

# Index

---

- absolute sense interpretation 25
- accidents 126–8, 130
- accuracy of estimates 70–3, 99–100
- ADL model *see* autoregressive distributed lag model
- advanced descriptive statistics 31–2
- advertising 54, 56, 61
- alcohol consumption 40, 107, 186–7
- Almon lag model 132
- $\alpha$  5, 50–4, 58, 66, 69–71, 82
  - OLS estimation 93
- alternative hypothesis 79
- American Statistical Association 17
- analysis of variance 110
- annual frequency 9
- ANOVA *see* analysis of variance
- applications of time series models 179–211
  - Granger causality 186–92
  - hypothesis testing involving more than one coefficient 205–9
  - vector autoregressions 193–204
  - volatility in asset prices 179–86
- approximation 51–2
- AR(1) model 142, 146–51, 156
- arbitrage 146, 180
- ARCH *see* autoregressive conditional heteroskedasticity
- arithmetic lag model 131
- AR( $p$ ) 151–5
- artificial data 101–3, 148
- asset prices 179–86
- association 67
- atheory 195
- augmented Dickey–Fuller test 157
- autocorrelated errors 214
- autocorrelation 138
- autocorrelation function 138–42
- autoregressive conditional heteroskedasticity 179–80
- autoregressive distributed lag model 159–60, 177, 187–90
- autoregressive model for univariate time series 142–4
- Bank of England 201
- bar chart 18
- base year 13, 26, 28
- base year weights 27–8
- basic data handling 9–33
- baskets 12
- bell-shaped distribution 23

- best fitting line 50–4, 65, 91–3
- $\beta$  5, 50–2, 54, 58, 66, 69–71
  - calculating a confidence interval for 73–9
  - factors affecting accuracy of 70–3
  - OLS estimation 93
  - testing whether  $\beta = 0$  79–83
- bi-directional causality 190–2
- bias 213, 220
  - omitted variables bias 99–100
- binary variables 11
  - see also* dummy variables
- bins 19
- built-in causality 51
  
- calculation of banana price index 25–6
- calculation of confidence interval for  $\beta$  73–9
- calculation of fruit price index 26–31
  - Laspeyres price index 27–8
  - Paasche price index 28–31
- calculus 53–4, 106
- canceling out 165
- capital asset pricing model 181
- CAPM *see* capital asset pricing model
- causality 38–41, 51, 186–92
  - in both directions 190–2
  - and correlation 40
  - direct and indirect 41
  - see also* Granger causality
- censored variables 212–13
- census and related data sets 16
- ceteris paribus* 94–7, 104, 115–17, 126–7, 160–1, 174
- change 138, 155, 175
- change in GDP 12
- change in price 13
- chaos 182
- ChartWizard<sup>®</sup> 18
- cigarette smoking 40–1, 107, 186–7
- class intervals 19, 31
- class width 19
- clustering in volatility 182
  
- co-trending 165
- coefficients 5, 51, 106
- cointegrating regression 167
- cointegration 121, 147, 165–71
  - between prices of two goods 166, 169–70, 173–4
  - estimation and testing with cointegrated variables 167–71
  - and Granger causality 192
- common trends 165
- comparative sense interpretation 25
- computer purchases 162–3
- confidence intervals 2, 70, 73–9
- confidence level 70
- consumer price index 30, 189, 199, 214
- consumption 138, 154, 216
- cookbooks 2
- correlation 35–47
  - between several variables 45–6
  - mathematical variables 46–7
  - understanding correlation 35–6
  - understanding correlation through verbal reasoning 36–9
  - understanding correlation through *XY*-plots 42–5
  - understanding why variables are correlated 39–42
- correlation matrix 45–6, 98, 105, 111
- cost of production in electric utility industry
  - multiple regression 103–5
  - regression 53–6, 60, 85–6
- count data model 213
- covariance 47
- CPI *see* consumer price index
- creating lagged periods 123–5
- critical value 81, 84, 87, 207
- cross-sectional data 10, 121–2, 125, 138
- current year weights 28–31
  
