

4.18	New developments in composite indicators applications	1044
4.18.1	Building composite indicators in the functional domain: a suggestion for an evolutionary HDI. <i>Francesca Fortuna, Alessia Naccarato and Silvia Terzi</i>	1045
4.18.2	Small Area Estimation of Inequality Measures via Simplex Regression. <i>Silvia De Nicolò, Maria Rosaria Ferrante and Silvia Pacei</i>	1051
4.18.3	Relational Well-Being and Poverty in Italy Benessere relazionale e povertà in Italia. <i>Elena Dalla Chiara and Federico Perali</i>	1057
4.18.4	A composite indicator to assess sustainability of agriculture in European Union countries. <i>Alessandro Magrini and Francesca Giambona</i>	1063
4.18.5	Interval-Based Composite Indicators with a Triplex Representation: A Measure of the Potential Demand for the “Ristori” Decree in Italy. <i>Carlo Drago</i>	1069
4.19	New developments in GLM theory and applications	1075
4.19.1	Variational inference for the smoothing distribution in dynamic probit models. <i>Augusto Fasano and Giovanni Rebaudo</i>	1076
4.19.2	Interpretability and interaction learning for logistic regression models. <i>Nicola Rares Franco, Michela Carlotta Massi, Francesca Ieva and Anna Maria Paganoni</i>	1082
4.19.3	Entropy estimation for binary data with dependence structures. <i>Linda Altieri and Daniela Cocchi</i>	1088
4.19.4	A Comparison of Some Estimation Methods for the Three-Parameter Logistic Model. <i>Michela Battauz and Ruggero Bellio</i>	1094
4.19.5	A statistical model to identify the price determinations: the case of Airbnb. <i>Giulia Contu, Luca Frigau, Gian Paolo Zammarchi and Francesco Mola</i>	1100
4.20	New developments in social statistics analysis	1106
4.20.1	Data-based Evaluation of Political Agents Against Goals Scheduling. <i>Giulio D'Epifanio</i>	1107
4.20.2	Local heterogeneities in population growth and decline. A spatial analysis for Italian municipalities. <i>Federico Benassi, Annalisa Busetta, Gerardo Gallo and Manuela Stranges</i>	1113
4.20.3	The assessment of environmental and income inequalities. <i>Michele Costa</i>	1119
4.20.4	Household financial fragility across Europe. <i>Marianna Brunetti, Elena Giarda and Costanza Torricelli</i>	1125
4.20.5	Refugees' perception of their new life in Germany. <i>Daria Mendola and Anna Maria Parroco</i>	1131
4.21	New perspectives in clinical trials	1137
4.21.1	Improved maximum likelihood estimator in relative risk regression. <i>Euloge C. Kenne Pagui, Francesco Pozza and Alessandra Salvan</i>	1138
4.21.2	Development and validation of a clinical risk score to predict the risk of SARS-CoV-2 infection. <i>Laura Savaré, Valentina Orlando and Giovanni Corrao</i>	1144
4.21.3	Functional representation of potassium trajectories for dynamic monitoring of Heart Failure patients. <i>Caterina Gregorio, Giulia Barbatì and Francesca Ieva</i>	1150
4.21.4	Effect of lung transplantation on the survival of patients with cystic fibrosis: IMaCh contribution to registry data. <i>Cristina Giudici, Nicolas Brouard and Gil Bellis</i>	1156
4.21.5	Categories and Clusters to investigate Similarities in Diabetic Kidney Disease Patients. <i>Veronica Distefano, Maria Mannone, Claudio Silvestri and Irene Poli</i>	1162

