Book of Short Papers SIS 2018

Editors: Antonino Abbruzzo - Eugenio Brentari

Marcello Chiodi - Davide Piacentino

Copyright © 2018

PUBLISHED BY PEARSON

WWW.PEARSON.COM

First printing, novembre 2018 - ISBN-9788891910233

SIS 2018 49 TH SCIENTIFIC MEETING OF THE ITALIAN STATISTICAL SOCIETY
PALERMO 20-22 JUNE

Contents

1 Preface 17

2 Plenary Sessions 19

2.1 A new paradigm for rating data models. *Domenico Piccolo* 19

2.2 Statistical challenges and opportunities in modelling coupled behaviour-disease dynamics of vaccine refusal. *Plenary/Chris T. Bauch* 32

3 Specialized Sessions 45

3.1 3.1 - Bayesian Nonparametric Learning 45

3.1.1 Bayesian nonparametric covariate driven clustering. *Raffaele Argiento, Ilaria Bianchini, Alessandra Guglielmi and Ettore Lanzarone* 46

3.1.2 A Comparative overview of Bayesian nonparametric estimation of the size of a population. *Luca Tardella and Danilo Alunni Fegatelli* 56

3.1.3 Logit stick-breaking priors for partially exchangeable count data. *Tommaso Rigon* 64

3.2 BDsports - Statistics in Sports 72

3.2.1 A paired comparison model for the analysis of on-field variables in football matches. *Gunther Schauberger and Andreas Groll* 72

3.2.2 Are the shots predictive for the football results?. *Leonardo Egidi, Francesco Pauli, Nicola Torelli* 81

3.2.3 Zero-inflated ordinal data models with application to sport (in)activity. *Maria Iannario and Rosaria Simone* 89

3.3 Being young and becoming adult in the Third Millennium: definition issues and processes analysis 97

3.3.1 Do Social Media Data predict changes in young adults' employment status? Evidence from Italy. *Andrea Bonanomi and Emiliano Sironi* 97

3.12.2 Bayesian inference for hidden Markov models via duality and approximate filtering distributions. *Guillaume Kon Kam King, Omiros Papaspiliopoulos and Matteo Ruggiero* . . . 248
3.12.3 K-means seeding via MUS algorithm. *Leonardo Egidi, Roberta Pappada`, Francesco Pauli, Nicola Torelli* . . . 256

4 Sollicited Sessions . . . 263

4.1 Advances in Discrete Latent Variable Modelling 263
4.1.1 A joint model for longitudinal and survival data based on a continuous-time latent Markov model. *Alessio Farcomeni and Francesco Bartolucci* . . . 264
4.1.2 Modelling the latent class structure of multiple Likert items: a paired comparison approach. *Brian Francis* . . . 273
4.1.3 Dealing with reciprocity in dynamic stochastic block models. *Francesco Bartolucci, Maria Francesca Marino, Silvia Pandolfi* . . . 281
4.1.4 Causality patterns of a marketing campaign conducted over time: evidence from the latent Markov model. *Fulvia Pennoni, Leo Paas and Francesco Bartolucci* 289

4.2 Complex Spatio-temporal Processes and Functional Data 297
4.2.1 Clustering of spatio-temporal data based on marked variograms. *Antonio Balzanella and Rosanna Verde* . . . 297
4.2.2 Space-time earthquake clustering: nearest-neighbor and stochastic declustering methods in comparison. *Elisa Varini, Antonella Peresan, Renata Rotondi, and Stefania Gentili* . . . 304
4.2.3 Advanced spatio-temporal point processes for the Sicily seismicity analysis. *Marianna Siino and Giada Adelfio* . . . 312
4.2.4 Spatial analysis of the Italian seismic network and seismicity. *Antonino D'Alessandro, Marianna Siino, Luca Greco and Giada Adelfio* . . . 320

4.3 Dimensional Reduction Techniques for Big Data Analysis 328
4.3.1 Clustering Data Streams via Functional Data Analysis: a Comparison between Hierarchical Clustering and K-means Approaches. *Fabrizio Maturo, Francesca Fortuna, and Tonio Di Battista* . . . 328
4.3.2 Co-clustering algorithms for histogram data. *Francisco de A.T. De Carvalho and Antonio Balzanella and Antonio Irpino and Rosanna Verde* . . . 338
4.3.3 A network approach to dimensionality reduction in Text Mining. *Michelangelo Misuraca, Germana Scepi and Maria Spano* . . . 344
4.3.4 Self Organizing Maps for distributional data. *Rosanna Verde and Antonio Irpino* 352

4.4 Enviromental Processes, Human Activities and their Interactions 353
4.4.1 Estimation of coral growth parameters via Bayesian hierarchical non-linear models. *Crescenza Calculli, Barbara Cafarelli and Daniela Cocchi* . . . 353
4.4.2 A Hierarchical Bayesian Spatio-Temporal Model to Estimate the Short-term Effects of Air Pollution on Human Health. *Fontanella Lara, Ippoliti Luigi and Valentini Pasquale* 361
4.4.3 A multilevel hidden Markov model for space-time cylindrical data. *Francesco Lagona and Monia Ranalli* . . . 367
4.4.4 Estimation of entropy measures for categorical variables with spatial correlation. *Linda Altieri, Giulia Roli* . . . 373

4.5 Innovations in Census and in Social Surveys 381
4.5.1 A micro-based approach to ensure consistency among administrative sources and to improve population statistics. *Gianni Corsetti, Sabrina Prati, Valeria Tomeo, Enrico Tucci* . . . 381
4.5.2 Demographic changes, research questions and data needs: issues about migrations. *Salvatore Strozza and Giuseppe Gabrielli* . . . 392

3.3.2 Parenthood: an advanced step in the transition to adulthood. *Cinzia Castagnaro, Antonella Guarneri and Eleonora Meli* 106

3.4 Economic Statistics and Big Data 114

3.4.1 Improvements in Italian CPI/HICP deriving from the use of scanner data. *Alessandro Brunetti, Stefania Fatello, Federico Polidoro, Antonella Simone* 114

3.4.2 Big data and spatial price comparisons of consumer prices. *Tiziana Laureti and Federico Polidoro* 123

3.5 Financial Time Series Analysis 131

3.5.1 Dynamic component models for forecasting trading volumes. *Antonio Naimoli and Giuseppe Storti* 131

3.5.2 Conditional Quantile-Located VaR. *Giovanni Bonaccolto, Massimiliano Caporin and Sandra Paterlini* 140

3.6 Forensic Statistics 146

3.6.1 Cause of effects: an important evaluation in Forensic Science. *Fabio Corradi and Monica Musio* 146

3.6.2 Evaluation and reporting of scientific evidence: the impact of partial probability assignments. *Silvia Bozza, Alex Biedermann, Franco Taroni* 155

3.7 Missing Data Handling in Complex Models 161

3.7.1 Dependence and sensitivity in regression models for longitudinal responses subject to dropout. *Marco Alfo' and Maria Francesca Marino* 161

3.7.2 Multilevel analysis of student ratings with missing level-two covariates: a comparison of imputation techniques. *Maria Francesca Marino e Carla Rampichini* 170

3.7.3 Multilevel Multiple Imputation in presence of interactions, non-linearities and random slopes. *Matteo Quartagno and James R. Carpenter* 175

3.8 Monitoring Education Systems. Insights from Large Scale Assessment Surveys 183

3.8.1 Educational Achievement of Immigrant Students. A Cross-National Comparison Over-Time Using OECD-PISA Data. *Mariano Porcu* 183

3.9 New Perspectives in Time Series Analysis 192

3.9.1 Generalized periodic autoregressive models for trend and seasonality varying time series. *Francesco Battaglia and Domenico Cucina and Manuel Rizzo* 192

3.10 Recent Advances in Model-based Clustering 201

3.10.1 Flexible clustering methods for high-dimensional data sets. *Cristina Tortora and Paul D. McNicholas* 201

3.10.2 A Comparison of Model-Based and Fuzzy Clustering Methods. *Marco Alfo', Maria Brigida Ferraro, Paolo Giordani, Luca Scrucca, and Alessio Serafini* 208

