## Abstracts

Testing whether volatility can be written as a function of the asset price Yacine Ait-Sahalia 2

Quantile forecasting based on a multivariate hysteretic autoregressive model with GARCH Errors and timevarying correlations Cathy W.S. Chen 3

Systematic and discretionary hedge funds: classiffication and performance comparison Hui-Ching Chuang 4

Proxy CDS curves for individual corporates globally Jin-Chuan Duan5

Tutorial on the first and second fundamental theorem of quantitative risk management Paul Embrechts 6

Risk-sharing, robustness and regulation Paul Embrechts 7

Some risk measures on Wiener space Hans Föllmer 8

Model-uncertain value-at-risk, expected shortfall and sharpe ratio, using stochastic approximation Emmanuel Gobet 9

Modeling and tracking bubbles Christian Hafner 10

Pricing and Hedging Crypto options Wolfgang Karl Härdle 11

Limits to arbitrage in markets with stochastic settlement latency Nikolaus Hautsch 12

Tutorial on high-frequency econometrics Nikolaus Hautsch 13

Mean-risk portfolio choice with weighted VaR and lawinvariant coherent risk measures Hanqing Jin 15

Simultaneous volatility and skewness risk in asset pricing Kian Guan Lim 16

Basket credit derivatives pricing in a Markov chain model with interacting intensities and contagion risk Xiaosong Qian 17 Optimal make-take fees for market making regulation Mathieu Rosenbaum 18

Mean-variance hedging without information Martin Schweizer 19

Backtesting, prequential analysis and prediction process Hideatsu Tsukahara 20

Return cross-predictability in firms with similar employee satisfaction Jun Tu 21

Using generalized estimating equations to estimate nonlinear models with spatial data Weining Wang 22

Swing pricing for mutual funds Marko Weber 23

Autoencoder asset pricing models Dacheng Xiu 24

Fund flows and performance under unobservable dynamic managing ability Jingrui Xu 25

Double machine learning with gradient boosting and its application to the big N Jui-Chung Yang, Ray 26

Hermite expansion for transition densities of irreducible diffusions with an application to option pricing
Nian Yang 27

Convex duality in portfolio theory Qiji Zhu 28

# Testing whether volatility can be written as a function of the asset price

## YACINE AIT-SAHALIA

Princeton University, USA

### ABSTRACT

There is broad agreement that the volatility of most economic and financial state variables is time-varying. One way to achieve this in continuous-time models is to make volatility its own latent stochastic variable, leading to stochastic volatility models. Another, more restrictive approach, consists in making the volatility a function of the state variables themselves, leading to local volatility models. This paper tests whether a semimartingale with stochastic volatility can be written as one with local volatility. The proposed test is based on high frequency asymptotics, is robust to the presence of jumps and noise in the data, and requires only the computation of sums of increments of the observed state variables. The test is applied to stock price and interest data and suggests that stochastic volatility is a necessary feature of models for the former, while less so for the latter.

Joint work with Jean Jacod.

# Quantile forecasting based on a multivariate hysteretic autoregressive model with GARCH Errors and time-varying correlations

<u>Cathy W.S. Chen</u><sup>a</sup>, Hong Than-Thi<sup>a</sup>, Mike K.P. So<sup>b</sup>, and Songsak Sriboonchitta<sup>c</sup>

<sup>a</sup>Feng Chia University, Taiwan <sup>b</sup>Hong Kong University of Science and Technology, Hong Kong <sup>c</sup>Chiang Mai University, Thailand

### ABSTRACT

To understand and predict chronological dependence in the second-order moments of asset returns, we consider a multivariate hysteretic autoregressive (HAR) model with GARCH specification and time-varying correlations, by providing a new way to describe a nonlinear dynamic structure of the target time series. The hysteresis variable governs the nonlinear dynamics of the proposed model in which the regime switch can be delayed if the hysteresis variable lies in a hysteresis zone. The proposed model combines three useful model components for modeling economic and financial data: (1) the multivariate HAR model, (2) the multivariate hysteretic volatility models, and (3) a dynamic conditional correlation structure. We incorporate an adapted multivariate Student-t innovation based on a scale mixture normal presentation in the HAR model to tolerate for dependence and different shaped innovation components. We carry out multivariate volatilities, Value-at-Risk, and marginal expected shortfall based on a Bayesian sampling scheme through adaptive Markov chain Monte Carlo (MCMC) methods, which allow us to statistically estimate all unknown model parameters and forecasts simultaneously. We illustrate the proposed methods herein by using both simulated and real examples and measure for industry downside tail risk jointly.

