Deutsche Forschungsgemeinschaft

New Developments in Econometrics and Time Series

Rome, November 16-17, 2017

Venue:
Einaudi Institute for Economics and Finance (EIEF)
Via Sallustiana, 62- Rome

Organization Committee:
Holger Dette Ruhr-Universitaet Bochum)
Marc Hallin (ECARES-ULB)
Marco Lippi (EIEF)

PRELIMINARY PROGRAM

(Presentations: 50 minutes)

Thursday, November 16th

08:45 – 09:15 Registration

09:15 – 09:30 Opening Sessions – Welcoming address: EIEF Representative and Holger Dette

Morning Session - Chair: Holger Dette

09:30 – 10:20 **Manfred Deistler** (EOS - TUWien)
"High Frequency Linear Time Series Models and Mixed Frequency Data: Identifiability and Estimation" joint with Brian D.O. Anderson and Mag. A. Braumann

10:20 – 11:10 **Esther Ruiz** (Universidad Carlos III de Madrid)
"Resampling Uncertainty of Principal Components Factors" joint with Javier de Vicente

11:10 – 11:40 Coffee break

11:40 – 12:30 **Majid M. Al-Sadoon** (UPF and Barcelona GSE)
"A Unifying Theory of Tests of Rank"

12:30 – 14:00 Lunch break

Afternoon Session- Chair: Chair: Marc Hallin

14:00 – 14:50 **Edgar Dobriban** (Wharton School – UPenn)
"Parallel Analysis Selects the Significant Factors"

14:50 – 15:40 **Dacheng Xiu** (University of Chicago)
"Inference on Risk Premia in the Presence of Omitted Factors" joint with Stefano Giglio

15:40 - 16:10 Coffee break

16:10 – 17:00 **Piotr S. Kokoszka** (Colorado State University)
"Some Methods for Functional Econometric Time Series"

17:00 – 17:50 **Anne Van Delft** (Université Catholique de Louvain)
"Testing for Stationarity of Functional Time Series in the Frequency Domain" joint with Alexander Aue

19:30 - 21:30 Conference Dinner

Friday, November 17th

Morning Session - Chair: Manfred Deistler

09:30 – 10:20 **Moritz Jirak** (TU Braunschweig)
"Transition between Berry-Esséen Bounds and Edgeworth Expansions under Dependence"

10:20 – 11:10 **Denis Chetverikov** (UCLA)
"On Cross-Validated Lasso" joint with Zhipeng Liao and Victor Chernozhukov

11:10 – 11:40 Coffee break

11:40 – 12:30 **David Preinerstorfer** (ECARES-ULB)
"Controlling the Size of Autocorrelation Robust Tests" joint with Benedikt M. Pöstscher

12:30 – 14:00 Lunch break

Afternoon Session- Chair: Marco Lippi

14:00 – 14:50 **Germain Van Bever** (ECARES)
"Halfspace Depths for Scatter, Concentration and Shape Matrices" joint work with Davy Paindaveine

14:50 – 15:40 **Thomas V. Mikosch** (University of Copenhagen)
"The Eigenstructure of Sample Covariance Matrices for High-Dimensional Heavy-tailed Stochastic Volatility Models" joint with Johannes Heiny

15:40 – 16:10 Coffee break

16:10 – 17:00 **Alexander Aue** (UC Davis)
"Bootstrapping Linear Spectral Statistics in High Dimensions" joint with Andrew Blandino and Miles Lopes

17:00 – 17:50 **Michael Wolf** (University of Zürich)
"Beyond Sorting: A More Powerful Test for Cross-Sectional Anomalies" joint with Olivier Ledoit and Zhao Zhao

17:50 – 18:00 Closing address: Marco Lippi and Holger Dette

09:30

Manfred Deistler (EOS - TUWien)

"High Frequency Linear Time Series Models and Mixed Frequency Data: Identifiability and Estimation"

joint with Brian D.O. Anderson (ANU, Canberra) and Mag A. Braumann (TU Braunschweig)

The presentation considers VAR and VARMA models, as well as generalized linear dynamic factor models in case of mixed frequency (MF) data. The MF data are obtained by linear aggregation; special emphasis is given to the case of stock variables. The first part of the presentation deals with identifiability of the parameters of the underlying high frequency models from MF data. Here in particular generic identifiability for the VAR case as well as for the VARMA case where the AR order is larger than or equal to the MA order is shown. In the second part we consider several approaches to parameter estimation and we discuss and compare the properties of these estimators. Here special emphasis is given to the EM algorithm.