4.22	New perspectives in models for multivariate dependency	1168
4.22.1	Parsimonious modelling of spectroscopy data via a Bayesian latent variables approach. <i>Alessandro Casa, Tom F. O'Callaghan and Thomas Brendan Mur</i>	1169
4.22.2	Bias reduction in the equicorrelated multivariate normal. <i>Elena Bortolato and Euloge Clovis Kenne Pagui</i>	1175
4.22.3	Some results on identifiable parameters that cannot be identified from data. <i>Christian Hennig</i>	1181
4.23	Novel approaches for official statistics	1187
4.23.1	Web data collection: profiles of respondents to the Italian Population Census. <i>Elena Grimaccia, Gerardo Gallo, Alessia Naccarato, Novella Cecconi and Alessandro Fratoni</i>	1188
4.23.2	Trusted Smart Surveys: architectural and methodological challenges at a glance. <i>Mauro Bruno, Francesca Inglese and Giuseppina Ruocco</i>	1194
4.23.3	On bias correction in small area estimation: An M-quantile approach. <i>Gaia Bertarelli, Francesco Schirripa Spagnolo, Raymond Chambers and David Haziza</i>	1200
4.23.4	The address component of the Statistical Base Register of Territorial Entities. <i>Davide Fardelli, Enrico Orsini and Andrea Pagano</i>	1206
4.23.5	A well-being municipal indicator using census data: first results. <i>Massimo Esposito</i>	1212
4.24	Prior distribution for Bayesian analysis	1218
4.24.1	On the dependence structure in Bayesian nonparametric priors. <i>Filippo Ascolani, Beatrice Franzolini, Antonio Lijoi, and Igor Prünster</i>	1219
4.24.2	Anisotropic determinantal point processes and their application in Bayesian mixtures. <i>Lorenzo Ghilotti, Mario Beraha and Alessandra Guglielmi</i>	1226
4.24.3	Bayesian Screening of Covariates in Linear Regression Models Using Correlation Thresholds. <i>Ioannis Ntzoufras and Roberta Paroli</i>	1232
4.25	Recent advances in clustering methods	1238
4.25.1	Biclustering longitudinal trajectories through a model-based approach. <i>Francesca Martella, Marco Alfò and Maria Francesca Marino</i>	1239
4.25.2	Monitoring tools for robust estimation of Cluster Weighted models. <i>Andrea Cappelletto and Francesca Greselin</i>	1245
4.25.3	Co-clustering Models for Spatial Transcriptomics: Analysis of a Human Brain Tissue Sample. <i>Andrea Sottosanti and Davide Risso</i>	1251
4.25.4	Graph nodes clustering: a comparison between algorithms. <i>Ilaria Bombelli</i>	1257
4.26	Social demography	1263
4.26.1	Childcare among migrants: a comparison between Italy and France. <i>Eleonora Trappolini, Elisa Barbiano di Belgiojoso, Stefania Maria Lorenza Rimoldi and Laura Terzera</i>	1264
4.26.2	Employment Uncertainty and Fertility in Italy: The Role of Union Formation. <i>Giammarco Alderotti, Valentina Tocchioni and Alessandra De Rose</i>	1270
4.26.3	Determinants of union dissolution in Italy: Do children matter? <i>Valentina Tocchioni, Daniele Vignoli, Eleonora Meli and Bruno Arpino</i>	1276
4.26.4	Working schedules and fathers' time with children: A Sequence Analysis. <i>Annalisa Donno and Maria Letizia Tanturri</i>	1282
4.26.5	Correlates of the non-use of contraception among female university students in Italy. <i>Annalisa Busetta, Alessandra De Rose and Daniele Vignoli</i>	1288

4.27	Social indicators applications and methods	1294
4.27.1	A logistic regression model for predicting child language performance. <i>Andrea Briglia, Massimo Mucciardi and Giovanni Pirrotta</i>	1295
4.27.2	Subject-specific measures of interrater agreement for ordinal scales. <i>Giuseppe Bove</i>	1301
4.27.3	A Tucker3 method application on adjusted-PMRs for the study of work-related mortality. <i>Vittoria Carolina Malpassuti, Vittoria La Serra and Stefania Massari</i>	1307
4.27.4	Two case-mix adjusted indices for nursing home performance evaluation. <i>Giorgio E. Montanari and Marco Doretti</i>	1313
4.27.5	The ultrametric covariance model for modelling teachers' job satisfaction. <i>Carlo Cavicchia, Maurizio Vichi and Giorgia Zaccaria</i>	1319
4.28	Some recent developments in compositional data analysis	1325
4.28.1	A Robust Approach to Microbiome-Based Classification Problems. <i>Gianna Serafina Monti and Peter Filzmoser</i>	1326
4.28.2	What is a convex set in compositional data analysis? <i>Jordi Saperas i Riera, Josep Antoni Martín Fernández</i>	1332
4.28.3	Compositional Analysis on the Functional Distribution of Extended Income. <i>Elena Dalla Chiara and Federico Perali</i>	1338
4.28.4	Evaluating seasonal-induced changes in river chemistry using Principal Balances. <i>Caterina Gozzi and Antonella Buccianti</i>	1344
4.28.5	Compositional Data Techniques for the Analysis of the Ragweed Allergy. <i>Gianna S. Monti, Maira Bonini, Valentina Ceriotti, Matteo Pelagatti and Claudio M. Ortolani</i>	1350
4.29	Spatial data analysis	1356
4.29.1	Spatial multilevel mixed effects modeling for earthquake insurance losses in New Zealand. <i>F. Marta L. Di Lascio and Selene Perazzini</i>	1357
4.29.2	Weighted distances for spatially dependent functional data. <i>Andrea Diana, Elvira Romano, Claire Miller and Ruth O'Donnell</i>	1363
4.29.3	Spatial modeling of childcare services in Lombardia. <i>Emanuele Aliverti, Stefano Campostrini, Federico Caldura and Lucia Zanotto</i>	1369
4.29.4	On the use of a composite attractiveness index for the development of sustainable tourist routes. <i>Claudia Cappello, Sandra De Iaco, Sabrina Maggio and Monica Palma</i>	1375
4.30	Statistical applications in education	1381
4.30.1	Does self-efficacy influence academic results? A separable-effect mediation analysis. <i>Chiara Di Maria</i>	1382
4.30.2	Statistics Knowledge assessment: an archetypal analysis approach. <i>Bruno Adabbo, Rosa Fabbricatore, Alfonso Iodice D'Enza and Francesco Palumbo</i>	1388
4.30.3	Exploring drivers for Italian university students' mobility: first evidence from AlmaLaurea data. <i>Giovanni Boscaïno and Vincenzo Giuseppe Genova</i>	1394
4.30.4	Can Grading Policies influence the competition among Universities of different sizes? <i>Gabriele Lombardi and Antonio Pio Distaso</i>	1400
4.30.5	The class A journals and the Italian academic research outcomes in Statistical Sciences. <i>Maria Maddalena Barbieri, Francesca Bassi, Antonio Irpino and Rosanna Verde</i>	1406
4.31	Statistical methods for finance	1412
4.31.1	Hypotheses testing in mixed-frequency volatility models: a bootstrap approach. <i>Vincenzo Candila and Lea Petrella</i>	1413

