

Subject index

A

abbreviations 11
ado-files 656
aggregate data, binary outcome models
..... 487
AIC ... see Akaike information criterion
Akaike information criterion 359
 count-data example 598, 605
AME see marginal effects, average
 marginal effect
append command 58
areg command 259
arithmetic operators 12
asclogit command 505
asmprobit command 519
average marginal effect ... see marginal
 effects

B

batch mode for Stata 14
Bayesian information criterion 359
 count-data example 598, 605
BIC see Bayesian information
 criterion
binary outcome models 459–489
 Bernoulli trial 459
 binomial regression 459
 classification test 473
 coefficient interpretation 465
 complementary log-log model . 460
 endogenous regressors 479
 generalized logit model 468
 goodness-of-fit measures 471
 goodness-of-fit test 472
 grouped-data analysis 486
 heteroskedastic probit model .. 469

binary outcome models, *continued*

 hypothesis tests 466
 instrumental-variables probit .. 481
 latent-variable model 461
 linear probability model .. 460, 462
 logit model 460, 464
 marginal effects 476
 maximum likelihood estimation ...
 461
 model comparison 465, 470
 panel-data estimators 621
 predicted outcome 473
 predicted probabilities 474
 probit model 460
 pseudo- R^2 471
 robust variance estimation 462
 specification analysis 468
 two-stage least squares 485
biprobit command 530
bootstrap command 438
bootstrap methods 429–457
 asymptotic refinement ... 430, 445
 asymptotically pivotal statistic ...
 445
 BCa confidence interval ... 436, 445
 bias estimation 437
 bias-corrected confidence interval ..
 436
 bootstrap pairs 431
 bootstrap- t method 445
 case bootstrap 432
 clustered bootstrap 434
 confidence-intervals example .. 437
 Hausman test 443
 jackknife 455
 misuse of bootstrap 430
 nonparametric bootstrap 432

- bootstrap methods, *continued*
 - normal-based confidence interval..
 - 435
 - number of bootstraps 433
 - paired bootstrap 432
 - parametric bootstrap 452
 - percentile-*t* confidence interval....
 - 436, 447
 - percentile-*t* method 445
 - percentile-*t* Wald test 446
 - quantile regression variance matrix
 - 214
 - residual bootstrap 453
 - standard-error estimation.... 429,
 - 432
 - subsampling 455
 - two-step estimator 441
 - variance of standard error 440
 - wild bootstrap 454
 - boxcox command 99
 - bsample command 448
 - bsqreg command 214
 - by prefix command 49
 - bysort prefix command 50
- C**
- capture command 652
 - case sensitivity 11
 - cd command 13
 - censored data see tobit model
 - centile command 215
 - chi-squared goodness-of-fit test see
 - specification tests
 - cii command 421
 - clad command 545
 - clear command 24, 34
 - clogit command 506
 - cloglog command 462
 - clustered data 312–317
 - cluster-robust variance matrix 84,
 - 335
 - coefficient interpretation
 - in marginal effects 344
 - in multinomial outcome models...
 - 500, 507
 - coefficient interpretation, *continued*
 - in single-index models 345
 - collapse command 488
 - comma-separated values data ... 32, 37
 - comments in Stata commands 16
 - condivreg command 202
 - confidence intervals
 - asymmetric confidence interval ...
 - 412
 - bootstrap confidence intervals
 - 435
 - bootstrap confidence-intervals
 - example 437
 - percentile-*t* confidence interval
 - 447
 - Wald confidence-interval definition
 - 410, 411
 - Wald confidence-interval example..
 - 412
 - constraint command 89, 168
 - constraints() option 89
 - continue command 27
 - correlate command 86, 252
 - count-data models 567–613
 - comparison of four distributions ..
 - 589
 - endogenous regressors 605
 - equidispersion 568
 - excess zeros 584
 - finite-mixture model 589
 - fitted probabilities 579, 594
 - frequency distribution 571
 - generalized negative binomial
 - regression 581
 - goodness of fit 579
 - hurdle model 583
 - interpretation of coefficients... 576
 - main features of count data ... 567
 - marginal effects 576, 578, 588,
 - 594, 601
 - model selection 598, 604
 - NB1 model 569
 - NB2 model 569
 - negative binomial (NB) model
 - 569, 577