- daily frequency 9
- data 1–2
- data collection on consumer prices 13

- data directory 229–31
- data manipulation 168
- data transformations 11–12
- data as units 10
- Datastream 17
- de-seasonalized data 150
- deciles 32
- deforestation 125–6
  - correlation with population density 36–40
  - regression on population density 53, 55, 78, 81–2, 85
  - $XY$ -plots depicting population density 21–2
- $\Delta Y$  136, 139
- dependent variable 51, 119–20, 212–13
- depreciation 18
- descriptive statistics 23–5, 31–2
  - advanced 31–2
- deterministic trend 148–54, 162, 173, 177
  - model 149, 151–5
- devaluation 18
- deviations from mean 66
- Dickey–Fuller critical value 153–4
- Dickey–Fuller test 153–5, 164, 168–70, 178, 183
  - see also*  $t$ -statistic
- difference stationary series 145, 148
- differenced data 175
- differencing 136
- direct causality 41
- direct variance 74
- discrete dependent variable model 212
- dispersion 23–4, 32, 74
- distributed lag models 121–33
  - lagged variables 123–5
  - notation 125–58
  - other distributed lag models 131–3
  - selection of lag order 128–31
  - see also* regression with time lags
- distribution 20
- Dow Jones Industrial Average 13, 27, 33
- drift 180–1
- dummy dependent variables 119–20
  - model 212
- dummy variables 11, 109–20, 212
  - interacting dummy and non-dummy variables 117–19
  - multiple regression 112–14
  - multiple regression with dummy and non-dummy variables 115–17
  - simple regression 111–12
  - what if dependent variable is a dummy? 119–20
- duration 213
- dynamic effects 122
- $e$  51–2
- E-views 4, 212
- ECM *see* error correction model
- econometrics 1–2
- effect of advertising on sales: regression 54, 56, 61
- effect of interest rates on exchange rate 101
- elasticity 62
- election polls: regression 75–6
- empirical project topics 226–8
- empirical project writing 223–8
- empirical VAR model 195
- endogenous variable 216–20
- Engle–Granger test 168–9, 175, 178
- equation of a straight line 5
- equilibrium 161, 165–6, 171–4, 217
- equilibrium error 165
- equity pricing puzzle 226–7
- error 33, 52, 168, 213–15
  - autocorrelated 214
  - equilibrium error 165
- error correction model 171–4
- error message 67
- error variance 71–3
- estimation and testing with cointegrated variables 167–71
- European Monetary Union 181

- Excel 3–7, 15, 18, 31, 87, 110
  - note to users 164
- exchange rates 49, 101, 175
- exogenous variable 216–20
- explanatory variable 51, 122
- extensions of the AR(1) model 146–51
  
- F*-distribution 89
- F*-statistic 84–6, 194, 206–8
- factors affecting accuracy of  $\beta$  70–3
- Federal Reserve Bank of St Louis 17
- finance *see* applications of time series models
- financial derivatives 181
- fit 57–61
- fitted values and  $R^2$  57–61
- forecasting with VARs 200–3
- forms of error 213–15
- frequency 9, 19, 32
- frequency tables 19
- FTSE index 13
- Function Wizard<sup>®</sup> 35
- functions 5, 132
  
- Gauss 212
- GDP *see* gross domestic product
- GDP deflator 14
- general considerations when writing
  - empirical project 225
- generalized least squares estimator 214
- geometric lag model 131
- GLS *see* generalized least squares estimator
- goodness of fit 58
- government and other continuous surveys 16
- Granger causality 186–204
  - in an ADL model with  $p$  and  $q$  lags 188–90
  - causality in both directions 190–2
  - with cointegrated variables 192
  - in a simple ADL model 187–8
  - test 188
  - see also* causality; vector autoregressions
- Granger non-causality 188–9
- Granger Representation Theorem 171, 178
- graphical methods of working with data 18–22
- gross domestic product 5, 9, 11–14, 19, 216
- growth of GDP 11–12
  
- hedging 181
- hedonic pricing approach 92
- heteroskedasticity 179, 214
- high volatility 181
- histograms 19–20
- holding all other explanatory variables
  - constant 94, 96, 106, 127
  - see also ceteris paribus*
- house prices
  - and causality 38
  - explaining 92, 95–7
  - multiple regression 92, 95–7
  - and regression 78, 83
  - regression with dummy variables 109–13, 115–18
- hypothesis testing 70, 79–86, 205–9
  - involving more than one coefficient 205–9
  - involving  $R^2$  84–6
- identification 219
- immediate causality 41
- income 152
  - autocorrelation 139–41
  - autoregression 146, 149
  - trend 135–7
- index numbers 12–14, 25–31
  - calculating a banana price index 25–6
  - calculating a fruit price index 26–31
- indirect causality 41
- industrial accidents *see* accidents
- inequality 23
- inflation 13, 49, 181, 187
- initial public offerings 226
- instrumental variables estimator 215

