Journal of Econometrics 239 (2024) 105627

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Data science in economics and finance: Introduction

1. Introduction

Data science is a thriving, broad discipline that combines the best of classical and modern statistics, as well as applied and theoretical econometrics, machine learning, probability, and optimization. The data science revolution has transformed the way that quantitative research is conducted in the social, behavioral, and biomedical sciences, and in finance and industry more generally. Professor Jianqing Fan is a towering figure in data science, and this themed issue on "Data Science in Economics and Finance" was organized to celebrate his 60th birthday. It complements the conference "Statistical Foundations of Data Science and their Applications: A Conference in Celebration of Jianqing Fan's 60th Birthday" (https://fan60.princeton.edu/), which took place on May 8–10, 2023, in Princeton, NJ.[1]

Professor Jianqing Fan received his Bachelor of Science from Fudan University in 1982, his Master's degree in Probability and Statistics from the Chinese Academy of Sciences in 1985, and his Ph.D. in Statistics from the University of California at Berkeley in 1989. He was appointed as assistant, associate, and full professor at the University of North Carolina at Chapel Hill (1989–2003), as professor at the University of California at Los Angeles (1997–2000), and as Professor of Statistics and Chairman at the Chinese University of Hong Kong (2000–2003). Since 2003, he has been a professor in the Department of Operations Research and Financial Engineering at Princeton University, where he was named the Frederick L. Moore'18 Professor of Finance in 2006, served as Chair from 2012 to 2015, and was the Director for the Committee on Statistical Studies from 2005 to 2017. In addition, he has been the director of both the statistics and financial econometrics labs that he created since 2008. Fig. 1 displays a recent portrait of Professor Fan.

The quality and scope of Professor Fan's work position him at the pinnacle of statisticians, financial econometricians, and data scientists. He is widely recognized as one of the greatest statisticians and financial econometricians of our time in many metrics: breadth, depth, productivity, mentoring, and service. He has made seminal contributions to numerous fields, and his research has been published in the best journals in statistics, finance, econometrics, probability, operations research, and many other disciplines. His research is distinguished by an innovative and powerful mathematical treatment, always coupled with a distinctly modern perspective. He has often broken ground for new fields, which has inspired much follow-up research, both in theoretical and methodological developments. Professor Fan has consistently been ranked as one of the most highly cited data scientists in the world. He holds a prominent international standing and is actively engaged with the statistical, econometrics, and data science communities, as demonstrated by his involvement in numerous international initiatives.

Professor Fan has coauthored four highly regarded books: *Local Polynomial Modeling and Its Applications* (1996), *Nonlinear time series: Parametric and Nonparametric Methods* (2003), *Elements of Financial Econometrics* (2015), and *Statistical Foundations of Data Science* (2020). He has also (co-)authored over 300 articles spanning finance, economics, statistical machine learning, computational biology, spectral methods for data science, nonparametric and semiparametric modeling, nonlinear time series, survival analysis, longitudinal data analysis, and other aspects of theoretical and methodological statistics. His wide-ranging and far-reaching contributions to statistics, financial econometrics, computational biology, and statistical machine learning have been recognized by the 2000 COPSS Presidents' Award, given annually to one outstanding statistician under the age of 40. Additionally, he was invited as a speaker at the 2006 International Congress for Mathematicians, a highly regarded honor in mathematics once every 4 years. He received the Humboldt Research Award in 2006, a distinguished award given annually to recognize the lifetime achievements of up to 100 internationally renowned scientists and scholars worldwide, the Morningside Gold Medal of Applied Mathematics in 2007, honoring triennially an outstanding applied mathematician of Chinese descent under age 45, and was selected as a Guggenheim

[1] Cattaneo gratefully acknowledges financial support from the National Science Foundation through grant DMS-2304646.

https://doi.org/10.1016/j.jeconom.2023.105627

Available online 27 December 2023

Fig. 1. Professor Jianqing Fan (范剑青), 2023.