- count-data models, *continued*
 - nonlinear least squares 582
 - overdispersion 569
 - overfitting 599
 - panel-data estimators 633
 - Poisson model 321, 568
 - quantile regression 226
 - robust variance estimation for Poisson 574
 - test of endogeneity 609
 - test of overdispersion 575
 - two-part model 583
 - unobserved heterogeneity 569
 - zero-inflated models 599
 - zero-truncated models 586
- countfit** command 579, 604
- creturn** command 17
- critical values
 - chi-squared compared with F 400
 - computation of 402
 - standard normal compared with t 400
- D**
- data-generating process 119
- data management 31–58
 - appending datasets 58
 - common pitfalls 40
 - data transformations 48
 - data types 31
 - demeaned variables 52
 - dictionary file 39
 - imputing missing values 47
 - indicator variables 50
 - inputting data 34
 - interaction variables 51
 - labeling variables 43
 - long form 55
 - merging datasets 56
 - missing values 45
 - naming variables 43
 - ordering data 55
 - preserving and restoring data 55
 - PSID example 40
 - saving data 52
- data management, *continued*
 - selecting sample 53
 - special data formats 37
 - wide form 55
- dataset description
 - fishing-mode multinomial data 494
 - HRS private health insurance 463
 - MEPS 2001 ambulatory
 - expenditures 538
 - MEPS 2002 doctor visits 320
 - MEPS 2003 doctor visits 571
 - MEPS emergency room visits 600
 - MEPS drug expenditures 164, 184
 - MEPS medical expenditures 73
 - NHANES II survey design data 170
 - PSID earnings example 40
 - PSID panel wage data 240
 - Rand Health Insurance
 - Experiment 618
 - U.S. state panel cigarette data 272
 - Vietnam clustered village data 312
- delimit** command 16
- describe** command 74
- descriptive statistics 75–81
- destring** command 33
- do** command 14
- do-file 13
- drawnorm** command 135
- drop** command 54
- E**
- e-class commands 19, 441, 654
- egen** command 49
- elasticities
 - in linear regression model 107
 - in nonlinear models 344, 351
- endogeneity test *see* specification tests
- endogenous regressors *also see* instrumental variables
 - binary outcome model 479
 - definition 178

- endogenous regressors, *continued*
- dynamic panel-data model 295
 - fixed-effects model 237, 257
 - Hausman test 426, 443
 - Hausman–Taylor model 290
 - in count-data models 605
 - in tobit model 544
 - instrumental-variables regression 180
 - nonlinear instrumental-variables regression 392
 - panel instrumental variables 288
 - simulation example 148
 - tobit model 544
- ereturn command 651
- ereturn display command 116
- ereturn list command 19
- error messages 12, 658
- estadd command 91
- estat abond command 300
- estat alternatives command 508
- estat bootstrap command 436
- estat classification command 473
- estat commands 417
- estat correlation command 522
- estat covariance command 522
- estat endogenous command 189, 191
- estat firststage command 197
- estat gof command 426, 472
- estat hettest command 100, 159
- estat ic command 359
- estat imtest command 102, 426
- estat mfx command 509, 522
- estat override command 426
- estat ovtest command 99
- estat sargan command 301
- estat vce command 405
- estimates drop command 92
- estimates stats command 89
- estimates store command 89
- estimates table command 89
- estimation commands
- summary of linear panel commands 240
- estimation commands, *continued*
- summary of multinomial commands 494
 - summary of nonlinear panel commands 618
 - summary of Stata commands 319
- estout command 92
- eststo command 92
- esttab command 91
- F**
- factor variables 9–11, 50–52, 92–94, 354–357
- findit command 5
- finite-mixture model see count-data models
- fixed effects see panel data
- floating-point data 32
- fmm command 592
- foreach command 25, 353, 551
- formats to display data 33
- forvalues command 26, 131, 252
- fsun command 77
- G**
- generalized least squares 153–169
- cluster–robust variance matrix 155
 - efficient estimation 154
 - feasible GLS estimation 153
 - FGLS heteroskedastic errors example 160
 - FGLS systems example 164
 - GLS estimation 153
 - heteroskedastic errors example 156
 - robust variance matrix 155
 - seemingly unrelated regressions 162
 - weighted LS estimator 155, 162
 - working matrix 155
- generalized linear models
- definition 327
 - Poisson example 328