3.10.3 Covariate measurement error in generalized linear models for longitudinal data: a latent Markov approach. *Roberto Di Mari, Antonio Punzo, and Antonello Maruotti* 216

3.11 Statistical Modelling 224

3.11.1 A regularized estimation approach for the three-parameter logistic model. *Michela Battauz and Ruggero Bellio* 224

3.11.2 Statistical modelling and GAMLSS. *Mikis D. Stasinopoulos and Robert A. Rigby and Fernanda De Bastiani* 233

3.12 Young Contributions to Statistical Learning 239

3.12.1 Introducing spatio-temporal dependence in clustering: from a parametric to a nonparametric approach . *Clara Grazian, Gianluca Mastrantonio and Enrico Bibbona* 239

4.5.3 Towards more timely census statistics: the new Italian multiannual dissemination programme. *Simona Mastroluca and Mariangela Verrascina* 400

4.6 Living Conditions and Consumption Expenditure in Time of Crises 409

4.6.1 Household consumption expenditure and material deprivation in Italy during last economic crises. *Ilaria Arigoni and Isabella Siciliani* 409

4.7 Network Data Analysis and Mining 418

4.7.1 Support provided by elderly Italian people: a multilevel analysis. *Elvira Pelle, Giulia Rivellini and Susanna Zaccarini* 418

4.7.2 Data mining and analysis of comorbidity networks from practitioner prescriptions. *Giancarlo Ragozini, Giuseppe Giordano, Sergio Pagano, Mario De Santis, Pierpaolo Cavallo* 426

4.7.3 Overlapping mixture models for network data (manet) with covariates adjustment. *Saverio Ranciati and Giuliano Galimberti and Ernst C. Wit and Veronica Vinciotti* 434

4.8 New Challenges in the Measurement of Economic Insecurity, Inequality and Poverty 440

4.8.1 Social protection in mitigating economic insecurity. *Alessandra Coli* 440

4.8.2 Changes in poverty concentration in U.S. urban areas. *Francesco Andreoli and Mauro Mussini* 450

4.8.3 Evaluating sustainability through an input-stateoutput framework: the case of the Italian provinces. *Achille Lemmi, Laura Neri, Federico M. Pulselli* 458

4.9 New Methods and Models for Ordinal Data 466

4.9.1 Weighted and unweighted distances based decision tree for ranking data. *Antonella Plaia, Simona Buscemi, Mariangela Sciandra* 466

4.9.2 A dissimilarity-based splitting criterion for CUBREMOT. *Carmela Cappelli, Rosaria Simone and Francesca Di Iorio* 474

4.9.3 Constrained Extended Plackett-Luce model for the analysis of preference rankings. *Cristina Mollica and Luca Tardella* 480

4.9.4 A prototype for the analysis of time use in Italy. *Stefania Capecchi and Manuela Michelini* 487

4.10 New Perspectives in Supervised and Unsupervised Classification 493

4.10.1 Robust Updating Classification Rule with applications in Food Authenticity Studies. *Andrea Cappozzo, Francesca Greselin and Thomas Brendan Murphy* 493

4.10.2 A robust clustering procedure with unknown number of clusters. *Francesco Dotto and Alessio Farcomeni* 500

4.10.3 Issues in joint dimension reduction and clustering methods. *Michel van de Velden, Alfonso Iodice D'Enza and Angelos Markos* 508

4.11 New Sources, Data Integration and Measurement Challenges for Estimates on Labour Market Dynamics 514

4.11.1 The development of the Italian Labour register: principles, issues and perspectives . *C. Baldi, C. Ceccarelli, S. Gigante, S. Pacini* 514

4.11.2 Digging into labour market dynamics: toward a reconciliation of stock and flows short term indicators. *F. Rapiti, C. Baldi, D. Ichim, F. Pintaldi, M. E. Pontecorvo, R. Rizzi* 523

4.11.3 How effective are the regional policies in Europe? The role of European Funds. *Gennaro Punzo, Mariateresa Ciommi, and Gaetano Musella* 531

4.11.4 Issues in joint dimension reduction and clustering methods. *Lucio Masserini and Matilde Bini* 539

4.12 Quantile and Generalized Quantile Methods 547

4.12.1 Multiple quantile regression for risk assessment. *Lea Petrella and Valentina Raponi* 547

4.12.2 Parametric Modeling of Quantile Regression Coefficient Functions. *Paolo Frumento and Matteo Bottai* 550

4.12.3 Modelling the effect of Traffic and Meteorology on Air Pollution with Finite Mixtures of M-quantile Regression Models. *Simone Del Sarto, Maria Francesca Marino, Maria Giovanna Ranalli and Nicola Salvati* 552

4.12.4 Three-level M-quantile model for small area poverty mapping. *Stefano Marchetti and Nicola Salvati* 560

4.13 Recent Advances on Extreme Value Theory 560

4.13.1 Extremes of high-order IGARCH processes. *Fabrizio Laurini* 560

4.14 Spatial Economic Data Analysis 569

4.14.1 Spatial heterogeneity in principal component analysis: a study of deprivation index on Italian provinces. *Paolo Postiglione, M. Simona Andreano, Roberto Benedetti, Alfredo Cartone* 569

4.15 Spatial Functional Data Analysis 578

4.15.1 Object oriented spatial statistics for georeferenced tensor data. *Alessandra Menafoglio and Davide Pigoli and Piercesare Secchi* 578

4.15.2 A Spatio-Temporal Mixture Model for Urban Crimes. *Ferretti Angela, Ippoliti Luigi and Valentini Pasquale* 585

4.16 Statistical Methods for Service Quality 591

4.16.1 Cumulative chi-squared statistics for the service quality improvement: new properties and tools for the evaluation. *Antonello D'Ambra, Antonio Lucadamo, Pietro Amenta, Luigi D'Ambra* 591

4.16.2 A robust multinomial logit model for evaluating judges' performances. *Ida Camminatiello and Antonio Lucadamo* 600

4.16.3 Complex Contingency Tables and Partitioning of Three-way Association Indices for Assessing Justice CourtWorkload. *Rosaria Lombardo, Yoshio Takane and Eric J Beh* 607

4.16.4 Finding the best paths in university curricula of graduates to improve academic guidance services. *Silvia Bacci and Bruno Bertaccini* 615

4.17 Statistical Modelling for Business Intelligence Problems 623

4.17.1 A nonlinear state-space model for the forecasting of field failures. *Antonio Pievatolo* 623

4.17.2 Does Airbnb affect the real estate market? A spatial dependence analysis. *Mariangela Guidolin and Mauro Bernardi* 632

4.17.3 Bayesian Quantile Trees for Sales Management. *Mauro Bernardi and Paola Stolfi* 640

4.17.4 Discrimination in machine learning algorithms. *Roberta Pappadá and Francesco Pauli* 648

4.18 Statistical models for sports data 656

4.18.1 Exploring the Kaggle European Soccer database with Bayesian Networks: the case of the Italian League Serie A. *Maurizio Carpita and Silvia Golia* 656

4.18.2 A data-mining approach to the Parkour discipline. *Paola Pasca, Enrico Ciavolino and Ryan L. Boyd* 665

4.18.3 Players Movements and Team Shooting Performance: a Data Mining approach for Basketball. *Rodolfo Metulini* 673

4.19 Supporting Regional Policies through Small Area Statistical Methods 681
4.19.1 Survey-weighted Unit-Level Small Area Estimation. *Jan Pablo Burgard and Patricia Dörr* 681
4.20 The Second Generation at School 681
4.20.1 Resilient students with migratory background. *Anna Di Bartolomeo and Giuseppe Gabrielli* 681
4.20.2 Residential Proximity to Attended Schools among Immigrant-Origin Youths in Bologna. *Federica Santangelo, Debora Mantovani and Giancarlo Gasperoni* 690
4.20.3 From school to ... future: strategies, paths and perspectives of immigrant immediate descendants in Naples . *Giustina Orientale Caputo and Giuseppe Gargiulo* 698
4.21 Tourism Destinations, Household, Firms 706
4.21.1 The Pricing Behaviour of Firms in the On-line Accommodation Market: Evidence from a Metropolitan City. *Andrea Guizzardi and Flavio Maria Emanuele Pons* 706
4.21.2 The Migration-Led-Tourism Hypothesis for Italy: A Survey. *Carla Massidda, Romano Piras and Ivan Etzo* 716
4.21.3 Tourism Statistics: development and potential uses. *Fabrizio Antolini* 724
4.21.4 Tourism attractiveness in Italy. Some empirical evidence comparing origin-destination domestic tourism flows. *Francesca Giambona, Emanuela Dreassi, and Alessandro Magrini* 732
4.22 What's Happening in Africa 740
4.22.1 Environmental shocks and internal migration in Tanzania. *Maria Francesca Marino, Alessandra Petrucci, and Elena Pirani* 740
4.22.2 Determinants and geographical disparities of BMI in African Countries: a measurement error small area approach. *Serena Arima and Silvia Polettini* 748