Fellow in 2009, a highly prestigious recognition in the arts and sciences. He also won the Pao-Lu Hsu Prize (2013), presented every three years by the International Chinese Statistical Association to individuals under the age of 50, the Guy Medal in Silver (2014), presented once a year by the Royal Statistical Society, the Noether Senior Scholar Award (2018), presented once a year by American Statistical Association, the Le Cam Award and Lectures (2021), presented by the Institute of Mathematical Statistics once every 3 years. He was elected as a Fellow of the American Association for the Advancement of Science, the Institute of Mathematical Statistics, and the American Statistical Association. He was also honored as an Academician of Academia Sinica in 2012, and was elected to the Royal Flemish Academy of Belgium in 2023.

We briefly overview Professor Fan's substantial contributions to statistics, econometrics, and to the profession and society, in terms of research and service.

Research. Professor Fan's research can broadly be categorized as follows.

- **Statistics**. (a) *Deconvolution*. Professor Fan was among the first to demonstrate that the difficulty of nonparametric deconvolution depends critically on the smoothness of the convolution kernel. He was the first to solve the problems of nonparametric error-in-variable regression and extended the techniques to deconvolve functions in the Besov space using wavelets. His pioneering work has had a strong impact on the contemporary development of deconvolution. (b) *Local polynomial modeling*. The term local modeling was coined by Professor Fan, and he is widely recognized as the architect of contemporary local polynomial modeling. His two original papers revolutionized the field of nonparametric modeling and forever changed the classical view of nonparametric function estimation. His pioneering contributions are summarized in his highly regarded book on *Local Polynomial Modeling* (1996). His impact on nonparametric modeling is enormous. (c) *Multivariate nonparametric and semiparametric modeling*. Professor Fan has made fundamental contributions to multivariate nonparametric and semiparametric modeling. An important work in this direction was by him and his collaborators on the sophisticated generalized partial-linear single-index model frequently used in econometrics and biostatistics. In a seminal paper, Professor Fan discovered surprisingly the estimability of the additive components in additive models. He has also contributed fundamentally to the development of the varying-coefficient models. (d) *Generalized likelihood ratio theory*. Professor Fan was instrumental in developing the generalized likelihood ratio test for nonparametric inference. He demonstrated surprisingly that the generalized Wilks phenomenon continues to hold for many nonparametric testing problems under very different statistical contexts. Since his path-breaking work, there have been numerous follow-ups in both econometrics and statistics. (e) *Nonlinear time series*. Professor Fan was among the first to introduce contemporary techniques to solve nonlinear time series problems. Before his contributions, the area was generally perceived as theoretical exercises. His work in this area is exemplified in his highly regarded book on *Nonlinear Time Series* (2003). Professor Fan's many contributions have significantly changed the outlook on nonlinear time series modeling. (f) *Wavelet*. Professor Fan was the first to successfully apply wavelets transform to compare signals of high-dimension, and to conduct efficient nonparametric inference, which is the theoretical foundation for high-dimensional classification. He introduced a novel approach to overcome the drawbacks of commonly used wavelet approaches, and demonstrated surprisingly that all penalized least-squares possess the optimal properties for the models with heteroskedasticity, non-equispaced designs, and non-normal noises.
- **High- and Ultrahigh-dimensional statistical learning**. Professor Fan is a leading expert in high- and ultrahigh-dimensional statistical learning, a rapidly developing research area. He advocated a family of folded concave penalized likelihood procedures to ensure model stability and asymptotic unbiasedness, and demonstrated the oracle properties of his proposal. His pioneering work

in 2001 laid out a solid foundation for modern statistical learning and has stimulated a surged interest in the field with nearly 10,000 citations. In an another seminal paper read before the Royal Statistical Society in 2008, Professor Fan introduced the concept of sure independence screening to reduce dimensionality. He demonstrated that one can safely reduce the number of variables from exponentially high $p = \exp(na)$ to $d = o(n)$, where n is the sample size. Therefore, the ultrahigh dimensional variable selections become moderate-scale ones so that the penalized likelihood methods can be used with expeditious and efficient computation. In addition, he pioneered the research on endogeneity, heterogeneity, heavy tailedness, robustness, missingness, dependence among covariates, dependence adjustments and augmentations that are permeated in big data applications, as summarized in his paper on "Challenges in Big Data Analysis" published in 2014. In addition, he addressed critically large-scale multiple testing, covariance regularization, graphical models, high-dimensional inference, distributed learning, and over-parametrization issues. He gave an in-depth theoretical analysis of deep Q-learning, the algorithm that powered many modern successful AIs using neural networks, quantified the stability of neural networks, and unlocked the power of structured neural networks for high-dimensional big data modeling.