Keywords: Hysteresis; MCMC method; scale mixture of normal distributions; multivariate Student-t distribution; Marginal expected shortfall; Value-at-Risk; Out-ofsample forecasting.

# Systematic and discretionary hedge funds: classification and performance comparison

## Hui-Ching Chuang

Yuan Ze University, Taiwan

### ABSTRACT

In this paper we introduce an approach to building classifiers that bifurcate hedge funds into systematic and discretionary funds and evaluate their performance. This approach makes use of textual analysis and statistical learning methods that are free from subjective judgement of investment strategies. In our empirical study, we find that a random forest classifier yields the highest accuracy ratio, and that the funds classified as systematic, on average, result in higher raw returns, Sharpe ratios, and factor-adjusted returns than their discretionary counterparts. A bootstrap analysis also shows that the standardized alphas of a large portion of systematic and discretionary funds are statistically significantly different from zero, suggesting such performance is due to fund managers' authentic investment skills, rather than their luck. Nonetheless, we find that systematic Equity Hedge funds are to be preferred to their discretionary counterparts because the standardized alphas of the former stochastically dominate those of the latter.

# Proxy CDS curves for individual corporates globally

## JIN-CHUAN DUAN

National University of Singapore, Singapore

### ABSTRACT

Corporate credit default swap (CDS) premium is the market price of credit risk posed by a corporate obligor. Although corporate CDS are commonly used for risk benchmarking in accounting and credit risk management, liquid CDS are limited to under 500 corporate names globally. CDS users must either confine their usage to this limited subset or resort to aggregates derived from the liquid CDS in different industry/rating combinations. This paper offers an intuitive, practical and robust predictive regression model linking liquid USD-denominated CDS premiums of different tenors to a set of obligor-specific attributes, and with the model one can generate proxy CDS curves for corporates without liquid or traded CDS. One key attribute is the actuarial spread that reflects the actuarial value of a CDS contract and is made available by the Credit Research Initiative of National University of Singapore for all exchange-listed firms globally. Other attributes in the predictive regression model include investment vs. speculative grades based on an obligor's credit rating, and some general credit environment variables such as the April 2009 CDS Big Bang, among others. This predictive regression, constructed with the historical record on 405 corporate CDS names, enables daily production of proxy CDS curves on over 35,000 currently active exchange-listed corporates globally.

# Tutorial on the first and second fundamental theorem of quantitative risk management

## PAUL EMBRECHTS

ETH Zürich, Switzerland

### ABSTRACT

I very much hope that participants to this 4-hour tutorial will think beforehand what their candidates for a First and Second FT of QRM would/could be. Rather than lifting the "veil of secrecy" at this point, I can say so much as that I will guide the participants through some key results of the book: "A.J. McNeil, R. Frey and P. Embrechts (2015) Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press", see also www.qrmtutorial.org for an accompanying website. For this tutorial, I more specifically think of "a methodological excursion through the land of risk management while keeping practical relevance always in sight".

# Risk-sharing, robustness and regulation

## PAUL EMBRECHTS

ETH Zürich, Switzerland

## ABSTRACT

In this talk I will summarize some recent work on the concept of risk-sharing, discuss in particular properties like robustness, moral hazard and equilibrium pricing and finally highlight consequences for capital regulation for the financial and insurance industry.

# Some risk measures on Wiener space

## HANS FÖLLMER

Humboldt-Universität zu Berlin, Germany

### ABSTRACT

We discuss some convex risk measures on Wiener space that are defined in terms of different Wasserstein metrics and different notions of relative entropy. Their comparison involves Talagrand's optimal transport inequality for Gaussian measures, in particular a new refined version in terms of specific relative entropy. The focus will be on the sensitivity of such risk measures to the fine structure of Brownian motion, in particular to the behaviour of quadratic variation.