5 Contributed Sessions 757
5.1 Advanced Algorithms and Computation 758
5.1.1 Brexit in Italy. *Francesca Greco, Livia Celardo, Leonardo Salvatore Alaimo* . 758
5.1.2 Distance based Depth-Depth classifier for directional data. *Giuseppe Pandolfo and Giovanni C. Porzio* 765
5.1.3 Approximate Bayesian Computation for Forecasting in Hydrological models. *Jonathan Romero-Cuéllar, Antonino Abbruzzo, Giada Adelfio and Félix Francés* 769
5.1.4 Customer Churn prediction based on eXtreme Gradient Boosting classifier. *Mohammed Hassan Elbedawi Omar and Matteo Borrotti* 775
5.1.5 HPC-accelerated Approximate Bayesian Computation for Biological Science. *Ritabrata Dutta* 781
5.1.6 PC Algorithm for Gaussian Copula Data. *Vincenzina Vitale and Paola Vicard* 789
5.2 Advances in Clustering Techniques 795
5.2.1 On the choice of an appropriate bandwidth for modal clustering. *Alessandro Casa, José E. Chacón and Giovanna Menardi* 795
5.2.2 Unsupervised clustering of Italian schools via non-parametric multilevel models. *Chiara Masci, Francesca Ieva and Anna Maria Paganoni* 802
5.2.3 Chiara Masci, Francesca Ieva and Anna Maria Paganoni. *Laura Bocci and Donatella Vicari* 808
5.2.4 Robust Reduced k-Means and Factorial k-Means by trimming. *Luca Greco and Antonio Lucadamo and Pietro Amenta* 813
5.2.5 Dirichlet processes, posterior similarity and graph clustering. *Stefano Tonellato* 819
5.2.6 Bootstrap ClustGeo with spatial constraints. *Veronica Distefano, Valentina Mameli, Fabio Della Marra* 825

5.3 Advances in Statistical Models 831

5.3.1 Regression modeling via latent predictors. *Francesca Martella and Donatella Vicari* 831

5.3.2 Analysis of dropout in engineering BSc using logistic mixed-effect models. *Luca Fontana and Anna Maria Paganoni* 838

5.3.3 dgLARS method for relative risk regression models. *Luigi Augugliaro and Angelo M. Mineo* 844

5.3.4 A Latent Class Conjoint Analysis for analysing graduates profiles. *Paolo Mariani, Andrea Marletta, Lucio Masserini and Mariangela Zenga* 850

5.3.5 A longitudinal analysis of the degree of accomplishment of anti-corruption measures by Italian municipalities: a latent Markov approach. *Simone Del Sarto, Michela Gnaldi, Francesco Bartolucci* 856

5.3.6 Modelling the effect of covariates for unbiased estimates in ecological inference methods. *Venera Tomaselli, Antonio Forcina and Michela Gnaldi* 862

5.4 Advances in Time Series 868

5.4.1 Filtering outliers in time series of electricity prices. *Ilaria Lucrezia Amerise* . . . 868

5.4.2 Time-varying long-memory processes. *Luisa Bisaglia and Matteo Grigoletto* 875

5.4.3 Statistical Analysis of Markov Switching DSGE Models. *Maddalena Cavicchioli* 881

5.4.4 Forecasting energy price volatilities and comovements with fractionally integrated MGARCH models. *Malvina Marchese and Francesca Di Iorio* 886

5.4.5 Improved bootstrap simultaneous prediction limits. *Paolo Vidoni* 892

5.5 Data Management 898

5.5.1 Using web scraping techniques to derive co-authorship data: insights from a case study. *Domenico De Stefano, Vittorio Fuccella, Maria Prosperina Vitale, Susanna Zaccarin* 898

5.5.2 Dealing with Data Evolution and Data Integration: An approach using Rarefaction. *Luca Del Core, Eugenio Montini, Clelia Di Serio, Andrea Calabria* 905

5.5.3 Monitoring event attendance using a combination of traditional and advanced surveying tools. *Mauro Ferrante, Amit Birenboim, Anna Maria Milito, Stefano De Cantis* 911

5.5.4 Indefinite Topological Kernels. *Tullia Padellini and Pierpaolo Brutti* 917

5.5.5 Data Integration in Social Sciences: the earnings intergenerational mobility problem. *Veronica Ballerini, Francesco Bloise, Dario Briscolini and Michele Raitano* 923

5.5.6 An innovative approach for the GDPR compliance in Big DAta era. *M. Giacalone, C. Cusatelli, F. Fanari, V. Santarcangelo, D.C. Sinitó* 929

5.6 Developments in Graphical Models 935

5.6.1 An extension of the glasso estimator to multivariate censored data. *Antonino Abbruzzo and Luigi Augugliaro and Angelo M. Mineo* 935

5.6.2 Bayesian Estimation of Graphical Log-Linear Marginal Models. *Claudia Tarantola, Ioannis Ntzoufras and Monia Lupparelli* 942

5.6.3 Statistical matching by Bayesian Networks. *Daniela Marella and Paola Vicard and Vincenzina Vitale* 948

5.6.4 Sparse Nonparametric Dynamic Graphical Models. *Fabrizio Poggioni, Mauro Bernardi, Lea Petrella* 954

5.6.5 Non-communicable diseases, socio-economic status, lifestyle and well-being in Italy: An additive Bayesian network model. *Laura Maniscalco and Domenica Matranga* 960

5.6.6 Using Almost-Dynamic Bayesian Networks to Represent Uncertainty in Complex Epidemiological Models: a Proposal. *Sabina Marchetti* 966

5.7 Educational World 972

5.7.1 How to improve the Quality Assurance System of the Universities: a study based on compositional analysis . *Bertaccini B., Gallo M., Simonacci V., and Menini T.* . 972

5.7.2 Evaluation of students' performance at lower secondary education. An empirical analysis using TIMSS and PISA data.. *G. Graziosi, T. Agasisti, K. De Witte and F. Pauli* 977

5.7.3 Testing for the Presence of Scale Drift: An Example. *Michela Battauz* 983

5.7.4 The evaluation of Formative Tutoring at the University of Padova. *Renata Clerici, Lorenza Da Re, Anna Giraldo, Silvia Meggiolaro* 988

5.7.5 Benefits of the Erasmus mobility experience: a discrete latent variable analysis. *Silvia Bacci, Valeria Caviezel and Anna Maria Falzoni* 993

5.7.6 University choice and the attractiveness of the study area. Insights from an analysis based on generalized mixed-effect models. *Silvia Columbu, Mariano Porcu and Isabella Sulis* 999

5.8 Environment 1005

5.8.1 The climate funds for energy sustainability: a counterfactual analysis. *Alfonso Carfora and Giuseppe Scandurra* 1005

5.8.2 Exploratory GIS Analysis via Spatially Weighted Regression Trees. *Carmela Iorio, Giuseppe Pandolfo, Michele Staiano, and Roberta Siciliano* 1012

5.8.3 A functional regression control chart for profile monitoring. *Fabio Centofanti, Antonio Lepore, Alessandra Menafoglio, Biagio Palumbo and Simone Vantini* 1018

5.8.4 Understanding pro-environmental travel behaviours in Western Europe. *Gennaro Punzo, Rosalia Castellano, and Demetrio Panarello* 1023