- intelligence test 42, 47
- interest rates 9, 49, 101, 122, 175
- Internet 14–17
- interpretation of OLS estimates 54–7, 94–7
- interquartile range 32
- interval estimates 70
- intuition 3, 24, 56, 156, 214
  - economic 165
  - mathematical 144
- inverse variance 74
- IPOs *see* initial public offerings
- IS–LM model 216–17
  
- Keynesian views of the economy 197
- Koyck model 131
  
- lack of bias 220
- lag length 122, 129
  - selection in VARs 198–9
- lag order 128–31
- lagged variables 122–5
  - definition 123–5
- Laspeyres price index 27–8
- least squares estimate 53
- levels of significance 80, 83–4, 152
- levels versus growth rates 11–12
- LIMDEP 120, 212
- limitations and extensions 211–21
  - problems calling for multiple equation models 216–20
  - problems with particular forms of dependent variable 212–13
  - problems with particular forms of error 213–15
- limited dependent variable model 120, 212–13
- linear function 5
- linear regression model 51
- linearity assumption 62
- logarithmic transformation 62
- logarithms 6, 62
- logit estimator 120, 212
  
- logs *see* natural logarithms
- long memory behavior 140, 145, 156
- long run multiplier 161–3, 167, 172, 175, 194
- low power 155, 169, 178
- LPI *see* Laspeyres price index
- lung cancer 40–1, 107, 186–7
  
- macro-economic time series databanks 17
- macroeconomics *see* applications of time series models
- magnitude of correlation 43, 46
- Manchester Information & Associated Services *see* MIMAS
- marginal effect 54, 66, 161
- market efficiency 209
- mathematical details for the AR(1) model 156
- mathematical details of correlation 46–7
- mathematical details of regression
  - coefficients 106
- mathematical details of simple regression 65–6
- matrix algebra 107
- mean 23, 32, 66
  - deviations from 66
- measure of location 31
- measuring the fit of a regression model 57–61
- median 31–2
- medical statistics 110
- microeconomics 122
- MicroFit 4, 84, 87, 124, 154, 168, 212
  - forecasting with VARs 201–3
- MIMAS 15–17
  - census and related data sets 16
  - government and other continuous surveys 16
  - macro-economic time series databanks 17
- mode 31
- model fit 224
- money demand 167

- monthly frequency 9
- multicollinearity 100–5, 132–3
- multicollinearity problem 101–6, 132, 149, 160, 175
- multiple equation models 216–20
- multiple regression 91–108, 112–17
  - with dummy and non-dummy variables 115–17
  - with dummy variable 112–14
  - interpreting OLS estimates 94–7
  - mathematical interpretation of regression coefficients 106
  - multicollinearity 100–5, 132–3
  - omitted variables bias 99–100
  - pitfalls of using simple regression in
    - multiple regression context 97–9
    - regression as a best fitting line 92–3
    - statistical aspects of multiple regression 93–4
- multiplier 161
  
- National Bureau of Economic Research 15
- natural logarithms 6, 62, 135
- near-zero correlation 138, 140
- negative correlation 36, 44, 181
- negative relationships 22, 42
- New York Stock Exchange 186
- nominal GDP 14
- non-dummy variables 115–19
  - interacting dummy and non-dummy variables 117–19
  - see also* dummy variables
- non-negative integers 213
- nonlinearity 61–4
- nonstationarity 138, 144
  - see also* unit root
- nonstationary regression 121
- nonstationary time series 145–6, 148
- nonstationary variables 121, 133
- notation 125–8
- null hypothesis 79, 84, 169
- NYSE *see* New York Stock Exchange
- observations 123–5, 138
- obtaining data 14–17
  - MIMAS 16–17
  - Resources for Economists on the Internet 15
- OLS *see* ordinary least squares estimates
- omitted variables bias 99–100, 112
- OPEC oil embargo 155
- ordinary least squares estimates 53–7, 69, 94–7
- ordinary least squares estimation of the multiple regression model 93
- organic oranges 166–7, 169–70, 173–4, 195, 204
- other distributed lag models 131–3
- outliers 22, 58, 224
  