- generalized method of moments
 count-data example 610
 nonlinear example 392, 393
 panel Arellano–Bond estimator . . .
 295
- generate** command 48
- gettoken** command 655
- Gibbs sampler 137
- GLM see generalized linear models
- glm** command 327
- global** command 21
- global** macros 21, 24
- glogit** command 489
- gmm** command 328–330, 610–611
- gnbreg** command 582
- gologit2** command 529
- goodness-of-fit measures 357
- graph box** command 61
- graph combine** command . . 60, 80, 158
- graph export** command 60
- graph matrix** command 69
- graph save** command 60
- graph twoway** command 59
- graphs 58–70
 binary outcome plot 474
 box-and-whisker plot 61
 combining graphs 60, 158
 graph formats 61
 graphing commands 59
 histogram 63
 kernel density plot 63, 80
 line plot 66
 multiple scatterplots 69
 nonparametric regression 67
 outlier detection 96
 panel-data plots 247
 plotting known density 400
 residual plots 95, 158
 scatterplot 66
 Stata Graph Editor 61
- grouped data, binary outcome models
 486
- grqreg** command 221
- H**
- hausman** command 189, 267, 427
- Hausman test see specification tests
- heckman** command 558
- help** command 3
- help mata** command 663
- hetprob** command 469
- hierarchical models 316
- histogram** command 63
- hnblogit** command 587
- hplogit** command 585
- hsearch** command 5
- hurdle model see count-data models
- hypothesis tests 399–428
 auxiliary regression computation . .
 417
 binary outcome models 466
 bootstrap methods 429
 chi-squared compared with F
 400, 403
 computation of p -values 402
 computation of critical values . . 402
 delta method 409
 invariance under transformation . .
 413
 Lagrange multiplier statistic defini-
 tion 416
 Lagrange multiplier test example . .
 101, 468
 likelihood-ratio statistic definition
 413
 likelihood-ratio statistic direct
 computation 416
 likelihood-ratio statistic example . .
 415
 linear hypotheses 403
 linear regression model 88
 nonlinear hypotheses 409
 one-sided tests 408
 percentile- t Wald test 446
 power definition 421
 power in more than one direction
 102
 power using noncentral chi-squared
 424

hypothesis tests, *continued*

- power using simulation ... 146, 422
- score test 416
- size definition 420
- size using simulation 145, 420
- standard normal compared with t
..... 400, 403
- test of cross-equation restrictions..
..... 167
- test of heteroskedasticity 158,
219
- Wald statistic definition .. 405, 409
- Wald statistic examples .. 406, 409

I

- `idcluster()` option 435
- `if` qualifier 6
- `in` qualifier 6
- indicator variables 50
- `infile` command 38
- `infix` command 38
- influential observations 96
- information matrix test see
specification tests
- `input` command 36
- `insheet` command 37
- instrumental variables 177–209
 - 2SLS example 186
 - 3SLS example 207
 - basic theory 177
 - bias due to weak instruments.. 196
 - binary outcome model 479
 - diagnostics for weak instruments..
..... 195
 - finite-sample properties 194
 - first-stage F statistic 196, 198
 - first-stage equation 179
 - generalized method of moments
(GMM) estimator 181
 - Hansen test 191
 - in binary outcome models 485
 - in count-data models 605
 - inconsistency of OLS 178

instrumental variables, *continued*

- IV estimator 180
- jackknife IV (JIVE) estimator .. 204,
205
- just-identified model 180, 186,
197
- limited-information ML (LIML) es-
timator 204,
205
- optimal GMM estimator .. 181, 183
- optimal GMM example 187
- overidentified model 180, 187,
199
- panel Arellano–Bond estimator ...
..... 293
- panel Hausman–Taylor estimator ..
..... 290
- panel IV estimator 287
- partial R^2 196, 198
- robust variance matrix 182
- Sargan test 191
- simple example 179
- size distortion due to weak instru-
ments 197
- structural equation 179
- test of endogeneity 188
- test of overidentifying restrictions
..... 191
- test of weak instruments.. 196, 198
- three-stage least-squares (3SLS) es-
timator 206
- two-stage least-squares (2SLS) esti-
mator 180, 183,
295
- underidentified model 180
- valid instruments 181
- weak instruments... 181, 194, 202
- interaction variables 51
- interactive use of Mata 662
- interactive use of Stata 1
- intraclass correlation 253
- `intreg` command 545
- `invnormal()` function 122
- `ipshin` command 279
- `iqreg` command 214