- **Spectral Methods for Data Science**. Professor Fan gave an in-depth analysis and revealed critical insights on the empirical eigenstructure for high dimensional spiked covariance. He was among the first to unveil an entrywise eigenvector perturbation bound, and applied it successfully to the analysis of high-dimensional robust covariance estimation. He pioneered the first and even second order expansions on the singular vectors that facilitate uncertainty quantification in modern data science. In addition, he gave both in-depth analyses of algorithmic and statistical properties for various challenges in matrix completion, community detection, item ranking, synchronization, and blind deconvolution as summarized in the monograph on *Spectral Methods for Data Science* (2021). He pioneered the inference of the membership profiles in network models and preference scores in item ranking, as well as their associated uncertainty quantification on the estimated ranks.
- **Financial Econometrics**. Professor Fan has made many fundamental contributions to financial econometrics, from time-dependent modeling of financial data and inferences of stochastic diffusion models to estimating and managing market risks, pricing financial derivatives, and managing large portfolios. He was among the first to introduce nonparametric techniques to diffusion models, to develop innovative statistical techniques to answer the question of whether the interest rate and/or stock dynamics follow certain famous econometric models, whether the dynamic is time-dependent or Markovian, to formally characterize the advantage of factor model in large covariance matrix estimation, to introduce time-dependent nonparametric models for asset pricing and risk management, and to introduce portfolio optimization with gross exposure constraints. His work provided important insights into financial puzzles concerning why incorrect constraints such as the optimal no-short-sale portfolio outperform the Markowitz optimal portfolio, why the leverage effects disappear in high-frequency finance, and how to robustly measure earning surprises when some analysts are possibly biased.
- **Biostatistics and Bioinformatics**. Professor Fan was among the first to use nonparametric methods to analyze censored regression, to establish the sampling properties of nonparametric Cox regression, and to systematically develop penalized likelihood inferences for variable selection in the analysis of univariate and multivariate survival data. His many fundamental contributions to functional and longitudinal data are exemplified by his work on functional ANOVA, varying-coefficient model, semiparametric modeling, and semiparametric estimation of the covariance matrix. Teamed up with scientists worldwide, Professor Fan's scientific "vocabulary" keeps steadily expanding. He has made many fundamental contributions to computational biology, as demonstrated in his interdisciplinary papers on the normalization and significant analysis of DNA microarray, including a JASA discussion paper where a high-dimensional semiparametric model was used for understanding the normalization of microarray data, and his recent contributions on the analysis of scRNA-seq data.

Service and Leadership. Professor Fan has an impressive record of public service and leadership in the profession and the community at large.

- **Professional Service**. Professor Fan was the President of the Institute of Mathematical Statistics (IMS) from 2006 to 2009 (President-Elect, President, and Past President). One of the most important aspects of his accomplishments as the IMS President was the reach out of IMS to many other societies in the world. This was largely due to his very broad research scope and vision that statistics as a discipline can only grow when it is partnered with other research communities, and that the field can be further energized when statisticians all around the world see the values and engage with each other. His accomplishment in this area is reflected in the massive increased list of conferences and workshops around the world sponsored by IMS in the IMS Bulletin. IMS-China was launched in 2008 and the IMS-Asia Pacific Rim (APR) Conference Series was established to promote statistical activities in the APR region in 2009. In addition, he proposed and chaired an ad-hoc committee on the creation of the *The Annals of Applied Statistics* in 2005, and launched the journal during his presidency with Professor Bradley Efron as the inaugural editor-in-chief. With the same energy and vision, as the President of the International Chinese Statistician Association from 2008 to 2010, Professor Fan transformed ICSA into a more encompassing and diverse society that builds much stronger ties with ASA, IMS, and other statistical and professional societies and bridges across academia, industry, and government.
- **Editorial Service**. Professor Jianqing Fan is currently a joint editor for the *Journal of American Statistical Association* (2023–2026), and an associate editor for *Management Science* (2018–). He was the co-editor(-in-chief) of the *Annals of Statistics* (2004–2006), an editor of *Probability Theory and Related Fields* (2003-2005), co-editor of *Econometrics Journal* (2007–2012) and *Journal of Econometrics* (2012–2018; managing editor 2014–2018), and a joint editor of *Journal of Business & Economic Statistics* (2018–2021). He has also served on the editorial boards of a number of journals, including *Journal of the American Statistical Association* (1996–2017), *Econometrica* (2010–2013), *Annals of Statistics* (1998–2003), *Statistica Sinica* (1996–2002), and *Journal of Financial Econometrics* (2009–2012).