10:20

Esther Ruiz (Universidad Carlos III)

"Resampling Uncertainty of Principal Components Factors"

joint with Javier de Vicente (Universidad Carlos III)

In the context of Dynamic Factor Models (DFMs), one of the most popular procedures for factor extraction is Principal Components (PC). Measuring the uncertainty associated to PC factor estimates should be part of interpreting them. Although the asymptotic distribution of PC factors is known, it could not be an appropriate approximation to the finite sample one for the sample sizes and cross-sectional dimensions usually encountered in practice. The main problem is that it does not take into account parameter uncertainty. Alternatively, several bootstrap procedures have been proposed in DFM with goals related to inference. We show that these procedures are not appropriate to measure the uncertainty of PC factor estimates and propose an alternative resampling procedure designed with this purpose. The asymptotic and finite sample properties of the proposed procedure are analyzed and compared with those of the asymptotic and alternative extant bootstrap procedures. The results are empirically illustrated obtaining confidence intervals of the underlying factor in a system of Spanish macroeconomic variables and in a system of house prices of advanced and emerging markets.

11:40

Majid M. Al-Sadoon (UPF and Barcelona GSE)

"A Unifying Theory of Tests of Rank"

The general principles underlying tests of matrix rank are investigated. It is demonstrated that statistics for such tests can be seen as implicit functions of null space estimators. In turn, the asymptotic behaviour of the null space estimators is shown to determine the asymptotic behaviour of the statistics through a plug-in principle. The theory simplifies the asymptotics under a variety of alternatives of empirical relevance as well as misspecification, clarifies the relationships between the various existing tests, makes use of important results in the numerical analysis literature, and motivates numerous new tests. A brief Monte Carlo study illustrates the results.

14:00

Edgar Dobriban (Wharton School – UPenn)

"Parallel Analysis Selects the Significant Factors"

Scientists and social scientists routinely perform Factor Analysis (FA) for unsupervised discovery of factors in the data. Selecting the number of factors is an important but hard problem. A permutation method known as parallel analysis (PA), proposed by Horn (1965), and refined by Buja and Eyuboglu (1992), is one of the most popular approaches. PA selects the factors whose singular values are larger than those of matrices obtained by permuting each feature independently. Despite extensive empirical evidence that PA works better than Kaiser's "eigenvalues larger than one" criterion and Bartlett's likelihood ratio test, as well as being more objective than the scree plot, it currently has no theoretical justification. In this paper, we provide a theoretical study of PA. We establish precise conditions under which PA consistently \emph{selects the significant factors} for large datasets. We call a factor significant if its contribution to the data is larger than the intrinsic noise level. We allow the sample size n and the dimension p to grow to infinity such that $p/n\to\gamma>0$ or $p/n\to\infty$. Our analysis rests on new bounds for operator norms of permutation random matrices. For practitioners we provide simple guidelines for when PA is likely to work: (1) when the dimension of the data is large, and (2) when the factors load on more than just a few variables.

14:50

Dacheng Xiu (University of Chicago)

"Inference on Risk Premia in the Presence of Omitted Factors"

joint with Stefano Giglio (University of Chicago)

We propose a three-pass method to estimate the risk premia of observable factors in a linear asset pricing model, which is valid even when the observed factors are just a subset of the true factors that drive asset prices or they are measured with error. We show that the risk premium of a factor can be identified in a linear factor model regardless of the rotation of the other control factors as long as they together span the space of true factors. Motivated by this rotation invariance result, our approach uses principal components to recover the factor space and combines the estimated principal components with each observed factor to obtain a consistent estimate of its risk premium. Our methodology also accounts for potential measurement error in the observed factors and detects when such factors are spurious or even useless. The methodology exploits the blessings of dimensionality, and we therefore apply it to a large panel of equity portfolios to estimate risk premia for several workhorse linear models. The estimates are robust to the choice of test portfolios within equities as well as across many asset classes.