- iterative methods 363–397
 - checking analytical derivatives
 - 378
 - checking program 377
 - generalized method-of-moments
 - example in Mata 394
 - gradient methods 368
 - linear program simplex method ...
 - 213
 - Mata example 365
 - Mata optimization 389
 - maximization options 367
 - messages during iterations 369
 - m1 method d0 386
 - m1 method d1 386
 - m1 method d2 388
 - m1 method lf 372
 - m1 method lf0 383
 - m1 method lf1 387–388
 - m1 method lf2 388–389
 - multicollinearity 379
 - multiple optimums 369, 380
 - Newton–Raphson algorithm... 364
 - Newton–Raphson example in Mata
 - 365
 - NLIV example in Mata 394
 - NLS example 376
 - numerical derivatives 370
 - optimize() method d2 395
 - optimize() method gf2 390
 - Poisson example in Mata 390
 - Poisson gradient and Hessian.. 364
 - robust estimate of the VCE.... 387
 - single-index model example ... 373
 - stopping criteria 369
 - two-index model example 375
 - ivprobit command 482
 - ivreg2 command 205, 208
 - ivregress command 183, 485
 - ivtobit command 544
- J**
- jackknife
 - IV estimator 204
 - definition 455
 - jackknife command 456
 - jive command 205
- K**
- kdensity command 64, 80
 - keep command 54
 - kernreg command 67
 - knnreg command 67
 - kurtosis measure 76
- L**
- label command 43
 - Lagrange multiplier test see
 - hypothesis tests
 - latent-class model ... see finite-mixture
 - model
 - LEF see linear exponential family
 - levinlin command 279
 - likelihood-ratio test see hypothesis
 - tests
 - lincom command 410
 - linear exponential family
 - definition 327
 - examples 322
 - Poisson example 327
 - linear regression model 81–118
 - basic theory 81
 - Box–Cox transformation 98
 - cluster–robust variance matrix.. 84
 - endogenous regressors 148, 177
 - feasible GLS estimation 153
 - FGLS heteroskedastic errors exam-
 - ple 160
 - FGLS system of equations example
 - 164
 - GLS estimation 153
 - hypothesis tests 80, 88
 - influential observations 96
 - information matrix test 102
 - instrumental variables 177
 - linear constraints 168
 - linear versus log regression 85
 - marginal effects 106
 - measurement-error example ... 148
 - omnibus test 102

- linear regression model, *continued*
 - panel-data basics 235
 - panel-data extensions 287
 - prediction 104
 - prediction in logs 108
 - quantile regression 211
 - regression in logs 85
 - residual analysis 95
 - retransformation problem 108
 - robust variance matrix 84
 - sampling weights 111
 - simulations to verify properties 141
 - specification analysis 94
 - survey data 173
 - test interpretation 102
 - test of heteroskedasticity 100, 158
 - test of normality 102
 - test of omitted variables 98
 - three-stage least squares 206
 - two-stage least squares 180
 - using Mata 115
 - variance matrix estimation 84
 - weighted marginal effects 114
 - weighted prediction 114
 - weighted regression 111, 113, 173
- linktest** command 100
- list** command 36, 44
- local** command 22
- local macros 22, 24, 26, 653, 656
- log** command 15
- log file 14
- logical operators 12
- logistic** command 462
- logit** command 462
- logit model see binary outcome models
- lognormal data
 - linear regression model 85
 - tobit model 545
- long command lines 16
- long-form data 280, 283, 503
- longitudinal data see panel data
- looping commands
 - foreach** 25
 - forvalues** 26
 - while** 26
- lowess** command 68
- lpoly** command 68
- lrtest** command 415
- M**
- macros
 - compared with scalars 23
 - global 21
 - local 22
- margeff** command 346
- marginal effects 343–357
 - average marginal effect 349
 - binary outcome models 476
 - calculus method 343
 - comparison of calculus and finite-differences method 348
 - count-data models 576
 - elasticities 344, 351
 - finite-differences method 343
 - in linear regression model 106
 - in multinomial outcome models 502, 509
 - in panel logit model 630
 - interacted regressors 355
 - log model 86
 - marginal effect at representative value 348
 - marginal effect at the mean 346
 - polynomial regressors 354
 - summary 344
 - weighted 114
- margins** command 106, 341–347, 477, 478
- Mata 661–674
 - ado-files using Mata 674
 - colon operators 666
 - combining matrices 669
 - commands in Mata 661
 - commands in Stata 662
 - declaration of program arguments 672

