### INDEX

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accelerated failure time models</td>
<td>188</td>
</tr>
<tr>
<td>Adjusted coefficient of determination</td>
<td>10</td>
</tr>
<tr>
<td>calculation</td>
<td></td>
</tr>
<tr>
<td>Adjusted coefficient of variation</td>
<td>7</td>
</tr>
<tr>
<td>definition</td>
<td></td>
</tr>
<tr>
<td>Airlines data</td>
<td>112, 115, 117, 119, 121, 122, 124, 126, 127, 128, 206</td>
</tr>
<tr>
<td>fixed/random effects model, covariance matrices</td>
<td>128</td>
</tr>
<tr>
<td>fixed time effects analysis</td>
<td>121</td>
</tr>
<tr>
<td>using Proc GLM</td>
<td></td>
</tr>
<tr>
<td>firm effects analysis</td>
<td>124</td>
</tr>
<tr>
<td>using Proc GLM</td>
<td></td>
</tr>
<tr>
<td>groupwise heteroscedasticity estimators</td>
<td>209</td>
</tr>
<tr>
<td>HCCME estimators</td>
<td>220</td>
</tr>
<tr>
<td>least squares residuals</td>
<td>206</td>
</tr>
<tr>
<td>comparison</td>
<td>207</td>
</tr>
<tr>
<td>likelihood ratio test</td>
<td>206</td>
</tr>
<tr>
<td>LSDV estimation</td>
<td></td>
</tr>
<tr>
<td>using Proc GLM</td>
<td>119</td>
</tr>
<tr>
<td>using Proc IML</td>
<td>115</td>
</tr>
<tr>
<td>using Proc panel</td>
<td>117</td>
</tr>
<tr>
<td>using OLS calculations</td>
<td>117</td>
</tr>
<tr>
<td>mean of residuals</td>
<td>127</td>
</tr>
<tr>
<td>pooled regression model</td>
<td>112</td>
</tr>
<tr>
<td>summary statistics</td>
<td>254</td>
</tr>
<tr>
<td>temporary SAS data set</td>
<td>208</td>
</tr>
<tr>
<td>time series plot</td>
<td>206</td>
</tr>
<tr>
<td>regression, dummy variables</td>
<td>48</td>
</tr>
<tr>
<td>Analysis of variance (ANOVA) techniques</td>
<td>2, 5, 6, 12, 13, 45, 58</td>
</tr>
<tr>
<td>table</td>
<td>12, 13, 58</td>
</tr>
<tr>
<td>ARCH(1) model</td>
<td>88</td>
</tr>
<tr>
<td>process</td>
<td>89, 91</td>
</tr>
<tr>
<td>ARCH(q) process</td>
<td>89</td>
</tr>
<tr>
<td>unconditional variance</td>
<td>89</td>
</tr>
<tr>
<td>Arellano–Bond GMM estimator</td>
<td>224</td>
</tr>
<tr>
<td>first-step estimator</td>
<td>222</td>
</tr>
<tr>
<td>second-step estimator</td>
<td>222</td>
</tr>
<tr>
<td>Asymptotic covariance matrix</td>
<td>65, 114, 234</td>
</tr>
<tr>
<td>Asymptotic variance</td>
<td>116</td>
</tr>
<tr>
<td>Asymptotic variance-covariance matrix</td>
<td>57</td>
</tr>
<tr>
<td>Attrition models</td>
<td>153</td>
</tr>
<tr>
<td>Autocorrelation process</td>
<td>93–96</td>
</tr>
<tr>
<td>detection</td>
<td>96–101</td>
</tr>
<tr>
<td>Durbin–Watson test</td>
<td>96</td>
</tr>
<tr>
<td>Lagrange multiplier test</td>
<td>97</td>
</tr>
<tr>
<td>first-order</td>
<td>96</td>
</tr>
<tr>
<td>occurrence</td>
<td>93</td>
</tr>
<tr>
<td>ordinary least square (OLS) estimation problems</td>
<td>94–95</td>
</tr>
<tr>
<td>parameters</td>
<td>101, 102</td>
</tr>
<tr>
<td>FGLS estimation method</td>
<td>101</td>
</tr>
<tr>
<td>GLS estimation method</td>
<td>101</td>
</tr>
<tr>
<td>second-order</td>
<td>96, 102, 104</td>
</tr>
<tr>
<td>Autoregressive conditional heteroscedastic models (ARCH)</td>
<td>44, 87–92</td>
</tr>
<tr>
<td>generalized ARCH models</td>
<td>44, 89</td>
</tr>
<tr>
<td>process</td>
<td>88, 91</td>
</tr>
<tr>
<td>testing effects</td>
<td>90–92</td>
</tr>
<tr>
<td>Autoregressive autocorrelation model (AR)</td>
<td>101–109</td>
</tr>
<tr>
<td>AR(2) model, residuals</td>
<td>102–105, 108, 109</td>
</tr>
<tr>
<td>first-order autocorrelation model (AR1)</td>
<td>94</td>
</tr>
<tr>
<td>fitness procedure</td>
<td>101</td>
</tr>
<tr>
<td>Proc autoreg</td>
<td>101</td>