# Model-uncertain value-at-risk, expected shortfall and sharpe ratio, using stochastic approximation

## Emmanuel Gobet

École Polytechnique, France

### ABSTRACT

VaR, ES, SR have as a common point the property of being the limit of Stochastic Approximation algorithm. In this work we discuss how to quantity the model uncertainty in these types of limits when some model parameters are unknown and modelled as random variables. The random limits are decomposed into chaos which coefficients are computed recursively. The resulting algorithm is an SA algorithm in increasing dimension. Convergence results are established and experimental results are investigated.

Joint works with Stephane Crepey, Gersende Fort, Uladzislau Stazhynski and Linda Chamakh

# Modeling and tracking bubbles

## CHRISTIAN HAFNER

Université Catholique de Louvain, Belgium

### ABSTRACT

Markets with imperfect information are often characterized by bubble-like behaviour and returns that are not martingale differences. Classification into bubble and non-bubble periods is far from clear once heteroscedasticity and non-normality are taken into account. We propose and compare alternative score-driven models capable of tracking financial bubbles and typical volatility patterns. We discuss properties, estimation, inference and diagnostic tests of the models. In an application to a set of cryptocurrencies, we confirm the importance of a bubble component. Possible extensions include the multivariate case with a common bubble component, and models with links between the bubble and volatility.

Joint work with Andrew Harvey

# Pricing and Hedging Crypto options

WOLFGANG KARL HÄRDLE

Humboldt-Universität zu Berlin, Germany Singapore Management University, Singapore

### ABSTRACT

The CRIX (CRyptocurrency IndeX) has been constructed based on a number of cryptos and provides a high coverage of market liquidity, hu.berlin/crix. The crypto currency market is a new asset market and attracts a lot of investors recently. Surprisingly a market for contingent claims hat not been built up yet. A reason is certainly the lack of pricing tools that are based on solid financial econometric tools. Here a first step towards pricing of derivatives of this new asset class is presented. After a careful econometric pre-analysis we motivate an affine jump diffusion model, i.e., the SVCJ (Stochastic Volatility with Correlated Jumps) model. We calibrate SVCJ by MCMC and obtain interpretable jump processes and then via simulation price options. The jumps present in the cryptocurrency fluctuations are an essential component. Concrete examples are given to establish an OCRIX exchange platform trading options on CRIX.

Hedging OCRIX products requires strategies under jump-induced market incompleteness. The challenges and the difficulties in hedge strategy determination are therefore Market incompleteness, Unpredictability of jumps, creating a Trade-Off between hedging possibilities. We finally have to do Hedging under model misspecification and present P&L distributions for several calibrated models.

# Limits to arbitrage in markets with stochastic settlement latency

## NIKOLAUS HAUTSCH

University of Vienna, Austria

### ABSTRACT

Distributed ledger technologies rely on consensus protocols confronting traders with random waiting times until the transfer of ownership is accomplished. This time-consuming settlement process exposes arbitrageurs to price risk and imposes limits to arbitrage. We derive theoretical arbitrage boundaries under general assumptions and show that they increase with expected latency, latency uncertainty, spot volatility, and risk aversion. Using high-frequency data from the Bitcoin network, we estimate arbitrage boundaries due to settlement latency of on average 124 basis points, covering 88% of the observed cross-exchange price differences. We document cross-exchange flows chasing arbitrage opportunities only if we account for transaction cost and settlement latency. Settlement through decentralized systems thus induces non-trivial frictions affecting market efficiency and price formation. Professor Dr. Nikolaus Hautsch University of Vienna Page 1 of 2

## "High-Frequency Econometrics"

Singapore, August 29, 2019

## Abstract

This short course introduces to financial high-frequency data, presents the concepts of integrated variation and realized variance and discusses models for ex post variance estimators. In the second part of the course, the role and effects of market microstructure noise on variance estimators will be presented. This includes a deeper discussion of novel insights on local mispricing, noise endogeneity and concepts from market microstructure theory. In the last part of the course, high-frequency based estimators of large-dimensional asset return covariances and their use in portfolio allocation problems will be presented. Particular emphasis will be given on the role of transaction costs and the link between turnover penalization and covariance shrinkage.