- Mata, continued*
- element-by-element operators 666
 - generalized method of moments example 394
 - Gibbs sampler example 137
 - help** command 663
 - identity matrix 664
 - matrix cross products 669
 - matrix functions 667
 - matrix input 663
 - matrix inversion 668
 - matrix of constants 664
 - matrix operators 666
 - matrix subscripts 669
 - Newton–Raphson iterative
 - example 365
 - Newton–Raphson Poisson
 - example 365
 - OLS regression example 115
 - optimization in Mata 389
 - overview 661
 - Poisson example 390
 - program example 672
 - programming in Mata 672
 - Stata commands in Mata 662
 - Stata interface functions 666
 - Stata matrix from Mata matrix 671
 - Stata variables from Mata matrix 671
- mata** command 661
- matrices in Mata *see* Mata
- matrices in Stata 645–651
- combining matrices 648
 - matrix cross products 649
 - matrix functions 648
 - matrix input 645
 - matrix operators 648
 - matrix subscripts 647
 - OLS example 650
 - overview 645
- matrix accum** command 649
- matrix algebra** definitions 82
- matrix** command 18, 645
- matrix define** command 645
- matrix rownames** command 646
- matrix vecaccum** command 649
- maximization options *see* iterative methods
- maximum likelihood estimator (MLE)
 - definition 322
 - misspecified density 322
 - pseudo-MLE 322
 - quasi-MLE 322
- mean** command 143
- MEM *see* marginal effects, marginal effect at the mean
- MER *see* marginal effects, marginal effect at representative value
- merge** command 56
- mf** command 346
- missing-value codes 46
- missing-values imputation 47
- mixed linear models *see* panel data
- mixlogit** command 523
- ml check** command 373, 377
- ml maximize** command 372
- ml model** command 372, 383
- ml search** command 374
- ml trace** command 377
- mlevel** command 384, 386
- mlmatbysum** command 389
- mlmatsum** command 385, 388
- mlogit** command 498
- mlsum** command 384, 386
- mlvecsum** command 384, 387
- model comparison 359
- model diagnostics 357–361
- moment-based tests 425
- Monte Carlo methods *see* simulation
- mprobit** command 518
- multicollinearity 379
- multinomial outcome models 491–533
 - additive random-utility model (ARUM) 493, 511
 - alternative-specific regressors 493, 495, 497
 - basic theory 491
 - bivariate probit model 529

- multinomial outcome models, *continued*
 - case-specific regressors . . . 493, 495, 497, 505
 - coefficient interpretation . . . 500, 507
 - conditional logit model 503
 - conditional versus multinomial
 - logit 507
 - GHK simulator 519
 - independence of irrelevant alternatives 511
 - marginal effects 492, 502, 509, 516, 522, 528
 - maximum likelihood estimator 492
 - model comparison 517
 - multinomial logit model 498
 - multinomial probit model 517
 - multinomial regression 491
 - multivariate outcomes 529
 - nested logit model 511
 - ordered logit model 527
 - ordered outcome models 525
 - ordered probit model 527
 - predicted probabilities . . . 501, 508, 516, 522, 528, 531
 - prediction for new alternative . . . 508
 - probabilities 492
 - random-parameters logit 523
 - simulated maximum likelihood estimator 519
 - Stata commands summary 494

- N**
 - nbreg** command 577
 - nl** command 325
 - nlcom** command 411
 - nlogit** command 513
 - nlogitgen** command 513
 - nlogitree** command 513
 - nlstur** command 532
 - nonlinear instrumental variables . . . 393
 - count-data example 610
 - nonlinear least-squares estimator
 - definition 325
 - example 326
 - nonlinear regression 320–336
 - nonlinear seemingly unrelated
 - equations 532
 - nonparametric methods
 - kernel density estimation 63
 - kernel regression 67
 - local polynomial regression 68
 - lowess regression 68
 - nearest-neighbor regression 67