4.31.2	Quantile Regression Forest with mixed frequency Data. <i>Mila Andreani, Vincenzo Candila and Lea Petrella</i>	1419
4.31.3	Higher order moments in Capital Asset Pricing Model betas. <i>Giuseppe Arbia, Riccardo Bramante and Silvia Facchinetti</i>	1425
4.31.4	When Does Sentiment Matter in Predicting Cryptocurrency Bubbles? <i>Arianna Agosto and Paolo Pagnottoni</i>	1431
4.32	Statistical methods for high dimensional data	1437
4.32.1	Virtual biopsy in action: a radiomic-based model for CALI prediction. <i>Francesca Ieva, Giulia Baroni, Lara Cavinato, Chiara Masci, Guido Costa, Francesco Fiz, Arturo Chiti and Luca Viganò</i>	1438
4.32.2	Functional alignment by the "light" approach of the von Mises-Fisher-Procrustes model. <i>Angela Andreella and Livio Finos</i>	1444
4.32.3	A screening procedure for high-dimensional autologistic models. <i>Rodolfo Metulini and Francesco Giordano</i>	1450
4.32.4	Covariate adjusted censored gaussian lasso estimator. <i>Luigi Augugliaro, Gianluca Sottile and Veronica Vinciotti</i>	1456
4.32.5	Ranking-Based Variable Selection for ultra-high dimensional data in GLM framework. <i>Francesco Giordano, Marcella Niglio and Marialuisa Restaino</i>	1462
4.33	Statistical methods in higher education	1468
4.33.1	Effects of remote teaching on students' motivation and engagement: the case of the University of Modena & Reggio Emilia. <i>Isabella Morlini and Laura Sartori</i>	1469
4.33.2	A random effects model for the impact of remote teaching on university students' performance. <i>Silvia Bacci, Bruno Bertaccini, Simone Del Sarto, Leonardo Grilli and Carla Rampichini</i>	1475
4.33.3	Multinomial semiparametric mixed-effects model for profiling engineering university students. <i>Chiara Masci, Francesca Ieva and Anna Maria Paganoni</i>	1481
4.33.4	Evaluating Italian universities: ANVUR periodic accreditation judgment versus international rankings. <i>Angela Maria D'Uggento, Nunziata Ribecco and Vito Ricci</i>	1487
4.33.5	Women's career discrimination in the Italian Academia in the last 20. <i>Daniele Cuntrera, Vincenzo Falco and Massimo Attanasio</i>	1493
4.34	Statistical methods with Bayesian networks	1499
4.34.1	Statistical Micro Matching Using Bayesian Networks. <i>Pier Luigi Conti, Daniela Marella, Paola Vicard and Vincenzina Vitale</i>	1500
4.34.2	Modeling school managers challenges in the pandemic era with Bayesian networks. <i>Maria Chiara De Angelis and Flaminia Musella and Paola Vicard</i>	1506
4.34.3	Structural learning of mixed directed acyclic graphs: a copula-based approach. <i>Federico Castelletti</i>	1512
4.34.4	Inference on Markov chains parameters via Large Deviations ABC. <i>Cecilia Viscardi, Fabio Corradi, Michele Boreale and Antonietta Mira</i>	1518
4.34.5	A propensity score approach for treatment evaluation based on Bayesian Networks. <i>Federica Cugnata, Paola M.V. Rancoita, Pier Luigi Conti, Alberto Briganti, Clelia Di Serio, Fulvia Mecatti and Paola Vicard</i>	1524
4.35	Statistical modelling for the analysis of contemporary societies	1530
4.35.1	Social Network Analysis to analyse the relationship between 'victim-author' and 'motivation' of violence against women in Italy. <i>Alessia Forciniti</i>	1531
4.35.2	Satisfaction and sustainability propensity among elderly bike-sharing users. <i>Paolo Maranzano, Roberto Ascari, Paola Maddalena Chiodini and Giancarlo Manzi</i>	1537

