ADVANCES IN ECONOMETRICS

Series Editors: Thomas B. Fomby,
R. Carter Hill, Ivan Jeliazkov,
Juan Carlos Escanciano,
Eric Hillebrand, David Jacho-Chávez,
Daniel L. Millimet, Rodney Strachan

Recent Volumes:

ADVANCES IN ECONOMETRICS VOLUME 36

ESSAYS IN HONOR OF AMAN ULLAH

EDITED BY

GLORIA GONZÁLEZ-RIVERA
Department of Economics, University of California Riverside, Riverside, CA, USA

R. CARTER HILL
Department of Economics, Louisiana State University, Baton Rouge, LA, USA

TAE-HWY LEE
Department of Economics, University of California Riverside, Riverside, CA, USA

Emerald

United Kingdom – North America – Japan
India – Malaysia – China

Ximing Wu — Department of Agricultural Economics, Texas A&M University, College Station, TX, USA

Xin Yun — Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, People's Republic of China

Yonghui Zhang — School of Economics, Renmin University, Beijing, People's Republic of China

Yu Yvette Zhang — Department of Agricultural Economics, Texas A&M University, College Station, TX, USA

Yifeng Zhu — Department of Economics, Emory University, Atlanta, GA, USA

Victoria Zinde-Walsh — Department of Economics, McGill University, Montréal, QC, Canada

INTRODUCTION

The 36th volume of *Advances in Econometrics* in honor of Aman Ullah is a collection of 19 articles contributed by his former students, colleagues, and academic friends, covering a wide range of issues in econometrics. The volume mostly draws upon the works presented at the conference that took place at the historic Mission Inn in Riverside California on March 13–15, 2015.

The volume is organized into six parts. Part I. Tribute (one article) includes a review that pays tribute to the professional career of Aman Ullah. The remaining 18 articles of the volume are organized into five parts corresponding to five general econometrics topics to which Ullah has contributed extensively over the course of his career. These are: Part II. Panel Data Models (four articles), Part III. Finite Sample Econometrics (two articles), Part IV. Information and Entropy (three articles), Part V. Issues in Econometric Theory (three articles), and Part VI. Nonparametric and Semiparametric Methods (six articles). We follow a similar ordering of these six parts to that of the conference program. In each part, the order of articles is alphabetical according to the first author's last name.

"PART I. TRIBUTE" CONSISTS OF A REVIEW ARTICLE

The article "A Selective Review of Aman Ullah's Contributions to Econometrics" by Yong Bao, Yanqin Fan, Liangjun Su, and Victoria Zinde-Walsh provides a historical account of Aman Ullah's contributions to a wide spectrum of econometrics. The main specialization of each of the four author guides his/her own account on how Ullah has contributed to robust inference, finite sample theory, nonparametrics and semiparametrics, and panel and spatial models. On robust inference, the article reviews Aman's early works with his thesis advisor Professor Nagar that go back to the early 1970s, then discusses the double-k class estimators in the late 1970s, and his continuous effort in pursuing robustness of statistics with respect to distributional assumptions in the 1980s. The second section reviews Aman's contributions to finite sample theory in econometrics and his approaches to addressing the exact theory and approximations for a wide range of statistics in various econometrics and statistics models. The third section presents a brief review of some of Aman's works in several sub-areas: nonparametric regression and its derivatives, consistent model specification testing, nonparametric IV and simultaneous equations models, and combined estimation or model averaging. The last section discusses Aman's contributions to panel and spatial models such as seemingly unrelated regressions, parametric, semiparametric and nonparametric panel data models and spatial models.

"PART II. PANEL DATA MODELS" CONSISTS OF THE FOLLOWING FOUR ARTICLES

In "Semiparametric Estimation of Partially Linear Varying Coefficient Panel Data Models," the authors Yonghong An, Cheng Hsiao, and Dong Li estimate a partially linear varying coefficient fixed effect panel data model that consists of a parametric linear part and a nonparametric varying coefficient part. By estimating the model with the series method, the authors show that the parametric component can be estimated at the $\sqrt{N}$-rate and the nonparametric component can be consistently estimated at the standard nonparametric rate. They extend the model to allow endogeneity in the parametric component and establish the asymptotic properties of the semiparametric instrumental variable estimators.

In "Testing for Spatial Lag and Spatial Error Dependence in a Fixed Effects Panel Data Model Using Double Length Artificial Regressions," Badi H. Baltagi and Long Liu consider a battery of tests in a model where the fixed effects are first eliminated using the method in Lee and Yu (2010). The derivations of the tests are then similar to those derived for a cross-section spatial regression model (Baltagi & Li, 2001). This article revisits the joint and conditional Lagrange Multiplier tests derived by Debarsy and Ertur (2010) for a fixed effects spatial lag regression model with spatial auto-regressive error, and derives these tests using artificial double length regressions. These tests and their corresponding LM tests are compared using an empirical example and a Monte Carlo simulation.