</tr>
<tr>
<td>SAS usage</td>
<td>101</td>
</tr>
<tr>
<td>Autoregressive moving average process</td>
<td>88</td>
</tr>
<tr>
<td>Bartlett test statistic</td>
<td>207</td>
</tr>
<tr>
<td>variance comparison</td>
<td>207</td>
</tr>
</tbody>
</table>

*Applied Econometrics Using the SAS® System,* by Vivek B. Ajmani  
Copyright © 2009 John Wiley & Sons, Inc.
INDEX

BHHH methods, algorithm, 192
Binary response models, 154
Bootstrap estimation method, 262, 267
  calculating standard errors, 264
  cumulative distributions, plot diagram, 263
  estimation technique, 262
  in regression analysis, 265
  in SAS, 264
  lower/upper confidence limit, 263
OLS estimates, 271
Proc Univariate statements, 265
Bootstrapped regression analysis, 267–269
  gasoline consumption data, 269
  residuals method, 266, 268
SAS, 267
Breusch–Pagan Lagrange multiplier test, 76, 78–80, 129
  credit card expenditure data, 80
Central limit theorem, 5
Ceteris paribus condition, 2
Chi-square test, 191
  distribution, 97, 207
  table, 66
  values, 189
Chow test static, 40, 41, 42
  by Proc model, 42
  structural break in gasoline data, 42–43
  ρ value, 40, 41
Classical regression model, spherical disturbances assumption, 71
Cobb–Douglas model, 34, 35, 37
  production data, 35
  regression analysis, 35, 37
  SAS code, 35
Coefficient of determination, 10, 14
  calculation, 10
Coefficient of variation, 13, 51
  definition, 13
Combined gasoline consumption data,
  regression analysis, 41
Complex panel data models, 116
  autocorrelation violations, 116
  dynamic panel data models, 116
  heteroscedasticity violations, 116
Conditional probabilities, 173
  calculation, 173
Confidence interval, 7–8, 18, 189
Consumer price index (CPI), 9, 14
  inflation rate, 14
Cook’s D statistic, definition, 20
Cook’s statistic, See Cook’s D statistic
Correlation, 15, 16
  b/w variables, 16
  coefficients, 25
  matrix, 26
  nature of, 15
  scatter plots, 15
Covariance matrix, 95, 125, 128
  construction, 95
  diagonal elements, 125
Cox’s proportional hazard models, 190
CPI, See Consumer price index
Credit card expenditure data, 203
  ALPHA vs. likelihood value plot, 281
  Breusch–Pagan Lagrange multiplier test, 279
  FGLS estimators, 280
  GLS estimator, 283
  heteroscedasticity, 278
  iterative FGLS estimators, 203
  maximum likelihood estimations (MLEs) parameters, 284
  regression analysis, 205
  White’s test, 278
Cross-equation correlation, 1, 140
Cross-model covariance matrix, 140
  diagonal elements of, 140
Cumulative distribution function (CDF), 170
Cumulative hazard rate function, 171
CUSUM test, 41–45
  critical values, 43
  definition, 43
  gasoline consumption data, 44, 45
  plot, 45
  procedure, 41
Data matrix, 7, 10
Data plotting, 15–16
Data set, 47
Data testing, 38
  for structural breaks, 38
  linear restriction hypothesis test, 38
Davidson/MacKinnon’s estimator, 81
  DM1 versions, 83
  DM2 versions, 83
Definite matrix, 53
Degrees of freedom, 6, 13, 29, 65, 91
  model, 10, 50
  n–k, 29
Dependent variable, 6, 18
  predictions of, 18–21
Determinants, 241
  definition, 241
  properties of, 241
Direct marketing companies, 153
Discrete choice models, 153
  binary response models, 154
  parameters interpretation, 155
  shortcomings, 154
Discrete random variable, 153
Disturbance vector, 114
Dummy variables, 45
  estimators, 72, 113
  in models, 45–51
  model, 114
  vector, 114
Duration analysis, 169, 178
  distribution functions, 178–186
  exponential distribution, 179