4.35.3	Media and Investors' Attention. Estimating analysts' ratings and sentiment of a financial column to predict abnormal returns. <i>Riccardo Ferretti and Andrea Sciandra</i>	1543
4.35.4	Predictions of regional HCE: spatial and time patterns in an ageing population framework. <i>Laura Rizzi, Luca Grassetti, Divya Brundavanam, Alvisa Palese and Alessio Fornasin</i>	1549
4.36	Surveillance methods and statistical models in the Covid-19 crisis	1555
4.36.1	The Italian Social Mood on Economy Index during the Covid-19 Crisis. <i>Alessandra Righi and Diego Zardetto</i>	1556
4.36.2	Modeling the first wave of the COVID-19 pandemic in the Lombardy region, Italy, by using the daily number of swabs. <i>Claudia Furlan and Cinzia Mortarino</i>	1562
4.36.3	Analysing the Covid-19 pandemic in Italy with the SIPRO model. <i>Martina Amongero, Enrico Bibbona and Gianluca Mastrantonio</i>	1568
4.36.4	Intentions of union formation and dissolution during the COVID-19 pandemic. <i>Bruno Arpino and Daniela Bellani</i>	1574
4.37	Time series methods	1580
4.37.1	Bootstrap-based score test for INAR effect. <i>Riccardo Ievoli and Lucio Palazzo</i>	1581
4.37.2	Evaluating the performance of a new picking algorithm based on the variance piecewise constant models. <i>Nicoletta D'Angelo, Giada Adelfio, Antonino D'Alessandro and Marcello Chiodi</i>	1587
4.37.3	Conditional moments based time series cluster analysis. <i>Raffaele Mattera and Germana Scepi</i>	1593
4.37.4	On the asymptotic mean-squared prediction error for multivariate time series. <i>Gery Andrés Díaz Rubio, Simone Giannerini, and Greta Goracci</i>	1599
4.37.5	Spherical autoregressive change-point detection with applications. <i>Federica Spoto, Alessia Caponera and Pierpaolo Brutti</i>	1605
5	Posters	1611
5.1	A method for incorporating historical information in non-inferiority trials. <i>Fulvio De Santis and Stefania Gubbiotti</i>	1612
5.2	Optimal credible intervals under alternative loss functions. <i>Fulvio De Santis and Stefania Gubbiotti</i>	1618
5.3	Statistical learning for credit risk modelling. <i>Veronica Bacino, Alessio Zoccarato, Caterina Liberati and Matteo Borrotti</i>	1624
5.4	Evaluating heterogeneity of agreement with strong prior information. <i>Federico M. Stefanini</i>	1630
5.5	Analysis of the spatial interdependence of the size of endoreduplicated nuclei observed in confocal microscopy. <i>Ivan Sciascia, Andrea Crosino, Gennaro Carotenuto and Andrea Genre</i>	1636
5.6	A Density-Peak Approach to Clustering Graph-Structured Data. <i>Riccardo Giubilei</i>	1642
5.7	The employment situation of people with disabilities in Tuscany, A Survey on the workplace. <i>Paolo Addis, Alessandra Coli and Gianfranco Francese</i>	1648
5.8	Robustness of statistical methods for modeling paired count data using bivariate discrete distributions with general dependence structures. <i>Marta Nai Ruscone and Dimitris Karlis</i>	1654

