

Contents

Contributors	xi
Preface	xiii

Part I Finance

1. Financial econometrics and big data: A survey of volatility estimators and tests for the presence of jumps and co-jumps	3
<i>Arpita Mukherjee, Weijia Peng, Norman R. Swanson, and Xiye Yang</i>	
1 Introduction	4
2 Setup	6
3 Realized measures of integrated volatility	8
3.1 Realized volatility	9
3.2 Realized bipower variation	9
3.3 Tripower variation	10
3.4 Two-scale realized volatility	10
3.5 Multiscale realized volatility	11
3.6 Realized kernel	11
3.7 Truncated realized volatility	12
3.8 Modulated bipower variation	12
3.9 Threshold bipower variation	12
3.10 Subsampled realized kernel	13
3.11 <i>MedRV</i> and <i>MinRV</i>	14
4 Jump testing	14
4.1 Barndorff-Nielsen and Shephard test	15
4.2 Lee and Mykland test	15
4.3 Jiang and Oomen test	16
4.4 Aït-Sahalia and Jacod test	17
4.5 Podolskij and Ziggel (PZ) test	18
4.6 Corradi, Silvapulle, and Swanson test	18
5 Co-jump testing	19
5.1 BLT co-jump testing	20
5.2 JT co-jump testing	21
5.3 MG threshold co-jump test	22

5.4	GST co-exceedance rule	22
5.5	CKR co-jump testing	23
6	Empirical experiments	24
6.1	Data description	24
6.2	Methodology	25
6.3	Findings	26
7	Conclusion	48
	Appendix R code	49
	References	57
2.	Real time monitoring of asset markets: Bubbles and crises	61
	<i>Peter C.B. Phillips and Shuping Shi</i>	
1	Introduction	62
2	The PSY Procedure	64
2.1	The Augmented Dickey–Fuller test	64
2.2	The Recursive Evolving Algorithm	66
3	The PSY Test for Bubble Identification	68
3.1	The Rationale	68
3.2	Consistency	68
4	The PSY Test for Crisis Identification	70
4.1	The Rationale	70
4.2	Consistency	71
5	A New Composite Bootstrap	72
6	Empirical Applications with R	72
6.1	Example 1: The S&P 500 Market	72
6.2	Example 2: Credit Risk in the European Sovereign Sector	75
7	Conclusion	77
	References	77
	Further reading	80
3.	Component-wise AdaBoost algorithms for high-dimensional binary classification and class probability prediction	81
	<i>Jianghao Chu, Tae-Hwy Lee, and Aman Ullah</i>	
1	Introduction	82
2	AdaBoost	83
3	Extensions to AdaBoost algorithms	90
3.1	Real AdaBoost	91
3.2	LogitBoost	92
3.3	Gentle AdaBoost	94
4	Alternative classification methods	96
4.1	Deep Neural Network	96
4.2	Logistic regression with LASSO	102
4.3	Semiparametric single-index model	104
5	Monte Carlo	107

6 Applications	111
7 Conclusions	112
Acknowledgments	113
References	113

Part II

Macro Econometrics

4. Mixed data sampling (MIDAS) regression models	117
<i>Eric Ghysels, Virmantas Kvedaras, and Vaidotas Zemlys-Balevičius</i>	
1 Introduction	117
2 A stylized MIDAS regression model	119
2.1 A few examples of the constraint function h	121
2.2 Selection of h , d , and k	121
2.3 Statistical inference	122
3 Linear and quasi-linear MIDAS models (affine g)	123
3.1 Unconstrained MIDAS	124
3.2 MIDAS	125
3.3 MIDAS with nonparametric smoothing of weights	127
4 Nonlinear parametric MIDAS models	128
4.1 General considerations	128
4.2 Logistic smooth transition MIDAS (LSTR-MIDAS)	129
4.3 MIDAS with min–mean–max effects (MMM-MIDAS)	130
5 Semiparametric MIDAS models	131
5.1 MIDAS with partially (quasi)linear effects (PL-MIDAS)	133
5.2 The single index MIDAS model (SI-MIDAS)	134
6 Illustration with simulated data	135
6.1 Data generation	135
6.2 Estimation	137
7 Empirical examples	144
7.1 Okun’s law	145
7.2 Inflation and the effective federal funds rate	149
References	151
5. Encouraging private corporate investment in India	155
<i>Hrishikesh Vinod, Honey Karun, and Lekha S. Chakraborty</i>	
1 Introduction	155
2 Literature review	157
3 Interpreting data and model implications	161
4 Estimation and results	164
5 Data abbreviations and sources	173
6 Causality results	175
7 Conclusion	177
Appendix	178
References	182