  lognormal distribution, 184
  Weibull distribution, 179
Durbin–Watson statistic test, 91, 96, 97, 101, 102
Box and Pierce’s test (B&P), 97
error sums of squares, 90
Ljung’s modification, 97
mean sums of squares, 90
serial correlation, 90
Dynamic panel data models, 220
dynamic panel data estimation, 221
generalized methods of moments estimation (GMM), 220
estimation technique, 221
with explanatory variables, 223

Earning’s equation model, 47
data matrix, 47
dummy variable, 47
Elasticity, 2
definition, 2
vs. marginal effect, 2
Endogeneity, alternative hypothesis, 64
Engle’s ARCH model, See ARCH(1) model
Error sums of squares (SSE), 4, 78
Explanatory variables, 2, 3, 24, 45, 54, 55, 70, 71, 75, 110, 111, 114, 118, 129
categories, 45
estimation, 71
feasible generalized least squares (FGLS) estimators, 71
generalized least squares (GLS) estimators, 71
measurement errors, 54, 55
revisited estimation, 80
types, 110
observed/controllable, 110
unobserved/uncontrollable, 110
Exponential distribution, 179, 183
hazard function, 179, 183
probability density function, 179
survival function, 179, 183
Extra variable model, 7
sums of squares of error, 7

F-statistic value, 13, 25, 30, 64, 121
formula, 29, 34, 37, 39
critical value, 37
Proc IML use, 29
hypothesis tests, 13
Failure times/censoring, 169–170
Feasible generalized least squares (FGLS), 134
asymptotic covariance matrix, 134
cross-equation covariance, 140
estimation, 84, 87, 88
by credit card data, 87, 88
estimator, 86, 102, 232
cross-sectional correlation, 232
general procedure, 86
Proc Reg output, 86
SAS step, 86
standard errors, 233
Grunfeld’s investment data set, 134, 135
OLS residuals, 134
Fitted/full model, 93
degree of autocorrelation, 93
residuals, 93
Fixed effects model, 113–123
estimation methods, 113
between-group effects approach, 113
least squares-group effects approach, 113
within-group effects approach, 113
Proc GLM, 118
Frisch–Waugh theorem, 6, 114

GARCH model, 89–91
effects, 91
principle, 90
unconditional variance, 90
AR(1) model, 100
iterated FGLS estimates, 107
output, 100, 101
AR(2) model, 101, 104, 105
iterated FGLS estimates, 108
MLE estimates, 103, 105
output, 101
AR models, residuals comparison, 109
autocorrelation, 98, 99
Durbin–Watson statistics, 98
Proc Autoreg detecting method, 98, 99
full model residuals, 94
time series plot, 94
independent variables, 26
Proc Corr output, 26
model, 93
multicollinearity output, 25
OLS vs. AR(2) models, 109
residuals comparison, 109
reduced model residuals, 94
time series plot, 94
regression analysis, 39, 40
Gauss–Markov theorem, 5
Generalized least squares (GLS) estimation technique, 86, 133
estimator, 96
Generalized methods of moments estimation (GMM), 148, 220
Arellano–Bond, 224
cigar.txt panel data, 222
dynamic panel data models, 220, 221
estimators, 150
2SLS, 151
labor equation, 150
weight matrix, 151
White’s estimator, 151
explanatory variables, 223
optimal weight matrix, 221
General linear hypothesis, 27, 28, 29
  hypothetical model, 27
  least squares estimator, 28
  restriction equation, 27
  SAS use, 29
  testing, 33
  Proc Reg output, 34
  variance-covariance matrix, 28
General panel data model, 111, 120
GNP, See Gross national product
Goldfeld–Quandt tests, 78
  explanatory variable, 78
Good fit model, 25
Goodness-of-fit statistics, 6–7, 185
  adjusted coefficient of determination, 6
  assessment method, 185
  Proc Lifereg, 185
  coefficient of determination, 6
  definition, 6
  Good national product, 9