6	Satellite events	1660
6.1	Measuring uncertainty in key official economic statistics	1661
6.1.1	Uncertainty in production and communication of statistics: challenges in the new data ecosystem. <i>Giorgio Alleva and Piero Demetrio Falorsi</i>	1662
6.1.2	Uncertainty and variance estimation techniques for poverty and inequality measures from complex surveys: a simulation study. <i>Riccardo De Santis, Lucio Barabesi and Gianni Betti</i>	1668
6.1.3	Pandemics and uncertainty in business cycle analysis. <i>Jacques Anas, Monica Billio, Leonardo Carati, Gian Luigi Mazzi and Hionia Vlachou</i>	1674
6.2	Covid-19: the urgent call for a unified statistical and demographic challenge	1680
6.2.1	Environmental epidemiology and the Covid-19 pandemics	1681
6.2.1.1	The Covid-19 outbreaks and their environment: The Valencian human behaviour. <i>Xavier Barber, Elisa Espín, Lucía Guevara, Aurora Mula, Kristina Polotskaya and Alejandro Rabasa</i>	1682
6.2.2	Estimation of Covid 19 prevalence	1686
6.2.2.1	Optimal spatial sampling for estimating the SARS-Cov-2 crucial parameters. <i>Piero Demetrio Falorsi and Vincenzo Nardelli</i>	1687
6.2.2.2	Survey aimed to estimate the seroprevalence of SARS-CoV-2 infection in Italian population at national and regional level. <i>Stefano Falorsi, Michele D'Alò, Claudia De Vitiis, Andrea Fasulo, Danila Filipponi, Alessio Guandalini, Francesca Inglese, Orietta Luzi, Enrico Orsini and Roberta Radini</i>	1693
6.2.3	Measuring and modeling inequalities following the Covid-19 crisis	1699
6.2.3.1	COVID-19 impacts on young people's life courses: first results. <i>Antonietta Bisceglia, Concetta Scolorato and Giancarlo Ragozini</i>	1700
6.2.3.2	Exploring Students' Profile and Performance Before and After Covid-19 Lock-down. <i>Cristina Davino and Marco Gherghi</i>	1705
6.2.4	Nowcasting the Covid-19 outbreaks methods and applications	1711
6.2.4.1	Modeling subsequent waves of COVID-19 outbreak: A change point growth model. <i>Luca Greco, Paolo Girardi and Laura Ventura</i>	1712
6.2.4.2	The second wave of SARS-CoV-2 epidemic in Italy through a SIRD model. <i>Michela Baccini and Giulia Cereda</i>	1718
6.2.5	The impact of Covid-19 on survey methods	1724
6.2.5.1	Collecting cross-national survey data during the COVID-19 pandemic: Challenges and implications of data collection for the 50+ population in the Survey of Health, Ageing and Retirement in Europe (SHARE). <i>Michael Bergmann, Arne Bethmann, Yuri Pettinicchi and Borsch-Supan</i>	1725
6.2.5.2	Adapting a Long-Term Panel Survey to Pandemic Conditions. <i>Peter Lynn</i>	1731
6.2.6	Young contributions in Covid-19 statistical modelling	1737
6.2.6.1	Statistical communication of COVID-19 epidemic using widely accessible interactive tools. <i>Marco Mingione and Pierfrancesco Alaimo Di Loro</i>	1738
6.2.6.2	Modelling COVID-19 evolution in Italy with an augmented SIRD model using open data. <i>Vincenzo Nardelli, Giuseppe Arbia, Andrea Palladino and Luigi Giuseppe Atzeni</i>	1744

Hypotheses testing in mixed–frequency volatility models: a bootstrap approach

Test d'ipotesi nei modelli di volatilità a frequenza mista: un approccio bootstrap

Vincenzo Candila, Lea Petrella

Abstract It is widely recognized that standard likelihood–based inference suffers from the presence of nuisance parameters. This problem is particularly relevant in the context of Mixing–Data Sampling (MIDAS) models, when volatility forecasting is the research topic and where often covariates' data are sampled at a different (usually lower) frequency than the asset returns. In this framework, testing the significance of the MIDAS terms brings together the presence of nuisance parameters that under the null hypothesis are not identifiable. This circumstance interferes with the asymptotic distribution of the common statistical tests employed in this framework. In particular, the asymptotic distribution is no more a χ^2 distribution. The present paper proposes a bootstrap likelihood ratio (BLR) test to overcome this problem, simulating the likelihood ratio test distribution. Using a Monte Carlo experiment, the proposed BLR test presents quite good performances in terms of the test's size and power.

Abstract *E' ampiamente riconosciuto che gli approcci inferenziali basati sulla massima verosimiglianza soffrono della presenza di nuisance parameters. Questo problema è particolarmente rilevante nel contesto di modelli Mixing–Data Sampling (MIDAS), usati nell'ambito delle previsioni di volatilità. In questo framework, testare la significatività dei termini MIDAS comporta la gestione dei nuisance parameters che, sotto l'ipotesi nulla, sono non identificabili. Questa circostanza interferisce con la distribuzione asintotica dei test statistici comunemente usati in questo ambito. In particolare, la distribuzione asintotica non risulta più essere una χ^2 . Il presente lavoro propone un bootstrap likelihood ratio (BLR) test per superare questo problema, simulando la distribuzione del likelihood ratio test. Attraverso una simulazione Monte Carlo, il test BLR proposto presenta ottime performance, in termini di size e potenza.*

Key words: Likelihood ratio test, MIDAS, nuisance parameter, bootstrap.