## **Course Outline**

### 1. The Realized Variance

- 1.1. High-Frequency Data
- 1.2. The Concept of Realized Variance
- 1.3. Modelling Realized Variances

### 2. Noise-Adjusted Variance Estimators

- 2.1. Market Microstructure Noise
- 2.2. Kernel-Based Estimators
- 2.3. Local Mispricing and Noise

### 3. Large-Scale Portfolio Allocation

- 3.1. Realized Covariance Estimation
- 3.2. Realized Kernels
- 3.3. Portfolio Optimization under Transaction Costs

## Literature

Ait-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005): "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise", *Review of Financial Studies*, 18, 351-416. Andersen, T.G., Archakov, I., Cebiroglu, G., and Hautsch, N. (2019): "Local Mispricing and Microstructural Noise: A Parametric Perspective", Working Paper.

Andersen, T.G., and Benzoni, L. (2009) :"Realized Volatility", in: "Handbook of Financial Time Series", Editors: Andersen, T.G., and Benzoni, L. Davis, R.A., Kreiß, J.-P., and Mikosch, T., Springer. Barndorff-Nielsen, O. E., Hansen, P.R., Lunde, A., and Shephard, N. (2008): "Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise", *Econometrica*, 76, 1481-1536.

**Barndorff-Nielsen, O. E., Hansen, P.R., Lunde, A., and Shephard, N. (2011)**: "Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading", *Journal of Econometrics*, 162, 149-169.

**Barndorff-Nielsen, O.E., and Shephard, N. (2004)**: "Econometric Analysis of Realized Covariation: High Frequency Based Covariane, Regression, and Correlation in Financial Economics", *Econometrica*, 72, 885-925.

**Corsi, F. (2009):** "A Simple Approximate Long-Memory Model of Realized Volatility", *Journal of Financial Econometrics*, 7, 174-196

Hansen, P. R., Huang, Z., and Shek, H.S. (2012): "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility", *Journal of Applied Econometrics*, 27, 877-906.

Hansen, P. R., and Lunde, A. (2006): "Realized Variance and Market Microstructure Noise", *Journal of Business & Economics Statistics*, 24, 127-161.

Hautsch, N. (2012): "Econometrics of Financial High-Frequency Data", Springer.

Hautsch, N., Kyj, L., and Oomen, R.C.A. (2012): "A blocking and regularization approach to high dimensional realized covariance estimation", *Journal of Applied Econometrics*, forthcoming

Hautsch, N., Kyj, L., and Malec, P. (2015): "Do High-Frequency Data Improve High-Dimensional Portfolio Allocation?", *Journal of Applied Econometrics*, 30(2), 263-290.

Hautsch, N., and Voigt, S. (2019): "Large-Scale Portfolio Allocation Under Transaction Costs and Model Uncertainty", *Journal of Econometrics*, forthcoming

Hayashi, T., and Yoshida, N. (2005): "On covariance estimation of non-synchronously observed diffusion processes", *Bernoulli*, 11 (2), 359-379.

**Noureldin, D., Shephard, N., and Sheppard, K. (2012):** "Multivariate High-Frequency-Based Volatility (HEAVY) Models", *Journal of Applied Econometrics*, 27, 907-933.

**Shephard, N., and Sheppard, K. (2010):** "Realising the Future: Forecasting with High-Frequency-Based Volatility (HEAVY) Models", *Journal of Applied Econometrics*, 25, 197-231.

**Zhang, L., Mykland, P. A., and Ait-Sahalia, Y. (2005):** "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data", *Journal of the American Statistical Association*, 100, 1394-1411.

# Mean-risk portfolio choice with weighted VaR and law-invariant coherent risk measures

# HANQING JIN

University of Oxford, UK

## ABSTRACT

We study a continuous-time mean-risk portfolio choice problem in which an agent, with or without the bankruptcy constraint, chooses among the portfolios that achieve an exogenously given expected terminal wealth target with the objective of minimizing the risk of his portfolio. The risk is measured either by a so-called weighted value-at-risk risk measure, which is a generalization of value-at-risk and conditional value-at-risk, or by a law-invariant coherent risk measure.