Group-specific mean square errors, 207
Groupwise heteroscedasticity estimator, 205, 209
  airlines data analysis, 205
  airlines data set, 208, 209
  assumption for model, 208
  Chi-squared distribution, 207
  homoscedasticity assumption, 205
  likelihood ratio test, 206
  mean square error (MSE), 207
  using Harvey’s multiplicative heteroscedasticity approach, 210
Grunfeld data analysis, 136
  using Proc Syslin SUR, 136–140
Grunfeld data set, 134, 135, 228
  FGLS estimator, 229
  FGLS pooled estimators, 228
  pooled OLS regression, 135
  Proc Syslin SUR, 136
Grunfeld investment model, 1
Harvey’s multiplicative heteroscedasticity, 204, 208
  MLE estimates, 204
  single parameter, 204
  model parameters estimation, 204
  variance-covariance matrix, 205
Hausman analysis, 65
  by Proc IML, 65
  consumption data, 65
Hausman’s specification test, 61, 64–69, 128–130
  by Proc model, 66–69
  consumption data, 66
  generation, 129
Hausman–Taylor estimator, 210
  coefficients estimates, 212
  endogenous/exogenous variables, 210
  for random effects model, 210
  instrumental variables, 212
  Proc IML, 218
  Proc model output, 216–217
PSID, 212
  random effects and LSDV model, 212
  standard errors, 218
  steps, 211
  wages equation, 219
Hazard function, 170–178
  definition, 170
HCCME estimators, 82
  credit card data, 82
  OLS estimate of covariance matrix, 219
Heteroscedasticity, 70, 71, 72, 74, 76, 78, 91
  detection, 72, 74
  formal hypothesis tests, 74–80
  least squares residuals, 72
  residual plots, 78
  testing, 91
  nominal exchange data, 91
  variance-covariance matrix, 71
Heteroscedastic variance, 22
  funnel-shaped graph, 22
Homoscedasticity, 80, 207
  null hypothesis, 80, 207
Human’s specification test, 151
  exogenous/endogenous variable, 151
  OLS/2SLS estimates, 152
  Proc model procedure, 151
Hypothesis testing, 7–8, 27, 28, 39
  confidence intervals, 8
  linear restrictions, 28
  regression coefficient, 8
Idempotent matrices, 243
  definition, 243
  econometrics, 243
Identity matrix, 240
  definition, 240
  properties, 240
Independent disturbances, 93
  assumption, 93
Independent regressions, 78
Independent variables, See Explanatory variables
Inflation rate, 14, 15
  definition, 14
Instrumental variables, 52, 55, 56
  estimation of, 55–60
  covariance matrix, 56
  data matrix, 60
  standard error, 58, 60
  data matrix, 56
  least squares model, 55
  matrix, 56
  exper, 56
  exper², 56
  mothereduc, 56
  regression, 58
Instrumental variables analysis, 58
  Proc Syslin, 58
  earning data, 58
Multiple linear regression model, 1, 3, 4
matrix form, 3
Newton–Raphson method, 157–163, 192
algorithm, 157
for Logit model, 157–163
Nonspherical disturbances, 70, 71
autocorrelation, 70
heteroscedasticity, 70
Null hypothesis, 6, 8, 13, 29, 30, 32, 61, 62, 65, 80, 101, 128, 129
Off-diagonal elements, 140
Omitted variable model, 53
bias, 53
Ordinary least squares (OLS) analysis, 58, 59, 72, 86, 149
earning data, 59
estimator, 4, 31, 33, 53, 56, 83, 102, 255, 256
consistency, 56
histogram of, 257–260
labor equation, 149
mean and standard deviation, 256
probability, 53
simulated type, 1 error rate, 256
estimation, 58, 82, 234
covariance matrix, 82
critical assumption, 142
equation-by-equation, 140
estimation techniques, 144
keynesian model, 143
problems, 142–144
standard errors, 82
structural equation, 144
model, 58, 97, 104, 154
of credit card expenses data, 72
regression statistics, 25
residuals, 75, 86
