Workshop on Modelling Complex Data: New Methodology and Theory

6 and 7 June 2022

P/T/006, Exhibition Centre, Campus West, University of York

Day 1: Monday, 6 June

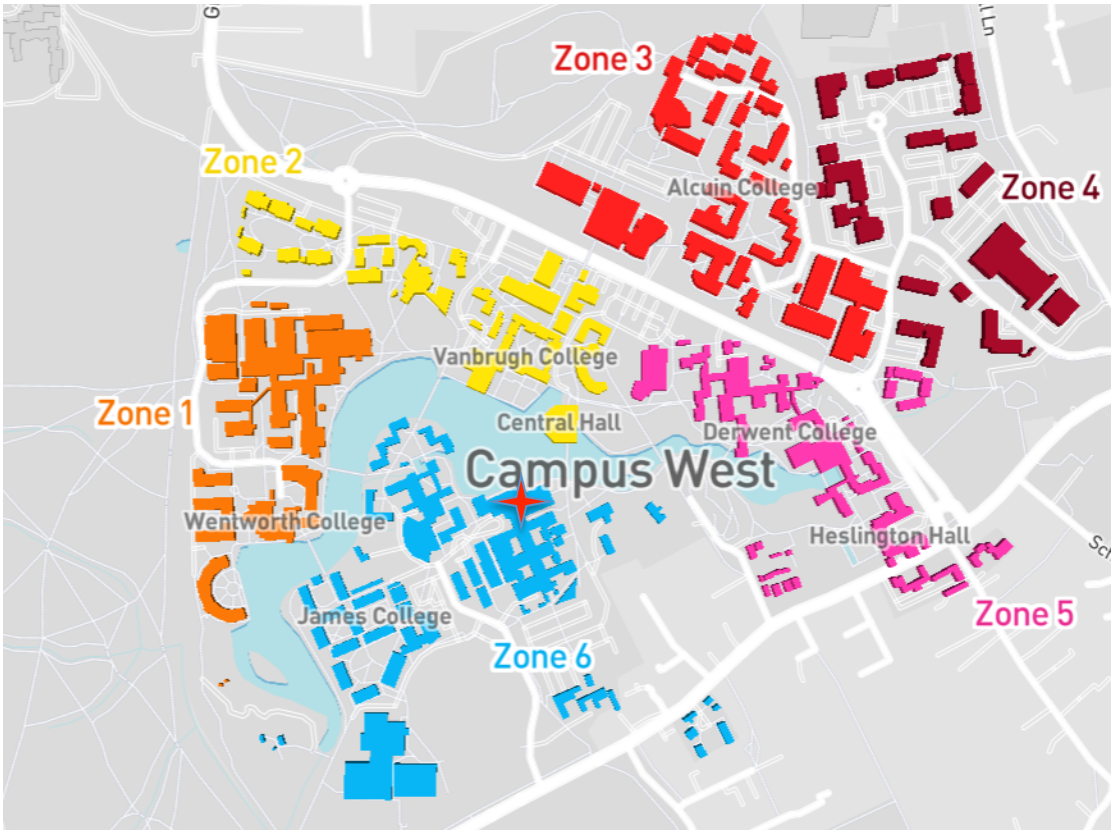
10:30—11:00	**Welcome & refreshments**	
11:00—12:30	**Session 1 (Chair: Degui Li)**	
11:00—11:45	**Oliver Linton** University of Cambridge	Dynamic Autoregressive Liquidity (DArLiQ)
11:45—12:30	**Alastair Hall** University of Manchester	Bootstrap Methods for Impulse Responses Obtained by Local Projections
12:30—14:00	**Lunch**	
14:00—15:30	**Session 2 (Chair: Ben Powell)**	
14:00—14:45	**Haeran Cho** University of Bristol	High-dimensional Time Series Segmentation via Factor-adjusted Vector Autoregressive Modelling
14:45—15:30	**Clifford Lam** London School of Economics	Rank and Factor Loadings Estimation in Time Series Tensor Factor Model by Pre-averaging
15:30—16:00	**Tea break**	
16:00—17:30	**Session 3 (Chair: Yue Zhao)**	
16:00—16:45	**Ruijun Bu** University of Liverpool	Nonparametric Estimation of Large Spot Volatility Matrices for High-Frequency Financial Data
16:45—17:30	**Francesco Bravo** University of York	Estimation of Nonsmooth Nonparametric Moment Conditions Models
18:30	**Workshop dinner (by invitation)**	