5.9 Family & Economic issues 1029

5.9.1 Measuring Economic Uncertainty: Longitudinal Evidence Using a Latent Transition Model. *Francesca Giambona, Laura Grassini and Daniele Vignoli* 1029

5.9.2 Intentions to leave Italy or to stay among foreigners: some determinants of migration projects. *Ginevra Di Giorgio, Francesca Dota, Paola Muccitelli and Daniele Spizzichino* 1036

5.9.3 Wages differentials in association with individuals, enterprises and territorial characteristics. *S. De Santis, C. Freguja, A. Masi, N. Pannuzi, F. G. Truglia* 1042

5.9.4 The Transition to Motherhood among British Young Women: Does housing tenure play a role?. *Valentina Tocchioni, Ann Berrington, Daniele Vignoli and Agnese Vitali* 1048

5.10 Finance & Insurance 1054

5.10.1 Robust statistical methods for credit risk. *A. Corbellini, A. Ghiretti, G. Morelli and A. Talignani* 1054

5.10.2 Depth-based portfolio selection. *Giuseppe Pandolfo, Carmela Iorio and Antonio D'Ambrosio* 1061

5.10.3 Estimating large-scale multivariate local level models with application to stochastic volatility. *Matteo Pelagatti and Giacomo Sbrana* 1067

5.11 Health and Clinical Data 1073

5.11.1 Is retirement bad for health? A matching approach. *Elena Pirani, Marina Ballerini, Alessandra Mattei, Gustavo De Santis* 1073

5.11.2 The emergency department utilisation among the immigrant population resident in Rome from 2005 to 2015. *Eleonora Trappolini, Laura Cacciani, Claudia Marino, Cristina Giudici, Nera Agabiti, Marina Davoli* 1080

5.11.3 Multi-State model with nonparametric discrete frailty. *Francesca Gasperoni, Francesca Ieva, Anna Maria Paganoni, Chris Jackson and Linda Sharples* 1087

5.11.4 A Functional Urn Model for CARA Designs. *Giacomo Aletti, Andrea Ghiglietti, and William F. Rosenberger* 1093

5.11.5 Assessment of the INLA approach on gerarchic bayesian models for the spatial disease distribution: a real data application. *Paolo Girardi, Emanuela Bovo, Carmen Stocco, Susanna Baracco, Alberto Rosano, Daniele Monetti, Silvia Rizzato, Sara Zamberlan, Enrico Chinellato, Ugo Fedeli, Massimo Rugge* 1099

5.12 Medicine 1105

5.12.1 Hidden Markov Models for disease progression. *Andrea Martino, Andrea Ghiglietti, Giuseppina Guatteri, Anna Maria Paganoni* 1105

5.12.2 A simulation study on the use of response-adaptive randomized designs. *Anna Maria Paganoni, Andrea Ghiglietti, Maria Giovanna Scarale, Rosalba Miceli, Francesca Ieva, Luigi Mariani, Cecilia Gavazzi and Valeria Edefonti* 1112

5.12.3 The relationship between health care expenditures and time to death: focus on myocardial infarction patients. *Luca Grassetti and Laura Rizzi* 1118

5.12.4 A multivariate extension of the joint models. *Marcella Mazzoleni and Mariangela Zenga* 1124

5.12.5 Multipurpose optimal designs for hypothesis testing in normal response trials. *Marco Novelli and Maroussa Zagoraiou* 1130

5.12.6 Additive Bayesian networks for an epidemiological analysis of swine diseases. *Marta Pittavino and Reinhard Furrer* 1136

5.13 Population Dynamics 1142

5.13.1 Employment Uncertainty and Fertility: a Meta-Analysis of European Research Findings. *Giammarco Alderotti, Daniele Vignoli and Michela Baccini* 1142

5.13.2 What Shapes Population Age Structures in the Long Run. *Gustavo De Santis and Giambattista Salinari* 1148

5.13.3 The impact of economic development on fertility: a complexity approach in a cross-country analysis. *NiccolóInnocenti, Daniele Vignoli and Luciana Lazzeretti* 1154

5.13.4 A Probabilistic Cohort-Component Model for Population Fore-casting - The Case of Germany. *Patrizio Vanella and Philipp Deschermeier* 1159

5.13.5 Mortality trends in Sardinia 1992-2015: an ecological study. *Vanessa Santos Sanchez, Gabriele Ruiu Marco Breschi, Lucia Pozzi* 1163

5.14 Recent Developments in Bayesian Inference 1169

5.14.1 Posterior distributions with non explicit objective priors. *Erlis Ruli, Nicola Sartori and Laura Ventura* 1169

5.14.2 A predictive measure of the additional loss of a non-optimal action under multiple priors. *Fulvio De Santis and Stefania Gubbiotti* 1169

5.14.3 Bayesian estimation of number and position of knots in regression splines. *Gioia Di Credico, Francesco Pauli and Nicola Torelli* 1169

5.14.4 The importance of historical linkages in shaping population density across space. *Ilenia Epifani and Rosella Nicolini* 1169

5.15 Recent Developments in Sampling 1169

5.15.1 Species richness estimation exploiting purposive lists: A proposal. *A. Chiarucci, R.M. Di Biase, L. Fattorini, M. Marcheselli and C. Pisani* 1169

5.15.2 Design-based exploitation of big data by a doubly calibrated estimator. *Maria Michela Dickson, Giuseppe Espa and Lorenzo Fattorini* 1176

5.15.3 Design-based mapping in environmental surveys. *L. Fattorini, M. Marcheselli and C. Pisani* 1182

5.15.4 Testing for independence in analytic inference. *Pier Luigi Conti and Alberto Di Iorio* 1188

5.15.5 On the aberrations of two-level Orthogonal Arrays with removed runs. *Roberto Fontana and Fabio Rapallo* 1194

5.16 Recent Developments in Statistical Modelling 1200

5.16.1 Quantile Regression Coefficients Modeling: a Penalized Approach. *Gianluca Sottile, Paolo Frumento and Matteo Bottai* 1200

5.16.2 Simultaneous calibrated prediction intervals for time series. *Giovanni Fonseca, Federica Giummolé and Paolo Vidoni* 1207

5.16.3 Reversibility and (non)linearity in time series. *Luisa Bisaglia and Margherita Gerolimetto* 1213

5.16.4 Heterogeneous effects of subsidies on farms' performance: a spatial quantile regression analysis. *Marusca De Castris and Daniele Di Gennaro* 1219

5.16.5 On the estimation of high-dimensional regression models with binary covariates. *Valentina Mameli, Debora Slanzi and Irene Poli* 1226

5.17 Social Indicators 1232

5.17.1 Can a neighbour region influence poverty? A fuzzy and longitudinal approach. *Gianni Betti, Federico Crescenzi and Francesca Gagliardi* 1232

5.17.2 Weight-based discrimination in the Italian Labor Market: how do ethnicity and gender interact? *Giovanni Busetta, Maria Gabriella Campolo, and Demetrio Panarello* 1239

5.17.3 The Total Factor Productivity Index as a Ratio of Price Indexes. *Lisa Crosato and Biancamaria Zavanella* 1245

5.17.4 Monetary poverty indicators at local level: evaluating the impact of different poverty thresholds. *Luigi Biggeri, Caterina Giusti and Stefano Marchetti* 1251

5.17.5 A gender inequality assessment by means of the Gini index decomposition. *Michele Costa* 1257

5.18 Socio-Economic Statistics 1263

5.18.1 The NEETs during the economic crisis in Italy, Young NEETs in Italy, Spain and Greece during the economic crisis. *Giovanni De Luca, Paolo Mazzocchi, Claudio Quintano, Antonella Rocca* 1263

5.18.2 Camel or dromedary? A study of the equilibrium distribution of income in the EU countries. *Crosato L., Ferretti C., Ganugi P.* 1270

5.18.3 Small Area Estimation of Inequality Measures. *Maria Rosaria Ferrante and Silvia Pacei* 1276

5.18.4 Testing the Learning-by-Exporting at Micro-Level in light of influence of "Statistical Issues" and Macroeconomic Factors. *Maria Rosaria Ferrante and Marzia Freo* 1281