16:10

Piotr Kokoszka (Colorado State University)

"Some Methods for Functional Econometric Time Series"

The talk will introduce fundamental ideas of the analysis of time series of functions. Examples of such time series are yield curves and intraday return curves. Within the framework of functional data analysis, fundamental concept of long-run covariance, autocorrelations and their estimators will be introduced. Three specific inferential problems will be discussed: 1) testing if a functional time series is stationary, 2) testing if it is a functional weak white noise, 3) detecting change points in its mean structure in the presence of changing variability. An asymptotic framework as well as numerical implementation will be outlined.

17:00

Anne Van Delft (Université Catholique de Louvain)

"Testing for Stationarity of Functional Time Series in the Frequency Domain"

joint with Alexander Aue (UC Davis)

Interest in functional time series has spiked in the recent past with papers covering both methodology and applications being published at a much increased pace. This article contributes to the research in this area

by proposing stationarity tests for functional time series based on frequency domain methods. Setting up the tests requires a delicate understanding of periodogram- and spectral density operators that are the functional counterparts of periodogram- and spectral density matrices in the multivariate world. Two sets of statistics are proposed. One is based on the eigendecomposition of the spectral density operator, the other on a fixed projection basis. Their properties are derived both under the null hypothesis of stationary functional time series and under the smooth alternative of locally stationary functional time series. The methodology is theoretically justified through asymptotic results. Evidence from simulation studies and an application to annual temperature curves suggests that the tests work well in finite samples.

Friday, November 17th

09:30

Moritz Jirak (TU Braunschweig)

"Transition between Berry-Esséen bounds and Edgeworth Expansions under Dependence"

Given a stationary, weakly dependent Bernoulli-shift sequence, we establish the transition between a Berry-Esséen bound and a second order Edgeworth expansion. Unlike to previous results in the literature, we do not require a conditional Cramér type criterium, only an unconditional tail bound is required. This tail bound is sharp: If a second order Edgeworth expansion exists, such a bound necessarily holds. If it fails to be present, we still get an optimal Berry-Esséen bound, thus describing the exact transition. We also obtain (fractional) expansions given $3 < p < 4$ moments, where a similar transition occurs. Corresponding results also hold for the Wasserstein metric, where a related, integrated tail bound turns out to be optimal. In all results, the imposed weak-dependence is very mild. As an application, (studentized) second order expansions are obtained for many well-known time series models, including (augmented)-Garch models of any order.

10:20

Denis Chetverikov (UCLA)

"On Cross-validated Lasso"

joint with Zhipeng Liao (UCLA), and Victor Chernozhukov (MIT)

In this paper, we derive a rate of convergence of the Lasso estimator when the penalty parameter λ for the estimator is chosen using K-fold cross-validation; in particular, we show that in the model with the

Gaussian noise and under fairly general assumptions on the candidate set of values of λ, the prediction norm of the estimation error of the cross-validated Lasso estimator is with high probability bounded from above up to a constant by $(s \log p/n)^{1/2} \cdot \log^{7/8}(p\ n)$, where n is the sample size of available data, p is the number of covariates, and s is the number of non-zero coefficients in the model. Thus, the cross-validated Lasso estimator achieves the fastest possible rate of convergence up to a small logarithmic factor $\log^{7/8}(pn)$. In addition, we derive a sparsity bound for the cross-validated Lasso estimator; in particular, we show that under the same conditions as above, the number of non-zero coefficients of the estimator is with high probability bounded from above up to a constant by $s \log^5(pn)$. Finally, we show that our proof technique generates non-trivial bounds on the prediction norm of the estimation error of the cross-validated Lasso estimator even if the assumption of the Gaussian noise fails; in particular, the prediction norm of the estimation error is with high-probability bounded from above up to a constant by $(s \log^2 (pn)/n)^{1/4}$ under mild regularity conditions.

11:40

David Preinerstorfer (ECARES-ULB)

"Controlling the Size of Autocorrelation Robust Tests"

joint with Benedikt M. Pöstscher (University of Vienna)

Autocorrelation robust tests are notorious for suffering from size distortions and power problems. We investigate under which conditions the size of autocorrelation robust tests can be controlled by an appropriate choice of critical value.