# Simultaneous volatility and skewness risk in asset pricing

## KIAN GUAN LIM

Singapore Management University, Singapore

### ABSTRACT

The contribution in this paper is two-fold. Firstly, we propose a multiple criteria framework (MCF) for generalizing the traditional two-moment CAPM to a market portfolio that contains optimal balance of variance and skewness. Only in the extreme case of zero skewness does this portfolio converge to the Sharpe-Lintner market portfolio. Our model decomposes effective portfolio variance into two parts, an undesirable total variance part and a desirable positive variability part. Minimization is taken with respect not to total variance as in the Sharpe-Lintner CAPM. but to the total variance less the positive variability. Positive covariability reduces risk premium, and has a similar effect compared to positive skewness or upper partial moment. The empirical results in our MCF model indicate good performance of the model. Secondly, and more importantly, the model produces easily computable optimal portfolios accounting simultaneously for negative variance contribution but positive skewness contributions to a risk-averse representative agent. We employ the model to explain the idiosyncratic volatility and negative premium puzzle in the literature. This is accomplished under rational equilibrium asset pricing without imposing additional linear regression specifications.

# Basket credit derivatives pricing in a Markov chain model with interacting intensities and contagion risk

## XIAOSONG QIAN

Soochow University, China

### ABSTRACT

We analyze basket credit derivative contracts (BCDS and BCLN) with counterparty risk using a markov chain interacting intensities with contagion model. We assume the default intensities of the protection seller and the references are affected by an external shock event and contagion risk. The arrival of the shock event is a Cox process whose stochastic intensity is an affine diffusion process with jumps. We derive a recursive formula for the joint default probability of reference assets. The pricing formulas of BCDS and BCLN are presented. We also examine how the correlated default risks between the protection seller and the underlying entity may affect the premium rates.

# Optimal make-take fees for market making regulation

## MATHIEU ROSENBAUM

École Polytechnique, France

### ABSTRACT

We consider an exchange who wishes to set suitable make-take fees to attract liquidity on its platform. Using a principal-agent approach, we are able to describe in quasi-explicit form the optimal contract to propose to a market maker. This contract depends essentially on the market maker inventory trajectory and on the volatility of the asset. We also provide the optimal quotes that should be displayed by the market maker. The simplicity of our formulas allows us to analyze in details the effects of optimal contracting with an exchange, compared to a situation without contract. We show in particular that it leads to higher quality liquidity and lower trading costs for investors. This is joint work with Omar El Euch, Thibaut Mastrolia and Nizar Touzi.

# Mean-variance hedging without information

## MARTIN SCHWEIZER

ETH Zürich, Switzerland

### ABSTRACT

Our goal is to solve the problems of mean-variance hedging (MVH) and meanvariance portfolio selection (MVPS) under restricted information. We work in a setting where the underlying price process S is a semimartingale, and a smaller filtration G models the information available for constructing trading strategies. We concentrate on the zero-information case where  $\mathbb{G} = \mathbb{F}^{det}$ , which means that trading strategies must be deterministic functions. If we assume that S is a time-dependent affine transformation of a square-integrable martingale, we can give explicit solutions to the MVH and MVPS problems. Explicit formulas are obtained for hedging European call options in the Bachelier and Black-Scholes models. This is based on joint work with Danijel Zivoi and Mario Sikic.

# Backtesting, prequential analysis and prediction process

## HIDEATSU TSUKAHARA

Seijo University, Japan

### ABSTRACT

In the prequential framework, data are observed sequentially in time, and at each time it is required to make a prediction about the distribution of the next observation conditional on the past, based on the data up to that point. Sometime later, we need to check whether the sequence of our forecasts are consistent with that of observations. This ex-post examination of predictive performance is called a backtesting in finance. Through the recent controversy on backtestability issue, we now know that, depending on which aspect of the conditional distribution we try to predict, we are faced with varying difficulty in devising backtesting procedures. In this talk, some attempts are made to extend Davis' calibration concept to larger classes of statistical functions (with values in an abstract space). Comparison of two probability forecasting systems under absolute continuity condition may be interpreted in terms of the corresponding prediction processes which always possess Markov property, and we explore its implications. Computation for a few simple examples from time series analysis will be shown to exemplify the theory. Finally, the possibility of extensions to the case with auxiliary random variables (covariates) and to the continuous-time case will be discussed.