5.18.5 The mobility and the job success of the Sicilian graduates *Ornella Giambalvo and Antonella Plaia and Sara Binassi* 1287

5.19 Statistical Analysis of Energy Markets 1293

5.19.1 Forecasting Value-at-Risk for Model Risk Analysis in Energy Markets. *Angelica Gianfreda and Giacomo Scandolo* 1293

5.19.2 Prediction interval of electricity prices by robust nonlinear models. *Lisa Crosato, Luigi Grossi and Fany Nan* 1300

5.19.3 Bias Reduction in a Matching Estimation of Treatment Effect. *Maria Gabriella Campolo, Antonino Di Pino and Edoardo Otranto* 1305

5.20 Statistical Inference and Testing Procedures 1311

5.20.1 Comparison of exact and approximate simultaneous confidence regions in nonlinear regression models. *Claudia Furlan and Cinzia Mortarino* 1311

5.20.2 Tail analysis of a distribution by means of an inequality curve. *E. Taufer, F. Santi, G. Espa and M. M. Dickson* 1318

5.20.3 Nonparametric penalized likelihood for density estimation. *Federico Ferraccioli, Laura M. Sangalli and Livio Finos* 1324

5.20.4 Rethinking the Kolmogorov-Smirnov Test of Goodness of Fit in a Compositional Way. *G.S. Monti, G. Mateu-Figueras, M. I. Ortego, V. Pawlowsky-Glahn and J. J. Egozcue* 1330

5.20.5 Stochastic Dominance for Generalized Parametric Families. *Tommaso Lando and Lucio Bertoli-Barsotti* 1336

5.21 Statistical Models for Ordinal Data 1341

5.21.1 A comparative study of benchmarking procedures for interrater and intrarater agreement studies. *Amalia Vanacore and Maria Sole Pellegrino* 1341

5.21.2 Measuring the multiple facets of tolerance using survey data. *Caterina Liberati and Riccarda Longaretti and Alessandra Michelangeli* 1348

5.21.3 Modified profile likelihood in models for clustered data with missing values. *Claudia Di Caterina and Nicola Sartori* 1352

5.21.4 Worthiness Based Social Scaling. *Giulio D'Epifanio* 1358

5.21.5 Direct Individual Differences Scaling for Evaluation of Research Quality. *Gallo M., Trendafilov N., and Simonacci V.* 1363

5.21.6 A test for variable importance. *Rosaria Simone* 1367

5.22 Statistical Models New Proposals 1373

5.22.1 Decomposing Large Networks: An Approach Based on the MCA based Community Detection. *Carlo Drago* 1373

5.22.2 On Bayesian high-dimensional regression with binary predictors: a simulation study. *Debora Slanzi, Valentina Mameli and Irene Poli* 1380

5.22.3 On the estimation of epidemiological parameters from serological survey data using Bayesian mixture modelling. *Emanuele Del Fava, Piero Manfredi, and Ziv Shkedy* 1386

5.22.4 An evaluation of KL-optimum designs to discriminate between rival copula models. *Laura Deldossi, Silvia Angela Osmetti, Chiara Tommasi* 1392

5.22.5 Variational Approximations for Frequentist and Bayesian Inference. *Luca Maestrini and Matt P. Wand* 1398

5.22.6 Node-specific effects in latent space modelling of multidimensional networks. *Silvia D'Angelo and Marco Alfó and Thomas Brendan Murphy* 1404

5.23 Statistics for Consumer Research 1410

5.23.1 A panel data analysis of Italian hotels. *Antonio Giusti, Laura Grassini, Alessandro Viviani* 1410

5.23.2 A Bayesian Mixed Multinomial Logit Model for Partially Microsimulated Data on Labor Supply. *Cinzia Carota and Consuelo R. Nava* 1417

5.23.3 Comparison between Experience-based Food Insecurity scales. *Federica Onori, Sara Viviani and Pierpaolo Brutti* 1423

5.23.4 Sovereign co-risk measures in the Euro Area. *Giuseppe Arbia, Riccardo Bramante, Silvia Facchinetti, Diego Zappa* 1429

5.23.5 Simultaneous unsupervised and supervised classification modeling for clustering, model selection and dimensionality reduction. *Mario Fordellone and Maurizio Vichi* 1435

5.23.6 Consumers' preference for coffee consumption: a choice experiment including organoleptic characteristics and chemical analysis *Rossella Berni, Nedka D. Nikiforova and Patrizia Pinelli* 1442

5.24 Statistics for Earthquakes 1449

5.24.1 How robust is the skill score of probabilistic earthquake forecasts? *Alessia Caponera and Maximilian J. Werner* 1449

5.24.2 Functional linear models for the analysis of similarity of waveforms. *Francesca Di Salvo, Renata Rotondi and Giovanni Lanzano* 1456

5.24.3 Detection of damage in civil engineering structure by PCA on environmental vibration data. *G. Agró, V. Carlisi, R. Mantione* 1462

5.25 Statistics for Financial Risks 1468
5.25.1 Conditional Value-at-Risk: a comparison between quantile regression and copula functions. *Giovanni De Luca and Giorgia Rivieccio* 1468
5.25.2 Systemic events and diffusion of jumps. *Giovanni Bonaccolto, Nancy Zambon and Massimiliano Caporin* 1474
5.25.3 Traffic Lights for Systemic Risk Detectio. *Massimiliano Caporin, Laura Garcia-Jorcano, Juan-Angel Jiménez-Martin* 1480
5.25.4 Bayesian Quantile Regression Treed. *Mauro Bernardi and Paola Stolfi* 1487
5.25.5 Model Selection in Weighted Stochastic Block models. *Roberto Casarin, Michele Costola, Erdem Yenerdag* 1492
5.26 Tourism & Cultural Participation 1496
5.26.1 The determinants of tourism destination competitiveness in 2006-2016: a partial least square path modelling approach. *Alessandro Magrini, Laura Grassini* 1496
5.26.2 Participation in tourism of Italian residents in the years of the economic recession. *Chiara Bocci, Laura Grassini, Emilia Rocco* 1503
5.26.3 Cultural Participation in the digital Age in Europe: a multilevel cross-national analysis. *Laura Bocci and Isabella Mingo* 1509
5.26.4 Tourist flows and museum admissions in Italy: an integrated analysis. *Lorenzo Cavallo, Francesca Petrei, Maria Teresa Santoro* 1516
5.26.5 Posterior Predictive Assessment for Item Response Theory Models: A Proposal Based on the Hellinger Distance. *Mariagiulia Matteucci and Stefania Mignani* . . . 1522
5.27 Well-being & Quality of Life 1528
5.27.1 Is Structural Equation Modelling Able to Predict Well-being? *Daniele Toninelli and Michela Cameletti* 1528
5.27.2 The well-being in the Italian urban areas: a local geographic variation analysis. *Eugenia Nissi and Annalina Sarra* 1535
5.27.3 Comparing Composite Indicators to measure Quality of Life: the Italian "Sole 24 Ore" case. *Gianna Agró, Mariantonietta Ruggieri and Erasmo Vassallo* 1541
5.27.4 Quality of working life in Italy: findings from Inapp survey. *Paolo Emilio Cardone* 1547
5.27.5 Well-being indices: what about Italian scenario? *Silvia Facchinetti and Elena Siletti* 1554
5.27.6 How can we compare rankings that are expected to be similar? An example based on composite well being indicators. *Silvia Terzi e Luca Moroni* 1560

6 Poster Sessions 1567
6.0.1 A distribution curves comparison approach to analyze the university moving students performance. *Giovanni Boscaino, Giada Adelfio, Gianluca Sottile* . . . 1567
6.0.2 A Partial Ordering Application in Aggregating Dimensions of Subjective Well-being. *Paola Conigliaro* 1574
6.0.3 A note on objective Bayes analysis for graphical vector autoregressive models. *Lucia Paci and Guido Consonni* 1580
6.0.4 Bayesian Population Size Estimation with A Single Sample. *Pierfrancesco Alaimo Di Loro and Luca Tardella* 1586
6.0.5 Classification of the Aneurisk65 dataset using PCA for partially observed functional data. *Marco Stefanucci, Laura Sangalli and Pierpaolo Brutti* 1592
6.0.6 Deep Learning to the Test: an Application to Traffic Data Streams. *Nina Deliu and Pierpaolo Brutti* 1597
6.0.7 Estimating the number of unseen species under heavy tails. *Marco Battiston, Federico Camerlenghi, Emanuele Dolera and Stefano Favaro* 1603
6.0.8 How to measure cybersecurity risk. *Silvia Facchinetti, Paolo Giudici and Silvia Angela Osmetti* 1609