- $p$  and  $q$  lags 188–90
- P-values 2, 80–3, 94–5, 99–100, 103–5, 168
- Paasche price index 28–31
- panel data 11
- partial derivative 106
- PcGive 124, 212
- Penn World Table 15
- percentage change 12, 175
- percentiles 32
- perfect correlation 43–4
- perfect multicollinearity 151
- permanent change 161
- permanent income hypothesis 167
- personal income *see* income
- pie chart 18
- pitfalls of using simple regression in
  - multiple regression context 97–9
- point estimates 70, 75
- polynomial distributed lag model 132
- positive correlation 36, 43–4
- positive relationships 22, 42
- pound/dollar exchange rate 181–2
  - see also* volatility in asset prices
- PPI *see* Paasche price index
- predicted value 58
- price index 12–13

- price inflation 189–91, 208–9
- price levels 12–13
- prices of two goods: cointegration 166, 169–70, 173–4
- probability theory 69
- probit estimator 120, 212
- problems calling for multiple equation models 216–20
- problems with particular forms of dependent variable 212–13
- problems with particular forms of error 213–15
- proper weighted average 28
- properties of correlation 36
- proximate causality 41
- purchasing power parity 19, 167
- PWT *see* Penn World Table
- quadratic function 132
- qualitative choice model 212–13
- qualitative data 10–11, 109, 120
- quantitative data 10–11, 109
- quarterly frequency 9
- quartiles 32
- $r$  35–7, 46
- R. Zelig 7
- random walk with drift 180–1
- random walk model 145–6, 186
- randomly fluctuating behavior 143
- raw data 11, 18–20, 66
- real GDP 11–12, 14, 135
- recession 138
- regression as a best fitting line 50–4, 92–3
- regression with dummy variables 109–20
  - interacting dummy and non-dummy variables 117–19
  - multiple regression 112–14
  - multiple regression with dummy and non-dummy variables 115–17
  - simple regression 111–12
  - what if the dependent variable is a dummy? 119–20
- regression line 50
- regression model 50–1
- regression with time lags 121–34
  - lagged variables 123–5
  - notation 125–8
  - other distributed lag models 131–3
  - selection of lag order 128–31
- regression with time series variables 159–78
  - when  $X$  and  $Y$  are stationary 160–4
  - when  $Y$  and  $X$  are cointegrated 171–4
  - when  $Y$  and  $X$  have unit roots:
    - cointegration 165–71
    - when  $Y$  and  $X$  have unit roots: spurious regression 164–5
    - when  $Y$  and  $X$  have unit roots but are not cointegrated 175–6
- replication 110
- report elements 223–4
- residuals 52, 58–9, 168, 173
- Resources for Economists on the Internet 15
- restricted regression models 206
- retail price index 30
- RMPY variables 196–8, 202
- robustness 214
- RPI *see* retail price index
- $R^2$  57–61, 93–4, 100–4, 224
  - hypothesis testing 84–6
- $r_p$  139
- S&P500 13, 33
- safety training 126–8, 130, 161
- sales and effect of advertising 54, 56, 61
- sample size 23
- scatter diagram 21
- seasonal dummies 150–1
- seasonality 150
  - see also* dummy variables
- seasoned equity offerings 226–7
- seemingly unrelated regression equations estimator 218
- selection of lag length in VARs 198–9
- selection of lag order 128–31