Overidentifying restrictions testing, 63
in earning data, 63
Panel data method, 110, 111
advantages, 110–111
definition, 110
overview, 110
Panel data models, 111–112, 219, 224
autocorrelation, 224, 227
covariance matrices, robust estimation of, 219
FGLS methods estimators, 225
fixed effects, 111
GLS estimation, 225
heterogeneity, 224
homoscedastic disturbances, 219
ordinary least squares estimation method, 111
pooled regression, 111
PROC IML code, 226
random effects, 111
Poisson regression, 163–165
estimation, 165–168
parameters interpretation, 165
Pooled regression model, 112–113, 118
coefficient of determination, 113
expression equation, 113
parameters estimation, 112
OLS, 112
root mean square error, 113, 118
Prais–Winsten Method, 234
transformations, 96
usage, 96
Prediction intervals graphs, 21
Proc Reg output, 21
Price index of gasoline ($P_g$), 38
Probability distribution function, 171
Probability of failure (PDF), 174
calculation, 174
Probability plots, 22
Probit and Logit models, 155
estimation/inference, 156
Proc autoreg, 96, 102
CUSUMLB procedure, 44
CUSUMUB procedure, 44
usage, 102
reference guide, 91
Proc Corr procedure, 16
general form, 16
Proc GLM, 49, 121
airlines data regression, 49–50
data analysis, 49
dummy variables, 49
Proc Gplot, 16, 20
confidence intervals, 20
Proc IML analysis, 11, 47, 57, 114, 204, 248, 272
$1 \times 1$ matrices, 249
Anderson–estimator, 289
Arellano–Bond method, 224, 290
code computes, 286
concatenate matrices, 252
cost time, 251
determinants of matrices, 241
diag command, 252
diagonal matrix, 251
DO-END statement, 253
dynamic panel data data analysis, 47, 57
data matrix, 57
dynamic panel data data analysis, 47, 57
data matrix, 57
econometric analysis, 251
elementary matrix operations, 250
addition/subtraction, 250
inverses, eigenvalues, and eigenvectors, 250
Kronecker products, 250
GMM calculations, 222
GNP and Invest time series, 272
groupwise heterogeneity, 286
Grunfeld’s data analysis, 229
INDEX 309

Hausman’s specification test, 212, 276
Hausman–Taylor’s estimates, 218
heteroscedasticity, 277
identity matrix, 240
IF-THEN/ELSE statement, 253
Kronecker products calculation, 244
linear equations, 248
linear hypothesis, 275
matrix-generating functions, 251
block diagonal matrices, 251
diagonal matrix, 251
identity matrix, 251
J matrix, 251
matrix inverses, 243
matrix multiplications, 239
max(min) commands, 251
of investment data, 11
Proc IML code, 273
restricted least squares estimator, 274
robust variance-covariance matrices, 277
SAS procedures, 212
standard errors of estimator, 274
statements, 30
SUMMARY command, 253
trace of a matrix, 242
transpose matrix, 240
VecDiag function, 252
White’s test, 277–278
within-group mean residuals estimates, 215
Proc IML code, 226
FGLS estimator, 226
general linear hypothesis, 273
Kronecker product, 244
Proc IML command, 65, 237
create row and column vectors, 237
identity matrix, 238
matrix multiplications, 239
trace of a matrix, 241
Proc IML diag command, 252
diagonal matrix, 252
Proc Import statement, 9
Proc Life procedure, 173
Proc Lifereg models, 178, 191
Proc Lifetest analysis, 173, 175, 178
tabular presentation, 175
Proc Lifetime, 177
Proc model, 151, 215
HCCME option, 219
instrumental variable regression, 215
OLS/2SLS models, 151
procedure, 76
Proc Panel, 81, 114, 116–118, 120, 121, 123, 125, 128–131, 212, 219–221
documentation, 81
procedure, 114
Proc Plot procedure, 15
statements for, 15
Proc Reg analysis, 12, 15, 32, 47, 62, 101, 151, 255