6.	High-mixed frequency forecasting methods in R—With applications to Philippine GDP and inflation	185
	<i>Roberto S. Mariano and Suleyman Ozmuçur</i>	
1	Introduction	186
2	Alternative forecasting models in this study	187
2.1	Quarterly models	188
2.2	Monthly models	189
2.3	MIDAS regressions	189
2.4	Mixed-frequency dynamic latent factor models (MF-DLFM)	190
3	Application to forecasting Philippine GDP and inflation and computer implementation in R	191
3.1	Getting data with R	191
3.2	Statistical properties of quarterly real GDP and GDP deflator	192
4	Estimated models in R	196
4.1	Box–Jenkins (ARIMA(p,d,q)) univariate time-series models	196
4.2	Vector autoregressive models	197
4.3	Bridge equations	201
4.4	Principal components using monthly indicators and bridge equations using principal components	203
4.5	MIDAS models	205
4.6	ADL-MIDAS	207
4.7	MIDAS-VAR	209
4.8	VARX and VARXM	211
4.9	ARIMA and MIDAS	213
4.10	Factor MIDAS	214
4.11	Dynamic factor models—Small number of indicators	216
4.12	Dynamic factor models—Large number of indicators	219
5	Comparison of forecasts and concluding remarks	223
	Acknowledgments	224
	References	224
	Further reading	227

7. Nonlinear time series in R: Threshold cointegration with tsDyn **229**

Matthieu Stigler

1	Introduction: Linear and threshold cointegration	230
2	Estimation, testing, and interpretation	233
2.1	Estimation and testing for cointegration	233
2.2	Estimation and testing for threshold models	234
2.3	Estimation and testing for threshold cointegration models	237
2.4	Estimation of a threshold estimated model	240
2.5	(Generalized) impulse response functions	240

3	The tsDyn package	243
3.1	Linear models: AR, VAR, and VECM in tsDyn	243
3.2	Univariate models: SETAR	244
3.3	Multivariate models: TVAR and TVECM	245
4	Empirical application	246
4.1	Unit roots and linear cointegration	247
4.2	Threshold cointegration in the univariate residual-based approach	250
4.3	Threshold cointegration in the multivariate system-based approach	256
5	Conclusion	260
	Acknowledgment	260
	References	260
	Further reading	264

Part III

Micro Econometrics

8.	Econometric analysis of productivity: Theory and implementation in R	267
	<i>Robin C. Sickles, Wonho Song, and Valentin Zelenyuk</i>	
1	Introduction	267
2	Why estimate production (technical) efficiency?	268
3	Regression-based methods to estimate production efficiency	270
3.1	The stochastic frontier paradigm	270
4	Envelopment estimators	280
4.1	The origins of DEA	280
4.2	The basic DEA model	280
4.3	The myriad of DEA models	283
4.4	Statistical analysis of DEA and FDH	286
5	SFA efficiency software in R	287
5.1	Basic model setup	288
5.2	Figures and tables	288
5.3	Different estimators	288
6	DEA efficiency software in R	289
7	Summary and final remarks	290
	Acknowledgments	290
	References	290
9.	Stochastic frontier models using R	299
	<i>Giancarlo Ferrara</i>	
1	Introduction	299
2	Methods	301
2.1	Contextual variables	305
2.2	Spatial external factors	307
2.3	P-splines: Computational aspects	308

3	Numerical illustrations	312
3.1	Example 1: Linear homoscedastic model	312
3.2	Example 2: Nonlinear exponential homoscedastic model	313
3.3	Example 3: Nonmonotone model	315
3.4	Example 4: Quadratic polynomial model with heteroscedasticity	317
4	Empirical application to crops data	318
5	Conclusions	323
	References	323
	Index	327