- O**
 - ologit** command 527
 - operators 12
 - oprobit** command 529
 - optimize()** Mata function 389
 - order** command 55
 - outfile** command 53
 - outsheet** command 53
 - outsum** command 77
 - overidentifying restrictions test see specification tests

- P**
 - panel data . . 235–285, 287–312, 615–643
 - Arellano–Bond estimator 293
 - between estimator 260
 - between variation 264, 620
 - binary outcome models . . . 621, 631
 - cluster–robust inference 239
 - cointegration 279
 - command summary 240
 - comparison of estimators 264
 - count-data models 633
 - dynamic model 293, 304
 - endogeneity 237, 621
 - exogeneity 271
 - first-difference estimator . . 269, 271
 - first-difference model 294
 - fixed versus random effects . . . 266
 - fixed-effects estimator . . . 257, 627, 638
 - fixed-effects model . . 237, 257, 616
 - generalized estimating equations 617

panel data, *continued*

generalized method-of-moments estimator 295
 Hausman test 266, 267
 Hausman–Taylor estimator . . . 290, 293
 individual-effects model . . 237, 257, 277, 304
 individual-invariant regressor 236, 245, 260
 instrumental-variables estimator 287, 290, 294
 least-squares dummy-variable estimator 259
 linear Stata commands summary 240
 logit model 623
 long panel 236, 271, 279
 management of data 280, 283
 marginal effects 630
 mixed linear models 239, 305, 312
 mixed nonlinear model 630
 multiplicative-effects model . . . 615
 negative binomial model 641
 nonlinear models overview . . . 615
 nonlinear Stata commands summary 618
 Poisson model 633
 pooled estimator 250, 251, 254, 256, 273, 276, 622, 634
 pooled model 238, 254, 616
 population-averaged estimator 624, 635
 population-averaged model . . . 238, 255, 617
 prediction 268, 630
 R^2 264
 random-coefficients estimator . . 310
 random-coefficients model . . . 239, 310, 616
 random-effects estimator 261, 263, 626, 636
 random-effects model 238, 262, 306, 616, 631

panel data, *continued*

random-intercept model 306
 short panel 235, 293, 615
 spatial correlation 274
 summary of data 240, 253
 time-invariant regressor . . 236, 245, 250, 251, 257, 258, 263, 266, 270, 289, 290, 293, 298
 time-series autocorrelations . . . 251, 253
 time-series plots 247
 two-stage least-squares estimator 295
 two-way-effects model . . . 238, 273, 311, 312
 unbalanced panel . . . 236, 243, 264, 620
 unit roots 279
 variance components 264
 within estimator see panel data, fixed-effects estimator
 within scatterplot 249
 within variation 244, 246, 264, 620
 percentiles 76
 poisson command 323, 573
 Poisson model see count-data models
 post command 130
 postclose command 130
 postestimation commands summary 324
 postfile command 130, 151, 423
 prchange command 479
 prcounts command 602
 predict command 104, 268, 337
 prediction 336–341
 at specified regressor value 339
 binary outcome models 474
 in count-data models 581
 in levels from log model 108, 563–565
 in linear panel-data model 268
 in linear regression model 104
 in panel logit model 630