Alexander Chudik, Kamiar Mohaddes, M. Hashem Pesaran, and Mehdi Raissi contribute the article "Long-Run Effects in Large Heterogeneous Panel Data Models with Cross-Sectionally Correlated Errors." They develop

a cross-sectionally augmented distributed lag (CS-DL) approach to the estimation of long-run effects in large dynamic heterogeneous panel data models with cross-sectionally dependent errors, as a complementary method to cross-sectionally augmented ARDL specifications. The asymptotic distribution of the CS-DL estimator is derived under coefficient heterogeneity in the case where the time dimension (T) and the cross-section dimension (N) are both large. The CS-DL approach is compared with more standard panel data estimators that are based on autoregressive distributed lag (ARDL) specifications. It is shown that unlike the ARDL-type estimator, the CS-DL estimator is robust to misspecification of dynamics and error serial correlation. The theoretical results are illustrated with small sample evidence obtained by means of Monte Carlo simulations, which suggest that the performance of the CS-DL approach is often superior to the alternative panel ARDL estimates, particularly when T is not too large and lies in the range of 30–50.

In "Semiparametric Estimation of Partially Linear Dynamic Panel Data Models with Fixed Effects," Liangjun Su and Yonghui Zhang consider a semiparametric model where the linear part includes either exogenous or endogenous variables or both, and the nonparametric part includes the lagged-dependent variable and some other exogenous variables. Two types of estimation methods are proposed for the first-differenced model. The asymptotic properties of the estimators are investigated when N tends to infinity and the time period T is fixed. Moreover, they propose a specification test for the linearity of the nonparametric component. The performance of the proposed estimators is evaluated via Monte Carlo simulations. The proposed semiparametric estimator is applied to model the relationship between economic growth, intellectual property right protection, and initial economic conditions.

"PART III. FINITE SAMPLE ECONOMETRICS" CONSISTS OF THE FOLLOWING TWO ARTICLES

In "Finite-Sample Bias of the Conditional Gaussian Maximum Likelihood Estimator in ARMA Models," Yong Bao derives the finite-sample bias of the conditional Gaussian maximum likelihood estimator in ARMA models when the error follows some unknown non-normal distribution. The general procedure relies on writing down the score function and its higher-order derivative matrices in terms of quadratic forms in the non-normal error vector. Evaluation of the bias can then be straightforwardly conducted. For the special case of ARMA(1,1), the bias results are further simplified and compared with the existing results in Tanaka (1984). Simulations results confirm that the new bias formula works well in approximating the true finite-sample bias of the Gaussian maximum likelihood estimator under both normal and non-normal distributions.

In "Finite Sample Bias Corrected IV Estimation for Weak and Many Instruments," Matthew Harding, Jerry Hausman and Christopher J. Palmer explore the use of double k-class estimators in scenarios involving weak and/or many instruments and show that there is a scope for gain. They consider the finite sample distribution of the 2SLS estimator and derive bounds on its exact bias in the presence of weak and/or many instruments. They contrast the behavior of the exact bias expressions and the asymptotic expansions popular in the literature, including a consideration of the no-moment problem exhibited by many Nagar-type estimators. After deriving a finite sample unbiased k-class estimator, they introduce two new IV estimators, both in the double k-class estimator. These estimators require choosing the k_2 parameter in a data-driven way. For one of the estimators, k_2 is chosen to correct for the 2SLS bias, and for the other, k_2 is such that minimizes the MSE. Monte Carlo simulations show that these new estimators outperform 2SLS and the Fuller estimators in terms of mean bias and MSE.

"PART IV. INFORMATION AND ENTROPY" CONTAINS THE FOLLOWING THREE ARTICLES

Although in principle prior information can significantly improve inference, incorporating incorrect prior information will bias the estimates of any inferential analysis. This fact deters many scientists from incorporating prior information into their inferential analyses. In the article "On the Construction of Prior Information – An Info-Metrics Approach," Amos Golan and Robin L. Lumsdaine discuss three different, yet complimentary, approaches for constructing prior information with an application to the social sciences where prior information is often hard to come by and very hard to justify or validate. Within an info-metrics framework, they borrow concepts and philosophical reasoning from the natural sciences, where experiments are more regularly conducted and prior information is often used in inferential analysis, despite being nontrivial to specify what that information is and how to quantify it.

