 273

10 Linear by categorical interactions 275

10.1 Chapter overview 275
10.2 Linear and two-level categorical: No interaction 275
10.2.1 Overview 275
10.2.2 Examples using the GSS 278
10.3 Linear by two-level categorical interactions 283
10.3.1 Overview 283
10.3.2 Examples using the GSS 285
10.4 Linear by three-level categorical interactions 290
10.4.1 Overview 290
10.4.2 Examples using the GSS 293
10.5 Summary 299

11 Polynomial by categorical interactions 301

11.1 Chapter overview 301
11.2 Quadratic by categorical interactions 301
11.2.1 Overview 302
11.2.2 Quadratic by two-level categorical 305
11.2.3 Quadratic by three-level categorical 312

11.3 Cubic by categorical interactions 318
11.4 Summary 323
12 Piecewise by categorical interactions 325
12.1 Chapter overview 325
12.2 One knot and one jump 328
12.2.1 Comparing slopes across gender 332
12.2.2 Comparing slopes across education 333
12.2.3 Difference in differences of slopes 333
12.2.4 Comparing changes in intercepts 334
12.2.5 Computing and comparing adjusted means 334
12.2.6 Graphing adjusted means 337
12.3 Two knots and two jumps 341
12.3.1 Comparing slopes across gender 346
12.3.2 Comparing slopes across education 347
12.3.3 Difference in differences of slopes 348
12.3.4 Comparing changes in intercepts by gender 349
12.3.5 Comparing changes in intercepts by education 350
12.3.6 Computing and comparing adjusted means 351
12.3.7 Graphing adjusted means 354
12.4 Comparing coding schemes 356
12.4.1 Coding scheme #1 356
12.4.2 Coding scheme #2 358
12.4.3 Coding scheme #3 360
12.4.4 Coding scheme #4 361
12.4.5 Choosing coding schemes 363
12.5 Summary 364
13 Continuous by continuous by categorical interactions 365
13.1 Chapter overview 365
13.2 Linear by linear by categorical interactions 366
13.2.1 Fitting separate models for males and females 366

13.2.2 Fitting a combined model for males and females 368
13.2.3 Interpreting the interaction focusing in the age slope 370
13.2.4 Interpreting the interaction focusing on the educ slope . . . 372
13.2.5 Estimating and comparing adjusted means by gender 374
13.3 Linear by quadratic by categorical interactions 376
13.3.1 Fitting separate models for males and females 376
13.3.2 Fitting a common model for males and females 378
13.3.3 Interpreting the interaction 379
13.3.4 Estimating and comparing adjusted means by gender 380
13.4 Summary 382
14 Continuous by categorical by categorical interactions 383
14.1 Chapter overview 383
14.2 Simple effects of gender on the age slope 387
14.3 Simple effects of education on the age slope 388
14.4 Simple contrasts on education for the age slope 389
14.5 Partial interaction on education for the age slope 389
14.6 Summary 390
IV Beyond ordinary linear regression 391
15 Multilevel models 393
15.1 Chapter overview 393
15.2 Example 1: Continuous by continuous interaction 394
15.3 Example 2: Continuous by categorical interaction 397
15.4 Example 3: Categorical by continuous interaction 401
15.5 Example 4: Categorical by categorical interaction 404
15.6 Summary 408
16 Time as a continuous predictor 411
16.1 Chapter overview 411
16.2 Example 1: Linear effect of time 412
16.3 Example 2: Linear effect of time by a categorical predictor 416

16.4 Example 3: Piecewise modeling of time 421
16.5 Example 4: Piecewise effects of time by a categorical predictor . . . 426
16.5.1 Baseline slopes 430
16.5.2 Change in slopes: Treatment versus baseline 431
16.5.3 Jump at treatment 432
16.5.4 Comparisons among groups 433
16.6 Summary 434
17 Time as a categorical predictor 437
17.1 Chapter overview 437
17.2 Example 1: Time treated as a categorical variable 438
17.3 Example 2: Time (categorical) by two groups 443
17.4 Example 3: Time (categorical) by three groups 447
17.5 Comparing models with different residual covariance structures . . . 452
17.6 Summary 454
18 Nonlinear models 455
18.1 Chapter overview 455
18.2 Binary logistic regression 456
18.2.1 A logistic model with one categorical predictor 456
18.2.2 A logistic model with one continuous predictor 463
18.2.3 A logistic model with covariates 465
18.3 Multinomial logistic regression 470
18.4 Ordinal logistic regression 475
18.5 Poisson regression 478
18.6 More applications of nonlinear models 481
18.6.1 Categorical by categorical interaction 481
18.6.2 Categorical by continuous interaction 487
18.6.3 Piecewise modeling 492
18.7 Summary 498
19 Complex survey data 499

V Appendices **505**

A The margins command **507**

A.1 The predict() and expression() options 507

A.2 The at() option 510

A.3 Margins with factor variables 513

A.4 Margins with factor variables and the at() option 517

A.5 The dydx() and related options 519

B The marginsplot command **523**

C The contrast command **535**

D The pwcompare command **539**

References **545**

Author index **549**

Subject index **551**