- prediction, *continued*
 of multinomial model probabilities
 501
 out of sample 338
 summary 336
 weighted 114
predictnl command 337
preserve command .. 55, 110, 339, 352
probit command 462
probit model see binary outcome
 models
program define command 651
program drop command 652
programs 651–659
 bootstrap example .. 441, 446, 449,
 450
 central limit theorem example .. 128
 checking 377
 checking parameter estimates .. 381
 checking standard-error estimates ..
 382
 debugging 657
 in Mata 672
 Monte Carlo integration example ..
 140
 named positional arguments .. 653
 OLS with chi-squared errors ... 141
 overview 651
 parsing syntax 655
 positional arguments 653
 r-class example 654
 temporary variables 653
prvalue command 476, 581
pseudo- R^2 358, 471
p-values
 bootstrap 446
 chi-squared compared with F .. 400
 computation of 402
 standard normal compared with t
 400
pwcorr command 87
- Q**
qcount command 228
qplot command 215
- qreg command 214
quadrature methods 139, 518, 625
quantile regression 211–234
 bootstrap robust variance matrix ..
 214
 coefficient comparison across quan-
 tiles 221
 coefficient interpretation 216
 computation 213
 conditional quantiles 211
 data example 215
 definition of quantiles 211
 for count data 226
 generated data example 222
 loss functions 212
 marginal effects 217
 QR estimator 213
 retransformation 217
 test of equality 220
 test of heteroskedasticity 219
 variance matrix estimation 213
query command 17
- R**
 R^2 357
random effects see panel data
random-number generation see
 simulation
rbeta() function 124
rbinomial() function 124
rchi2() function 123
r-class commands 18, 128, 654
recode command 49
reformat command 92
reg3 command 207
regress command 7, 20, 87
relational operators 12
rename command 43
replace command 46, 49
reshape command ... 55, 280, 283, 504
residuals
 definitions of various residuals
 359
 residual plots 95
restore command ... 55, 110, 339, 352

- return code 12, 659
- return command 142
- return list command 18
- rgamma() function 124, 569
- _rmcoll command 380
- rnbinomial() function 126, 570
- rnormal() function 122
- robust variance matrix 84, 334
- rpoisson() function 124, 568, 569
- rt() function 123
- runiform() function 120
- rvfplot command 95

- S**
- sample-selection model...see selection model
- sampling weights in linear regression model 111
- save command 52
- saveold command 53
- scalars 17, 23
- score test see hypothesis tests, Lagrange multiplier test
- search command 4
- seemingly unrelated regression
 - equations 162–169
 - feasible GLS estimation 163
 - FGLS example 164
 - imposing cross-equation restrictions 168
 - robust variance matrix 166
 - SUR estimator 163
 - SUR model 162
 - test of cross-equation restrictions... 167
 - test of error independence 166
- selection model 556–565
 - exclusion restrictions 561
 - identification 558
 - inverse of the Mills' ratio 560
 - model definition 556
 - model likelihood 557
 - prediction 565
 - two-step estimation 442, 560
- set command 17
- set memory command 34
- set more off command 14
- set seed command 120, 214, 431
- simulate command 129
- simulated maximum likelihood estimator 519
- simulation 119–151
 - bias of estimator 144
 - bias of standard-error estimator 144
 - bootstrap example 449
 - central limit theorem example.. 127
 - Cholesky decomposition 136
 - computing integrals 138
 - direct transformation 133
 - distribution of sample mean... 127
 - draws from multivariate normal 135
 - draws from negative binomial.. 569
 - draws from normal 122
 - draws from Poisson 568
 - draws from truncated normal.. 134
 - draws from uniform 120
 - draws of continuous variates... 123
 - draws of discrete variates..... 124
 - draws using Markov chain Monte Carlo method 136
 - endogenous regressors example 148
 - FGLS heteroskedastic errors examples 156
 - Gibbs sampler example 137
 - inconsistency of estimator 148
 - interpreting simulation output 144
 - inverse probability transformation 132
 - measurement-error example ... 148
 - mixture of distributions 133
 - Monte Carlo integration example.. 140
 - Monte Carlo methods 119
 - OLS with chi-squared errors ... 141
 - power of test computation ... 146, 422