6.0.9 Implementation of an innovative technique to improve Sauvignon Blanc wine quality. *Filippa Bono, Pietro Catanaia and Mariangela Vallone* 1613
6.0.10 Investigating the effect of drugs consumption on survival outcome of Heart Failure patients using joint models: a case study based on regional administrative data. *Marta Spreafico, Francesca Gasperoni, Francesca Ieva* 1619
6.0.11 Mapping the relation between University access test and student's university performance. *Vincenzo Giuseppe Genova, Antonella Plaia* 1625
6.0.12 Multivariate analysis of marine litter abundance through Bayesian space-time models. *C. Calculli, A. Pollice, L. Sion, and P. Maiorano* 1631
6.0.13 Power Priors for Bayesian Analysis of Graphical Models of Conditional Independence in Three Way Contingency Tables. *Katerina Mantzouni, Claudia Tarantola and Ioannis Ntzoufras* 1635
6.0.14 Random Garden: a Supervised Learning Algorithm. *Ivan Luciano Danesi, Valeria Danese, Nicolo' Russo and Enrico Tonini* 1641
6.0.15 Spatiotemporal Prevision for Emergency Medical System Events in Milan. *Andrea Gilardi, Riccardo Borgoni, Andrea Pagliosa, Rodolfo Bonora* 1647
6.0.16 Spatial segregation immigrant households in Messina. *Angelo Mazza and Massimo Mucciardi* 1653
6.0.17 Supervised Learning for Link Prediction in Social Networks. *Riccardo Giubilei, Pierpaolo Brutti* 1657
6.0.18 Women's empowerment and child mortality: the case of Bangladesh. *Chiara Puglisi, Annalisa Busetta* 1663

SIS 2018
49 TH SCIENTIFIC MEETING OF THE ITALIAN STATISTICAL SOCIETY
PALERMO 20-22 JUNE

1. Preface

This book includes the papers presented at the "49th Meeting of the Italian Statistic Society". The conference has registered 445 participants, 350 reports divided into 4 plenary sessions, 20 specialised sessions, 25 sessions solicited, 27 sessions spontaneous, 2 poster sessions. The high number of participants, the high quality of the interventions, the productive spirit of the conference, the ability to respect the time table, are the main indices of the full success of this conference. The meeting hosted also, as plenary sessions, the ISTAT annual report 2018, and a round table on statistics and job markets. Methodological plenary sessions concerned with ordinal data, the dynamics of climate change and models in biomedicine.
Moreover, two related events were held: Start-up Research (SUR) and Stats Under the Stars (SUS4). The SUS4 event attracted many sponsors of statistical, financial, editorial fields as well as numerous students, not only from Italy but also from abroad (Groningen, Tyumen, Barcelona, and Valencia): 98 students for a total of 25 teams. The SUR was a 2-day meeting where small research groups of young scholars, advised by senior researchers with a well-established experience in different areas of Statistics, was asked to develop innovative methods and models to analyse a common dataset from the Neurosciences.

Labour market condition in Italy during and after the financial crises: a segmented regression analysis approach of interrupted time series

Le condizioni del mercato del lavoro in Italia nel periodo durante e successivo alle crisi finanziarie: un'analisi basata sulle serie temporali interrotte

Lucio Masserini and Matilde Bini

Abstract
One of the most widely recognized indicators of the labour market condition is a rising unemployment rate. In Italy, after the 2008 global financial crisis and the 2012 European sovereign debt crisis, this indicator continuously increased over time until late 2014, after which it seems to happen a trend reversal. The aim of this paper is to assess the existence a significant trend reversal in the unemployment rate after 2014, by analysing quarterly data collected from the Italian National Institute of Statistics using a segmented regression analysis approach of interrupted time series. In particular, the analysis is carried out considering some subpopulations of interest, by stratifying unemployment rate for age groups, in order to examine youth unemployment, gender and macro-regions. Moreover, a focus is given to the analysis of the percentage of people Not Engaged in Education, Employment or Training, to provide a more in-depth analysis of the labour market.
Abstract *Uno degli indicatori più ampiamente riconosciuti come misura della recessione è il tasso di disoccupazione. A partire dalle due recenti crisi mondiali, quella finanziaria globale del 2008 e quella del debito sovrano europeo del 2012, questo indicatore in Italia è costantemente aumentato fino a quando al termine del*

[1] Lucio Masserini, Statistical Observatory – University of Pisa; email: lucio.masserini@unipi.it
Matilde Bini, European University of Rome; email: matilde.bini@unier.it

2014 subisce un'inversione di tendenza, ciò forse dovuto ai recenti interventi locali introdotti nel mercato del lavoro. Lo scopo di questo lavoro è di valutare l'esistenza di un'inversione di tendenza significativa dell'indicatore dopo il 2014, analizzando i dati trimestrali raccolti dall'Istituto Nazionale di Statistica Italiano, utilizzando un approccio di analisi di regressione segmentato delle serie temporali interrotte. In particolare, l'analisi è stata realizzata considerando alcune sottopopolazioni di interesse, stratificando il tasso di disoccupazione per fasce di età, al fine di esaminare la disoccupazione giovanile, sul genere e sulle macro-regioni. Inoltre, un'attenzione particolare è stata data all'analisi della percentuale di persone non impegnate nell'istruzione, nell'occupazione o nella formazione per fornire un'analisi più approfondita del mercato del lavoro.

Key words: unemployment rate, interrupted time series analysis, segmented regression, NEET

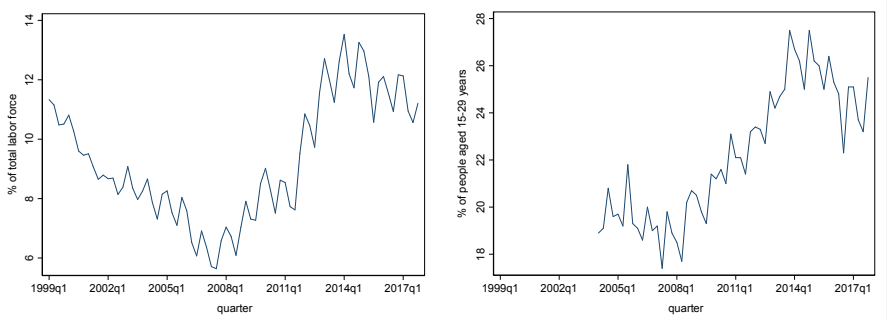
1 Introduction

The recent two economic crises occurred around the world during the period from late 2000s to late 2011, produced negative effects on countries' economies, particularly on GDP growth, on labour productivity and on labour market. As revealed by the International Labour Organization (ILO, 2001), already since the 2009 about 22 million people were unemployed worldwide in particular in developed economies and in the European Union. The unemployment rate continued towards a dramatically increase with high and persistent levels of unemployment until 2014 after which, it seems to happen a trend reversal, maybe due to the recent domestic interventions introduced in the labour market. The aim of this paper is to assess the existence of a significant trend reversal in the unemployment rate after 2014, using a segmented regression analysis approach of interrupted time series. Quarterly data were collected from the website of the Italian National Institute of Statistics. the analysis is carried out considering some subpopulations of interest, by stratifying unemployment rate for age groups, in order to examine youth unemployment, gender and macro-regions. Changes in the trend are also evident in the percentage of a particular sub group of young people, defined as neither employed nor in education or training (NEET). This could mean that local interventions produced effects in some extent also to this category.