Day 2: Tuesday, 7 June

9:30—11:00	**Session 4 (Chair: Jessica Hargreaves)**	
9:30—10:15	**Indeewara Perera** University of Sheffield	Specification Tests for GARCH Processes with Nuisance Parameters on the Boundary
10:15—11:00	**Marina Knight** University of York	Forecasting UK Gross Value Added Time Series using a Locally Stationary Approach
11:00—11:30	**Tea break**	
11:30—13:00	**Session 5 (Chair: Jia Chen)**	
11:30—12:15	**Zudi Lu** University of Southampton	Learning from Nonlinear Dynamic Spatio-Temporal Network Modelling: Theory with Application
12:15—13:00	**Dennis Kristensen** University College London	Fast Estimation of Dynamic Structural Models with Fixed Effects
13:00	**Lunch**	

Acknowledgement: This workshop is funded by BA/Leverhulme Research Grants and supported by the Department of Mathematics, University of York.

Direction for workshop venue: Exhibition Centre, ground floor (with). The room is in the wing nearest to James College and the Student Centre (YUSU).



Dynamic Autoregressive Liquidity (DArLiQ)

Speaker: Oliver Linton (Day 1, 11:00—11:45)

Abstract

We introduce a new class of semiparametric dynamic autoregressive models for the Amihud illiquidity measure, which captures both the long-run trend in the illiquidity series with a nonparametric component and the short-run dynamics with an autoregressive component. We develop a GMM estimator based on conditional moment restrictions and an efficient semiparametric ML estimator based on an iid assumption. We derive large sample properties for both estimators. We further develop a methodology to detect the occurrence of permanent and transitory breaks in the illiquidity process. Finally, we demonstrate the model performance and its empirical relevance on two applications. First, we study the impact of stock splits on the illiquidity dynamics of the five largest US technology company stocks. Second, we investigate how the different components of the illiquidity process obtained from our model relate to the stock market risk premium using data on the S&P 500 stock market index.

Bootstrap Methods for Impulse Responses Obtained by Local Projections

Speaker: Alastair Hall (Day 1, 11:45—12:30)

Abstract

We propose a method based on the dependent wild bootstrap for the calculation of confidence intervals for impulse response functions that have been obtained from local projection regressions estimated via instrumental variables. We show the first-order asymptotic validity of the bootstrap in this context. In a simulation study, we compare the finite sample coverage of our proposed confidence intervals with those based on asymptotic theory and the block bootstrap. The methods are illustrated via an application to the analysis of monetary policy shocks on macroeconomic variables.

High-dimensional Time Series Segmentation via Factor-adjusted Vector Autoregressive Modelling

Speaker: Haeran Cho (Day 1, 14:00—14:45)

Abstract

Piecewise stationarity is a widely adopted assumption for modelling non-stationary time series. However fitting piecewise stationary vector autoregressive (VAR) models to high-dimensional data is challenging as the number of parameters increases as a quadratic of the dimension. Recent approaches to address this have imposed sparsity assumptions on the parameters of the VAR model, but such assumptions have been shown to be inadequate when datasets exhibit strong (auto)correlations. We propose a piecewise stationary time series model that accounts for pervasive serial and cross-sectional correlations through a factor structure, and only assumes that any remaining idiosyncratic dependence between variables can be modelled by a sparse VAR model. We propose an accompanying two-stage change point detection methodology which fully addresses the challenges arising from not observing either the factors or the idiosyncratic VAR process directly. Its consistency in estimating both the total number and the locations of the change points in the latent components, is established under conditions considerably more general than those in the existing literature. We demonstrate the competitive performance of the proposed methodology on simulated datasets and an application to US blue chip stocks data.