- simulation, *continued*
 - pseudorandom numbers 120
 - size of test computation . . 145, 420
 - using `postfile` command . . . 130, 151
 - using `simulate` command 141
 - single-index model
 - coefficient interpretation 345
 - definition 345
 - numerical derivatives for 371
 - skewness measure 76
 - `sort` command 55
 - specification tests 425–427
 - binary outcome models 468
 - chi-squared goodness-of-fit test 426
 - classification test 473
 - for serially uncorrelated panel errors 300
 - goodness-of-fit test 472
 - Hausman test 267, 426, 443
 - information matrix test 425
 - moment-based tests 425
 - RESET test 99
 - test for fixed effects 267
 - test of endogeneity 188, 443
 - test of heteroskedasticity 100, 158
 - test of omitted variables 98
 - test of overidentifying restrictions 191, 300, 426
 - test of weak instruments 196
 - `sqreg` command 214
 - `st_addvar()` Mata function 671
 - `st_data()` Mata function 665
 - `st_matrix()` Mata function . . 116, 665, 671
 - `st_store()` Mata function 671
 - `st_view()` Mata function 115, 665
 - standard errors *see* variance–covariance matrix
 - Stata Journal* 3
 - Stata manuals 2
 - Stata Technical Bulletin* 3
 - `stata()` Mata function 662
 - `statsby` prefix command 276
 - string data 33, 36
 - `summarize` command 6, 19, 75
 - SUR model *see* seemingly unrelated regression equations
 - `sureg` command 163
 - survey data 169–175
 - clustering 170
 - commands for survey data . . . 171
 - complex survey data 169
 - stratification 170
 - weighted mean 173
 - weighted regression 173
 - weighting 169
 - `svy` prefix command 170
 - `svy: regress` command 173
 - `svydescribe` command 172
 - `svymean` command 173
 - `svyset` command 171
 - syntax
 - basic command syntax 5
 - parsing syntax 655
 - `syntax` command 655
 - systems of equations
 - linear regressions 162
 - linear simultaneous equations . . 206
 - nonlinear regressions 532
- T**
- `tab2` command 78
 - `table` command 77
 - tables for data 77
 - tables of output 89
 - in Word or \LaTeX 91
 - `tabstat` command 79
 - `tabulate` command 44, 51, 75
 - `tempvar` command 653
 - `test` command 88, 405
 - test of heteroskedasticity 158
 - test of overidentifying restrictions . . *see* specification tests
 - `testnl` command 409
 - `tobcm` command 552
 - `tobit` command 538

tobit model 535–553
 censored data 535
 density 537
 diagnostic tests 549
 endogenous regressor 544
 generalized residuals 549
 lognormal data 545
 marginal effects 541
 panel-data estimators 631
 partial observability 535
 prediction 540, 563
 selection model 556
 test of heteroskedasticity 552
 test of normality 551
 truncated mean 536
 two-limit tobit 548
 unknown censoring point 537
 tobit regression, two-part model 553
 tokens() Mata function 115
 tostring command 33
 trace command 659
 treatreg command 193
 tsset command 121
 ttest command 80
 two-part model
 for count data see count-data
 models
 for log expenditure data 553
 prediction 564
 two-stage least-squares estimator see
 instrumental variables
 two-step estimator, bootstrap standard
 error 441
 twoway command 66

U

uniform() function see runiform()
 function
 update command 2
 use command 35

V

variance–covariance matrix 330–336
 bootstrap estimate 431

variance–covariance matrix, *continued*
 cluster–robust variance matrix
 335
 default estimate 83, 334
 definition 331
 jackknife estimate 455
 of two-step estimator 441
 panel data 239
 robust 334
 vce(bootstrap) option 239, 336,
 432
 vce(cluster *clustvar*) option
 84, 335
 vce(hac *kernel*) option 336
 vce(jackknife) option 336, 456
 vce(oim) option 334
 vce(opg) option 334
 vce(robust) option 84, 239, 334
 VCE see variance–covariance matrix
 vce() option see variance–covariance
 matrix

W

Wald confidence intervals see
 confidence intervals
 Wald test see hypothesis tests
 weak instruments test see
 specification tests
 weight 6
 weighted mean 112
 weighted regression 113
 while command 26
 wide-form data 280, 283, 503
 wildcards 11

X

xi prefix command 51
 xtabond command 295, 301
 xtabond2 command 301
 xtclolog command 621
 xtdata command 249
 xtdescribe command 243, 315
 xtddpd command 303
 xtddpsys command 301
 xtfrontier command 633

xtgee command 256, 265, 624
xtgls command 274, 275
xthtaylor command 291, 292
xtintreg command 632
xtivreg command 288, 289
xtline command 247
xtlogit command 623
xtmixed command 306, 312, 316
xtnbreg command 633
xtpcse command 273, 275
xtpmg command 279
xtpoisson command 633
xtpqml command 638
xtprobit command 621
xtrc command 310
xtreg command 257, 269, 278, 315
xtregar command 277, 278
xtscc command 274, 275, 278
xtset command 243, 314
xtsum command 244
xttab command 246
xttobit command 631
xttrans command 246

Z

zero-inflated models see count-data
models
zero-truncated models .. see count-data
models
zinb command 601
zip command 601
ztnb command 586
ztp command 586