- selectivity 225
- SEOs *see* seasoned equity offerings
- series 135
- SHAZAM 87, 124, 154, 212
- short run behavior 172, 174
- significance 80, 83–4, 152
- simple ADL model 187–90
  - with  $p$  and  $q$  lags 188–90
- simple random walk model 180
- simple regression 49–68, 111–12
  - with dummy variable 111–12
  - fitted values and  $R^2$  57–61
  - interpreting OLS estimates 54–7
  - mathematical details 65–6
  - nonlinearity in regression 61–4
  - regression as a best fitting line 50–4
- simultaneity 210
- simultaneous equations model 215, 218
- single factor ANOVA 110
- splicing 30–1
- spreadsheet 88, 124, 131
- spurious regression 121, 164–5, 168, 175–6
- square 183
- $s_b$  73–4, 80–1, 88, 94
- SSR *see* sum of squared residuals
- stability 161, 177
- stagnation 23
- standard deviation 23–4, 32, 74
- standard error 74
- Stata 4, 212
- state of the economy 216
- stationarity 144
- stationarity condition 145
- stationary regression 121
- stationary time series 145–6, 149
- stationary variables 122
- statistical aspects of multiple regression 93–4
- statistical aspects of regression 69–89
  - calculating confidence interval for  $\beta$  73–9
  - factors affecting accuracy of  $\beta$  70–3
  - hypothesis testing involving  $R^2$  84–6
  - testing whether  $\beta = 0$  79–83
  - using statistical tables to test whether  $\beta = 0$  87–8
- statistical tables 87–8
- Statistics in Sport 17
- stochastic trend 148, 156, 160
- stock prices 179–80, 183–6
- straight-line relationship 111, 157
- structural breaks 155
- Student- $t$  distribution 75, 80, 87
- sum of squared residuals 53, 57
- summation notation 5–6
- supply–demand model 217–18
- SURE *see* seemingly unrelated regression
  - equations estimator
- systems of equations 216
- $t$ -statistic 80–1, 84, 95, 100, 168
  - see also* Dickey–Fuller test
- tendency 22, 37, 55
- test statistic 80
- testing in the AR( $p$ ) 151–5
  - testing involving  $\rho$  153–5
  - testing involving  $\alpha$ ,  $\gamma$ ,  $\dots$ ,  $\gamma_{p-t}$ , and  $\delta$  151–3
- testing whether  $\beta = 0$  79–83, 87–8
- theoretical macroeconomic model 195
- time series data 9–10, 32
- time series graphs 18–19
- time series regression when  $X$  and  $Y$  are stationary 160–4
- time series regression when  $Y$  and  $X$  have unit roots but are not cointegrated 175–6
- total derivative 106
- total multiplier 161
- total sum of squares 59
- Treasury bill rate 101, 196
- trend behavior 143–4, 147–9, 156, 165
- trend stationary 149
- trends 136
- $t_b$  73, 75–6, 87
- TSP 212

- TSS *see* total sum of squares
- “twin peaks” 20, 23
- two-factor ANOVA 110
- types of economic data 9–14
- cross-sectional data 10
  - data transformations 11–12
  - distinction between qualitative and quantitative data 10–11
  - index numbers 12–14
  - panel data 11
  - time series data 9–10
  - typical empirical project 223–4
- $u$  52–3
- understanding correlation 35–6
- properties of correlation 36
- unemployment 49, 122, 213
- unit root 147, 151, 164–5, 168–9, 175
- see also* nonstationarity
- unit root nonstationarity 145
- units 10, 54–5, 62
- univariate time series analysis 135–58
- autocorrelation function 138–42
  - autoregressive model 142–4
  - extensions of the AR(1) model 146–51
  - mathematical intuition for the AR(1) model 156
  - nonstationary versus stationary time series 145–6
  - testing in the AR( $p$ ) with deterministic trend model 151–5
- unrestricted regression models 206
- urbanization 22
- using few explanatory variables 99
- using statistical tables to test whether  $\beta = 0$  87–8
- value spread 72–3
- VAR *see* vector autoregressions
- VAR(1) with RMPY variables 196–7
- VAR(2) with RMPY variables 198, 202
- variables 9, 37, 39–42
- and correlation 39–42, 45–6
  - estimation and testing with cointegrated variables 167–71
  - lagged 122–5
  - nonstationary 121
  - stationary 122
- variance 24
- vector autoregressions 193–204, 216
- with cointegrated variables 203–4
  - forecasting 200–3
  - lag length selection 198–9
  - see also* Granger causality
- verbal reasoning and correlation 36–9
- volatility in asset prices 179–86
- wage inflation 189–91, 208–9
- wage-setting behavior 227–8
- weekly frequency 9
- weighted average 27–9
- proper 28
- Wharton Research Data Services 17
- working with data 18–25
- descriptive statistics 23–5
  - histograms 19–20
  - time series graphs 18–19
  - XY-plots 21–2
- World Wide Web 14–15
- writing an empirical project 223–8
- general considerations 225
  - project topics 226–8
  - typical empirical project 223–4
- XY-plots 21–2, 42–5, 50, 54, 61–4
- and correlation 42–5
  - factors affecting estimate accuracy 70–3
  - measuring fit of a regression model 57–61
- Yahoo! 17