data analysis, 47
endogeneity, 62
investment equation data, 15
of investment data, 12
OLS estimates, 32
tableout option of, 255
test, 32
Proc Reg module, 21
Proc Reg statements, 268
OLS regression, 268
Proc Syslin, 60, 151
earning data output, 60
procedure, 148
Proc Univariate, 255
data, 213
histogram option, 255
module, 22
Production data-translog model, 36
regression analysis, 36
Quadratic form, derivative of, 247
Quarterly investment data, 31, 33
Proc IML output, 31
Proc Reg output, 33
Random effects model, 123–131, 210
estimation, 130–131
Hausman–Taylor estimator, 210
random disturbances, assumptions, 211
tabular presentation, 130
tests, 125–128
Hausman specification, 125
Lagrange multiplier (LM), 125
Wages data, 213
Rank, 245
definition, 245
equivalence, 246
Proc IML, 245
properties of, 245
Real _Invest scatter plot, 18
vs. time plot, 17
vs. time trend, 18
vs. Real GNP plot, 16
Real_GNP coefficient, 14
RECID data, 172, 173, 175–177, 179, 180, 185, 186, 188, 190, 192
exponential distribution, 188
Kaplan Meier survival function plot, 176
lifetime hazard function plot, 179
lifetime survival function plot, 176, 177
normal distribution fit, 186, 192
Proc Lifetest analysis, 173
survival functions testing, 180–182
Weibull distribution fit, 190
Standard error(s), 18, 86, 87
   column, 14
   definition, 8
Strike data duration analysis, 196–200
Subject-specific heterogeneity, 111, 113, 123
   effects, 128, 129
Sums of squares for error (SSE), 5, 7
Sums of squares for regression (SSR), 5
Survival function, 170–178
   definition, 170
   Kaplan Meier method, 172
   life table method, 172
   plot, 176
   standard error, 172
Temporary SAS data set, 9, 10, 61, 65, 90, 172, 186, 208
Test statistic, 11, 28–30, 65, 66, 78, 79
   Proc IML statements, 79
   value, 14, 79
Time series data, 93
Translog model, 35, 36
Transpose matrix, 240
   definition, 240
   properties, 240
True population model, 54
   OLS estimate, 54
   probability limits, 54
Two-stage least squares estimator (2SLS) analysis, 56, 62, 148
   assumption of homoscedastic disturbances, 148
   labor equation, 149
   weight matrix, 151
Two-way fixed effects model, 123
   Proc GLM estimation method, 123
Unbiased estimator, 5, 8, 33, 71, 266
Unknown coefficients, 1, 2
U.S. gasoline consumption \( (G) \), 38
   time series plot, 38
Variance, 5, 95
   Variance-covariance matrix, 95, 113
Variance inflation factors (VIF), 24
   values, 25
Wage equation, 56, 64
   regression equation, 64
Wages data, 213
   random effects model, 213
   within-group effects model, 214
Wald's chi-square test, 161, 192
   values, 161, 188
Weak instruments analysis, 64
   in earning data, 64
Weibull distribution, 179, 183, 184, 190
   cumulative density function, 179
   hazard functions, 183
   probability density function, 179
   survival function, 184
   bar graph, 184
Weighted least squares regression
   methods, 84, 85
   credit card expenditure data, 85
   Proc Reg option, 84
   SAS statements, 84
White's estimator, 80, 81, 83, 148, 219, 220
   Proc model statements, 81
   HCCME option, 81
   variance-covariance matrix, 80
White's general test, 74–78
   credit card expense data, 76–78
   Proc IML programme, 74
   test statistic value, 76
Within-group model, 113, 120, 211, 212
   disadvantages, 113
   disturbance variances, 207, 208
   GLS estimation, 207
   merge data, 208
   OLS model, 208
   time-invariant disturbance, 211
   residuals vector, 207, 215
   root mean square, 120
   Wages data, 214