Rank and Factor Loadings Estimation in Time Series Tensor Factor Model by Pre-averaging

Speaker: Clifford Lam (Day 1, 14:45—15:30)

Abstract

Tensor time series data appears naturally in a lot of fields, including finance and economics. As a major dimension reduction tool, similar to its factor model counterpart, the idiosyncratic components of a tensor time series factor model can exhibit serial correlations, especially in financial and economic applications. This rules out a lot of state-of-the-art methods that assume white idiosyncratic components, or even independent/Gaussian data.

While the traditional higher order orthogonal iteration (HOOI) is proved to be convergent to a set of factor loading matrices, the closeness of them to the true underlying factor loading matrices are in general not established, or only under some strict circumstances like having i.i.d. Gaussian noises (Zhang and Xia, 2018). Under the presence of serial and cross-correlations in the idiosyncratic components and time series variables with only bounded fourth order moments, we propose a pre-averaging method that accumulates information from tensor fibres for better estimating all the factor loading spaces. The estimated directions corresponding to the strongest factors are then used for projecting the data for a potentially improved re-estimation of the factor loading spaces themselves, with theoretical guarantees and rate of convergence spelt out. We also propose a new rank estimation method based on the projected data which utilizes correlation information from the projected data, in the same spirit as Fan et. al. (2020) for factor models with independent data. Extensive simulation results reveal competitive performance of our rank and factor loading estimators relative to other state-of-the-art or traditional alternatives. A set of real data is also analysed.

Nonparametric Estimation of Large Spot Volatility Matrices for High-Frequency Financial Data

Speaker: Ruijun Bu (Day 1, 16:00—16:45)

Abstract

In this paper, we consider estimating spot/instantaneous volatility matrices of high-frequency data collected for a large number of assets. We first combine classic nonparametric kernel-based smoothing with a generalised shrinkage technique in the matrix estimation for noise-free data under a uniform sparsity assumption, a natural extension of the approximate sparsity commonly used in the literature. The uniform consistency property is derived for the proposed spot volatility matrix estimator with convergence rates comparable to the optimal minimax one. For the high-frequency data contaminated by the microstructure noise, we introduce a localised pre-averaging estimation method in the high-dimensional setting which first pre-whitens data via a kernel filter and then uses the estimation tool developed in the noise-free scenario, and further derive the uniform convergence rates for the developed spot volatility matrix estimator. In addition, we also combine the kernel smoothing with the shrinkage technique to estimate the time-varying volatility matrix of the high-dimensional noise vector, and establish the relevant uniform consistency result. Numerical studies are provided to examine performance of the proposed estimation methods in finite samples.

Estimation of Nonsmooth Nonparametric Moment Conditions Models

Speaker: Francesco Bravo (Day 1, 16:45—17:30)

Abstract

This paper considers estimation of possibly nonsmooth nonparametric moment conditions models with weakly dependent data. The paper considers local versions of GMM and GEL estimators and derives their asymptotic distributions. The results are fairly general and can be applied, for example, to an instrumental variables quantile varying coefficients model. Monte Carlo results show that the proposed local estimators perform well in finite samples, with the local GEL estimators having an edge over the asymptotically equivalent efficient local GMM estimator.