# Return cross-predictability in firms with similar employee satisfaction

## Jun Tu

Singapore Management University, Singapore

## ABSTRACT

We study the return predictability of similar employee satisfaction (SES) firms using new firm-ranking data of employee satisfaction from Glassdoor. We find that the returns of firm peers with SES have a predictive power for focal firm returns. A long-short portfolio sorted on the lagged returns of SES firm peers yields a significant Fama and French (2018) six-factor alpha of 135 bps per month. This result is distinct from industry and inter-firm momentum effects and cannot be explained by riskbased arguments. Our tests suggest that investors' limited attention is the primary reason of firms' underreaction to their SES firm returns.

# Using generalized estimating equations to estimate nonlinear models with spatial data

## WEINING WANG

City, University of London, UK

### ABSTRACT

In this paper, we study estimation of nonlinear models with cross sectional data using two-step generalized estimating equations (GEE) in the quasi-maximum likelihood estimation (QMLE) framework. In the interest of improving efficiency, we propose a grouping estimator to account for the potential spatial correlation in the underlying innovations. We use a Poisson model and a Negative Binomial II model for count data and a Probit model for binary response data to demonstrate the GEE procedure. Under mild weak dependency assumptions, results on estimation consistency and asymptotic normality are provided. Monte Carlo simulations show efficiency gain of our approach in comparison of different estimation methods for count data and binary response data. Finally we apply the GEE approach to study the determinants of the inflow foreign direct investment (FDI) to China.

## Swing pricing for mutual funds

## Marko Weber

National University of Singapore, Singapore

### ABSTRACT

We develop a model of the feedback between mutual fund outflows and asset illiquidity. Following a market shock, alert investors anticipate the impact on a fund's net asset value (NAV) of other investors' redemptions and exit first at favorable prices. This first-mover advantage may lead to fund failure through a cycle of falling prices and increasing redemptions. Our analysis shows that (i) the firstmover advantage introduces a nonlinear dependence between a market shock and the aggregate impact of redemptions on the fund's NAV; (ii) as a consequence, there is a critical magnitude of the shock beyond which redemptions brings down the fund; (iii) properly designed swing pricing transfers liquidation costs from the fund to redeeming investors and, by removing the nonlinearity stemming from the first-mover advantage, it reduces these costs and prevents fund failure. Achieving these objectives requires a larger swing factor at larger levels of outflows. The swing factor for one fund may also depend on policies followed by other funds.

## Autoencoder asset pricing models

DACHENG XIU

The University of Chicago Booth School of Business, USA

#### ABSTRACT

We propose a new latent factor conditional asset pricing model. Like Kelly, Pruitt, and Su (2019), our model allows for latent factors and factor exposures that depend on covariates such as asset characteristics. But, unlike the linearity assumption of KPS, we model factor exposures as a flexible nonlinear function of covariates. Our model retrofits the workhorse unsupervised dimension reduction device from the machine learning literature - autoencoder neural networks - to incorporate information from covariates along with returns themselves. This delivers estimates of nonlinear conditional exposures and the associated latent factors. Furthermore, our machine learning framework imposes the economic restriction of no-arbitrage. Our autoencoder asset pricing model delivers out-of-sample pricing errors that are far smaller (and generally insignificant) compared to other leading factor models.