14:00

Germain Van Bever (ECARES)

"Halfspace Depths for Scatter, Concentration and Shape Matrices"

joint work with Davy Paindaveine (ECARES)

In this talk, I will propose halfspace depth concepts for scatter, concentration and shape matrices. For scatter matrices, the concept is similar to those from Chen, Gao and Ren (2017) and Zhang (2002). Rather than focusing, as in these earlier works, on deepest scatter matrices, the focus is put on thoroughly investigating the properties of the proposed depth and the corresponding depth regions. This is done under minimal assumptions and, in particular, we do not restrict to elliptical nor absolutely continuous distributions. Interestingly, fully understanding scatter halfspace depth requires considering different geometries/topologies on the space of scatter matrices. We also discuss, in the spirit of Zuo and Serfling (2000) the structural properties a scatter

depth should satisfy and investigate whether or not these are met by scatter halfspace depth. Companion concept of depth for concentration and shape matrices are also proposed and studied. We show the practical relevance of the depth concept by studying the volatility in multivariate financial time series and illustrate how the depth concepts can be used to detect days with atypical behaviors in scatter, shape or scale.

14:50

Thomas V. Mikosch (University of Copenhagen)

"The Eigenstructure of Sample Covariance Matrices for High-Dimensional Heavy-tailed Stochastic Volatility Models"

joint with Johannes Heiny (Aarhus University)

We are interested in the asymptotic behavior of the eigenvalues of the sample covariance matrix XX' where the data matrix has the structure of a stochastic volatility model, i.e., $X_{it} = \sigma_{it} Z_{it}$ and the stationary volatility field (σ_{it}) is independent of the iid field (Z_{it}). We assume that Z_{it} has regularly varying tails with index $\alpha \in (0,4)$ and that σ_{it} has lighter tails. We are interested in the case when the dimension $p = p_n$ converges to infinity. In this case the eigenvalues of XX' are essentially determined by the diagonal elements of the matrix X. We consider limit theory of Poisson-type for these eigenvalues and also discuss the structure of the corresponding eigenvectors.

16:10

Alexander Aue (UC Davis)

"Bootstrapping Linear Spectral Statistics in High Dimensions"

joint with Andrew Blandino (UC Davis) and Miles Lopes (UC Davis)

Spectral statistics play a central role in many multivariate testing problems. It is therefore of interest to approximate the distribution of functions of the eigenvalues of sample covariance matrices. For low-dimensional settings, these approximations can be based both on large-sample results, such as central limit theorems, and on bootstrap methods. The situation is decidedly different in high-dimensional settings, for example in the case of sample size and dimensionality diverging proportionally considered in this talk. In this high-dimensional context, central limit theorems typically have a complicated analytical form impeding their straightforward application. Bootstrap methods are relatively unexplored and simple forms have even been shown to be invalid. The aim of this paper is to introduce a version of the bootstrap for the class of linear spectral statistics that is both theoretically justified and easy to apply in practice. It is in particular shown that the proposed bootstrap provides consistent in-law approximations in the high-dimensional setting that work well in finite-

sample scenarios, as evidenced through results from a simulation study that also uses eigenvalue distributions appearing in actual data. Additionally, the simulations indicate that the bootstrap appears to be informative for the distribution of the largest eigenvalue, even though this case is not covered by the theory.

17:00

Michael Wolf (University of Zürich)

"Beyond Sorting: A More Powerful Test for Cross-Sectional Anomalies"

joint with Olivier Ledoit (University of Zürich) and Zhao Zhao (Huazhong University of Science and Technology)

Many researchers seek factors that predict the cross-section of stock returns. The standard methodology sorts stocks according to their factor scores into quantiles and forms a corresponding long-short portfolio. Such a course of action ignores any information on the covariance matrix of stock returns. Historically, it has been difficult to estimate the covariance matrix for a large universe of stocks. We demonstrate that using the recent DCC-NL estimator of Engle et al. (2016) substantially enhances the power of tests for cross-sectional anomalies: On average, 'Student' t-statistics more than double.