Vincenzo Candila, MEMOTEF Department, Sapienza University of Rome, Italy, e-mail: vincenzo.candila@uniroma1.it and Lea Petrella, MEMOTEF Department, Sapienza University of Rome, Italy, e-mail: lea.petrella@uniroma1.it

1 Introduction

The financial econometrics literature has paid particular attention to the estimation of asset returns volatility during the last four decades. In this framework, empirical evidences suggest that the volatility has a slow-moving feature around which the conditional second moments of returns oscillate. Starting from this characteristic, a new type of volatility models, based on the decomposition of volatility into two components, namely a short and a long-run component, has been proposed (for more details, see the review of Amado et al., 2019). At the same time, it is quite common in financial data analysis that observations came at a different frequency (usually lower) than the returns' ones. The Mixing-Data Sampling (MIDAS) methods proposed by Ghysels et al. (2007) are designed to solve this problem. When the MIDAS techniques are applied within the GARCH framework, the long-run component of the models can depend on variables observed at different frequencies than daily (see, for example, Engle et al. (2013) and Conrad and Kleen (2020)). Recently, the MIDAS methods have also been applied in the quantile regression framework to forecast the Value-at-Risk (Candila et al., 2020).

Unfortunately, as stated in Ghysels et al. (2007), testing the null hypothesis of no influence of the MIDAS component can be problematic since the weights associated with each realization of the low-frequency variable, seen as nuisance parameters, are not identifiable. This circumstance has a fundamental impact on the asymptotic distribution of the commonly used tests, like the Wald or the Likelihood Ratio (LR) tests (see Hansen (1996) and Andrews (2001) for a complete survey on this topic).

In the context MIDAS variables within a volatility model, our paper aims at investigating the profitability of using a bootstrap LR (BLR) test where the distribution of the test is obtained using a suitable bootstrap procedure. Resorting to the bootstrap to derive the LR test distribution is not new at all: see, for instance, the contributions of Di Sanzo (2009) and Buseti and Di Sanzo (2012). But this is the first time the BLR test is used within the volatility models employing MIDAS components.

In terms of results, the size and power of the proposed BLR are calculated through an extensive Monte Carlo experiment in a GARCH model framework. Comparing the results with the standard LR test, the BLR appears to have an empirical size closer to the nominal one and quite good empirical power.

The rest of the paper is as follows: Section 2 illustrates the models and the proposed BLR test, while Section 3 presents the Monte Carlo experiment.

2 Bootstrap Likelihood Ratio test

Let $r_{i,t}$ be the log-return of an asset representing the first log-difference of the closing prices for the day i in the period (week or month) t . Then, let us consider the formalization of the GARCH-MIDAS model proposed by Engle et al. (2013):

Hypotheses testing in mixed–frequency volatility models: a bootstrap approach

$$r_{i,t} = \sigma_{i,t} \varepsilon_{i,t} = \sqrt{\tau_t \times g_{i,t}} \varepsilon_{i,t}, \quad \text{with } i = 1, \dots, N_t \quad \text{and } t = 1, \dots, T, \quad (1)$$

where, $\sigma_{i,t}$ representing the conditional standard deviation at day i and period t , consists of two (multiplicative) components: τ_t and $g_{i,t}$. In particular, τ_t is defined as the long–run component of the volatility at period t and $g_{i,t}$ the short–run term at day i for period t . Moreover, $N = \sum_{t=1}^T N_t$ is the total number of days considered with N_t being the number of days in the period t . In Eq. (1), $\varepsilon_{i,t}$ is the *iid* innovation term, with $E(\varepsilon_{i,t}) = 0$ and $E(\varepsilon_{i,t}^2) = 1$, and with a finite fourth moment.

Following the common dynamics specifications of the short– and the long–run components proposed in the GARCH–MIDAS literature, we consider for $g_{i,t}$ the unit–mean reverting GJR–GARCH(1,1) process given by:

$$g_{i,t} = (1 - \alpha - \gamma/2 - \beta) + \left(\alpha + \gamma \cdot \mathbf{1}_{(r_{i-1,t} < 0)} \right) \frac{(r_{i-1,t})^2}{\tau_t} + \beta g_{i-1,t}, \quad (2)$$

where $\mathbf{1}_{(\cdot)}$ is an indicator function and $\alpha > 0$; $\beta \geq 0$; $\gamma \geq 0$; $\alpha + \beta + \gamma/2 < 1$.

The component τ_t is:

$$\tau_t = \exp \left(m + \theta \sum_{k=1}^K \delta_k(\omega) MV_{t-k} \right), \quad (3)$$

where $m \in R$, $\theta \in R$ represents the response to the one–sided filter of the past K realizations of the MIDAS terms i.e. the low–frequency variable MV_t through the weighting function $\delta_k(\omega)$. The most common used $\delta_k(\omega)$ in this context is the Beta function:

$$\delta_k(\omega) = \frac{(k/K)^{\omega_1-1} (1-k/K)^{\omega_2-1}}{\sum_{j=1}^K (j/K)^{\omega_1-1} (1-j/K)^{\omega_2-1}}. \quad (4)$$

Under this configuration, the parameter space is then $\Theta = \{\alpha, \gamma, \beta, m, \theta, \omega_1, \omega_2\}$. Given K and a distributional assumption for $\varepsilon_{i,t}$ in (1) it is possible to calculate the maximum likelihood (ML) estimator for Θ .