# Fund flows and performance under unobservable dynamic managing ability

David Feldman<sup>a</sup> and  $\underline{Jingrui Xu}^b$ 

<sup>a</sup>UNSW Business School, Australia <sup>b</sup>Xiamen University, China

### ABSTRACT

We introduce a continuous-time rational model of fund flows and performance. Following Berk and Green's (2004) decreasing returns to scale framework, we allow dynamic unobservable fund managers' abilities, and risk-averse investors. In contrast to current models where constant managers' abilities induce consistently improving ability estimates, to asymptotic (precision) perfection, our investors face, forever, a tracking problem. Depending on parameter values, dynamic managers' inferred abilities and their precisions, may increase, decrease, or stay constant in fund returns' shocks, leading to increasing, decreasing or constant fund flows sensitivities to performance. Our empirical evidence supports dynamic managers' abilities over constant ones. We offer further insights into current empirical findings.

# Double machine learning with gradient boosting and its application to the big N audit quality effect

Jui-Chung (Ray) Yang<sup>a</sup>, Hui-Ching Chuang<sup>b</sup>, and Chung-Ming Kuan<sup>c</sup>

> <sup>a</sup>National Tsing Hua University, Taiwan <sup>b</sup>Yuan Ze University, Taiwan <sup>c</sup>National Taiwan University, Taiwan

### ABSTRACT

In this paper, we study the double machine learning (DML) approach of Chernozhukov et al. (2018) for estimating average treatment effect and apply this approach to examine the Big N audit quality effect in the accounting literature. This approach relies on machine learning methods and is suitable when a high dimensional nuisance function with many covariates is present in the model. This approach does not suffer from the "regularization bias" when a learning method with a proper convergence rate is used. We demonstrate by simulations that, for the DML approach, the gradient boosting method is to be preferred to other learning methods, such as regression tree and random forest. We then apply this approach with gradient boosting to estimate the Big N effect.

We find that Big N auditors have a positive effect on audit quality and that this effect is not only statistically significant but also economically important. We also show that, in contrast to the results of propensity score matching, our estimates of said effect are quite robust to the hyper-parameters in the gradient boosting algorithm.

Keywords: Audit Quality, Average Treatment Effect, Big N Effect, Double Machine Learning, Gradient Boosting, Performance-Matched Discretionary Accruals

JEL Classification: C14, C31, M42

# Hermite expansion for transition densities of irreducible diffusions with an application to option pricing

## NIAN YANG

Nanjing University, China

### ABSTRACT

A diffusion is said to be reducible if there exists a one-to-one transformation of the diffusion into a new one whose diffusion matrix is the identity diffusion matrix, otherwise it is irreducible (Ait-Sahalia, 2008). Ait-Sahalia (2002) uses Hermite polynomials as orthogonal basis to expand the transition density of the transformed process, because it is "closer" to a standard normal than that of the original process. However, most multivariate diffusions such as the stochastic volatility models are irreducible. As pointed out by Ait-Sahalia (2008), the straight Hermite expansion will not in general converge for irreducible diffusions.

In this paper we manage to develop the convergent Hermite expansion for transition densities of irreducible diffusions. By introducing a quasi-Lamperti transform unitizing the process' diffusion matrix at the initial time, we can expand the transition density of the transformed process using Hermite polynomials as the orthogonal basis. Then we derive explicit recursive formulas for the expansion coefficients using the Ito-Taylor expansion method, and prove the convergence of the expansion as the time interval shrinks to zero. Moreover, we show that the newly derived Hermite expansion unifies some existing methods including the pathwise expansion of Li (2013) and the Ito-Taylor expansion of Yang, Chen, and Wan (2019). Hence we theoretically justify the conjecture of Yang, Chen, and Wan (2019) that the pathwise and Ito-Taylor expansions are equivalent up to an appropriate way collecting terms. In addition, we demonstrate the advantage of Hermite expansion by deriving explicit recursive expansion formulas for European option prices under irreducible diffusions. Numerical experiments illustrate the accuracy and effectiveness of our approach.

This is a joint work with Xiangwei Wan from Shanghai Jiao Tong University.

# Convex duality in portfolio theory

## Qiji Zhu

Western Michigan University, USA

## ABSTRACT

Portfolio problems can be phased in terms of convex programing problems. The related dual problems often have interesting financial meanings. We illustrate the role of convex duality in portfolio theory using a simple one period model and points to many interesting directions for further exploration.