- **Advising and Mentoring**. Professor Fan has attracted and trained a remarkable number of statisticians and data scientists, and has done outstanding services to the scientific community at large as well as to the statistical society in particular through his leadership roles in professional societies, editorial positions, study panels, and international exchanges. Professor Fan has nurtured many young scholars and professionals to thrive in our profession and industry at large. Over the years, in addition to supervising numerous master theses and senior theses at Princeton University, he has officially supervised more than 100 Ph.D. students and postdocs, who have continued to become productive academics or industry professionals, and some of them are already well recognized leaders in their respective fields. Moreover, Professor Fan is extremely generous with his time with young scholars other than his own students; it will be a very long list if we wanted to name all of whom have benefited from his wisdom and encouragement.

2. The theme issue

As a colleague and trainees of Professor Jianqing Fan, we organized this themed issue to honor the scope of his research on the occasion of his 60th birthday. This themed issue collects 27 papers on high-dimensional data science with applications to economics and finance. These papers can be organized along five primary areas: (i) factor models and time series data, (ii) functional and high-frequency data, (iii) portfolio selection and credit risk, (iv) many covariates and robust methods, and (v) network data.

2.1. Factor models and times series data

Barigozzi et al. (2023) studies the theoretical foundations of the Generalized Dynamic Factor Model (GDFM), a crucial tool for analyzing time series data in economics and finance. This paper offers a comprehensive asymptotic distribution theory for GDFMs in high dimensional setting, bridging the gap in the theoretical analysis of these models and enabling accurate inference. Yu et al. (2023) focuses on testing multi-factor asset pricing models in potentially high-dimensional settings, where traditional quadratic-form tests can perform poorly against the sparse alternative hypothesis. The authors propose a new power-enhanced testing procedure and show formally that the new power-enhanced test retains the desired nominal significance level and achieves asymptotically consistent power against more general alternatives.

Wan et al. (2023) studies model specification in the context of latent factor model estimation. The authors propose a hybrid between manual domain knowledge and automatic multivariate analysis, such as principal component analysis to select observed covariates as factor proxies, allowing the number of factor proxies and the number of factors to diverge with the sample size. The proposed method relies on a penalized reduced rank regression to combine information, which is robust to heavy-tailed data. The authors study the theoretical properties of the new proposed estimation methods and provide extensive simulations. Zhu and Müller (2023) introduces a new class of autoregressive models for spherical time series, where the dimension of the spheres on which the observations of the time series are situated may be finite-dimensional or infinite-dimensional. The challenge in modeling such time series lies in the intrinsic non-linearity of spheres and Hilbert spheres, where conventional arithmetic operations are no longer available. To address this problem, this paper considers rotation operators to map observations on the sphere.

2.2. Functional and high-frequency data

Hao et al. (2023) considers functional longitudinal data and studies a time-dynamic model with interactions, focusing on the derivatives of these processes. Starting with a concurrent linear model, the paper introduces methods to estimate regression coefficients for irregular, possibly noisy data. Chang et al. (2023) also considers high-dimensional functional time series where functional variables are more than serially dependent observations, containing both dynamic and white noise components. A three-step method is proposed: dimension reduction based on autocovariance, a new block-regularized minimum distance estimation framework for block sparse estimates, and then the final functional sparse estimates are derived.

Aït-Sahalia and Sağlam (2023) proposes a model where a strategic high-frequency market maker exploits his speed and informational advantages to place quotes that interact with the market orders of low-frequency traders. The model has testable implications regarding the impact of speed on the provision of liquidity. Researchers analyze high-frequency data to better understand jumps, the sharp discontinuities in asset prices, and carrying important information, especially for financial crises. Ding et al. (2023) proposes a degree-corrected block model with dependent multivariate Poisson edges to study stock co-jump dependence and propose an algorithm to estimate the community structure. Stochastic frontier analysis is regularly used in empirical studies to evaluate the productivity and efficiency of companies. A typical stochastic frontier model involves a parametric frontier subject to a composite error term consisting of an inefficiency and a random error. Cheng et al. (2022) develops new tests for the specification of distribution of the inefficiency. The authors show that the new tests are asymptotic distribution-free tests with neither parametric frontier nor homoscedasticity assumptions, in both cross-sectional and panel settings, and further propose a novel bootstrap to carry out the new tests in practice.