In order to test the significance of the MIDAS component in (3), the following null hypothesis is considered:

$$H_0 : \theta = 0. \quad (5)$$

Typically, one can evaluate such a null using the a Wald or a LR test. We focus on this latter case. Let $\hat{\Theta}_0$ be the ML estimate of Θ under the null $\theta = 0$, that is the “restricted” model. The correspondent log–likelihood at $\hat{\Theta}_0$ is denoted by $\ell(\hat{\Theta}_0)$. Let $\hat{\Theta}$ be the ML estimate of Θ under the alternative $\theta \neq 0$ i.e. in the “unrestricted” model. The corresponding log–likelihood at $\hat{\Theta}$ is denoted by $\ell(\hat{\Theta})$. The LR test is:

$$LR = 2 \left[\ell(\hat{\Theta}) - \ell(\hat{\Theta}_0) \right]. \quad (6)$$

Assuming a significance level α , test statistic in (6) should reject H_0 when

$$LR > CV_\alpha, \tag{7}$$

where CV_α is the $(1 - \alpha)th$ quantile of the LR distribution under the null. Under some regularity conditions, it can be shown that the LR test follows asymptotically a Chi-square (χ^2) distribution. In our context, since under the null hypothesis in (5) the parameters ω_1 and ω_2 in (4) are not identified, the distribution of LR in (6) is no more a χ^2 distribution. For this reason, here we propose a bootstrap procedure to simulate the distribution of LR test (6) under the null (5). The proposed BLR procedure is as follows:

1. Estimate the unrestricted and restricted models. Compute the LR statistic as in Eq. (6).
2. Let $\widehat{\sigma}_{i,t}$ be the estimated volatility obtained from the restricted model. Compute the standardized residuals $\widehat{\varepsilon}_{i,t}$ under the null, for $i = 1, \dots, N_t$ and $t = 1, \dots, T$, that is:

$$\widehat{\varepsilon}_{i,t} = \frac{r_{i,t}}{\widehat{\sigma}_{i,t}}.$$

Let $\widehat{\varepsilon}_{i,t}^*$, be the bootstrap residual, obtained from resampling with replacement from the standardized residual series $\widehat{\varepsilon}_{i,t}$.

3. Compute the bootstrap replicates of $r_{i,t}$, denoted by $r_{i,t}^*$, through:

$$r_{i,t}^* = \widehat{\sigma}_{i,t}^* \widehat{\varepsilon}_{i,t}^*, \text{ for } i = 1, \dots, N_t \text{ and } t = 1, \dots, T,$$

where $\widehat{\sigma}_{i,t}^*$ is the bootstrap volatility:

$$\widehat{\sigma}_{i,t}^* = \sqrt{\widehat{\tau}_t^* \times \widehat{g}_{i,t}^*},$$

with the long-run term under the null identified as $\widehat{\tau}_t^* = \exp(\widehat{m})$ and the short-run term as

$$\widehat{g}_{i,t}^* = \left(1 - \widehat{\alpha} - \widehat{\gamma}/2 - \widehat{\beta}\right) + \left(\widehat{\alpha} + \widehat{\gamma} \cdot \mathbf{1}_{(r_{i-1,t}^* < 0)}\right) \frac{(r_{i-1,t}^*)^2}{\widehat{\tau}_t^*} + \widehat{\beta} \widehat{g}_{i-1,t}^*,$$

where $\widehat{\alpha}, \widehat{\gamma}, \widehat{\beta}$ and \widehat{m} are the ML estimates of the restricted model. In order to obtain (recursively) the bootstrap realizations of $r_{i,t}^*$ for the N days, start the procedure with $\widehat{\sigma}_{1,t}^* = \widehat{\sigma}_{1,t}$. Finally, estimate the restricted and unrestricted models on the series $r_{i,t}^*$. Hence, calculate the LR on the bootstrap returns $r_{i,t}^*$, denoted by LR^* .

4. Repeat the previous step B times, obtaining $(LR^{*(1)}, \dots, LR^{*(B)})$, which is the bootstrap distribution of LR.
5. The estimate of (the bootstrap) CV_α , based on $(LR^{*(1)}, \dots, LR^{*(B)})$ and labelled as \widehat{CV}_α , is obtained as the $1 - \alpha$ quantile of the bootstrap distribution of LR.

Finally, the null in (5) is rejected through the BLR test if $LR > \widehat{CV}_\alpha$.