2 Data and empirical strategy

Data were collected from I.Stat, the warehouse of statistics currently produced by the Italian National Institute of Statistics (ISTAT) which provides an archive of about 1,500 time series (http://dati.istat.it/). Quarterly data on two different kind of

indicators were downloaded from the theme 'Labour and wages': UR for the period 1993–four quarter of 2017, overall and stratified by gender, age groups and macro-regions; and the percentage of NEET for the period 2004– four quarter of 2017, overall and stratified by gender. Such data are derived from the official estimates obtained in the Labour force survey, carried out on a quarterly basis interviewing a sample of nearly 77,000 households representing 175,000 individuals. According to the Eurostat definition (Eurostat, 2017), UR is given by the number of people unemployed as a percentage of the labour force. The youth unemployment rate (YUR) is the number of unemployed 15–24 years-old expressed as a percentage of the youth labour force, and the NEET refers to the percentage of people aged between 15 and 29 years who currently do not have a job, are not enrolled in training or are not classified as a student. Figure 1a illustrates the trend in the overall UR in Italy from 1999 to the end of 2017. The choice of such a long period allows for a more accurate estimate of the secular trend, and this will be useful for the subsequent analysis. Looking at the graph, UR steadily declines until the third quarter of 2007 (2007q3). Then, starting from the fourth quarter of 2007 (2007q4), period in which the effects of the financial crisis following the bankruptcy of Lehman Brothers begin to appear, UR undergoes a first shock and shows a trend reversal. Afterwards, UR increases even more dramatically up to 13.26% at the end of 2014 (2014q4), after the European sovereign-debt crisis occurred in the late 2011 (2011q4). After this peak, UR seems to show a new trend reversal, showing a possible structural change.



(a) (b)

Figure 1: Total UR (a) and percentage of NEET (b) in Italy over the observed period

Figure 1b illustrates the trend of the overall percentage of NEET in Italy from the first quarter of 2004 (2004q1) to the four quarter of 2017 (2017q4). Here the series is shorter because data from previous years are not available. This trend seems, at least partly, to be similar to that of UR. Indeed, after a slight decrease in the period before the onset of the global financial crisis (2007q3), the percentage of NEET starts a steeply and steady growth that continues unchanged also after the occurrence of the sovereign debt crisis (2011q3). And as in UR, a trend reversal occurred starting from the end of 2014 (2014q4).

In the light of the previous considerations, the analysis period was divided into the following four sub-periods: the period before the so-called 2008 global financial crisis (until 2007q3); the subsequent three-year period known as the Great Recession aftermath of the financial crisis, characterised by a general economic decline observed in world markets (2007q4–2011q3); the period following the European sovereign debt crisis, which resulted in a second economic recession (2011q4–2014q4); and finally, the last three years (2015q1–2017q4), during which it seems to glimpse a slight decrease in both the UR and the percentage of NEET. Consequently, the three breaks in the series were set in 2007q4, 2011q4 and 2015q1. Therefore, the analysis period is limited to the years 1999q1–2017q4 for the UR and to 2004q1–2017q4 for the percentage of NEET. The identified interruptions allow to highlight: the severity of the two financial crises, respectively, the continuation of their effects in the subsequent years of recession, and the more recent trend reversal occurred after the end of 2014.

3 Interrupted time series analysis

In this study, a segmented regression approach of interrupted time series (ITS) analysis was carried out in order to assess and measure, in statistical terms, whether and how much the two financial crises have changed the level and trend in the outcome variables, immediately and over time, and to see if these changes are short- or long-term (Wegner, Soumerai and Zhang, 2002).

ITS analysis (Shadish, Cook and Campbell, 2002) is a simple but powerful tool used in quasi-experimental designs for estimating the impact of population-level or large scale interventions on an outcome variable observed at regular intervals before and after the intervention. In such circumstances, ITS allows to examine any change on the outcome variable in the post-intervention period given the trend in the pre-intervention period (Bernal, Cummins and Gasparrini, 2016). In this respect, the underlying secular trend in the outcome before the intervention is determined and used to estimate the counterfactual scenario, which represents what would have happened if the intervention had not taken place and serves as the basis for comparison. For the purposes of our study the interventions are given by two unplanned and real-world events, the aforementioned and well recognized financial crises, and by the subsequent already mentioned trend reversal. In segmented regression of ITS, each sub-period of the series is allowed to exhibit its own level and trend, which can be represented by the intercept and slope of a regression model, respectively. The intercept indicates the value of the series at the beginning of an observation sub-period; and the slope is the rate of change during a segment (or sub-period). Therefore, by following this approach it is possible to compare the pre-crisis level and trend with their post-crisis counterpart, in order to estimate the magnitude and statistical significance of any differences.

The ITS regression model with a single group under study (here, the Italian population), two interventions, which in this study are given by the two economic recessions in 2007q4 and 2011q4, and a possible UR trend reversal in 2015q1, can

be represented as it follows (Linden and Adams, 2011; Bernal, Cummins and Gasparrini, 2016):

$$y_t = \beta_0 + \beta_1 T_t + \beta_2 x_{t2007q4} + \beta_3 T_{t2007q4} x_{t2007q4} + \beta_4 x_{t2011q4} + \beta_5 T_{t2011q4} x_{t2011q4} + \\ + \beta_6 x_{t2015q1} + \beta_7 T_{t2015q1} x_{t2015q1} + \varepsilon_t .$$

In particular, y_t is the aggregated outcome variable at each equally-spaced time-points t, here represented by quarters; T_t is the time elapsed since the start of the study, where t varies between 1999q1 to 2017q4 for UR and between 2004q1 to 2017q4 for NEET, respectively; $x_{t2007q4}$ is a dummy variable indicating the onset of the global financial crisis in fourth quarter of 2007, coded as 0 (pre-crisis period) and 1 (post-crisis period); $T_{t2007q4}x_{t2007q4}$ is the interaction term between time and the 2007q4 global financial crisis; $x_{t2011q4}$ is a dummy variable indicating the onset of the 2011q4 European sovereign debt crisis, coded as 0 (pre-crisis period) and 1 (post-crisis period); and $T_{t2011q4}x_{t2011q4}$ is the interaction term between time and 2011q4 European sovereign debt crisis. Finally, $x_{t2015q1}$ is a dummy variable indicating the time in which a new trend reversal occurred, coded as 0 (before the trend reversal) and 1 (after the trend reversal); and $T_{t2015q1}x_{t2015q1}$ is the usual interaction term. Accordingly, β_0 is the intercept and represents the starting level of the outcome variable at T = 1999q1 for UR and T = 2004q1 for NEET, respectively; β_1 is the slope and represents the trajectory (or secular trend) of the outcome variable until the 2007q4 global financial crisis; β_2 is the level change that occurs immediately following the 2007q4 global financial crisis (compared to the counterfactual); β_3 is the difference between the slope pre and post the global financial crisis; β_4 is the level change that occurs immediately following the 2011q4 European sovereign debt crisis; β_5 is the difference between the slope pre and post the European sovereign debt crisis; β_6 is the level change that occurs immediately following the 2014q4 (compared to the counterfactual); β_6 is the difference between the slope pre and post the trend reversal; and ε_t represents the random error term which is assumed to follow a first auto-regressive (AR1) process. The regression coefficients are estimated by using Ordinary least-squares (OLS) method with the Newey-West (1987) standard errors.