Chen et al. (2023) focuses on developing a novel regression methodology tailored for high-frequency financial data, addressing both fixed and increasing dimensions while accommodating microstructure noise and asynchronous observations. The primary contribution lies in presenting a robust methodology for estimating spot beta in the realm of high-frequency econometrics, tackling challenges posed by noise and asynchronicity in observations. Chang et al. (2022) studies the estimation of the covariance matrix for the high-dimensional noise in high-frequency data, emphasizing the importance of understanding noise properties and their applications in various fields like finance and functional data analysis. The main contribution lies in proposing a novel approach that addresses challenges like latency, asynchronicity, and serial dependence, providing theoretical insights and optimal convergence rates for the covariance matrix estimation in this specific context.

2.3. Portfolio selection and credit risk

Wang et al. (2023) focuses on volatility forecasting, which is key for risk management and portfolio construction. The challenge lies in devising a robust proxy for the true volatility. The paper demonstrates that the accuracy of volatility predictions depends on the proxy's deviation from the true volatility. The paper presents deviation bounds for three robust volatility proxies, two from clipped data, and one from exponentially weighted Huber loss minimization.

The Mean–variance (MV) portfolio theory proposed plays an important role in finance. This theory enjoys many appealing properties and it has an analytic solution depending on only two inputs, the mean of assets' excess return of the portfolios and the inverse of the covariance matrix of such returns. In the era of big data, the implementation of the MV portfolio theory encounters some challenges due to (a) heavy-tail distributed assets, (b) the presence of outliers in the observed data, and (c) the high dimensionality of assets. This calls for new statistical methods and theory for MV portfolio theory and numerical optimization algorithms for MV portfolio selection in practical implementation. Petukhina et al. (2023) tackles these challenges by developing a new strategy for robustifying the weights in the MV portfolio optimization. The authors first study the impact of the number of assets on the estimated optimal weights when the ratio of the number of assets to the sample size tends to a constant within $(0, 1)$. In robustifying the weights this work presents a toolbox for stabilizing costs and weights for global minimum Markowitz portfolios. Utilizing a projected gradient descent technique, the new strategy avoids the estimation and inversion of the covariance operator as a whole and concentrates on robust estimation of the gradient descent increment. Cai et al. (2022) develops an asset-splitting algorithm for ultrahigh-dimensional portfolio selection based on the MV portfolio theory. The authors further study the theoretical property of their proposed algorithm. Fan et al. (2022) studies the MV portfolio selection with time-varying variance. The authors model the covariance structure using a factor model with time-varying factor loadings, which facilitates the capture of the dynamics of the covariance structure of asset returns, and hence, the optimal investment strategy in a dynamic setting.

Credit risk analysis also plays a vital role in the era of digital finance and it is one of the primary interests to identify customers with similar types of risk categories so that personalized financial services can be offered accordingly. Pei et al. (2022) proposes a latent class Cox model for heterogeneous time-to-event data. The new model naturally extends the Cox proportional hazards model to flexibly take into account the heterogeneity of covariate effects as often manifested in real data. Motivated by the analysis of a data set on financial portfolio returns, Sun et al. (2022) proposes a multi-kink quantile regression (MKQR) model with latent homogeneous structure for panel data analysis. The MKQR model accounts for both homogeneity and heterogeneity among individuals and parameters in panel data analysis. The authors develop an estimation procedure for both the unknown parameters and the latent homogeneous structure in the proposed model and establish an asymptotic theory of the proposed estimators. They also address computational issues with the implementation of the estimation procedure.

2.4. Many covariates and robust methods

Guo et al. (2023) studies high-dimensional mediation models where both the outcome model and the mediator model are linear with high-dimensional mediators, and develops inference procedures in that context. The paper complements classical meditation inference methods, which are not applicable in high-dimensional settings. The authors propose methods based on partially penalized least squares for estimation and inference about indirect effects and establish their theoretical properties in large samples. Wald-type and F-type test statistics are considered. The new methods are shown to perform well in finite samples via simulations and are also deployed to study stock reaction to the arrival of the COVID-19 pandemic, focusing on mediation effects of financial metrics that bridge the company's sector and stock return.