3 Monte Carlo Experiment

In this section, we consider a Monte Carlo experiment to learn about the profitability of using the BLR test when testing MIDAS components. For this goal, we generate R samples of data from the following data generating process (DGP):

$$r_{i,t} = \sqrt{\tau_t} \times g_{i,t} \varepsilon_{i,t}, \quad \text{with } i = 1, \dots, N_t, \quad \text{and } t = 1, \dots, T, \quad (8)$$

where:

$$\varepsilon_{i,t} \sim \text{iid } t_{(7)}, \quad (9)$$

$$\tau_t = \exp \left(m_0 + \theta_0 \sum_{k=1}^K \delta_k(\omega) MV_{t-k} \right), \quad (10)$$

$$g_{i,t} = (1 - \alpha_0 - \gamma_0/2 - \beta_0) + \left(\alpha_0 + \gamma_0 \cdot \mathbf{1}_{(r_{i-1,t} < 0)} \right) \frac{(r_{i-1,t})^2}{\tau_t} + \beta_0 g_{i-1,t}. \quad (11)$$

In Eq. (9), the error term $\varepsilon_{i,t}$ follows a standardized Student's t distribution with 7 degrees of freedom which allows for fat tails of real financial asset returns. We assume that the simulated stationary variable MV_t follows an AR(1): $MV_t = \varphi MV_{t-1} + e_t$, with $\varphi = 0.7$.

Using the R package *rumidas* (Candila, 2021), the DGP in (9) is simulated $R = 250$ times, according to two sample sizes: $N = \{500, 1000\}$. The true values of the parameters are:

$$\{\alpha_0 = 0.01, \gamma_0 = 0.1, \beta_0 = 0.9, m_0 = -1, \omega_{2,0} = 1.1\}.$$

The parameter of interest θ has instead the following values: $\theta_0 = \{0, 0.5, 1\}$.

The results of our experiment are illustrated in Table 1, where the estimated probabilities of rejecting the null across the R replicates are reported. More in detail, Panel A shows the empirical sizes for the BLR and the LR tests that is the occurrence of null rejection when the null is true. The results of the LR test are evaluated using the χ^2 distribution. Independently of the significance level adopted (0.01, 0.05, and 0.1) and of the sample length, the empirical size of the BLR appears much more in line with the actual size. When the null is false, as in Panels B and C, both the tests appear to have reasonable powers. These results support the use of the proposed BLR test instead of the LR one when mixed frequency models are employed.

References

Amado, C., A. Silvennoinen, and T. Teräsvirta (2019). Models with multiplicative decomposition of conditional variances and correlations. In J. Chevallier, S. Goutte, D. Guerreiro, S. Saglio, and B. Sanhaji (Eds.), *Financial Mathematics, Volatility and Covariance Modelling*, Volume 2, pp. 217–260. London: Routledge.

Table 1 BLR and LR empirical sizes and powers

Signif. level	0.01	0.05	0.10	0.01	0.05	0.10
Panel A: $\theta = 0$		$N = 500$			$N = 1000$	
BLR	0.012	0.056	0.136	0.016	0.040	0.108
LR	0.060	0.132	0.200	0.060	0.160	0.220

Panel B: $\theta = 0.5$		$N = 500$			$N = 1000$	
BLR	0.980	0.980	1.000	0.980	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000

Panel C: $\theta = 1$		$N = 500$			$N = 1000$	
BLR	0.980	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000

Notes: Panel A reports the empirical sizes of the BLR and LR tests, for the null in (5), that is the number of times (across the $R = 250$ replications) that the null is rejected (given that the null is true). The other panels report the empirical powers, that is the number of times (across the $R = 250$ replications) that the null is rejected (given that the null is false).

- Andrews, D. W. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69(3), 683–734.
- Busetti, F. and S. Di Sanzo (2012). Bootstrap LR tests of stationarity, common trends and cointegration. *Journal of Statistical Computation and Simulation* 82(9), 1343–1355.
- Candila, V. (2021). *rumidas: Univariate GARCH-MIDAS, Double-Asymmetric GARCH-MIDAS and MEM-MIDAS models*. R package version 0.1.1.
- Candila, V., G. M. Gallo, and L. Petrella (2020). Using mixed-frequency and realized measures in quantile regression. Technical report, SSRN.
- Conrad, C. and O. Kleen (2020). Two are better than one: Volatility forecasting using multiplicative component GARCH-MIDAS models. *Journal of Applied Econometrics* 35(1), 19–45.
- Di Sanzo, S. (2009). Testing for linearity in Markov switching models: a bootstrap approach. *Statistical Methods and Applications* 18(2), 153–168.
- Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics* 95(3), 776–797.
- Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews* 26(1), 53–90.
- Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64(2), 413–430.