In "The Wage Premium of Naturalized Citizenship," Esfandiar Maasoumi and Yifeng Zhu analyze the earning gap between naturalized citizens and noncitizens. Over many years, they find a positive gap characterized by a tent shape across the wage distribution. They focus on a normalized entropy measure of the gap between distributions, and they compare it with conventional measures at the mean, median, and other quantiles. In addition, naturalized citizen earnings (at least) second-order stochastically dominate non-citizen earnings in many of the recent years. They construct two counterfactual distributions to further examine the potential sources of the earning gap, the wage structure effect and the composition effect. Both of these sources contribute to the gap, but the composition effect, while diminishing somewhat after 2005, accounts for about 3/4 of the gap. The unconditional quantile regression (based on the Recentered Influence Function), and conditional quantile regressions confirm that naturalized citizens have generally higher wages, although the gap varies for different income groups in many years.

Built on the earlier contributions by Geweke (1982), Gourieroux, Monfort and Renault (1987), and Schreiber (2000), who measure causality in the context of Markov processes, Eric Renault and Daniela Scida, in their article "Causality and Markovianity: Information Theoretic Measures," reconsider the problem of causality measurement using the Kullback measure of divergence of a casual model from a noncausal counterpart. They specially focus on the design of measures that disentangle the measurement of markovianity and the measurement of Granger causality. This disentangling leads to revisit the equivalence between the Sims and Granger concepts of non-causality and the Log-Likelihood Ratio tests for each of them. The article draws attention to an important point neglected in the literature; namely, that for finite-order Markovian processes, Granger causality implies testing for non-nested hypotheses.

"PART V. ISSUES IN ECONOMETRIC THEORY" CONSISTS OF THE FOLLOWING THREE ARTICLES

In "A Likelihood-Free Reverse Sampler of the Posterior Distribution," Jean-Jacques Forneron and Serena Ng examine properties of an optimization-based sampler that employs auxiliary simulation, summary statistics, and resampling to approximate a posterior distribution for which inference may be hindered by the intractability of the likelihood function. This article extends Forneron and Ng (2014), which focused on the case of exact identification, to the overidentified case. It further provides additional insights into the reverse sampler, which is shown to be an importance sampler when the importance ratio is equal to the determinant or volume of the Jacobian matrix of the importance density transformation. By a change of variable argument, the output is a number of draws from the desired posterior distribution. Optimization always results in acceptable draws. Hence, when the minimum distance problem is not too difficult to solve, combining importance sampling with optimization can be much faster than the method of Approximate Bayesian Computation that by-passes optimization.

In "A Vector Autoregressive Moving Average Model for Interval-Valued Time Series Data," Ai Han, Yongmiao Hong, Shouyang Wang, and Xin Yun propose an Interval-valued Vector Autoregressive Moving Average (IVARMA) model that captures cross-dependence dynamics among multivariate interval-valued time series. It is an important extension to the traditional point-valued VARMA model and univariate autoregressive conditional interval model. They develop a minimum distance estimation method to estimate the parameters of an IVARMA model, and establish consistency, asymptotic normality, and asymptotic efficiency of the proposed estimator. A two-stage minimum distance estimator is shown to be asymptotically most efficient among the class of minimum-distance estimators. Simulation studies showcase the superiority of their estimators in finite sample performance.

In "Inference in Near-Singular Regression," Peter C. B. Phillips considers stationary time series regression models with near-collinear regressors and introduces new normalization techniques to develop the asymptotic theory for regression estimates and test statistics in cases where the signal matrix is nearly singular in finite samples and is asymptotically degenerate. When the regressors are exogenous, he develops the limit theory for the OLS estimator and test statistics. When the regressors are endogenous, he derives the limit theory for the corresponding 2SLS estimator, Wald test statistic, and some overidentification test. It is shown that near-singular designs of the type considered in the article are not completely fatal to least squares inference, but will inevitably involve size distortion except in special Gaussian cases. In the endogenous case, IV estimation is inconsistent and both the block Wald test and the Sargan overidentification test are conservative, biasing these tests in favor of the null.

"PART VI. NONPARAMETRIC AND SEMIPARAMETRIC METHODS" CONSISTS OF THE FOLLOWING SIX ARTICLES

In "Multivariate Local Polynomial Estimators: Uniform Boundary Properties and Asymptotic Linear Representation," Yanqin Fan and Emmanuel Guerre study the asymptotic bias and variance of a general class of multivariate local polynomial estimators of M-regression functions, which include the conditional mean and the conditional quantile functions as special cases. The focus is on the asymptotic bias and variance properties on the boundary and over the whole compact support of the multivariate covariate under a minimal assumption on the support. The support assumption ensures that the vicinity of the support boundary will be visited by the multivariate covariate. They show that, like in the univariate case, multivariate local polynomial estimators have good bias and variance properties near the boundary. They also establish asymptotic normality near the boundary and the optimal uniform convergence rate over the whole support. They obtain a uniform linearization result for the local polynomial quantile regression, and demonstrate that, with the uniformity results, the trimming practice in the local polynomial regression to avoid the boundary effect is unnecessary.