Man et al. (2023) studies a penalized robust expectile regression estimation method in high-dimensional settings. It establishes large sample properties for the estimator in two alternative regimes: the low-dimensional regime where the dimension of the covariates is much smaller than the sample size, and the ultra-high-dimensional regime with sparsity. The authors pay special attention to some of the algorithmic features of the iteratively reweighted estimation method in the ultra-high-dimensional setting. Chen et al. (2024) addresses the testing challenges in high-dimensional quantile regression, a critical tool in econometrics, especially in the era of big data where traditional mean regressions might be inadequate. It introduces a novel test statistic based on the quantile regression score function, offering robustness to heavy-tailed and non-Gaussian data. Zhang et al. (2022a) proposes a new test for equal distributions of several high-dimensional samples in separable metric spaces, with its test statistic constructed based on maximum mean discrepancy. The authors also establish the asymptotic null and alternative distributions of the test statistic are established under some mild conditions, and illustrate the good performance of the new test via a real data example of Gini index curves.

Wei et al. (2023) focuses on inference for the most effective policy or treatment among many alternatives in a setting where potentially many covariates enter the analysis. The authors propose a resampling-based inference procedure that accounts for the winner's curse in evaluating the best policies observed in a random sample, while also being more robust to the presence of many included covariates in the model. Good finite sample performance of the proposed method is demonstrated via Monte Carlo experiments. Zhang et al. (2022b) offers a post-screening diagnostic study, focusing on regression analysis and with the aim of testing whether replacing the original ultrahigh dimensional covariates with a given number of linear combinations of them results in a loss of regression information. The main contribution lies in proposing a novel two-stage testing procedure based on sufficient dimension reduction methods, addressing challenges associated with ultrahigh-dimensional data using the post-screening inference idea. Zhou and Zou (2023) introduces a high-dimensional regression method using a nonparametric Box–Cox model with an unspecified monotone transformation in two steps. Initially, a new technique, composite probit regression (CPR), is proposed, utilizing folded concave penalized CPR for parameter estimation. This estimator enjoys a strong oracle property without knowing the nonparametric function. Then, a univariate monotone regression is used to estimate the function. The computation uses a coordinate-majorization-descent algorithm and is efficient. Liu et al. (2022) introduces a new method called the Generalized Knockoff Procedure (GKnockoff) for controlling the false discovery rate (FDR) in structural change detection, a critical aspect of various fields including finance and genomics. GKnockoff demonstrates superior performance in FDR control and power compared to existing methods, addressing the need for reliable signal detection and controlling false discoveries in structural change detections.

2.5. Network data

Jin et al. (2023) considers network data and focuses on understanding the network community structure. The authors consider the Degree-Corrected Mixed-Membership (DCMM) model, which allows for more flexibility in both heterogeneity and mixed-memberships and is more realistic and broader than the classical stochastic block model (SBM). A fast algorithm for estimating the mixed membership vectors of all nodes and other parameters of DCMM is then discussed. Empirical illustrations using data from a political blog network, trade networks, a co-authorship network, and a citee network, are given to demonstrate the wide applicability of the proposed methods. Wu et al. (2023) introduces a bipartite network influence model (BNIM) to assess nodal heterogeneity from an influence standpoint. The model is made estimable by parameterizing influence indices using nodal attributes and a predefined link function, with a quasi-maximum likelihood approach for parameter estimation. Score tests are proposed to check nodal influence heterogeneity, and a quasi-likelihood ratio test is used to evaluate the link function's adequacy, with asymptotic properties established under set conditions.