Specification Tests for GARCH Processes with Nuisance Parameters on the Boundary

Speaker: Indeewara Perera (Day 2, 9:30—10:15)

Abstract

This paper develops tests for the correct specification of the conditional variance function in GARCH models when the true parameter may lie on the boundary of the parameter space. The test statistics considered are of Kolmogorov-Smirnov and Cramer-von Mises type, and are based on a certain empirical process marked by centered squared residuals. The limiting distributions of the test statistics depend on unknown nuisance parameters in a non-trivial way, making the tests difficult to implement. We therefore introduce a novel bootstrap procedure which is shown to be asymptotically valid under general conditions, irrespective of the presence of nuisance parameters on the boundary. The proposed bootstrap approach is based on shrinking of the parameter estimates used to generate the bootstrap sample toward the boundary of the parameter space at a proper rate. It is simple to implement and fast in applications, as the associated test statistics have simple closed form expressions. Although the bootstrap test is designed for a data generating process with fixed parameters (i.e., independent of the sample size n), we also discuss how to obtain valid inference for sequences of DGPs with parameters approaching the boundary at the root-n rate. A simulation study demonstrates that the new tests: (i) have excellent finite sample behaviour in terms of empirical rejection probabilities under the null as well as under the alternative; (ii) provide a useful complement to existing procedures based on Ljung-Box type approaches. Two data examples illustrate the implementation of the proposed tests in applications.

Forecasting UK Gross Value Added Time Series using a Locally Stationary Approach

Speaker: Marina Knight (Day 2, 10:15—11:00)

Abstract

Accurate forecasts of gross value added are of great importance not only to economists, but to a wide range of users in e.g. business, government and the media. Raw gross value added time series are often transformed to be closer to stationarity to enable modelling and forecasting using classical time series methods. A common and obvious non-stationarity in gross value added time series is the increase over time of the series' variance due to inflation. However, even inflation-corrected versions of the series exhibit subtle and statistically significant nonstationarities, which can cause problems for classical stationarity-assuming forecasters. We adopt a different approach that works directly with the gross value added series by developing recent forecasting methods for locally stationary time series. Our approach results in more accurate and reliable forecasts, also demonstrated on a broad collection of time series that showcase the wider practical utility of our methods.

Learning from Nonlinear Dynamic Spatio-Temporal Network Modelling: Theory with Application

Speaker: Zudi Lu (Day 2, 10:15—11:00)

Abstract

Nonlinear modelling of spatio-temporal network data is often a challenge due to irregularly observed locations and location-wide non-stationarity. In this paper we propose a semiparametric family of Dynamic Functional-coefficient Autoregressive Spatio-Temporal (DyFAST) models to address the difficulties. First, we specify the dynamic autoregressive smooth coefficients depending on both a concerned regime and location so that the models can characterise not only the dynamic regime-switching nature but also the location-wide non-stationarity in real spatio-temporal data. Second, two semiparametric smoothing schemes are proposed to model the dynamic neighbouring-time interaction effects with irregular locations incorporated by (spatial) weight matrices. The first scheme popular in econometrics supposes that the weight matrix is pre-specified. In practice, many weight matrices can be generated differently by data location features. Model selection for an optimal one is popular, but may suffer from loss of features of different candidates. Our second scheme is thus to suggest a weight matrix fusion to let data combine or select the candidates. Accordingly, different semiparametric smoothing procedures are developed. Both theoretical properties and Monte Carlo simulations are investigated. The empirical application to an EU energy market dataset further demonstrates the usefulness of our DyFAST models.

Fast Estimation of Dynamic Structural Models with Fixed Effects

Speaker: Dennis Kristensen (Day 2, 12:15—13:00)

Abstract

We propose a novel approximate fixed effects (AFE) estimator that employs interpolation in the computation of its objective function. This feature greatly reduces the number of times the underlying economic model needs to be solved. In the case of dynamic programming models this can reduce the estimation time from days to minutes. We study the asymptotic behavior of the AFE estimator and derive the leading additional biases due to approximations under mild regularity conditions. We demonstrate that the Jack-knife removes both the usual incidental parameter bias and biases due to interpolation. Monte Carlo results highlights the attractive features of the AFE which is much faster than the exact FE estimator and with only small additional estimation errors. We apply the AFE to fit the buffer-stock consumption-saving model with unrestricted heterogeneity in the discount factor on Danish register data.