4 Results

Four periods of linear trend were considered to analyse UR and NEET, with interruptions at 2007q4, 2011q4 and 2015q1, respectively. Separate segmented regression models were then estimated for age groups, gender and macro-regions via ordinary least-squares by using Newy-West standard errors in order to handle one lag autocorrelation. To account for the correct autocorrelation structure, Cumby-Huizinga test (Cumby and Huizinga, 1992) was performed and results confirm that autocorrelation was present at lag 1, but not at higher orders (up to the 9 lags were

tested). Results are shown in Table 1 for the UR and in Table 2 for the NEET. As regards the UR, the 1999 base rate showed some variability in the considered sub-groups. In fact, starting from 10.754 at national level, its value was particularly higher for the age group 15–24 (26.726) and for the South macro-regions (20.211) but lower for the North East (4.538) and North West (5.831) macro-regions, as well as for the males (8.177) and for the 45–54 age group. Moreover, the trend prior to the 2008 global financial crisis (1999q1–2007q3) showed a significant and general decrease, both at national level (-0.137; $p < 0.001$) and for the different age groups, macro regions and gender. Such reduction was more pronounced for the sub-groups traditionally considered as the most vulnerable ones, namely South macro-regions (-0.266; $p < 0.001$), females (-0.202; $p < 0.001$) and YUR (-0.182; $p < 0.001$). The onset of the global financial crisis (2007q4) caused an immediate and substantial UR increase at national level (+0.788; $p < 0.05$) and in almost all the considered sub-groups but no significant change was detected for younger people (age groups 15–24 and 25–34) and the North-East macro region. In particular, the more severe direct consequences were observed among females (+1.061; $p < 0.001$), for people in the central (+1.047; $p < 0.001$) and southern regions (+0.978; $p < 0.05$) and for the intermediate age group 35–44 (+0.997; $p < 0.001$). The aftermath of the financial crisis was quite strong and resulted in the Great recession in the subsequent years during which a substantial and significant trend change was observed compared to the previous period (+0.253 $p < 0.001$). However, in this case, the most serious consequences occurred particularly for YUR (+0.717; $p < 0.001$) and, to a much lesser extent, for the South macro-regions (+0.365; $p < 0.001$). On the other hand, the immediate consequences of the second financial crisis, following the European sovereign debt crisis (2011q4) were even stronger when compared to the previous financial crisis and resulted in a second economic recession, with an UR increase almost double at the national level (+1.540; $p < 0.001$). Such increase was higher for YUR (+3.676; $p < 0.05$) and for the South macro region (+2.611; $p < 0.001$) while there was no significant increase again for North East macro region.

After this second financial shock, the UR seems to further accelerate its increase only in some sub-groups while at national level no significant trend difference was observed. In particular, such acceleration was particularly higher for the South macro-regions (+0.355; $p < 0.001$), age group 25–34 (+0.246; $p < 0.05$) and females (+0.187; $p < 0.001$) while no significant further rate increase was detected for YUR. However, it should be emphasized here that this further increase, although lower than the one highlighted during the Great Recession, where present has to be added to that already existing, thus making particularly critical the situation.

Table 1: Estimates of the impact of the 2007q4 and 2011q4 financial crises on the UR in Italy and trend reversal after 2015q1

…	*Base rate (1999)*	*Trend 1999q1-2007q3*	*Rate change 2007q4*	*Trend change 2007q4*	*Rate change 2011q4*	*Trend change 2011q4*	*Rate change 2015q1*	*Trend change 2015q1*
Overall	10.754***	-0.137***	0.788**	0.253***	1.540***	0.136**	-1.234**	-0.357***
Males	8.177***	-0.095***	0.602*	0.250***	1.426***	0.101	-1.018	-0.396***
Females	14.655***	-0.202***	1.061***	0.261***	1.675***	0.187***	-1.515***	-0.304***
15–24	26.726***	-0.182***	0.955	0.717***	3.676**	0.350	-3.371	-1.607***
25–34	11.074***	-0.068***	0.127	0.290***	1.566**	0.246**	-1.811*	-0.600***
35–44	7.771***	-0.095***	0.997***	0.177***	1.245***	0.166***	-1.225***	-0.285***
45–54	6.401***	-0.099***	0.725***	0.196***	1.111***	0.109**	-0.869*	-0.242***
55–64	7.901***	-0.166***	0.847***	0.219***	1.350***	-0.005	-0.143	-0.031
North West	5.831***	-0.067***	0.800**	0.221***	0.892**	0.008	-0.776	-0.320***
North East	4.538***	-0.037***	0.220	0.166***	0.855	0.007	-0.673	-0.269***
Center	8.540***	-0.097***	1.047***	0.188***	1.364***	0.134**	-1.021	-0.329***
South	20.211***	-0.266***	0.978**	0.365***	2.611***	0.355***	-2.290***	-0.501***

Table 2: Estimates of the impact of the 2007q4 and 2011q4 financial crises on the percentage of NEET in Italy and trend reversal after 2015q1

…	*Base rate (2004)*	*Trend 2004q1-2007q3*	*Rate change 2007q4*	*Trend change 2007q4*	*Rate change 2011q4*	*Trend change 2011q4*	*Rate change 2015q1*	*Trend change 2015q1*
Overall	19.866***	-0.062	-0.250	0.316***	0.103	0.094	-1.510**	-0.529***
Males	15.152***	0.014	-0.398	0.334***	-0.037	0.099	-1.787**	-0.680***
Females	24.584***	-0.125*	-0.221	0.285***	0.269	0.084	-1.135	-0.369***
North West	12.575***	-0.079	0.664	0.382***	-1.242	0.103	-1.632	-0.572***
North East	10.522***	-0.009	-0.558	0.364***	-0.268	-0.069	-1.169	-0.545***
Center	15.406***	-0.090	-0.995	0.404***	0.989	-0.001	-1.053	-0.561***
South	29.798***	-0.072	0.295	0.258***	0.564	0.204***	-1.801***	-0.482***

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

The consequences of this second financial crisis seem to run out by the end of 2014. Indeed, a trend reversal occurs staring from 2015q1. In particular, the rate change decreases significantly only overall (-1.234; $p < 0.05$), for females (-1.515; $p < 0.001$), age group 35–44 (-1.225; $p < 0.001$) and the South macro-region (-2.290; $p < 0.001$). On the other hand, the trend change is more evident and is significant for all categories except for age class 55–64 and shows a stronger reduction for young generation 25–34 (-1.607; $p < 0.001$) and, to a lesser extent, for the South macro-region (-0.501; $p < 0.001$).

As regards the percentage of NEET, a considerable heterogeneity was found in the 2004 base rate, which was 19.866 at national level. Its value was higher for the South macro-regions (29.798) but lower for the North East (10.522) and North West (12.575) macro-regions; moreover, it was higher for females (24.584) than males (15.152). On the other hand, its trend prior to the 2008 global financial crisis (2004q1–2007q3) was basically constant at national level. The onset of the of the global financial crisis (2007q4) did not cause an immediate impact on the percentage of NEET, overall and in any of the considered sub-groups. However, a significant trend change was found both at national level (+0.316; $p < 0.001$) and for all the other sub-groups; such change was particularly higher only for the macro-regions of Center (+0.404; $p < 0.001$). The European sovereign debt crisis (2011q4) does not seem to alter this situation, neither for the rate change nor for the trend change. This means that after this second financial crisis the rise of the percentage of NEET remains steady and equal to the previous period without showing any jump. But for the last trend change (2015q1), results reveal to be significant for all sub groups, particularly for females (-0.680; $p < 0.001$). This confirms that after the end of the two crises and local interventions, a new period of recovery has finally started again. Also for the percentage of NEET, a trend reversal occurs staring from 2015q1. In particular, analogously to the UR, this reduction is significant for all the considered sub-groups, but it is more pronounced for males (-0.680; $p < 0.001$) and lower for females (-0.369; $p < 0.001$) and South marco-region (-0.482; $p < 0.001$).

References

1. Cumby, R. E., and Huizinga, J.: Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. Econometrica, 60, 185—195 (1992).
2. Linden, A., and Adams, J. L.: Applying a propensity-score based weighting model to interrupted time series data: Improving causal inference in program evaluation. *Journal of Evaluation in Clinical Practice*, 17, 1231–1238 (2011).
3. Lopez Bernal, J., Cummins, S., and Gasparrini, A.: Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int. J. Epidemiol*., 1–8 (2016).
4. Newey, W. K., and West, K. D.: A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708 (1987).
5. Shadish, W. R., Cook, T. D., and Campbell, D. T.: Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin (2002).
6. Wagner, A. K., Soumerai, S. B., Zhang, F., and Ross-Degnan, D.: Segmented regression analysis of interrupted time series studies in medication use research. J. *Clin. Pharm. Ther*., 27(4), 299–309 (2002).
7. Eurostat: *Glossary*. Luxembourg: Eurostat (2017).
8. ILO: *World of work report 2011: Making markets work for jobs.* Geneva: International Labour Office (2011).