References

Aït-Sahalia, Y., Sağlam, M., 2023. High frequency market making: The role of speed. J. Econometrics 105421.
Barigozzi, M., Hallin, M., Luciani, M., Zaffaroni, P., 2023. Inferential theory for generalized dynamic factor models. J. Econometrics 105422.
Cai, Z., Li, C., Wen, J., Yang, S., 2022. Asset splitting algorithm for ultrahigh dimensional portfolio selection and its theoretical property. J. Econometrics 105291.
Chang, J., Chen, C., Qiao, X., Yao, Q., 2023. An autocovariance-based learning framework for high-dimensional functional time series. J. Econometrics 105385.
Chang, J., Hu, Q., Liu, C., Tang, C.Y., 2022. Optimal covariance matrix estimation for high-dimensional noise in high-frequency data. J. Econometrics 105329.
Chen, Z., Cheng, V.X., Liu, X., 2024. Hypothesis testing on high dimensional quantile regression. J. Econometrics 238 (1), 105525.
Chen, D., Mykland, P.A., Zhang, L., 2023. Realized regression with asynchronous and noisy high frequency and high dimensional data. J. Econometrics 105446.
Cheng, M.-Y., Wang, S., Xia, L., Zhang, X., 2022. Testing specification of distribution in stochastic frontier analysis. J. Econometrics 105280.
Ding, Y., Li, Y., Liu, G., Zheng, X., 2023. Stock co-jump networks. J. Econometrics 105420.
Fan, Q., Wu, R., Yang, Y., Zhong, W., 2022. Time-varying minimum variance portfolio. J. Econometrics 105339.
Guo, X., Li, R., Liu, J., Zeng, M., 2023. Statistical inference for linear mediation models with high-dimensional mediators and application to studying stock reaction to COVID-19 pandemic. J. Econometrics 235 (1), 166–179.
Hao, S., Lin, S.-C., Wang, J.-L., Zhong, Q., 2023. Dynamic modeling for multivariate functional and longitudinal data. J. Econometrics 105573.
Jin, J., Ke, Z.T., Luo, S., 2023. Mixed membership estimation for social networks. J. Econometrics 105369.
Liu, J., Sun, A., Ke, Y., 2022. A generalized knockoff procedure for FDR control in structural change detection. J. Econometrics 105331.
Man, R., Tan, K.M., Wang, Z., Zhou, W.-X., 2023. Retire: Robust expectile regression in high dimensions. J. Econometrics 105459.
Pei, Y., Peng, H., Xu, J., 2022. A latent class cox model for heterogeneous time-to-event data. J. Econometrics 105351.
Petukhina, A., Klochkov, Y., Härdle, W.K., Zhivotovskiy, N., 2023. Robustifying Markowitz. J. Econometrics 105387.
Sun, Y., Wan, C., Zhang, W., Zhong, W., 2022. A multi-kink quantile regression model with common structure for panel data analysis. J. Econometrics 105304.
Wan, R., Li, Y., Lu, W., Song, R., 2023. Mining the factor zoo: Estimation of latent factor models with sufficient proxies. J. Econometrics 105386.
Wang, W., An, R., Zhu, Z., 2023. Volatility prediction comparison via robust volatility proxies: An empirical deviation perspective. J. Econometrics.
Wei, W., Zhou, Y., Zheng, Z., Wang, J., 2023. Inference on the best policies with many covariates. J. Econometrics 105460.
Wu, Y., Lan, W., Fang, K., Fan, X., 2023. Bipartite network influence analysis for a two-mode network. J. Econometrics.
Yu, X., Yao, J., Xue, L., 2023. Power enhancement for testing multi-factor asset pricing models via Fisher's method. J. Econometrics.
Zhang, J.-T., Guo, J., Zhou, B., 2022a. Testing equality of several distributions in separable metric spaces: A maximum mean discrepancy based approach. J. Econometrics 105286.
Zhang, Y., Zhou, Y., Zhu, L., 2022b. A post-screening diagnostic study for ultrahigh dimensional data. J. Econometrics 105354.
Zhou, H., Zou, H., 2023. The nonparametric Box–Cox model for high-dimensional regression analysis. J. Econometrics 105419.
Zhu, C., Müller, H.-G., 2023. Spherical autoregressive models, with application to distributional and compositional time series. J. Econometrics 105389.

Matias D. Cattaneo *
Department of Operations Research and Financial Engineering, Princeton University, United States of America
E-mail address: cattaneo@princeton.edu.

Yingying Fan
Data Sciences and Operations Department, Marshall School of Business, University of Southern California, United States of America

Runze Li
Department of Statistics, The Pennsylvania State University at University Park, United States of America

Rui Song[1]
Amazon, Inc., United States of America

* Corresponding editor.

[1] This work does not relate to the author's position at Amazon.