In "Model Averaging over Nonparametric Estimators," Daniel Henderson and Christopher F. Parmeter consider model-averaging methods for nonparametric regression. They combine up to eight nonparametric estimators of the mean function that differ in terms of their choices of kernel (Gaussian or Epanechnikov), bandwidth (least squares cross validation or AIC), and polynomial order (local constant or local linear). The authors propose two averaging schemes based on least-squares weights and cross-validation weights. They provide Monte Carlo evidence showing that model averaging based on either scheme nearly always results in more efficient estimators of the mean function than estimation based on any given choice of kernel, bandwidth and polynomial order. They also provide an empirical application of the proposed estimators.

In "Smoothness: Bias and Efficiency of Nonparametric Kernel Estimators," Yulia Kotlyarova, Marcia M. A. Schafgans, and Victoria Zinde-Walsh provide smoothness conditions for kernel-based estimators that ensure that the asymptotic rate at which the bias goes to zero is determined by the kernel order. In a finite sample, the leading term in the expansion of the bias may provide a poor approximation. The authors explore the relationship between smoothness and bias in a kernel estimator framework. They provide estimators for the degree of the smoothness and the bias and demonstrate that a linearly combined estimator constructed from different kernel functions and bandwidths can lead to estimators with reduced asymptotic mean-square error relative to the individual estimator at the optimal bandwidth. The higher-order kernels can reduce bias given some smoothness assumptions, but when these are violated non-standard results can arise. They examine the finite-sample performance of a combined estimator that minimizes the trace of the MSE of a linear combination of individual kernel estimators for a multimodal density. The combined estimator provides a robust alternative to individual estimators that protects against uncertainty about the degree of smoothness.

In "A Class of Nonparametric Density Derivative Estimators Based on Global Lipschitz Conditions," Kairat Mynbaev, Carlos Martins-Filho, and Aziza Aipenova show that the estimators for derivatives associated with a density function can be useful in identifying its modes and inflection points. In addition, these estimators play an important role in plug-in methods associated with bandwidth selection in nonparametric kernel density estimation. They extend the nonparametric class of density estimators proposed by Mynbaev and Martins-Filho (2010) to the estimation of *m*-order density derivatives when some smoothness conditions (the higher order Lipschitz conditions) on the density are imposed. Contrary to some existing derivative estimators, the proposed estimators have a full asymptotic characterization, including uniform consistency and asymptotic normality. An expression for the optimal bandwidth that minimizes an asymptotic approximation for the estimators' integrated squared error is provided. A Monte Carlo study sheds light on the finite sample performance of the estimators and contrasts it with that of density derivative estimators based on the classical Rosenblatt–Parzen approach.

Recent developments in shape-constrained nonparametric regression allow practitioners to impose constraints on local polynomial estimators thereby ensuring that the resulting estimates are consistent with underlying theory. In "Local Polynomial Derivative Estimation: Analytic or Taylor?," Jeffrey S. Racine considers the impact of using local polynomial Taylor-based derivative estimates for shape-constrained estimation, and cautions practitioners that these derivatives may be internally inconsistent with the estimated regression function. He points out that the analytic derivative estimator has the advantage to maintain internal consistency between the regression function and derivative estimates when there is a shape constraint but the Taylor-based derivative estimator enjoys a smaller mean-squared error. In such cases, the author suggests that practitioners might prefer to instead use analytic derivatives along the lines of those proposed in the local constant setting by Rilstone and Ullah (1989).

In "A Simple Consistent Nonparametric Estimator of the Lorenz Curve," Yu Yvette Zhang, Ximing Wu, and Qi Li propose a nonparametric estimator of the Lorenz curve that satisfies its theoretical properties, including monotonicity and convexity. They extend the integral transformation idea proposed by Ramsay (1998) to the estimation of the Lorenz curve, where a constrained problem is successfully transformed into an unconstrained one, which is estimated nonparametrically. The authors utilize splines to facilitate the numerical implementation of their estimator and to provide a parametric representation of the constructed Lorenz curve. They conduct Monte Carlo simulations to demonstrate the superior performance of the proposed estimator. They apply the method to estimate the Lorenz curve of the U.S. household income distribution and calculate the Gini index based on the estimated Lorenz curve.

Gloria González-Rivera
R. Carter Hill
Tae-Hwy Lee
Editors